



Article An Estuarine Cyanophage S-CREM1 Encodes Three Distinct Antitoxin Genes and a Large Number of Non-Coding RNA Genes

Hongrui Zheng¹, Yuanfang Liu¹, Ruiyu Zhou¹, Jihua Liu¹, Yongle Xu^{1,*} and Feng Chen^{2,*}

- ¹ Institute of Marine Science and Technology, Shandong University, Qingdao 266000, China
- ² Institute of Marine and Environmental Technology, University of Maryland Center for Environmental Science, Baltimore, MD 21202, USA
- * Correspondence: xuyongle@sdu.edu.cn (Y.X.); chenf@umces.edu (F.C.)

Abstract: Cyanophages play important roles in regulating the population dynamics, community structure, metabolism, and evolution of cyanobacteria in aquatic ecosystems. Here, we report the genomic analysis of an estuarine cyanophage, S-CREM1, which represents a new genus of T4-like cyanomyovirus and exhibits new genetic characteristics. S-CREM1 is a lytic phage which infects estuarine Synechococcus sp. CB0101. In contrast to many cyanomyoviruses that usually have a broad host range, S-CREM1 only infected the original host strain. In addition to cyanophage-featured auxiliary metabolic genes (AMGs), S-CREM1 also contains unique AMGs, including three antitoxin genes, a MoxR family ATPase gene, and a pyrimidine dimer DNA glycosylase gene. The finding of three antitoxin genes in S-CREM1 implies a possible phage control of host cells during infection. One small RNA (sRNA) gene and three cis-regulatory RNA genes in the S-CREM1 genome suggest potential molecular regulations of host metabolism by the phage. In addition, S-CREM1 contains a large number of tRNA genes which may reflect a genomic adaption to the nutrient-rich environment. Our study suggests that we are still far from understanding the viral diversity in nature, and the complicated virus-host interactions remain to be discovered. The isolation and characterization of S-CREM1 further our understanding of the gene diversity of cyanophages and phage-host interactions in the estuarine environment.

Keywords: cyanophage; new phage genus; antitoxin genes; non-coding RNA genes

1. Introduction

Picocyanobacteria of the genera Synechococcus and Prochlorococcus are important primary producers in a wide range of marine environments [1,2]. In estuaries, picocyanobacteria contribute up to 56% of the primary production [3]. Cyanophages, viruses infecting cyanobacteria, are prevalent in marine ecosystems, lysing up to 40% of Synechococcus cells every day, playing a key role in regulating the population dynamics of the cyanobacteria [4,5]. To date, all cyanophages isolated from marine ecosystems belong to the class Caudovirales, including Myoviridae, Podoviridae, Siphoviridae, and the families Ackermannviridae, and Herelleviridae [6]. Cyanomyoviruses are the most frequently isolated cyanophages in marine ecosystems, with T4-like cyanophages as the most important group [7,8]. T4-like cyanophages usually contain a set of core genes including virion formation and DNA replication genes [7]. In addition, T4-like cyanophages generally encode various auxiliary metabolic genes (AMGs), and they can affect host photosynthesis, carbon metabolism, nutrient acquisition, stress tolerance, and nucleic acid synthesis during infection [9–13]. AMGs in T4-like cyanophages indicate the characteristics of phage-host interactions, and the gain or loss of AMGs in the cyanophage genomes is the adaptation to environmental selection pressures [14]. tRNA genes are widely present in the T4-like cyanophage genomes,



Citation: Zheng, H.; Liu, Y.; Zhou, R.; Liu, J.; Xu, Y.; Chen, F. An Estuarine Cyanophage S-CREM1 Encodes Three Distinct Antitoxin Genes and a Large Number of Non-Coding RNA Genes. *Viruses* 2023, *15*, 380. https:// doi.org/10.3390/v15020380

Academic Editor: Dann Turner

Received: 19 December 2022 Revised: 25 January 2023 Accepted: 25 January 2023 Published: 28 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). with the number varying greatly [7]. tRNAs may play roles in improving the phage crossinfectivity of different hosts and in replacing host tRNAs to maintain translation [15,16]. Furthermore, other non-coding RNA genes, including CFrI, PhotoRC-II, *wcaG*, and the glutamine riboswitch, are also found in T4-like cyanophage genomes, which are predicted to influence the host's metabolism [8,17,18].

Estuaries connect marine and freshwater systems. Environmental gradients in estuaries are dynamic and pose great selective pressures on microorganisms [6,19]. Metagenomic studies revealed significant differences between estuarine phages and open ocean phages [20]. Estuarine cyanophages have developed unique genomic characteristics to cope with the highly dynamic environment [21,22]. The estuarine *Synechococcus* phage S-CBWM1 has a unique set of AMGs, structural, and DNA replication genes, and it possesses the largest number of tRNA genes that have ever been found in cyanophage isolates [22]. Notably, phages with more than 20 tRNAs are mainly isolated from eutrophic environments [22–24]. In addition, cyanophages isolated from eutrophic waters may carry fewer and more unique AMGs [25,26]. The gain and loss of unique AMGs in cyanophage genomes can be attributed to the fluctuating selection pressures in the estuarine environment [14], resulting in different host host-phage interactions compared to those in the marine environment [6]. However, only a few estuarine cyanophages have been isolated and studied, which limits our understanding of estuarine cyanophages' evolution and ecological roles. The isolation, genomic and physiological characterizations of new viruses are still of great significance for further exploring genetic diversity, understanding virus host-phage interactions, and elucidating their ecological roles in estuarine ecosystems [27,28].

Here, we characterized a newly isolated estuarine *Synechococcus* phage, S-CREM1. S-CREM1 represents a new genus of the T4-like cyanomyoviruses, and it carries a unique set of AMGs and various non-coding RNA genes. The isolation and characterization of S-CREM1 provide new insights into phage–host interactions in the estuarine environment.

2. Materials and Methods

2.1. Host Incubation and Cyanophage Isolation

Host strain *Synechococcus* sp. CB0101 isolated from Chesapeake Bay is a model strain of picocyanobacteria in the estuarine environment and belongs to *Synechococcus* subcluster 5.2. *Synechococcus* sp. CB0101 was grown in SN medium with 15‰ salinity (SN15), at 22 °C, under a constant light intensity of 20 µmol photons $m^{-2} s^{-1}$ in an illumination incubator [29,30]. S-CREM1 was isolated from the surface seawater of the Changjiang River Estuary (30.8° N, 122.6° E) using *Synechococcus* sp. CB0101 as the host [21]. The phage-containing seawater sample was collected from the surface of the Changjiang River Estuary and filtered through a 0.22 µm pore-size polycarbonate membrane (Millipore, Bedford, MA, USA) to remove microbial cells. The filtered sample was stored at 4 °C until use. Phages were first enriched in a 96-well microtiter plate and then isolated by the double-layer agar method [22,31]. The individual plaque was picked from the double-layer agar plate, resuspended with 2 mL of TM buffer (20 mM Tris-Cl and 10 mM MgSO₄), and then used in another round of phage purification using the double-layer agar method. The S-CREM1 isolate was obtained after five rounds of double-layer agar purification.

2.2. Host Range Determination

The cross infectivity of S-CREM1 was tested using 11 *Synechococcus* strains: CB0101, A10-1-5-1, CBW1003, CBW1006, CBW1107, CBW1004, PCC 7002, CC9311, WH 8102, WH 7803, and WH 7805. These 11 *Synechococcus* strains were cultured in different mediums with salinities ranging from 15 to 35. Ten microliters of 0.22 μ m filtered phage lysates were added to 0.2 mL of exponentially growing *Synechococcus* cultures in a 96-well microtiter plate in triplicate, while control cultures only received SN medium. All plates were incubated in the illumination incubator at 22 °C under 20 μ mol photons m⁻² s⁻¹ continuous white light. The infectivity was observed by comparing the cell lysis of the phage-added and the control groups within two weeks.

