Characterization and genomic Analysis of a novel *Pseudomonas* phage vB_PaeM_M1, representing a new viral family, *Psaeviridae*

Linyi Ren[1†], Yundan Liu[1†], Yantao Liang[1]*, Baohong Liu[2]*, Andrew McMinn[1,3], Kaiyang Zheng[1], Ziyue Wang[1], Hongmin Wang[1], Hongbing Shao[1,4], Yeong Yik Sung[4,5], Wen Jye Mok[4,5], Li Lian Wong[4,5] and Min Wang[1,4,6,7]*

[1]College of Marine Life Sciences, Institute of Evolution and Marine Biodiversity, Frontiers Science Center for Deep Ocean Multispheres and Earth System, Center for Ocean Carbon Neutrality, Ocean University of China, Qingdao, China, [2]Department of Hospital Infection Management, Qilu Hospital, Shandong University, Qingdao, China, [3]UMT-OUC Joint Centre for Marine Studies, Qingdao, China, [4]Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS, Australia, [5]Institute of Marine Biotechnology, Universiti Malaysia Terengganu, Kuala Terengganu, Malaysia, [6]The Affiliated Hospital of Qingdao University, Qingdao, China, [7]Haide College, Ocean University of China, Qingdao, China

Pseudomonas is a ubiquitous and ambiguous opportunistic pathogen, and plays an important ecological role in the ocean. Here, a new species, *Pseudomonas* phage vB_PaeM_M1, is described, which was isolated from the surface coastal waters of Qingdao, China. vB_PaeM_M1 contains a linear, double-stranded 89,387-bp genome with a GC content of 41.04% and encoding 184 putative open reading frames (ORFs). There were 50 conservative domains were predicted with BLASTp, including two auxiliary metabolic genes (Phosphate-inducible gene *phoH* and signal peptide peptidase A, *sppA*). Phylogenetic analysis of whole genome amino acid sequence and comparative genomic analysis showed that vB_PaeM_M1 has a distant evolutionary relationship with previously isolated viruses and can be grouped into a family-level novel viral cluster (VC_61) with eleven uncultured, assembled viral genomes, named as *Psaeviridae. Psaeviridae* contains two ORFs (ORFs 117 and 127), which were not detected in the genomes of other viral families, confirming the proposal for a new family. Combined with its ability to infect *Pseudomonas* and its representation of an unstudied viral family, vB_PaeM_M1 may be an important and novel model system for the study of interactions between viruses and host cells in marine ecosystems.

KEYWORDS

bacteriophage, *Pseudomonas*, genomic and comparative genomic analysis, phylogenetic analysis, *Psaeviridae*

Introduction

Viruses are the most abundant "life forms" in the ocean. They include bacteriophages (phages), archaeal viruses and eukaryotic viruses, which infect bacteria, archaea and eukaryotes, respectively (Carlton, 1999; Pourcel et al., 2017). By infecting and lysing their hosts, viruses play an important role in the regulation of marine microbial communities and global biogeochemical cycles (Weinbauer, 2004). Genomic and metagenomic analyses show that viruses are also the most abundant biological entities in diverse marine environments. Furthermore, they are not only widely distributed, but they also have very high genetic diversity (Breitbart et al., 2002). Some viruses can manipulate the genome of marine organisms through horizontal gene transfer and promote biological evolution (Suttle, 2007). Moreover, the expression of phage-encoded auxiliary metabolic genes (AMGs), such as *psbA* and *psbD* genes, play important roles in maintenance of photosynthetic energy production of host cells during infection (Bragg and Chisholm, 2008; Thompson et al., 2011). However, it is estimated that more than 90% of the metagenomic assembled viral populations remain unknown (Gregory et al., 2019).

Pseudomonas is a common conditional pathogen, which belongs to non-fermenting Gram-negative bacteria (Graham-Mize and Rosser, 2004). It is widely distributed in different environments such as water, air, soil, normal human skin, respiratory tracts and intestinal tracts (Alam et al., 2021). It is one of the main pathogens of nosocomial infection and has a very complex antimicrobial resistance mechanism, resulting in the existence of many antimicrobial resistance strains (Poole, 2011). This is attributable to various factors, including horizontal gene transfer and mutational changes in genes, making their treatment options very limited (Poole, 2011).

Although *Pseudomonas* has important ecological and medical significance, our understanding of the interaction between phages and *Pseudomonas* is still inadequate, especially considering their potential applications in phage therapy. As of November 2022, 739 complete *Pseudomonas* phage genomes have been submitted to GenBank. Compared with the vigorous development of metagenome research, there are few studies describing new marine phages (Yang et al., 2021). Therefore, the isolation and analysis of more *Pseudomonas* phages will lead to a better understanding of the impact of metabolic regulation of phages on the host community and the ways they can affect biogeochemical cycles, and will provide more candidates for the biological control of pathogenic *Pseudomonas*.

In this study a new *Pseudomonas* phage vB_PaeM_M1 was isolated from coastal waters off Qingdao, China. Through the genome, phylogeny and comparative genome analysis of vB_PaeM_M1, the *Pseudomonas* phage database was expanded. Due to its unique evolutionary relationship, vB_PaeM_M1 can be combined with eleven uncultured phage contigs identified from metagenomics to form a new family-level virus cluster (VC). This phage will provide important information for the further study of the relationship between phages and *Pseudomonas* in the ocean and the role of phages in marine ecosystems.

Materials and methods

Location and sampling

Surface seawater samples were collected from the waters near Qingdao in the Yellow Sea (120° 19 ′ 23 ″ E, 36° 4 ′ 4 ″ N). The 50 L water samples were filtered through a 20 μm filter to remove larger plankton and particles and then passed through 0.22 μm pore size (micropore) filter. Using tangential flow filtration (laboratory scale, 50 kDa), the seawater containing the virus was concentrated 500 times to obtain a 100 ml sample. The sample and the original seawater were stored in the dark at 4°C until the experiment was carried out (Wang et al., 2015).

Phage isolation and purification

After passing through 0.22 μm microporous membrane filters, to remove bacteria and phytoplankton, the bacteriophages were separated from the same seawater sample. Plaques were obtained by the double agar layer method (0.5% low melting point top agar, solid 2216E medium: agar 15 wt.%, semi-solid 2216E medium 7.5 wt.%, agar brand: Solarbio, Beijing) to check for the presence of phages; these were then cultured at a constant 26°C overnight. Each plaque was cored, purified three times and resuspended in 500 μL of sterile SM buffer [100 mM NaCl, 81.2 mM MgSO$_4$ 7H$_2$O, 50 mM Tris–HCl (pH 7.5), 0.01% gelatin]. The purified phage was then stored in SM buffer at 4°C for several months (Liu et al., 2019).

Bacterial strain isolation and identification

The host bacterial strains were isolated from unfiltered seawater samples by serial dilution, and then cultured in liquid ZoBell medium at 28°C (Duhaime et al., 2011; Liu et al., 2018). 16S rRNA gene was amplified by PCR for molecular identification with the following universal primers: 5′ 27f: AGAGTTTGATCMTGGCTCAG 3′ and 1492R: 5′ TACGGYT ACCTTGTTACGACTT 3′ (BGI tech solutions co. Ltd., Beijing Liuhe). By analyzing the results of 16S rRNA gene sequence analysis, BLAST search the 16S rRNA gene sequence to determine the type of strain. Genomic DNA was extracted and purified according to the method of Marmur, using a

commercial genomic DNA extraction kit (TIANGEN) (Marmur, 1961).