2.3. One-Step Growth Curve

At a multiplicity of infection (MOI) of 0.01, the S-CREM1 were inoculated into 1 mL exponentially growing cultures of *Synechococcus* sp. CB0101 ($OD_{750} = 0.5$) and incubated for 1 h at 22 °C under 20 μ mol photons m⁻² s⁻¹ continuous light for phage adsorption to host cells. The unabsorbed phages were removed by centrifugation at $6000 \times g$ for 10 min. Precipitated cells were resuspended in 100 mL of fresh SN15 medium in triplicates and incubated under the same conditions described above. Subsamples were taken at 0 h, 4 h, 8 h, 10 h, 12 h, 14 h, 16 h, 20 h, and 24 h to determine the variation in phage concentration. Phage concentration was quantified by quantitative real-time PCR (qPCR) [32], using the portal protein gene (g20) as the marker gene. The primers F (5'-TTATGAGTATGCTTGAGGAC-3') and R (5'-ATGAAGGAACGTTGAGTG-3') used in the g20 quantification were designed using the Primer Premier 5 software. The qPCR reactions were performed in a 10 μ L qPCR mix, which contained 5 μ L of SYBR Premix Ex TaqTM II, 1 μ L of each primer, 1 μ L of nuclease-free water, and 2 µL of DNA template. Thermal cycling was conducted in a CFX Connect (TM) real-time PCR system (Bio-Rad Laboratories, Hercules, CA, USA) consisting of a 10 min denaturation at 94 °C, and 40 cycles of denaturation at 94 °C for 30 s, annealing at 48 °C for 30 s, and elongation at 72 °C for 30 s.

2.4. Phage Amplification and Purification

S-CREM1 phage suspensions were inoculated into 2 L of exponentially growing cultures of *Synechococcus* sp. CB0101 at an MOI of 0.1. After host cell lysis, RNase A and DNase I were added to the lysates both at a final concentration of 2 µg mL⁻¹, and they were treated at room temperature for 1 h. Afterward, the NaCl concentration of phage lysates was adjusted to 1 M, and the lysates were ice-bathed for 0.5 h. To remove the remaining cells and debris, the phage lysates were centrifuged at 12,000× *g* at 4 °C for 20 min and further filtered through 0.22 µm filters (Millipore, Bedford, MA, USA). The filtrates were treated with PEG8000 (*w*/*v* 10%) and kept at 4 °C for 24 h [22,33]. The PEG-treated phage suspensions were centrifuged at 12,000× *g* at 4 °C for 1 h to precipitate phage particles and then resuspended with 6 mL of TM buffer. Concentrated phage particles were then purified by CsCl density gradient ultracentrifugation (gradient density 1.45, 1.5, 1.55, and 1.6 g mL⁻¹, 200,000× *g* at 4 °C, 6 h) in a SW 41Ti rotor (Beckman Optima L-100XP, Beckman Coulter, CA, USA) [8,34]. The visible phage band was extracted and then desalted using a 30 kDa centrifugal ultrafiltration unit. The purified high-titer phages were stored at 4 °C.

2.5. Transmission Electron Microscopy (TEM) Observation

The carbon-coated copper grids (200-mesh) were subjected to glow discharge for 20 s. Ten microliters of purified S-CREM1 suspensions were adsorbed to carbon-coated copper grids for 5 min and negatively stained twice with 2% (w/v) uranyl acetate for 10 s and 30 s, respectively. The stained sample was dried for 10 min and observed using a Tecnai G2 Spirit BioTwin transmission electron microscope (FEI Tecnai G2 F20, Thermo Fisher Scientific, Waltham, MA, USA).

2.6. Phage Genome DNA Extraction and Sequencing

Phage DNA was extracted from the purified high-titer phage suspension using the phenol–chloroform method described previously [6,33]. The genomic DNA was sequenced using the Illumina HiSeq 4000 platform by Shanghai Majorbio Bio-pharm Technology Co., Ltd. After quality control and trimming, a total of 2,192,371,253 bp clean reads were obtained. Afterward, the clean reads were assembled using IDBA-UD version 1.1.1 to generate the final complete genome sequence [35].

2.7. Genomic and Phylogenetic Analyses

Putative open reading frames (ORFs) of S-CREM1 were predicted by the RAST (http://rast.nmpdr.org/ (accessed on 12 May 2021)), the Gene-MarkS online server (http://exon.gatech.edu/GeneMark/ (accessed on 12 May 2021)), and the Meta Gene An-

notator (http://metagene.nig.ac.jp/ (accessed on 12 May 2021)). ORFs of S-CREM1 were annotated using BLASTP search against the NCBI non-redundant (NR) database (e-value $< 10^{-3}$) and conserved domain search against the NCBI Conserved Domain Database (e-value < 10^{-3} , bitscore > 40). For S-CREM1 ORFs with no predicted functions based on sequence analyses, distant homolog searches using HHpred (probability > 90%) and Phyre2 (confidence > 80%) were performed to assist the annotation based on predicted structural properties [36,37]. The tRNA genes were predicted by tRNAscan-SE (http://lowelab.ucsc.edu/tRNAscan-SE (accessed on 10 July 2021)) [38]. Other non-coding RNA genes, such as small RNA (sRNA) and cis-regulatory RNA genes, were predicted by searching against the Rfam database (https://rfam.xfam.org/family/RF03085 (accessed on 17 September 2022)) [18]. To characterize the genomic similarities of S-CREM1 with other phages, a total of 11,510 viral genomes were downloaded from the NCBI Viral RefSeq database. The similarity score between each pair of viral genomes was calculated by vCon-TACT 2.0 [39]. The 20 cyanophages that are most closely related to S-CREM1 were selected and further characterized with aspect to the genomic nucleotide similarity using VIRIDIC (http://rhea.icbm.uni-oldenburg.de/VIRIDIC/ (accessed on 14 January 2022)) [40]. Phylogenomic analyses of S-CREM1 and 45 T4-like cyanophages were performed based on amino acid sequences of 30 core genes. The core genes among the 46 phages were identified by OrthoFinder, aligned by MAFFT, and trimmed by TrimAI [41–43]. The phylogenomic tree was conducted with RAxML (version 8) employing the maximum likelihood method with the PROTGAMMAJTT model (bootstrap replicates = 100) [44]. Phylogenetic analyses of 2-oxoglutarate (2OG)-Fe(II) oxygenase, MoxR ATPase, and pyrimidine dimer DNA glycosylase genes were performed using the MEGA 7.0 software package [45]. The maximum likelihood method with the Jones-Taylor-Thornton (JTT) model and the neighbor-joining method with the *p*-distance model were used in the phylogenetic tree construction with 1000 bootstrap replicates.

2.8. Identification of the S-CREM1 Virion Proteins by Mass Spectrometry

Proteomic analysis of the S-CREM1 virions was performed with CsCl-purified phage suspensions. Fifty microliters of phage suspensions were mixed with the same volume of SDT lysis buffer (4% SDS, 100 mM Tris-HCl, 1 mM dithiothreitol, pH 7.6) and incubated in boiling water for 10 min. Dithiothreitol was added into the suspension at a final concentration of 100 mM and incubated in boiling water for 5 min. Then, 200 μ L of UA buffer (8 M urea, 150 mM Tris-HCl, pH 8.0) was added into the suspension, and the detergent was removed through ultrafiltration. A total of 100 μ L of iodoacetamide (IAA) buffer (100 mM IAA in UA) was used to modify the UA-buffered sample for 30 min at 25 $^{\circ}$ C in the dark. The protein suspension was combined with 100 μ L of UA buffer and centrifuged at 14,000 \times g for 15 min twice; then, 100 mL of 25 mM NH₄HCO₃ was added and the suspension was centrifuged at $14,000 \times g$ for 15 min twice. Afterward, the protein suspension was digested with 40 μ L of trypsin buffer (2 μ g of trypsin in 40 μ L of 100 mM NH₄HCO₃) at 37 °C for 18 h. The tryptic peptides were analyzed using liquid chromatography-electrospray ionization-tandem mass spectrometry (LC-ESI-MS/MS) by Shanghai Applied Protein Technology Co., Ltd. The determination was performed on the Q-Exactive mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) that was connected to an Easy nLC (Thermo Fisher Scientific, Waltham, MA, USA). Peptides were fractionated by buffer A (0.1% aqueous formic acid) and buffer B (84% acetonitrile and 0.1% aqueous formic acid) using a C₁₈ reversed-phase analytical column (Thermo Fisher scientific EASY column). The custom composite protein database was established based on the S-CREM1 ORF amino acid sequences, and the Mascot 2.4 software (Matrix Science, London, UK) was used to search against the database to analyze the mass spectrometry (MS) data.

2.9. Codon Usage (CU) and Relative Synonymous Codon Usage (RSCU) Analyses

To evaluate the potential contribution of tRNA genes to the phage gene translation efficiency, the CUs of S-CREM1 and *Synechococcus* sp. CB0101 were analyzed using Count-codon v4 (http://www.kazusa.or.jp/codon/countcodon.html (accessed on 10 October 2022)). In addition, the RSCU analysis was performed to investigate whether the tR-NAs of S-CREM1 and *Synechococcus* sp. CB0101 matched the most used codons in their genomes. The RSCU value is the ratio of the usage frequency of a specific codon to all expected synonymous codons in amino acid synonymous codons [46]. RSCU values >1 or <1 indicate that the CU frequency is higher or lower than expected. The RSCU values of S-CREM1 and *Synechococcus* sp. CB0101 were calculated using the CodonW v1.4.2 software (https://sourceforge.net/projects/codonw/ (accessed on 18 October 2022)).

2.10. Motifs Prediction of sRNA and cis-Regulatory RNA Genes

DNA sequence conserved overlapping motifs of sRNA and *cis*-regulatory RNA genes were predicted by MEME Suite 5.5.0 (https://meme-suite.org/meme/doc/meme.html (accessed on 10 October 2022)) [47]. A total of 145 sequences of the *abiF* sRNA identified with solid bit scores (>40) in the Rfam database were selected for the sRNA motif analysis. The sequences of *wcaG*, *manA*, and *glnA cis*-regulatory RNAs identified in cyanophage genomes were selected for conserved motif analyses.

3. Results and Discussion

3.1. General Features of S-CREM1

Cyanophage S-CREM1 which infects *Synechococcus* sp. CB0101 was isolated from the surface seawater of the Changjiang River Estuary (30.8°N, 122.6°E) in July 2019 [21]. S-CREM1 is a myovirus with an isometric icosahedral head (approximately 94 nm in diameter) and a contractile tail (approximately 165 nm in length and 13 nm in width) (Figure 1). The one-step growth curve shows that S-CREM1 has a latent period of 10–12 h and a burst size of 11 (Figure 1). Unlike most of the previously identified cyanomyoviruses that usually had a broad host range [4,48], S-CREM1 only infected the original host strain CB0101, while it had no infectivity on other tested *Synechococcus* strains isolated from similar or distinct environments as CB0101 (Table 1). The *Synechococcus* strains used for the host range test include estuarine, coastal and oceanic isolates. The cross-infectivity of S-CREM1 resembles that of another cyanophage S-SZBM1, which also has a narrow host range [49].

Table 1. Host range of cyanophage S-CREM1.

Tested Strain	Phylogenetic Clade	Isolation Source	Medium Salinity	Reference	Result ^a
Estuarine strains					
Synechococcus sp. CB0101	CB4, subcluster 5.2	Chesapeake Bay	15	[21]	+
<i>Synechococcus</i> sp. A10-1-5-1	CB5, subcluster 5.2	Changjiang River Estuary	25	Xu et al. unpublished	_
<i>Synechococcus</i> sp. CBW1003	Bornholm Sea	Chesapeake Bay	15	[50]	_
<i>Synechococcus</i> sp. CBW1006	Bornholm Sea	Chesapeake Bay	15	[50]	_
<i>Synechococcus</i> sp. CBW1107	Subalpine C II	Chesapeake Bay	15	[50]	_
<i>Synechococcus</i> sp. CBW1004	Unclassified	Chesapeake Bay	15	[50]	_
Synechococcus sp. PCC 7002	Unclassified	Magueyes Island	22	[51]	_
Marine strains					

Tested Strain	Phylogenetic Clade	Isolation Source	Medium Salinity	Reference	Result ^a
Synechococcus sp. CC9311	I, subcluster 5.1	California Current	35	[52]	_
Synechococcus sp. WH 8102	III, subcluster 5.1	Tropical Atlantic	30	[53]	_
Synechococcus sp. WH 7803	V, subcluster 5.1	Sargasso Sea	15	[29]	_
<i>Synechococcus</i> sp. WH 7805	VI, subcluster 5.1	Sargasso Sea	35	[29]	_

Table 1. Cont.

^a +, infected; –, uninfected.





Figure 1. The morphology and growth of S-CREM1. (**A**) Transmission electron microscopy images of S-CREM1 with a contractile sheath (**left**) and a complete tail (**right**). (**B**) One-step growth curve of S-CREM1.

3.2. Genomic Features of S-CREM1 and Proposal of a New Viral Genus

The genome of S-CREM1 is assembled into a circularly permuted DNA molecule, with a length of 177,957 bp and G + C content of 39.7%. A total of 220 open reading frames (ORFs), 24 tRNA genes, one small RNA (sRNA) gene, and three *cis*-regulatory RNA genes are predicted in the genome of S-CREM1 (Figure 2, Tables S1 and S2). Among the 220 ORFs, 114 ORFs have predictable functions, and 27 ORFs have no homologs in the NR database. The 114 ORFs with predictable functions in S-CREM1 can be divided into four categories, i.e., DNA replication and metabolism (26 ORFs), structure and packaging (29 ORFs), regulation (54 ORFs), and lysis (five ORFs) (Figure 2), accounting for 13%, 29%, 21.5%, and 4.4% of the genome size, respectively. Among the 220 ORFs of S-CREM1 is a

member of T4-like cyanophages (Table S1). The phylogenomic analysis among S-CREM1 and 45 T4-like cyanophages based on the 30 core genes showed that S-CREM1 formed a new clade with *Synechococcus* phage S-H38 which was isolated from the Yellow Sea, China (Figure 3). A total of 14 S-CREM1-encoded proteins were detected in the virion proteome by mass spectrometry. Of the 14 phage proteins, four were related to viral structure, including the baseplate, major capsid, and tail proteins (Figure 2). Of the ten remaining proteins, nine have unknown functions and one has no matches in the NR database (Table S1). ORF157 is predicted to be a distant homolog of lipoprotein lipase (Table S2). Lipoprotein lipase is known to play an important role in systemic lipid partitioning and metabolism [54], which may be involved in the conversion of triacylglycerol to diacylglycerol in host glycerolipid metabolism. Since this protein is unlikely to be a structural protein, it could be a highly expressed functional protein carried by the virions. Functional proteins encapsulated by virions have been reported in some cyanophages, which may be important for cyanophage infection [8,22,49].



Figure 2. Genome organization of S-CREM1. ORFs with different predicted functional categories are shown in different colors. The direction of ORF transcription is indicated by an arrow. The tRNA, sRNA, and *cis*-regulatory RNA genes are marked underneath the ORF bar. The acronym of GIY-YIG stands for GlyIleTyr–TyrIleGly; ssDNA, single-stranded DNA; 2OG, 2-oxoglutarate; NMNAT, nicotinamide/nicotinate mononucleotide adenylyltransferase. Virion proteins detected in the virion proteome by mass spectrometry analysis are indicated by red dashed frames.

The genomic nucleotide sequence similarities calculated by VIRIDIC between S-CREM1 and the most closely related 20 cyanophages in the NCBI Viral RefSeq database were 19.9–35.9% (Figure 4). According to the recognized virus naming and classification guide, the same genus viruses should share >50% nucleotide sequence similarity [55]. Therefore, we propose that S-CREM1 represents a new cyanophage genus and name it *Symyovirus*. The new genus *Symyovirus* has been submitted to ICTV.



0.05

Figure 3. The maximum-likelihood phylogenomic tree based on the 30 core genes among S-CREM1 and 45 T4-like cyanophages. Bootstrap values are calculated based on 100 replicates (only values > 50% are shown).