Morphology study by transmission electron microscope (TEM)

The purified phage samples were negatively stained with phosphotungstic acid (2%, wt/vol, pH 7.2). Images of purified particles of bacteriophage vB_PaeM_M1 were examined by TEM (jeol JEM-1200EX; jeol, Japan) 100 kV (Middelboe et al., 2010) at a magnification of ×400,000. The phage size was estimated from the electron micrographs (Li et al., 2016).

Extraction of phage DNA and bioinformatics analysis

Bacteriophage genomic DNA was extracted using an OMEGA viral DNA kit according to the manufacturer's instructions. Nucleic acids were detected by electrophoresis (Pires et al., 2011). Purified phage genomic DNA was sequenced using Illumina Miseq 2×300 paired-end sequence by the Novogene Company (Tianjin), the raw data were then filtered to get clean data and assembled using SOAP *de novo* v2.04 (Beilstein and Dreiseikelmann, 2006). The generated reads were filtered for adapter sequences and low-quality regions using Trimmomatic v0.36. This process removed Illumina adapter sequences, regions longer than 4 nucleotides and an average phred value below 15 and a minimum length of 36 bp. The termini were identified by PhageTerm v1.0.11 (Garneau et al., 2017). The reads with the maximum coverage were considered as phage termini (Supplemental Figure S1). Coding DNA sequences of the phage were predicted using RAST v2.0 (http://rast.nmpdr.org/), all open reading frames (ORFs) were annotated by BLASTp against the nonredundant proteins (NR) NCBI database (http://blast.ncbi.nlm.nih.gov) with E-value <1e-5 (Schneiker et al., 2006; Aziz et al., 2008). Genomes were searched for antimicrobial resistance genes (ARGs) using the CARD (Alcock et al., 2020) and ResFinder (Bortolaia et al., 2020) databases as a reference. Pfam search (https://pfam.xfam.org/search/sequence) with default parameters and hhpred search (https://toolkit.tuebingen.mpg.de/tools/hhpred) were using as online servers, the genome was searched by the CLC main workbench 20 (Hyman and Abedon, 2010; Duhaime et al., 2011; Mistry et al., 2021).

The phylogeny of *Pseudomonas* ZM01

The reference sequences of 50 *Pseudomonas* families with 16S rRNA genes most similar to M1 were selected from GenBank (Supplemental Table S1), including the host strain *Pseudomonas* ZM01, and aligned by mafft using G-INS-1 of strategy with 1000 iterations (Minh et al., 2020). Positions with

90% or more gaps were removed from the alignments with trimal (Capella-Gutiérrez et al., 2009). The phylogenic tree was calculated from multiple sequence alignments using IQ-tree2 (Letunic and Bork, 2019), applying GTR + F + R4 as the suggested DNA model with 1000 iterations of bootstrap. The tree was visualized by iTOL v4 (Bolger et al., 2014).

Phylogenetic analysis and comparative genomic analyses

Based on the whole genome amino acid sequences of bacteriophage vB_PaeM_M1 and *Pseudomonas* phages, the proteome tree was generated by ViPTree v3.1 (https://www.genome.jp/viptree) (Nishimura et al., 2017). tBLASTx and ViPTree were used for genome comparison to describe the relationship between bacteriophage vB_PaeM_M1 and its close relatives. All-vs-all BLASTp analysis used orthofinder v2.5.4 to compute the percentage of shared genes between phage vB_PaeM_M1 and all complete *Pseudomonas* phages genomes in the NCBI RefSeq database (Emms and Kelly, 2019). vConTACT 2.0 uses a guilt-by-contig-association classification on the International Committee on Taxonomy of Viruses (ICTV) taxonomy dataset to cluster and provide classification background for sequencing data. To further study vB_PaeM_M1 classification information, BLASTp was used to expand phage population (Teeling et al., 2004; Bolduc et al., 2017). In order to find homologous phages related to vB_PaeM_M1, each coding sequence of vB_PaeM_M1 was queried against the Integrated Microbial Genome/Virus (IMG/VR, v.3) database (E value, <1e-10; identity, >30; and alignment region covering >50%) (Paez-Espino et al., 2017; Paez-Espino et al., 2019; Roux et al., 2021). In order to find the conserved proteins of family A, BLASTp was used for comparison in the NR database (E value, <1e-10; identity, >30; and alignment region covering >50%). In vConTACT analysis, the selected sequence was compared with vB_PaeM_M1 as a group to obtain more accurate results (similar sequences were selected by Diamond, all of which satisfied the following parameters: E value, <1e-5; alignment region covering more than 50% of the shorter sequence; and identity >30%) (Zhan and Chen, 2019). Gephi visualizes the edge weighted model network based on the vConTACT analysis (Bastian et al., 2009). The ANI (average nucleotide identity) value of vB_PaeM_M1 was calculated by VIRIDIC, and the results were visualized by pheatmap (Moraru et al., 2020).

Results and discussion

The phylogeny of *Pseudomonas* ZM01

The phylogenetic tree was constructed based on the 16S rRNA gene of marine *Pseudomonas* ZM01 and 50 other

reference sequences of similar marine species in the *Pseudomonas* family (Figure 1A; Supplemental Table S1). *Pseudomonas* ZM01 forms a distinct clade, suggesting that *Pseudomonas* ZM01 might represent a novel variant of *Pseudomonas aeruginosa*.

Morphology of vB_PaeM_M1

A marine phage, named vB_PaeM_M1, that can infect *Pseudomonas* ZM01 was isolated from a surface seawater sample collected from the coastal waters of Qingdao, Yellow Sea. TEM images show that phage vB_PaeM_M1 has an isometric head (diameter of 90 to 90.9 nm [average ± standard deviation, 55 ± 3 nm]) and a contractible tail (length of 88 to 91 nm [11 ± 3 nm]) (Figure 1B).

Overall genome features

vB_PaeM_M1 was a linear, 89,387-bp, double stranded DNA (dsDNA) molecule genome with a GC content of 41.04% (Supplemental Tables S3, S4). The open reading frames (ORFs) were determined by BLASTp, HHpred and Pfam search analysis, and a total of 184 ORFs were predicted. Among them, 50 ORFs had known functions, and were divided into four functional modules: phage structure (ORFs 88, 113, 114, 115, 117, 119, 120, 121, 122, 123, 124, 126, 127, 128, 129, 130, 133, 134, 135, 137, 141 and 169), DNA replication and regulation (ORFs 21, 33, 65, 72, 74, 79, 80, 92, 97, 100, 108, 111, 138, 148, 149, 153, 155, 158, 159, 160, 167, 168, 173, 182, 183 and 184), phage packaging (ORFs 86, 109 and 110) and lytic protein (ORFs 75, 77 and 140). In addition, two AMGs (ORF 87 and

ORF 112) were predicted (Supplemental Table S2; Figure 2). With regard to acquired resistance by horizontal gene transfer, ResFinder and CARD did not detect any acquired antibiotic resistance genes.