3.3. Various and Unique AMGs in the S-CREM1 Genome

3.3.1. Cyanophage-Featured AMGs

The AMGs shared by most cyanophages are present in the S-CREM1 genome, such as photosynthesis-related genes *hli* (ORF19), *psbA* (ORF21), and *speD* (ORF55), carbon metabolism-related gene CP12 (ORF50), and phosphorus-acquisition-related genes *phoH* (ORF35) and *mazG* (ORF216). These phage-encoded photosynthesis-related genes may maintain and enhance the host photosynthetic activity during the viral infection, thus providing a fitness advantage for viral replication and production [9,56]. Protein CP12 is an inhibitor of the Calvin cycle, and the expression of phage CP12 during infection will inhibit the Calvin cycle of the host and direct carbon flux from glucose synthesis to the pentose phosphate pathway, which will retain ATP and NADPH for the viral replication process [10,57]. In addition, S-CREM1 also encodes a tryptophan halogenase (PrnA) that is commonly found in other cyanophage genomes [7]. PrnA (ORF69) catalyzes free tryptophan to chlorotryptophan, which is the first step of antibiotic pyrrolnitrin biosynthesis [58], and it may provide antibiotic protection to the host during S-CREM1 infection. The S-

CREM1 genome contains ten ORFs encoding 2OG-Fe(II) oxygenase superfamily proteins (Table S1 and S2), which are predicted to function in DNA repair, protein modification, and lipid metabolism [59]. Moreover, 2OG-Fe(II) oxygenase superfamily proteins are widely distributed in cyanophage genomes, usually ranging from one to five in number [7]. In particular, 24 2OG-Fe(II) oxygenase superfamily protein genes belonging to four subfamilies were predicted in the genome of cyanophage S-SCSM1 [8]. The ten 2OG-Fe(II) oxygenase superfamily protein genes of S-CREM1 can be divided into three subfamilies based on the conserved domains, TIGR02466, pfam13759, and pfam13640, and they show low amino acid sequence identity (0–41.2%) with each other (Table S3, Figure S1). In addition, the ten ORFs clustered into nine clades with cyanophage or heterotrophic bacterial sequences in the phylogenetic analyses (Figure S1), indicating their divergence and potential diverse functions in reprogramming host metabolisms during phage infection.

																					- 200 КБ Се - 100 КБ Ре - 0 КБ Ре
100	16.4	21.9	16.6	20.6	35.9	20.1	19.9	22.4	22.3	22.3	22.0	21.6	22.9	22.2	21.2	21.3	21.1	22.4	24.4	24.3	Synechococcus phage S-CREM1
16.4	100	30.0	30.8	15.6	16.6	17.6	19.5	17.6	18.0	18.1	18.0	18.3	17.5	18.5	20.2	18.1	18.9	18.9	18.6	18.6	Cyanophage P-TIM40
21.9	30.0	100	<mark>53.</mark> 2	21.2	20.1	18.7	19.0	19.1	20.4	20.4	19.7	20.9	21.0	20.1	18.9	19.3	19.7	20.5	20.3	20.3	Synechococcus phage S-SZBM1
16.6	30.8	53.2	100	18.1	19.6	20.8	20.6	22.9	23.7	23.7	21.0	22.1	22.6	21.8	20.7	21.6	21.1	23.7	22.6	22.6	Synechococcus phage S-WAM1
20.6	15.6	21.2	18.1	100	19.3	17.6	17.4	20.4	20.4	20.5	18.6	18.8	19.9	18.9	18.4	18.9	18.6	20.8	18.3	18.3	Synechococcus phage S-CBM2
35.9	16.6	20.1	19.6	19.3	100	23.0	24.8	24.7	25.0	25.0	24.2	26.1	24.9	26.8	23.5	23.6	23.4	26.1	26.2	26.1	Synechococcus phage S-H38
20.1	17.6	18.7	20.8	17.6	23.0	100	32.7	32.8	32.8	32.8	38.2	35.4	35.0	35.3	35.4	38.0	38.1	34.3	34.6	34.5	Cyanophage S-SSM2
19.9	19.5	19.0	20.6	17.4	24.8	32.7	100	47.6	48.4	48.5	50.2	50.8	49.4	56.1	42.1	50.2	50.0	44.1	46.3	46.3	Cyanophage Syn30
22.4	17.6	19.1	22.9	20.4	24.7	32.8	47.6	100	77.0	77.1	51.1	53.9	59.8	57.1	44.8	53.6	50.3	52.6	47.2	47.2	Cyanophage S-RIM50
22.3	18.0	20.4	23.7	20.4	25.0	32.8	48.4	77.0	100	99.9	52.3	55.0	61.1	58.3	45.5	53.5	51.2	54.4	47.5	47.4	Cyanophage KBS-M-1A
22.3	18.1	20.4	23.7	20.5	25.0	32.8	48.5	77.1	99.9	100	52.3	55.1	61.1	58.3	45.6	53.5	51.2	54.6	47.5	47.4	Synechococcus phage S-RIM8 A.HR1
22.0	18.0	19.7	21.0	18.6	24.2	38.2	50.2	51.1	52.3	52.3	100	66.5	60.7	57.9	53.4	65.9	63.6	48.5	47.3	47.2	Prochlorococcus phage Syn33
21.6	18.3	20.9	22.1	18.8	26.1	35.4	50.8	53.9	55.0	55.1	66.5	100	63.2	59.9	45.7	58.0	53.0	48.6	48.0	48.0	Synechococcus phage S-RIM2
22.9	17.5	21.0	22.6	19.9	24.9	35.0	49.4	<u>59.8</u>	61.1	61.1	60.7	63.2	100	60.7	44.9	55.9	51.8	53.7	47.7	47.7	Synechococcus phage S-H25
22.2	18.5	20.1	21.8	18.9	26.8	35.3	56.1	57.1	58.3	58.3	57.9	59.9	60.7	100	47.6	57.3	55.3	50.4	52.0	51.8	Synechococcus phage S-IOM18
21.2	20.2	18.9	20.7	18.4	23.5	35.4	42.1	44.8	45.5	45.6	<u>53.4</u>	45.7	44.9	47.6	100	54.4	56.6	50.2	47.6	47.6	Synechococcus phage ACG-2014c
21.3	18.1	19.3	21.6	18.9	23.6	38.0	50.2	53.6	53.5	53.5	65.9	58.0	55.9	57.3	54.4	100	62.9	48.1	47.6	47.4	Synechococcus phage S-CAM8
21.1	18.9	19.7	21.1	18.6	23.4	38.1	50.0	50.3	51.2	51.2	63.6	53.0	51.8	55.3	56.6	62.9	100	46.3	46.4	46.2	Synechococcus phage S-MbCM100
22.4	18.9	20.5	23.7	20.8	26.1	34.3	44.1	52.6	54.4	54.6	48.5	48.6	53.7	50.4	50.2	48.1	46.3	100	53.5	53.6	Synechococcus phage S-H35
24.4	18.6	20.3	22.6	18.3	26.2	34.6	46.3	47.2	47.5	47.5	47.3	4 8.0	47.7	52.0	47.6	47.6	46.4	53.5	100	99.8	Cyanophage Syn10
24.3	18.6	20.3	22.6	18.3	26.1	34.5	46.3	47.2	47.4	47.4	47.2	4 8.0	47.7	51.8	47.6	47.4	46.2	53.6	99.8	100	Synechococcus phage syn9
Synechococcus phage S-CREM1	Cyanophage P-TIM40	Synechococcus phage S-SZBM1	Synechococcus phage S-CBM2	Synechococcus phage S-CBM2	Synechococcus phage S-H38	Cyanophage S-SSM2	Cyanophage Syn30	Cyanophage S-RIM50	Cyanophage KBS-M-1A	Synechococcus phage S-RIM8 A.HR1	Prochlorococcus phage Syn33	Synechococcus phage S-RIM2	Synechococcus phage S-H25	Synechococcus phage S-IOM18	Synechococcus phage ACG-2014c	Synechococcus phage S-CAM8	Synechococcus phage S-MbCM100	Synechococcus phage S-H35	Cyanophage Syn10	Synechococcus phage syn9	

Figure 4. Intergenomic similarities among S-CREM1 and 20 cyanophages calculated using VIRIDIC. Genomic similarities between S-CREM1 and the 20 cyanophages are displayed in bold.

3.3.2. Three Antitoxin Genes

Notably, the S-CREM1 genome encodes three antitoxin genes, YefM (ORF54), TacA (ORF106), and MazE (ORF155) (Table 2). Toxin–antitoxin (TA) systems are genetic modules

consisting of a stable toxin and an unstable antitoxin, which are widespread in many bacteria [60]. Toxin and antitoxin generally exist in pairs, in which toxin may inhibit cell growth or cause cell death, while antitoxin forms stable complexes with a toxin to prevent the toxin from exerting toxicity [61]. TA systems have been proven to play critical roles in protecting bacteria against phage infections [60]. Phage infection will lead to the release of toxins from some TA systems, which can kill the host cell and inhibit phage replication [60]. In addition, the TA system is associated with stress responses in microbes and is a successful survival strategy under various environmental stresses [62]. Many more TA systems are found in freshwater and coastal *Synechococcus* genomes than in open ocean *Synechococcus* strains [63].

ORF No.	Antitoxin	TA System	TA System in the Host Genome
54	YefM	YefM-YoeB	+
106	TacA	TacA–TacT	_
155	MazE	MazE–MazF	+

Table 2. Three antitoxin genes predicted in the S-CREM1 genome.

The host strain, *Synechococcus* sp. CB0101, was isolated from the Chesapeake Bay [30]. The genome sequencing of CB0101 led to a discovery of wide presence of TA systems in Synechococcus [63,64]. The enrichment of TA genes in freshwater and estuarine Synechococcus compared to the coastal and ocean counterparts suggests that the inheritance of TA genes helps Synechococcus better adapt to changing environments and resist the infection of cyanophages [62,63]. The two Type II TA Pairs, YefM–YoeB and MazE–MazF, are present and expressed in host CB0101 [62], while YefM and MazE in S-CREM1 and host CB0101 have no homology. MazE-MazF is the first TA system described as capable of regulating or causing programmed bacterial death [65]. Phage encoding antitoxin or antitoxin mimics may be the simplest way to overcome TA-mediated defense [60]. Encoding antitoxin genes (YoeB, MazF, and TacA) by phage S-CREM1 may be a mechanism to counteract the TA system of CB0101, which has the potential to reduce the virulent infection to a certain degree so that host cells can survive longer on behalf of phage. It is plausible that antitoxins encoded by S-CREM1 may bind to free toxins in host cells during infection, preventing host cells from being killed to facilitate the replication of S-CREM1. Interestingly, a pair of type II TA genes was found in cyanophage vB_AphaS-CL131, which infects filamentous diazotrophic cyanobacterium and was predicted to play a role in preventing the host from entering dormancy and ensuring the continuous replication of phages [66]. The role of phage-encoded antitoxin genes is interesting and should warrant further investigation.

3.3.3. A MoxR Family ATPase Gene

The S-CREM1 genome carries a gene (ORF198) encoding a homolog of MoxR family ATPase, which is widely found in various prokaryotic species [67,68]. MoxR proteins are important regulators of multiple stress response pathways and are proven to function under acid, oxidative, and heat stresses in different heterotrophic bacteria [69,70]. In addition, the MoxR ATPase has also been found to function as a chaperone and play a role in tail development of *Acidianus* two-tailed virus [71]. The MoxR ATPases are classified into seven major subfamilies: MRP, CGN, APE2220, PA2707, RavA, TM0930, and YehL [68]. The S-CREM1 MoxR ATPase clustered into the CGN subfamily with other cyanophages, bacteriophage, and heterotrophic bacteria, while the cyanobacterial MoxR ATPases fell into the MRP subfamily in the phylogenetic analyses (Figure S2), indicating the different MoxR ATPase gene pool of evolution between cyanophages and their hosts.

3.3.4. Overlooked Pyrimidine Dimer DNA Glycosylase Genes in Cyanophages

S-CREM1 ORF91 is predicted to be a pyrimidine dimer DNA glycosylase by searching the Conserved Domain database. The pyrimidine dimer induced by ultraviolet (UV) is the most common mechanism causing DNA damage in microbes [72]. Pyrimidine dimer DNA

glycosylase functions as a base excision repair protein by digesting the pyrimidine dimer of the damaged DNA molecule through hydrolyzing the glycosylic bond of the 5' pyrimidine and the phosphodiester bond of intra-pyrimidine [73–75]. Although pyrimidine dimer DNA glycosylases are frequently found in bacteriophages, only a few pyrimidine dimer DNA glycosylases have been found in isolated cyanophages. There are many homologous sequences of ORF91 in the NCBI NR database (amino acid identity 55.4–64.6%), which were previously predicted as hypothetical proteins (Figure S3). The high similarities of the ORF91 homologs predicted in the genomes of cyanophage isolates with pyrimidine dimer DNA glycosylase in the Conserved Domain database showed that these ORFs may be pyrimidine dimer DNA glycosylases that were previously overlooked. Therefore, we suggest that pyrimidine dimer DNA glycosylase is widely encoded in the cyanophage genomes and plays a role in the repair of damaged DNA during infection. The close phylogenetic relationship of the pyrimidine dimer DNA glycosylase gene among cyanophages and heterotrophic bacteria (Figure S3) indicated that the cyanophage pyrimidine dimer DNA glycosylase genes may exchange with homologs from heterotrophic bacteria rather than their cyanobacterial hosts.

3.4. A Large Number of tRNA Genes

Twenty-four bona fide tRNA genes were identified in the S-CREM1 genome, including all but tRNA^{Cys} amino acid specificities (Table S4). Although tRNA genes are frequently found in cyanophage genomes, only S-CREM1 and three cyanophages, S-PM2, S-CBWM1, and S-CRM01, contain more than 20 tRNA genes (Table S5) [22–24]. S-CRM01 and S-CBWM1 contain a full set of tRNAs, and S-PM2 only lacks tRNA genes for Cys and Phe [22–24]. With the exception of the tRNA gene for the TGA anticodon, S-CREM1 contains the same type of tRNA in the other three cyanophages (Table S6). The presence of a nearly full set of tRNA genes in these four cyanophage genomes could be important to the protein synthesis of both the phage and host (Table S6).

Phage protein synthesis and reproduction are highly dependent on host transcriptional and translational mechanisms. The reason for phage-carrying tRNAs has always been of interest. A study by Enav et al. revealed that the tRNAs in cyanophages may help to overcome the limitation of cyanophage translation caused by the differences in G + Ccontent among different hosts [15]. In addition, Yang et al. suggest that host tRNA will be degraded upon infection, and then, the tRNA of the phage will supplement the host tRNA function to sustain translation [16]. In addition to the tRNA as anticodon TAT, tRNA genes of the same type as those in S-CREM1 are present in the genome of *Synechococcus* sp. CB0101. If the tRNA genes in S-CREM1 are functional, they may play a more efficient role in supplementing the function of the host tRNAs during phage infection. The comparison of CU and RSCU between S-CREM1 and Synechococcus sp. CB0101 showed that S-CREM1 and the host prefer to use different codons (Table 3, Figure 5). Among the 23 codons corresponding to tRNA types shared by S-CREM1 and CB0101, 17 are used more frequently in S-CREM1 than in its Synechococcus host, which indicates that S-CREM1 selects and retains tRNAs to compensate for the codon usage difference with its host and facilitate the translation of its own genes [76]. Furthermore, phage tRNAs are found to play roles in regulating translation, packaging, and initiating reverse transcription during infection [77]. It is noteworthy that cyanophages, as well as bacteriophages infecting heterotrophic bacteria with numerous tRNA genes (i.e., >20), are mainly isolated from environments with high nutrient levels (coastal seawater, estuarine water, freshwater, wastewater, and soil) [22]. A large number of tRNA genes in the S-CREM1 genome might be a genomic adaption to the eutrophic environment, enabling S-CREM1 to increase its fitness in the estuary. It would be interesting to know how environmental eutrophication and the growth status of hosts affect the viral possession of tRNA genes. The metabolic activity of prokaryotes is generally higher in the eutrophic environment than that in the oligotrophic habitat, and the expression of phage tRNA genes likely increases the translation efficiency during infection with sufficient nutrients available [22,78].