ORF 21 encodes pyrimidine dimer DNA glycosylase, which mainly removes pyrimidine dimer by hydrolyzing the 5 'pyrimidine glycosyl bond and phosphodiester bond in pyrimidine (Morikawa et al., 1992). The pyrimidine dimer mainly results from UV damage to DNA, which it can effectively repair. ORFs 33, 65, 74, 111, 168, 182 and 184 encode endonuclease, which hydrolyze the phosphodiester bond in the molecular chain and participate in the post transcriptional cleavage, modification and degradation of DNA (Wang et al., 2005). ORFs 72 and 160 might be involved in the hydrolysis of a variety of lipids formed from phosphoric acid (Aravind and Koonin, 1998). ORF 97 encodes DNA polymerase; these enzymes catalyze the polymerization of dideoxy ribonucleotides with DNA strands, which is necessary for DNA replication. ORFs 79 and 92 encode DNA ligase and RNA ligase respectively, which connect the phosphodiester bond between the DNA strand and the RNA strand in the process of replication and transcription (Nandakumar et al., 2006). ORF 100 is expected to encode a zinc finger protein, which is mainly related to A2L transcription factors of some viruses and participates in the transcription process of DNA (Iyer et al., 2006). ORF 138 encodes DNA-directed RNA polymerase, also known as transcriptase. It is an enzyme that synthesizes RNA using a DNA strand or RNA as a template, ribonucleoside triphosphate as a substrate, and polymerizes through phosphodiester bonds (Spåhr et al., 2009). ORFs 148 and 149 encode ribonucleotide reductase, which catalyzes nucleotides into deoxyribonucleotides and uses them for DNA synthesis. ORF 153 encodes thymidylate synthase, which can
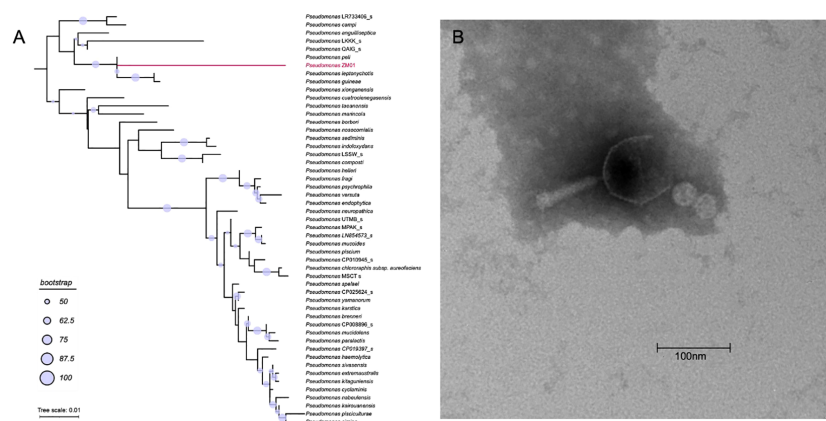

FIGURE 1
(A) The phylogenic tree based on the 16S rRNA gene of *Pseudomonas* ZM01 and other 120 reference 16S rRNA gene sequences of *Pseudomonas* ZM01. **(B)** Transmission electron micrograph of vB_PaeM_M1.
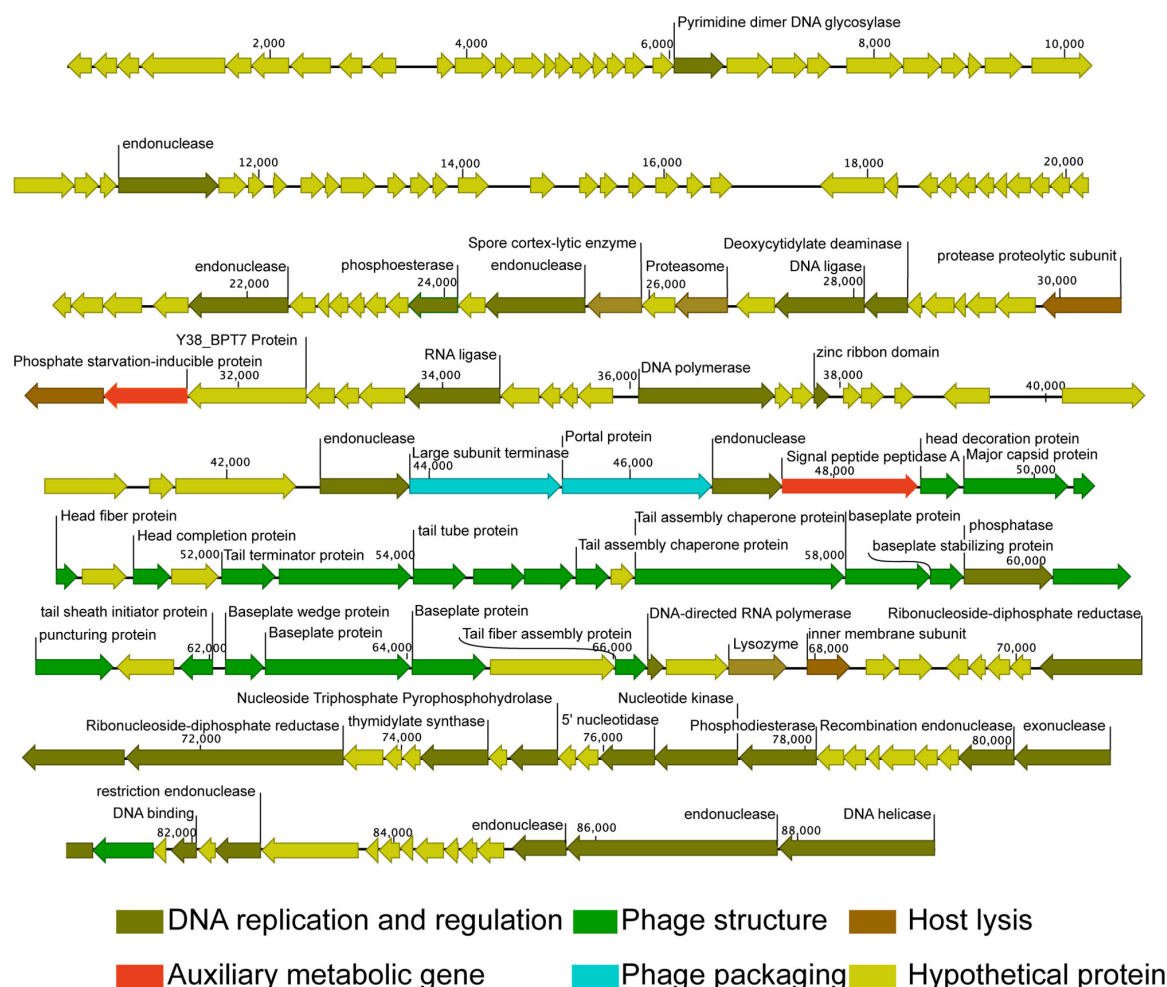
FIGURE 2

Genome map of *Pseudomonas* phage vB_PaeM_M1. Putative functional categories were defined according to annotation and are represented by different colors. The length of each arrow represents the length of each gene.

catalyze the methylation of uracil deoxynucleotides to synthesize 5'- thymine deoxyribonucleotides. It is an enzyme necessary for DNA synthesis (Graziani et al., 2006). ORF 155 encodes Nucleoside triphosphate pyrophosphohydrolase, which can cut off the nucleotide pyrophosphate bond (Gonçalves et al., 2011). ORF 184 encodes DNA helicase, which is similar to DNA Enterobacteria phage T3 (Beck et al., 1989).

Most of the genes related to phage packaging are located in the upper reaches of the vB_PaeM_M1 genome. ORF 86 encodes protease, which hydrolyzes the protein peptide chain (Kang et al., 2004). ORF 109 encodes the large TerL subunit, terminal enzyme and DNA recognition protein, which mediates packaging of the dsDNA virus copolymer, and requires interaction of the prohead with the virus DNA (Siehl et al., 1996). ORF 110 encodes portal protein, which controls the size of the assembled virus genome and can effectively prevent DNA

from escaping from the capsid during assembly (Wang et al., 2021). By using PhageTerm to determine the end of vB_PaeM_M1, it was found that the packaging process was similar to that of bacteriophage T7, and the cos site was found at the end of the M1 genome (Supplemental Figure S1). Different virus genes have different origins and vB_PaeM_M1 has a special phylogenetic state and evolutionary history. It is speculated that this phage may have originate from the T7 Virus but have subsequently acquired or lost some genes so that it no longer clusters with T7 viruses on the phylogenetic tree.

Three lytic-associated ORFs were encoded by the virus genome. ORF 75 encodes Spore cortex-lytic enzyme. These enzymes have been implicated in cell wall hydrolysis, most extensively in *Bacillus subtilis* (Boland et al., 2000). ORF 77 encodes proteasome. This protein plays a key role in the maintenance of protein homeostasis by removing misfolded or

damaged proteins, which can impair cellular functions, and by removing proteins whose functions are no longer required (Estrin et al., 2013). ORF 140 encodes lysozymes, which are 1,4-β-N-acetylmuramidases, that cleave glycosidic bonds, leading to the rupture of the cell wall with subsequent escape of the cell contents (Swaminathan et al., 2011).

The structure related genes are mainly located in the middle reaches of the vB_PaeM_M1 genome. ORF 114 encodes the main capsid protein (MCP), which is used to synthesize the protein shell of the virus encapsulating its genetic material (Katsura, 1989). ORFs 113, 115 and 117 encode the head decoration protein, head fiber protein and head completion protein respectively, which play important roles in the formation of the phage head structure (Tao et al., 1998; Yang et al., 2000). ORF 128, 129 and 135 encode the baseplate protein. This protein forms the central substrate hub of the phage and is considered to be related to horizontal gene transfer (Gloor and Chaconas, 1988; Haggård-Ljungquist et al., 1995; Christie and Calendar, 2016). ORFs 133 and 134 encode the baseplate wedge protein, which constitute a multi protein tubular organ that adheres to and penetrates the host cell membrane. It is a part of the conservative wedge in the bottom plate. It is the core of sheath polymerization and plays a key role in sheath assembly and contraction (Taylor et al., 2016). ORF 121, 122 and 123 are similar to ORF 10 of the *Pseudomonas* phage and are identified as a structure related gene with unknown function (Uchiyama et al., 2009).

The AMGs are phage encoded and host derived metabolic genes thought to be involved in regulating host metabolism to increase viral replication (Zimmerman et al., 2020). Two AMGs were predicted in the vB_PaeM_M1 genome, i.e. ORF 87 (Phosphate starvation-inducible protein) and ORF 112 (Signal peptide peptidase). ORF 87 encodes a *PhoH*-ike protein, which is found in *Escherichia coli* and belongs to the phosphate (*Pho*) regulator gene. *E. coli* responds to phosphate restriction by activating the transcription of this gene. Its products are involved in the use and transportation of various forms of combined phosphate or free phosphate. Among them, *PhoH* has two promoters, one can be induced by phosphate restriction and the other by constructive restriction. This gene has a putative nucleotide binding motif and the purified *Pho* protein has an ATP binding activity (Kim et al., 1993). ORF 112 encodes the signal peptide peptidase A (*SppA*). *SppA* is a membrane-bound enzyme that uses a serine/lysine catalyzed two-dimensional enzymatic mechanism to cleave residual signal peptides within the cell membrane (Nam and Paetzel, 2013).

Phylogeny and comparative genomic analysis of vB_PaeM_M1

In order to study the phylogenetic relationships between vB_PaeM_M1 and other isolated phages, a circular proteome tree was constructed using VipTree based on the whole-genome amino acid sequences of vB_PaeM_M1 and other reference phages (Figure 3). Thirty-five viruses with the highest similarity to vB_PaeM_M1 were selected to construct a rectangular proteome tree with vB_PaeM_M1 (Figure 4A) (Nishimura et al., 2017). The results of the phylogenetic analysis show that vB_PaeM_M1 represents a separate cluster far from the other isolated phages. To confirm this finding, six complete phage genome sequences were selected and a comparative genomic analysis was carried out based on tBLASTx with vB_PaeM_M1 (Figure 4B). Similar to the phylogenetic tree analysis, vB_PaeM_M1 shows a low similarity to other isolated phages and forms a new single clade by itself.

Intergenomic comparisons are useful in determining how viruses are related to each other. In order to further find the phylogenetic relationship between vB_PaeM_M1 and other isolated phages, the 35 viruses closest to vB_PaeM_M1 were selected from the phylogenetic tree for nucleotide-based intergenomic analysis (Figure 5). The heatmap shows that the ratio of shared genes between vB_PaeM_M1 and other isolated phages is very low. When the nucleotide sequence homology of bacteriophages is greater than 70%, ICTV regards them as members of the same genus (Bin Jang et al., 2019). All phylogenetic analyses showed that vB_PaeM_M1 was notably different from other isolated phages and should therefore be classified as a representative of an unknown virus family.

Although vB_PaeM_M1 clearly has the characteristics of a new genus/family, it is difficult to base the characteristics of a new genus/family on only one phage. In order to further explore the relationships between vB_PaeM_M1 and other phages, all-vs-all BLASTp were used to search the Integrated Microbial Genome/Virus (IMG/VR, v.3) database. It was found that 11 metagenomic-assembled uncultured viral genomes (UViGs) were similar to vB_PaeM_M1 (Supplemental Data 1). Comparative genomic analysis of vB_PaeM_M1 and vB_PaeM_M1-like viruses showed that metagenomic assembled genome viruses of vB_PaeM_M1-like viruses contained eight core genes. Most of the homologous ORFs throughout these phage genomes are located in the packaging module (portal protein, large subunit terminase and phage DNA packaging) and structural module (Baseplate protein and Structural protein). Based on the shared single white cluster (PC) between genomes, 354 viruses (336 reference sequences of NCBI, vB_PaeM_M1 and 17 UViGs similar to vB_PaeM_M1) were clustered by vConTACT 2.0. The classification of the virus genome is visualized by gephi (Figure 6A) (Bastian et al., 2009). Using this method, 60 virus clusters were identified in the whole dataset. vB_PaeM_M1 and 17 phages were grouped as a VC (VC_61) from the results of the genome-content-based analysis.