Codon ^a	Attribute	Usage	e (‰)	RSC	CU	Codon ^a	Attribute	Usage	e (‰)	RSC	CU	Codon ^a	Attribute	Usag	e (‰)	RS	CU	Codon ^a	Attribute	Usage	(‰)	RSC	CU
		Phage	Host	Phage	Host			Phage	Host	Phage	Host			Phage	Host	Phage	Host			Phage	Host	Phage	Host
TTT	Phe	18.7	6.3	0.9	0.7	TCT	Ser	22.8	9.2	1.6	0.6	TAT	Tyr	23.6	2.5	1.3	0.7	TGT	Cys	12.1	8.5	1.1	0.5
TTC	Phe	21.7	11.8	1.1	1.3	TCC	Ser	10.5	14.1	0.7	0.9	TAC	Tyr	13.9	4.2	0.7	1.3	TGC	Cys	10.9	24.4	1	1.5
TTA	Leu	17.7	2.1	1	0.2	TCA	Ser	21.6	15.3	1.5	1	TAA	Stop	17.2	1.9	1.1	0.3	TGA	Stop	19.7	15.3	1.3	2.2
TTG	Leu	20.8	12.8	1.2	1.1	TCG	Ser	9.5	17.6	0.7	1.2	TAG	Stop	10	3.5	0.6	0.5	TGG	Trp	17.1	25	1	1
CTT	Leu	18.1	11.3	1	0.9	CCT	Pro	11.6	18.8	1.4	0.8	CAT	His	13.9	12	1.2	0.7	CGT	Arg	6.1	11.9	0.6	0.6
CTC	Leu	11.7	17.5	0.7	1.4	CCC	Pro	4.8	21.9	0.6	0.9	CAC	His	10.1	21.3	0.8	1.3	CGC	Arg	3.2	31.9	0.3	1.6
CTA	Leu	19.9	4	1.1	0.3	CCA	Pro	13.3	27.3	1.6	1.1	CAA	Gln	27.6	14.9	1.2	0.7	CGA	Arg	7.7	19.4	0.8	1
CTG	Leu	19.8	29.1	1.1	2.3	CCG	Pro	3.1	27.7	0.4	1.2	CAG	Gln	16.8	29.2	0.8	1.3	CGG	Arg	3.1	27.7	0.3	1.4
ATT	Ile	26.1	5.5	1.2	0.6	ACT	Thr	24.4	6.3	1.6	0.6	AAT	Asn	24.3	5.5	1	0.6	AGT	Ser	16.7	5.9	1.1	0.4
ATC	Ile	23.2	18.5	1.1	2.1	ACC	Thr	10.7	18.4	0.7	1.6	AAC	Asn	22.3	11.5	1	1.4	AGC	Ser	7.3	27.9	0.5	1.9
ATA	Ile	16.6	2.5	0.8	0.3	ACA	Thr	18.5	9	1.2	0.8	AAA	Lys	26.9	6.8	1	0.7	AGA	Arg	26.6	9.6	2.8	0.5
ATG	Met	23.7	11.1	1	1	ACG	Thr	7.2	11.4	0.5	1	AAG	Lys	26.1	11.2	1	1.3	AGG	Arg	10.4	17.4	1.1	1
GTT	Val	20.3	11	1.3	1	GCT	Ala	12.2	28.6	1.3	0.9	GAT	Asp	26.8	19.3	1.4	1.4	GGT	Gly	23.1	18	1.8	0.8
GTC	Val	8.2	8.4	0.5	0.8	GCC	Ala	2.8	39.4	0.3	1.3	GAC	Asp	11	9.1	0.6	0.7	GGC	Gly	4.6	39.4	0.4	1.7
GTA	Val	17.8	4	1.2	0.4	GCA	Ala	17.7	26	1.9	0.8	GAA	Glu	28.6	12.5	1.3	0.8	GGA	Gly	17	14.7	1.4	0.6
GTG	Val	14.2	20.3	0.9	1.9	GCG	Ala	4.1	31.2	0.4	1	GAG	Glu	16.3	17.6	0.7	1.2	GGG	Glv	5.8	21.8	0.5	0.9

Table 3. CU and RSCU in S-CREM1 and Synechococcus sp. CB0101.

^a tRNA genes in S-CREM1 are denoted by their complementary codons in red font, and tRNA genes in *Synechococcus* sp. CB0101 are denoted by their complementary codons in black bold font.





3.5. One Small RNA (sRNA) and Three cis-Regulatory RNA Genes

One small RNA (sRNA) and three *cis*-regulatory RNA genes were identified in the S-CREM1 genome by searching against the Rfam database (Table 4). As major regulatory molecules in bacteria, sRNAs and *cis*-regulatory RNAs play important roles in nutrient uptake and metabolism [79], iron regulation [80], protein synthesis, RNA processing [17], biofilm matrix formation [81], and quorum sensing [82]. The *abiF* sRNA predicted in the S-CREM1 genome was not previously identified in cyanophage genomes in the Rfam database. The S-CREM1 abiF gene shares a conserved motif with 139 abiF genes that are identified from a variety of bacteria and three bacteriophage genomes in the Rfam database (Figure 6), indicating the high conservation of the *abiF* sRNA gene among microbial organisms and the high possibility that the S-CREM1 abiF sRNA is functional during phage infection. The three cis-regulatory elements predicted in the S-CREM1 genome are wcaG, manA, and glnA, which were also identified in other cyanophage genomes (Table S7). The wcaG, manA and glnA genes between S-CREM1 and other cyanophages share similar patterns of conserved motifs (Figure S4). cis-regulatory elements function as environmental change detectors, such as light or temperature variations, and regulate the message stability or translational efficiency of specific genes. The wcaG RNA may regulate the expression of genes related to the production of exopolysaccharides [17]. The manA RNA domains are usually located in the potential 5' untranslated regions of the genes related to nucleotide synthesis, mannose or fructose metabolism, and photosynthesis. Whether the manA RNA plays a regulatory role in the expression of these genes remains to be further verified [17]. The *glnA* RNA can regulate the expression of genes related to nitrogen metabolism, such as genes encoding nitrogen regulatory protein PII, glutamine synthetase, glutamate synthase, and ammonium transporters [17,83].

Table 4. One sRNA and three *cis*-regulatory RNA genes in the S-CREM1 genome.

Feature	Туре	Strand	Start	End	Rfam Accession No.	Score	E-Value
abiF	sRNA	+	10531	10567	RF03085	51.6	3.2E-06
wcaG	cis-regulatory	+	9745	9842	RF01761	91.4	2.3E-15
manA	cis-regulatory	+	10215	10429	RF01745	100.6	4.6E-23
glnA	cis-regulatory	_	164330	164232	RF01739	44.0	5.2E-06

Figure 6. The DNA sequence conserved motif in the *abiF* RNA genes among S-CREM1 and 139 sequences identified from 136 bacterial and three bacteriophage genomes in the Rfam database.

sRNA and *cis*-regulatory RNA have been frequently studied in bacteria and archaea. However, little is known about the types and functions of those regulatory elements in phages, especially in cyanophages. The phage–host system of S-CREM1-*Synechococcus* sp. CB0101 provides a good model for future research on the regulatory functions of phage-encoded sRNAs and *cis*-regulatory RNAs during viral infection.

4. Conclusions

Based on the phylogenomic analysis and comparative genomics, we proposed a new genus, *Symyovirus*, for cyanophage S-CREM1 which infects an estuarine *Synechococcus*. S-CREM1 exhibits several interesting genetic features including the possession of three antitoxin genes, the MoxR family ATPase and the pyrimidine dimer DNA glycosylase genes. We reported the presence of three antitoxin genes in the S-CREM1 genome and proposed a potential beneficial role of having antitoxin genes in cyanophages. The presence of a large number of tRNA genes suggests that S-CREM1 may have the capability to thrive in a nutrient-rich estuarine environment. The sRNA gene and three *cis*-regulatory RNA genes suggest that S-CREM1 has other functions in regulating host metabolism during infection. The isolation of cyanophage S-CREM1 and its genomic characterization provide new insights into phage taxonomy, evolution, and phage–host interactions.

Supplementary Materials: The following supporting information can be downloaded at: https://www.action.com/actionals //www.mdpi.com/article/10.3390/v15020380/s1, Figure S1: Unrooted maximum likelihood phylogenetic tree of the S-CREM1 predicted 2OG-Fe(II) oxygenase superfamily proteins. ORFs in brown font were identified based on predicted structural properties using HHpred and Phyre analyses. Bootstrap values (maximum-likelihood/neighbor-joining) are based on 1000 replicates, and only values > 50% are shown; Figure S2: Unrooted maximum likelihood phylogenetic tree of the MoxR ATPases. The number of bootstrap replicates = 1000. The bootstrap values (maximumlikelihood/neighbor-joining) are shown near each node, and only values >50% are shown; Figure S3: Unrooted maximum-likelihood phylogenetic tree of the S-CREM1 pyrimidine dimer DNA glycosylase. The bootstrap values (maximum-likelihood/neighbor-joining) of >50% are shown near each node. The number of bootstrap replicates = 1000; Figure S4: DNA sequence conserved motifs in wcaG(A), manA (B), and glnA (C) cis-regulatory RNA genes among S-CREM1 and other cyanophages in the Rfam database; Table S1: Predicted ORFs in the S-CREM1 genome with homologs in the NCBI nonredundant (NR) database; Table S2: Predicted ORFs in the S-CREM1 genome with distant homologs detected by using HHpred and Phyre2 search; Table S3: Amino acid sequence identities among the S-CREM1 2OG-Fe(II) oxygenase family proteins; Table S4: tRNA genes in the S-CREM1 genome; Table S5: Information of cyanophages with tRNA genes more than twenty; Table S6: Comparison of tRNA complementary codons among cyanophages with tRNA genes more than twenty; Table S7: wcaG, manA, and glnA cis-regulatory RNA genes found in cyanophages.