To find phylogenetic relationships between gene sequences of vB_PaeM_M1, 50 phages related to vB_PaeM_M1, based on the results of vConTACT, were selected to study their shared gene rate. Based on these results, a heatmap was drawn by orthofinder (Figure 7A). vB_PaeM_M1 and UViGs contain shared genes and
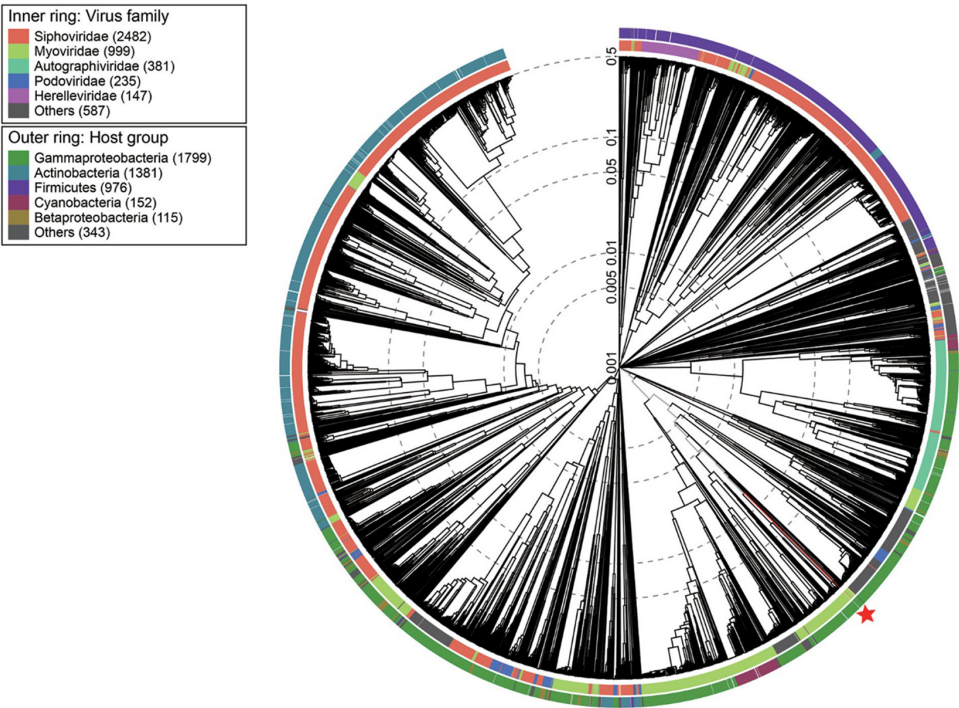
FIGURE 3
Viral phylogenetic tree of whole phage vB_PaeM_M1 genomes represented in the circular view. The colored rings represent the host groups (outer ring) and virus families (inner ring). These trees are calculated by BIONJ according to the genome distance matrix and take the midpoint as the root.

FIGURE 4
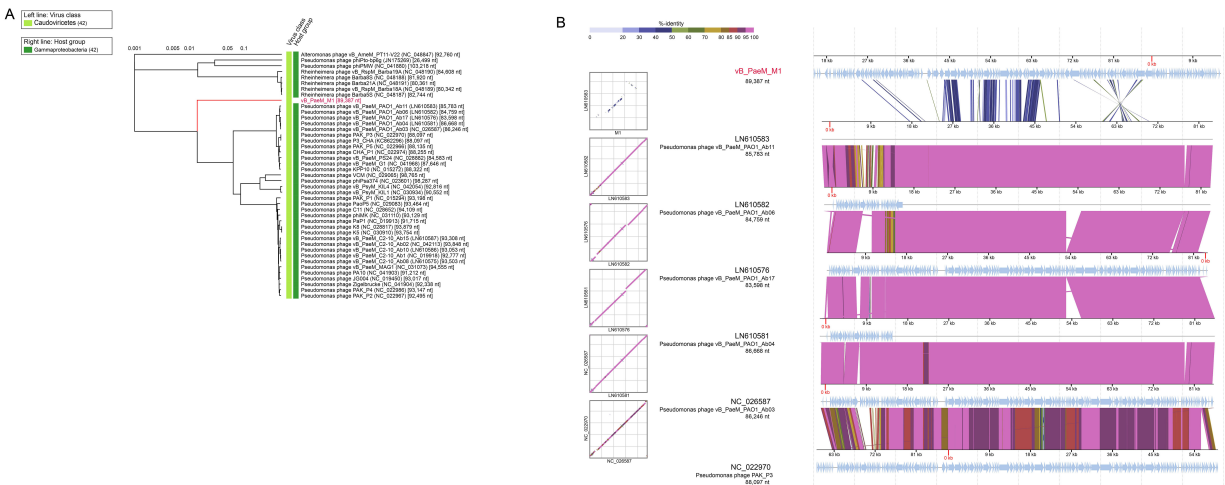Phylogenetic and comparative genomic analyses of *Pseudomonas* phage vB_PaeM_M1. **(A)** Phylogenetic tree of *Pseudomonas* phage vB_PaeM_M1 and the 36 closest virus genomes. **(B)** Genomic comparisons between *Pseudomonas* phage vB_PaeM_M1 and vB_PaeM_M1-like viruses. The shading below each genome indicates sequence similarities between the genomes, with different colors representing the levels of similarity.
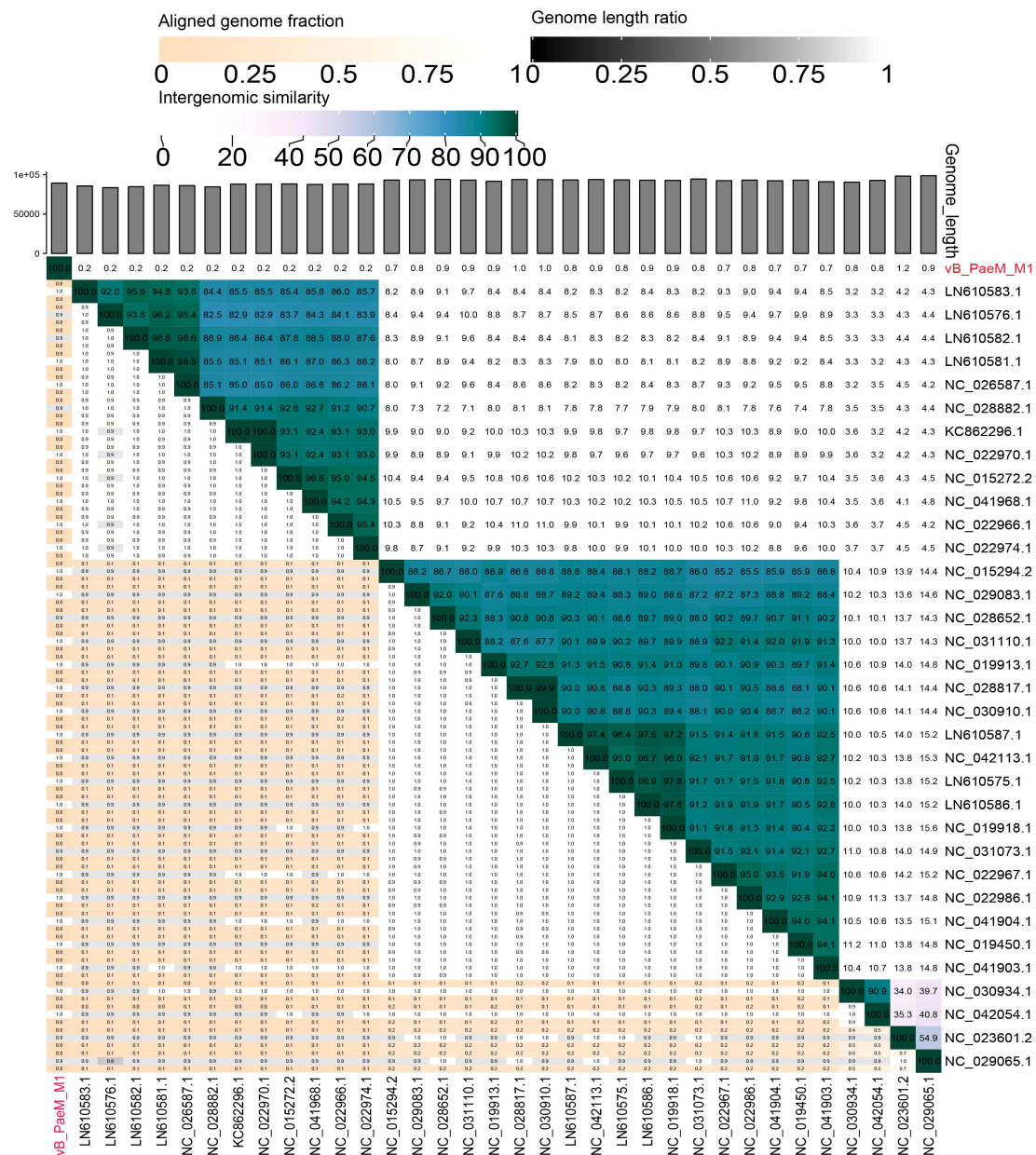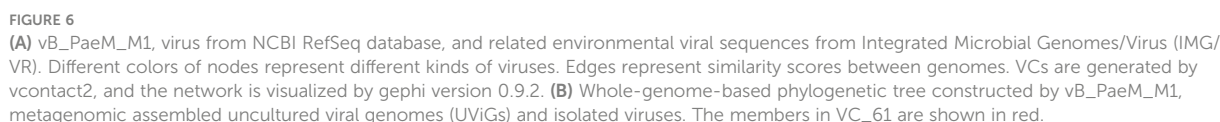
FIGURE 5

Shared gene heat map of vB_PaeM_M1 and vB_PaeM_M1 like virus in rectangular phylogenetic tree. In the right half, color coding allows rapid visualization of phage genome clustering based on intergenic similarity: the closer the relationship between genomes, the darker the color. These numbers represent the similarity values of each genome pair. In the left half, the three indicator values of each genome pair are expressed in order from top to bottom: aligned partial genome 1 (for the genome found in this row), genome length ratio (for two genomes in this pair), and aligned partial genome 2 (for the genome found in this column). Darker colors emphasize lower values, indicating that there is only a small number of genome aligned genome pairs (orange to white gradient), or there is a large difference in the length of the two genomes (black to white gradient). As the distance between phages increases, the aligned genomic fragments are expected to decrease.

have low association with other viruses, confirming that these phages (VC_61) might be identified as a new family. In addition, the whole-genome phylogeny results show that vB_PaeM_M1 and 11 other environmental viral sequences cluster together (shown in red) and are comparable to identified *podovirus* genera in ICTV

(Figure 6B). The gene map (Figure 7B) shows the core genome of VC_61, where ORFs 117 and 127 (shown in green) are conserved genes. ORF 117 is a head tail binding protein, which plays an important role in the packaging of phages found in *Pseudomonas* phage Zigelbrucke. ORF 127 is a baseplate organization protein,

FIGURE 6
(A) vB_PaeM_M1, virus from NCBI RefSeq database, and related environmental viral sequences from Integrated Microbial Genomes/Virus (IMG/VR). Different colors of nodes represent different kinds of viruses. Edges represent similarity scores between genomes. VCs are generated by vcontact2, and the network is visualized by gephi version 0.9.2. **(B)** Whole-genome-based phylogenetic tree constructed by vB_PaeM_M1, metagenomic assembled uncultured viral genomes (UViGs) and isolated viruses. The members in VC_61 are shown in red.

which is a structural protein in phages. It is also a conservative protein with specificity in VC_61. Using BLASTp to search these two conserved proteins in the NR database, it was found that no virus contained both these two proteins at the same time, so the two proteins can be regarded as conserved proteins. In conclusion, all results confirmed that vB_PaeM_M1 and related UViGs are a novel unassigned viral family, while vB_PaeM_M1 is the only isolate in this putative family which is here named *Psaeviridae*.

FIGURE 7
(A) Heat map of gene sharing rate between vB_PaeM_M1 and other phages. The sharing rate is displayed by color depth. **(B)** The gene map of VC_61. Core genomes are represented by yellow arrows and the conservative protein are represented by green arrows.

Conclusion

Viruses play a vital role in the control of marine microbial communities and are responsible for most prokaryote deaths. In view of the ecological significance of *Pseudomonas*, the study of its bacteriophages is still in its infancy. In this study, a novel *Pseudomonas* phage, vB_PaeM_M1, with unique genomic characteristics and phylogenetic position, is described. vB_PaeM_M1 represents a new viral family of *Caudoviricetes*, namely, *Psaeviridae*. Currently, databases of genes and proteins are extremely limited and more phages need to be isolated and identified from different environments. Bacteriophages play an important role not only in lysing the host, but also in manipulating the host to participate in the material circulation and energy flow of the marine environment. The establishment of *Psaeviridae* will undoubtedly deepen our understanding of *Pseudomonas* phages, enable us to better understand the physiological characteristics of bacteriophages in different aquatic environments and their relationship with hosts, and contribute to the data mining of a large number of metagenomic data sets.

Data availability statement

The complete genome sequence of *Pseudomonas* phage vB_PaeM_M1 is available in the GenBank database under accession number OP714164. In addition, the 16S rRNA gene sequence of *Pseudomonas* ZM01 has also been deposited in NCBI under accession number OP788125.

Author contributions

YTL, BL, and MW: conceptualization and project administration. LR: methodology, writing, and original draft preparation. YDL: phage isolation, genome analysis and software. KZ and ZW: data curation. HW and HS: validation. SY: data curation. WM and LW: visualization. YTL, MW, and AM: funding acquisition. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.1076885/full#supplementary-material

References

Alam, K., Islam, M. M., Li, C., Sultana, S., Zhong, L., Shen, Q., et al. (2021). Genome mining of *Pseudomonas* species: diversity and evolution of metabolic and biosynthetic potential. *Molecules* 26 (24), 7524. doi: 10.3390/molecules26247524

Alcock, B. P., Raphenya, A. R., Lau, T. T., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 48 (D1), D517–D525. doi: 10.1093/nar/gkz935

Aravind, L., and Koonin, E. V. (1998). Phosphoesterase domains associated with DNA polymerases of diverse origins. *Nucleic Acids Res.* 26 (16), 3746–3752. doi: 10.1093/nar/26.16.3746

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST server: rapid annotations using subsystems technology. *BMC Genomics* 9 (1), 1–15. doi: 10.1186/1471-2164-9-75

Bastian, M., Heymann, S., and Jacomy, M. (2009). "Gephi: An open source software for exploring and manipulating networks," in *Proceedings of the Third International Conference on Weblogs and Social Media, ICWSM 2009*, San Jose, California, USA, May 17-20, 2009.

Beck, P. J., Gonzalez, S., Ward, C. L., and Molineux, I. J. (1989). Sequence of bacteriophage T3 DNA from gene 2.5 through gene 9. *J. Mol. Biol.* 210 (4), 687–701. doi: 10.1016/0022-2836(89)90102-2

Beilstein, F., and Dreiseikelmann, B. (2006). Bacteriophages of freshwater brevundimonas vesicularis isolates. *Res. Microbiol.* 157 (3), 213–219. doi: 10.1016/j.resmic.2005.07.005

Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J. H., Roux, S., Adriaenssens, E. M., et al. (2019). Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat. Biotechnol.* 37 (6), 632–639. doi: 10.1038/s41587-019-0100-8

Boland, F. M., Atrih, A., Chirakkal, H., Foster, S. J., and Moir, A. (2000). Complete spore-cortex hydrolysis during germination of bacillus subtilis 168 requires SleB and YpeB. *Microbiology* 146, 57–64. doi: 10.1099/00221287-146-1-57

Bolduc, B., Jang, H. B., Doulcier, G., You, Z. Q., Roux, S., and Sullivan, M. B. (2017). vConTACT: An iVirus tool to classify double-stranded DNA viruses that infect archaea and bacteria. *PeerJ* 5, e3243. doi: 10.7717/peerj.3243