Author Contributions: H.Z., Y.L., and R.Z. performed the experiments and data analyses; H.Z. wrote the first draft of the manuscript; Y.X. revised the manuscript; Y.X., F.C., and J.L. contributed to the preparation of the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key Research and Development Program of China (2018YFA0605800), the National Natural Science Foundation of China (41706161), the Key Research and Development Program of Shandong Province (2020ZLYS04), and Joint Funds of the

National Natural Science Foundation of China (U1906216). The article processing charges (APC) were funded by 2018YFA0605800 in this section.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The complete genome sequence of *Synechococcus* phage S-CREM1 was submitted to the GenBank database under accession number OP535465.

Acknowledgments: We would like to thank Xiaoju Li and Haiyan Sui from Shandong University Core Facilities for Life and Environmental Sciences for their help with the transmission electron microscopy (TEM) observation. We thank Xuejin Feng for the help with data analysis. We thank Rui Wang and Zefeng Zhang for valuable suggestions.

Conflicts of Interest: The authors declare no competing interests.

References

- Partensky, F.B.J.; Vaulot, D. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: A review. *Bull. Inst. Océanogr.* 1999, 19, 457–475.
- Callieri, C.; Cronberg, G.; Stockner, J.G. Freshwater Picocyanobacteria: Single cells, Microcolonies and Colonial Forms. In *Ecology* of Cyanobacteria II; Whitton, B.A., Ed.; Springer: Amsterdam, The Netherlands, 2012; pp. 229–269.
- 3. Affronti, L.F.; Marshall, H.G. Using frequency of dividing cells in estimating autotrophic picoplankton growth and productivity in the Chesapeake Bay. *Hydrobiologia* **1994**, *284*, 193–203. [CrossRef]
- 4. Suttle, C.A. Viruses in the sea. Nature 2005, 437, 356–361. [CrossRef] [PubMed]
- 5. Fuhrman, J.A. Marine viruses and their biogeochemical and ecological effects. Nature 1999, 399, 541–548. [CrossRef]
- 6. Wang, K.; Chen, F. Prevalence of highly host-specific cyanophages in the estuarine environment. *Environ. Microbiol.* **2008**, *10*, 300–312. [CrossRef]
- Sullivan, M.B.; Huang, K.H.; Ignacio-Espinoza, J.C.; Berlin, A.M.; Kelly, L.; Weigele, P.R.; DeFrancesco, A.S.; Kern, S.E.; Thompson, L.R.; Young, S.; et al. Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environ. Microbiol.* 2010, *12*, 3035–3056. [CrossRef]
- 8. Wang, Q.; Cai, L.; Zhang, R.; Wei, S.; Li, F.; Liu, Y.; Xu, Y. A unique set of auxiliary metabolic genes found in an isolated cyanophage sheds new light on marine phage–host interactions. *Microbiol. Spectr.* **2022**, *10*, e02367-22. [CrossRef]
- 9. Lindell, D.; Sullivan, M.B.; Johnson, Z.I.; Tolonen, A.C.; Rohwer, F.; Chisholm, S.W. Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 11013–11018. [CrossRef]
- 10. Thompson, L.R.; Zeng, Q.; Kelly, L.; Huang, K.H.; Singer, A.U.; Stubbe, J.; Chisholm, S.W. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, E757–E764. [CrossRef]
- 11. Dwivedi, B.; Xue, B.; Lundin, D.; Edwards, R.A.; Breitbart, M. A bioinformatic analysis of ribonucleotide reductase genes in phage genomes and metagenomes. *BMC Evol. Biol.* **2013**, *13*, 33. [CrossRef]
- 12. Enav, H.; Mandel-Gutfreund, Y.; Béjà, O. Comparative metagenomic analyses reveal viral-induced shifts of host metabolism towards nucleotide biosynthesis. *Microbiome* **2014**, *2*, 9. [CrossRef]
- 13. Kelly, L.; Ding, H.; Huang, K.H.; Osburne, M.S.; Chisholm, S.W. Genetic diversity in cultured and wild marine cyanomyoviruses reveals phosphorus stress as a strong selective agent. *ISME J.* 2013, 7, 1827–1841. [CrossRef]
- 14. Crummett, L.T.; Puxty, R.J.; Weihe, C.; Marston, M.F.; Martiny, J.B. The genomic content and context of auxiliary metabolic genes in marine cyanomyoviruses. *Virology* **2016**, *499*, 219–229. [CrossRef]
- 15. Enav, H.; Béjà, O.; Mandel-Gutfreund, Y. Cyanophage tRNAs may have a role in cross-infectivity of oceanic *Prochlorococcus* and *Synechococcus* hosts. *ISME J.* **2012**, *6*, 619–628. [CrossRef]
- 16. Yang, J.Y.; Fang, W.; Miranda-Sanchez, F.; Brown, J.M.; Kauffman, K.M.; Acevero, C.M.; Bartel, D.P.; Polz, M.F.; Kelly, L. Degradation of host translational machinery drives tRNA acquisition in viruses. *Cell Syst.* **2021**, *12*, 771–779.e5. [CrossRef]
- 17. Weinberg, Z.; Wang, J.X.; Bogue, J.; Yang, J.; Corbino, K.; Moy, R.H.; Breaker, R.R. Comparative genomics reveals 104 candidate structured RNAs from bacteria, archaea, and their metagenomes. *Genome Biol.* **2010**, *11*, R31. [CrossRef]
- Griffiths-Jones, S.; Moxon, S.; Marshall, M.; Khanna, A.; Eddy, S.R.; Bateman, A. Rfam: Annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* 2005, 33, D121–D124. [CrossRef]
- 19. Field, C.B.; Behrenfeld, M.J.; Randerson, J.T.; Falkowski, P. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* **1998**, *281*, 237–240. [CrossRef]
- Sun, M.; Zhan, Y.; Marsan, D.; Páez-Espino, D.; Cai, L.; Chen, F. Uncultivated viral populations dominate estuarine viromes on the spatiotemporal scale. *mSystems* 2021, 6, e01020-20. [CrossRef]
- 21. Fucich, D.; Marsan, D.; Sosa, A.; Chen, F. Complete genome sequence of subcluster 5.2 *Synechococcus* sp. strain CB0101, isolated from the Chesapeake Bay. *Microbiol. Resour. Announc.* **2019**, *8*, e00484-19. [CrossRef]
- Xu, Y.; Zhang, R.; Wang, N.; Cai, L.; Tong, Y.; Sun, Q.; Chen, F.; Jiao, N. Novel phage–host interactions and evolution as revealed by a cyanomyovirus isolated from an estuarine environment. *Environ. Microbiol.* **2018**, *20*, 2974–2989. [CrossRef] [PubMed]