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* 30 (15), 2114–2120. doi: 10.1093/bioinformatics/btu170

Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., et al. (2020). ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* 75, 3491–3500. doi: 10.1093/jac/dkaa345

Bragg, J. G., and Chisholm, S. W. (2008). Modeling the fitness consequences of a cyanophage-encoded photosynthesis gene. *PloS One* 3 (10), e3550. doi: 10.1371/journal.pone.0003550

Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J. M., Segall, A. M., Mead, D., et al. (2002). Genomic analysis of uncultured marine viral communities. *Proc. Natl. Acad. Sci. United States America* 99 (22), 14250–14255. doi: 10.1073/pnas.202488399

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinforma. Oxf. Engl.* 25 (15), 1972–1973. doi: 10.1093/bioinformatics/btp348

Carlton, R. M. (1999). Phage therapy: Past history and future prospects. *Arch. Immunol. Ther. Exp. (Warsz)* 47 (5), 267–274. doi: 10.1007/978-981-19-1854-4_5

Christie, G. E., and Calendar, R. (2016). Bacteriophage P2. *Bacteriophage* 6 (1), e1145782. doi: 10.1080/21597081.2016.1145782

Duhaime, M. B., Wichels, A., Waldmann, J., Teeling, H., and Glöckner, F. O. (2011). Ecogenomics and genome landscapes of marine pseudoalteromonas phage H105/1. *Isme J.* 5 (1), 107–121. doi: 10.1038/ismej.2010.94

Emms, D. M., and Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20 (1), 238. doi: 10.1186/s13059-019-1832-y

Estrin, E., Lopez-Blanco, J. R., Chacón, P., and Martin, A. (2013). Formation of an intricate helical bundle dictates the assembly of the 26S proteasome lid. *Structure* 21, 1624–1635. doi: 10.1016/j.str.2013.06.023

Garneau, J. R., Depardieu, F., Fortier, L. C., Bikard, D., and Monot, M. (2017). PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. *Sci. Rep.* 7 (1), 1–10. doi: 10.1038/s41598-017-07910-5

Gloor, G., and Chaconas, G. (1988). Sequence of bacteriophage mu n and p genes. *Nucleic Acids Res.* 16 (11), 5211–5212. doi: 10.1093/nar/16.11.5211

Gonçalves, A. M. D., de Sanctis, D., and McSweeney, S. M. (2011). Structural and functional insights into DR2231 protein, the MazG-like nucleoside triphosphate pyrophosphohydrolase from deinococcus radiodurans. *J. Biol. Chem.* 286 (35), 30691–30705. doi: 10.1074/jbc.M111.247999

Graham-Mize, C. A., and Rosser, E. J.Jr. (2004). Comparison of microbial isolates and susceptibility patterns from the external ear canal of dogs with otitis externa. *J. Am. Anim. Hosp Assoc.* 40 (2), 102–108. doi: 10.5326/0400102

Graziani, S., Bernauer, J., Skouloubris, S., Graille, M., Zhou, C. Z., Marchand, C., et al. (2006). Catalytic mechanism and structure of viral flavin-dependent thymidylate synthase ThyX. *J. Biol. Chem.* 281 (33), 24048–24057. doi: 10.1074/jbc.M600745200

Gregory, A. C., Zayed, A. A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et al. (2019). Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177 (5), 1109–1123.e1114. doi: 10.1016/j.cell.2019.03.040

Haggård-Ljungquist, E., Jacobsen, E., Rishovd, S., Six, E. W., Nilssen, O., Sunshine, M. G., et al. (1995). Bacteriophage P2: genes involved in baseplate assembly. *Virology* 213 (1), 109–121. doi: 10.1006/viro.1995.1551

Hyman, P., and Abedon, S. T. (2010). Bacteriophage host range and bacterial resistance. *Adv. Appl. Microbiol.* 70, 217–248. doi: 10.1016/s0065-2164(10)70007-1

Iyer, L. M., Balaji, S., Koonin, E. V., and Aravind, L. (2006). Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res.* 117 (1), 156–184. doi: 10.1016/j.virusres.2006.01.009

Kang, S. G., Maurizi, M. R., Thompson, M., Mueser, T., and Ahvazi, B. (2004). Crystallography and mutagenesis point to an essential role for the n-terminus of human mitochondrial ClpP. *J. Struct. Biol.* 148 (3), 338–352. doi: 10.1016/j.jsb.2004.07.004

Katsura, I. (1989). Structure and inherent properties of the bacteriophage lambda head shell. VI. DNA-packaging-defective mutants in the major capsid protein. *J. Mol. Biol.* 205 (2), 397–405. doi: 10.1016/0022-2836(89)90350-1

Kim, S. K., Makino, K., Amemura, M., Shinagawa, H., and Nakata, A. (1993). Molecular analysis of the phoH gene, belonging to the phosphate regulon in escherichia coli. *J. Bacteriol* 175 (5), 1316–1324. doi: 10.1128/jb.175.5.1316-1324.1993

Letunic, I., and Bork, P. (2019). Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47 (W1), W256–w259. doi: 10.1093/nar/gkz239

Liu, Z., Li, H., Wang, M., Jiang, Y., Yang, Q., Zhou, X., et al. (2018). Isolation, characterization and genome sequencing of the novel phage SL25 from the yellow Sea, China. *Mar. Genomics* 37, 31–34. doi: 10.1016/j.margen.2017.09.008

Liu, Y., Zhao, L., Wang, M., Wang, Q., Zhang, X., Han, Y., et al. (2019). Complete genomic sequence of bacteriophage P23: A novel vibrio phage isolated from the yellow Sea, China. *Virus Genes* 55 (6), 834–842. doi: 10.1007/s11262-019-01699-3

Li, Y., Wang, M., Liu, Q., Song, X., Wang, D., Ma, Y., et al. (2016). Complete genomic sequence of bacteriophage H188: A novel vibrio kanaloae phage isolated from yellow Sea. *Curr. Microbiol.* 72 (5), 628–633. doi: 10.1007/s00284-015-0984-6

Marmur, J. (1961). A procedure for the isolation of deoxiribonucleic acid from microorganisms. *J. Mol. Biol.* 3, 208–218. doi: 10.1016/S0022-2836(61)80047-8

Middelboe, M., Chan, A. M., and Bertelsen, S. K. (2010). Isolation and life-cycle characterization of lytic viruses infecting heterotrophic bacteria and cyanobacteria. *Manual Aquat. Viral Ecol.* 118-133. doi: 10.4319/mave.2010.978-0-9845591-0-7.118

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., et al. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37 (5), 1530–1534. doi: 10.1093/molbev/msaa015

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., et al. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Res.* 49 (D1), D412–D419. doi: 10.1093/nar/gkaa913

Moraru, C., Varsani, A., and Kropinski, A. M. (2020). VIRIDIC-a novel tool to calculate the intergenomic similarities of prokaryote-infecting viruses. *Viruses* 12 (11), 1268-1272. doi: 10.3390/v12111268

Morikawa, K., Matsumoto, O., Tsujimoto, M., Katayanagi, K., Ariyoshi, M., Doi, T., et al. (1992). X-Ray structure of T4 endonuclease V: An excision repair enzyme specific for a pyrimidine dimer. *Science* 256 (5056), 523–526. doi: 10.1126/science.1575827

Nam, S. E., and Paetzel, M. (2013). Structure of signal peptide peptidase a with c-termini bound in the active sites: Insights into specificity, self-processing, and regulation. *Biochemistry* 52 (49), 8811–8822. doi: 10.1021/bi4011489

Nandakumar, J., Shuman, S., and Lima, C. D. (2006). RNA Ligase structures reveal the basis for RNA specificity and conformational changes that drive ligation forward. *Cell* 127 (1), 71–84. doi: 10.1016/j.cell.2006.08.038

Nishimura, Y., Yoshida, T., Kuronishi, M., Uehara, H., Ogata, H., and Goto, S. (2017). ViPTree: the viral proteomic tree server. *Bioinformatics* 33 (15), 2379–2380. doi: 10.1093/bioinformatics/btx157

Paez-Espino, D., Chen, I. A., Palaniappan, K., Ratner, A., Chu, K., Szeto, E., et al. (2017). IMG/VR: A database of cultured and uncultured DNA viruses and retroviruses. *Nucleic Acids Res.* 45 (D1), D457–d465. doi: 10.1093/nar/gkw1030

Paez-Espino, D., Roux, S., Chen, I. A., Palaniappan, K., Ratner, A., Chu, K., et al. (2019). IMG/VR v.2.0: an integrated data management and analysis system for cultivated and environmental viral genomes. *Nucleic Acids Res.* 47 (D1), D678–d686. doi: 10.1093/nar/gky1127

Pires, D., Sillankorva, S., Faustino, A., and Azeredo, J. (2011). Use of newly isolated phages for control of *Pseudomonas* aeruginosa PAO1 and ATCC 10145 biofilms. *Res. Microbiol.* 162 (8), 798–806. doi: 10.1016/j.resmic.2011.06.010

Poole, K. (2011). *Pseudomonas* aeruginosa: resistance to the max. *Front. Microbiol.* 2. doi: 10.3389/fmicb.2011.00065

Pourcel, C., Midoux, C., Vergnaud, G., and Latino, L. (2017). A carrier state is established in *Pseudomonas* aeruginosa by phage LeviOr01, a newly isolated ssRNA levivirus. *J. Gen. Virol.* 98 (8), 2181–2189. doi: 10.1099/jgv.0.000883

Roux, S., Páez-Espino, D., Chen, I. A., Palaniappan, K., Ratner, A., Chu, K., et al. (2021). IMG/VR v3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. *Nucleic Acids Res.* 49 (D1), D764–d775. doi: 10.1093/nar/gkaa946

Schneiker, S., Martins dos Santos, V. A., Bartels, D., Bekel, T., Brecht, M., Buhrmester, J., et al. (2006). Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium alcanivorax borkumensis. *Nat. Biotechnol.* 24 (8), 997–1004. doi: 10.1038/nbt1232

Siehl, D. L., Subramanian, M. V., Walters, E. W., Lee, S. F., Anderson, R. J., and Toschi, A. G. (1996). Adenylosuccinate synthetase: site of action of hydantocidin, a microbial phytotoxin. *Plant Physiol.* 110 (3), 753–758. doi: 10.1104/pp.110.3.753

Spåhr, H., Calero, G., Bushnell, D. A., and Kornberg, R. D. (2009). Schizosacharomyces pombe RNA polymerase II at 3.6-a resolution. *Proc. Natl. Acad. Sci. U.S.A.* 106 (23), 9185–9190. doi: 10.1073/pnas.0903361106

Suttle, C. A. (2007). Marine viruses–major players in the global ecosystem. *Nat. Rev. Microbiol.* 5 (10), 801–812. doi: 10.1038/nrmicro1750

Swaminathan, R., Ravi, V. K., Kumar, S., Kumar, M. V., and Chandra, N. (2011). Lysozyme: A model protein for amyloid research. *Adv. Protein Chem. Struct. Biol.* 84, 63–111. doi: 10.1016/B978-0-12-386483-3.00003-3

Tao, Y., Olson, N. H., Xu, W., Anderson, D. L., Rossmann, M. G., and Baker, T. S. (1998). Assembly of a tailed bacterial virus and its genome release studied in three dimensions. *Cell* 95 (3), 431–437. doi: 10.1016/s0092-8674(00)81773-0

Taylor, N. M., Prokhorov, N. S., Guerrero-Ferreira, R. C., Shneider, M. M., Browning, C., Goldie, K. N., et al. (2016). Structure of the T4 baseplate and its function in triggering sheath contraction. *Nature* 533 (7603), 346–352. doi: 10.1038/nature17971

Teeling, H., Meyerdierks, A., Bauer, M., Amann, R., and Glöckner, F. O. (2004). Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environ. Microbiol.* 6 (9), 938–947. doi: 10.1111/j.1462-2920.2004.00624.x

Thompson, L. R., Zeng, Q., Kelly, L., Huang, K. H., Singer, A. U., Stubbe, J., et al. (2011). Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 108 (39), E757–E764. doi: 10.1073/pnas.1102164108

Uchiyama, J., Maeda, Y., Takemura, I., Chess-Williams, R., Wakiguchi, H., and Matsuzaki, S. (2009). Blood kinetics of four intraperitoneally administered therapeutic candidate bacteriophages in healthy and neutropenic mice. *Microbiol. Immunol.* 53 (5), 301–304. doi: 10.1111/j.1348-0421.2009.00125.x

Wang, J., Jiang, Y., Vincent, M., Sun, Y., Yu, H., Wang, J., et al. (2005). Complete genome sequence of bacteriophage T5. *Virology* 332 (1), 45–65. doi: 10.1016/j.virol.2004.10.049

Wang, D. B., Sun, M. Q., Shao, H. B., Li, Y., Meng, X., Liu, Z. Y., et al. (2015). Characterization and genome sequencing of a novel bacteriophage PH101 infecting pseudoalteromonas marina BH101 from the yellow Sea of China. *Curr. Microbiol.* 71 (5), 594–600. doi: 10.1007/s00284-015-0896-5

Wang, Z., Zhang, F., Liang, Y., Zheng, K., Gu, C., Zhang, W., et al. (2021). Genome and ecology of a novel alteromonas podovirus, ZP6, representing a new viral genus, mareflavirus. *Microbiol. Spectr.* 9 (2), e0046321. doi: 10.1128/Spectrum.00463-21

Weinbauer, M. G. (2004). Ecology of prokaryotic viruses. *FEMS Microbiol. Rev.* 28 (2), 127–181. doi: 10.1016/j.femsre.2003.08.001

Yang, F., Forrer, P., Dauter, Z., Conway, J. F., Cheng, N., Cerritelli, M. E., et al. (2000). Novel fold and capsid-binding properties of the lambda-phage display platform protein gpD. *Nat. Struct. Biol.* 7 (3), 230–237. doi: 10.1038/73347

Yang, M., Xia, Q., Du, S., Zhang, Z., Qin, F., and Zhao, Y. (2021). Genomic characterization and distribution pattern of a novel marine OM43 phage. *Front. Microbiol.* 12. doi: 10.3389/fmicb.2021.651326

Zhan, Y., and Chen, F. (2019). Bacteriophages that infect marine roseobacters: Genomics and ecology. *Environ. Microbiol.* 21 (6), 1885–1895. doi: 10.1111/1462-2920.14504

Zimmerman, A. E., Howard-Varona, C., Needham, D. M., John, S. G., Worden, A. Z., Sullivan, M. B., et al. (2020). Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat. Rev. Microbiol.* 18 (1), 21–34. doi: 10.1038/s41579-019-0270-x