- Dreher, T.W.; Brown, N.; Bozarth, C.S.; Schwartz, A.D.; Riscoe, E.; Thrash, C.; Bennett, S.E.; Tzeng, S.C.; Maier. C.S. A freshwater cyanophage whose genome indicates close relationships to photosynthetic marine cyanomyophages. *Environ. Microbiol.* 2011, 13, 1858–1874. [CrossRef] [PubMed]
- Mann, N.H.; Clokie, M.R.J.; Millard, A.; Cook, A.; Wilson, W.H.; Wheatley, P.J.; Letarov, A.; Krisch, H.M. The Genome of S-PM2, a "Photosynthetic" T4-Type Bacteriophage That Infects Marine *Synechococcus* Strains. *J. Bacteriol.* 2005, 187, 3188–3200. [CrossRef] [PubMed]
- 25. Wang, M.; Gao, C.; Jiang, T.; You, S.; Guo, C.; He, H.; Liu, Y.; Zhang, X.; Shao, H.; Liu, H.; et al. Genomic analysis of Synechococcus phage S-B43 and its adaption to the coastal environment. *Virus Res.* **2020**, *289*, 198155. [CrossRef]
- Jiang, T.; Guo, C.; Wang, M.; Wang, M.; Zhang, X.; Liu, Y.; Liang, Y.; Jiang, Y.; He, H.; Shao, H.; et al. Genome Analysis of Two Novel *Synechococcus* Phages that Lack Common Auxiliary Metabolic Genes: Possible Reasons and Ecological Insights by Comparative Analysis of Cyanomyoviruses. *Viruses* 2020, 12, 800. [CrossRef]
- 27. Marston, M.F.; Martiny, J. Genomic diversification of marine cyanophages into stable ecotypes. *Environ. Microbiol.* **2016**, *18*, 4240–4253. [CrossRef]
- Stoddard, L.I.; Martiny, J.B.H.; Marston, M.F. Selection and Characterization of Cyanophage Resistance in Marine Synechococcus Strains. Appl. Environ. Microbiol. 2007, 73, 5516–5522. [CrossRef]
- 29. Waterbury, J.; Watson, S.; Valois, F.; Franks, D. Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus. Can. J. Fish. Aquat. Sci.* **1986**, *214*, 71–120.
- Chen, F.; Wang, K.; Kan, J.; Bachoon, D.S.; Lu, J.; Lau, S.; Campbell, L. Phylogenetic diversity of *Synechococcus* in the Chesapeake Bay revealed by ribulose-1, 5-bisphosphate carboxylase-oxygenase (RuBisCO) large subunit gene (rbcL) sequences. *Aquat. Microb. Ecol. Int. J.* 2004, *36*, 153–164. [CrossRef]
- 31. Wilson, W.H.; Joint, I.R.; Carr, N.G.; Mann, N.H. Isolation and molecular characterization of five marine cyanophages propagated on *Synechococcus* sp. strain WH7803. *Appl. Environ. Microbiol.* **1993**, *59*, 3736–3743. [CrossRef]
- 32. Ciglenečki, U.J.; Grom, J.; Toplak, I.; Jemeršić, L.; Barlič-Maganja, D. Real-time RT-PCR assay for rapid and specific detection of classical swine fever virus: Comparison of SYBR Green and TaqMan MGB detection methods using novel MGB probes. *J. Virol. Methods* **2008**, *147*, 257–264. [CrossRef]
- Zheng, H.; Liu, B.; Xu, Y.; Zhang, Z.; Man, H.; Liu, J.; Chen, F. An Inducible *Microbacterium* Prophage vB_MoxS-R1 Represents a Novel Lineage of Siphovirus. *Viruses* 2022, 14, 731. [CrossRef]
- 34. Ma, R.; Lai, J.; Chen, X.; Wang, L.; Yang, Y.; Wei, S.; Jiao, N.; Zhang, R. A Novel Phage Infecting *Alteromonas* Represents a Distinct Group of Siphophages Infecting Diverse Aquatic Copiotrophs. *Msphere* **2021**, *6*, e0045421. [CrossRef]
- 35. Peng, Y.; Leung, H.C.M.; Yiu, S.M.; Chin, F.Y.L. IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* **2012**, *28*, 1420–1428. [CrossRef]
- Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 2015, 10, 845–858. [CrossRef]
- 37. Söding, J.; Biegert, A.; Lupas, A.N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **2005**, *33*, W244–W248. [CrossRef]
- Chan, P.P.; Lowe, T.M. tRNAscan-SE: Searching for tRNA genes in genomic sequences. *Methods Mol. Biol.* 2019, 1962, 1–14. [CrossRef]
- Jang, H.B.; Bolduc, B.; Zablocki, O.; Kuhn, J.H.; Roux, S.; Adriaenssens, E.M.; Brister, J.R.; Kropinski, A.M.; Krupovic, M.; Lavigne, R.; et al. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nat. Biotechnol.* 2019, *37*, 632–639. [CrossRef]
- Moraru, C.; Varsani, A.; Kropinski, A.M. VIRIDIC—A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. *Viruses* 2020, 12, 1268. [CrossRef]
- 41. Emms, D.M.; Kelly, S. OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **2019**, *20*, 238. [CrossRef]
- 42. Katoh, K.; Asimenos, G.; Toh, H. Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* 2009, 537, 39–64. [CrossRef] [PubMed]
- 43. Capella-Gutiérrez, S.; Silla-Martínez, J.M.; Gabaldón, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, *25*, 1972–1973. [CrossRef] [PubMed]
- 44. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, *30*, 1312–1313. [CrossRef] [PubMed]
- 45. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]
- 46. Sharp, P.M.; Tuohy, T.M.; Mosurski, K.R. Codon usage in yeast: Cluster analysis clearly differentiates highly and lowly expressed genes. *Nucleic Acids Res.* **1986**, *14*, 5125–5143. [CrossRef]
- 47. Bailey, T.L.; Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1994**, *2*, 28–36.
- 48. Doron, S.; Fedida, A.; Hernández-Prieto, M.A.; Sabehi, G.; Karunker, I.; Stazic, D.; Feingersch, R.; Steglich, C.; Futschik, M.; Lindell, D.; et al. Transcriptome dynamics of a broad host-range cyanophage and its hosts. *ISME J.* **2015**, *10*, 1437–1455. [CrossRef]

- 49. Rong, C.; Zhou, K.; Li, S.; Xiao, K.; Xu, Y.; Zhang, R.; Yang, Y.; Zhang, Y. Isolation and Characterization of a Novel Cyanophage Encoding Multiple Auxiliary Metabolic Genes. *Viruses* **2022**, *14*, 887. [CrossRef]
- 50. Xu, Y.; Jiao, N.; Chen, F. Novel psychrotolerant picocyanobacteria isolated from Chesapeake Bay in the winter. *J. Phycol.* 2015, *51*, 782–790. [CrossRef]
- 51. Mackey, K.R.; Paytan, A.; Caldeira, K.; Grossman, A.R.; Moran, D.; McIlvin, M.; Saito, M.A. Effect of Temperature on Photosynthesis and Growth in Marine *Synechococcus* spp. *Plant Physiol.* **2013**, *163*, 815–829. [CrossRef]
- 52. Toledo, G.; Palenik, B. Synechococcus diversity in the California current as seen by RNA polymerase (rpoC1) gene sequences of isolated strains. *Appl. Environ. Microbiol.* **1997**, *63*, 4298–4303. [CrossRef]
- 53. Palenik, B.; Brahamsha, B.; Larimer, F.W.; Land, M.; Hauser, L.; Chain, P.; Lamerdin, J.; Regala, W.; Allen, E.E.; McCarren, J.; et al. The genome of a motile marine *Synechococcus*. *Nature* **2003**, *424*, 1037–1042. [CrossRef]
- Wu, S.A.; Kersten, S.; Qi, L. Lipoprotein Lipase and Its Regulators: An Unfolding Story. *Trends Endocrinol. Metab.* 2021, 32, 48–61. [CrossRef]
- 55. Adriaenssens, E.; Brister, J.R. How to Name and Classify Your Phage: An Informal Guide. Viruses 2017, 9, 70. [CrossRef]
- 56. Gao, E.-B.; Huang, Y.; Ning, D. Metabolic Genes within Cyanophage Genomes: Implications for Diversity and Evolution. *Genes* **2016**, *7*, 80. [CrossRef]
- Callaghan, M.M.; Amador-Noguez, D. Hostile Takeover: How Viruses Reprogram Prokaryotic Metabolism. *Annu. Rev. Microbiol.* 2021, 75, 515–539. [CrossRef]
- Karabencheva-Christova, T.G.; Torras, J.; Mulholland, A.J.; Lodola, A.; Christov, C.Z. Mechanistic Insights into the Reaction of Chlorination of Tryptophan Catalyzed by Tryptophan 7-Halogenase. *Sci. Rep.* 2017, 7, 17395. [CrossRef]
- 59. van Staalduinen, L.M.; Jia, Z. Post-translational hydroxylation by 2OG/Fe(II)-dependent oxygenases as a novel regulatory mechanism in bacteria. *Front. Microbiol.* **2015**, *5*, 798. [CrossRef]
- 60. LeRoux, M.; Laub, M.T. Toxin-Antitoxin Systems as Phage Defense Elements. Annu. Rev. Microbiol. 2022, 76, 21–43. [CrossRef]
- 61. Yamaguchi, Y.; Park, J.-H.; Inouye, M. Toxin-Antitoxin Systems in Bacteria and Archaea. *Annu. Rev. Genet.* 2011, 45, 61–79. [CrossRef]
- 62. Marsan, D.; Place, A.; Fucich, D.; Chen, F. Toxin-antitoxin systems in estuarine *Synechococcus* strain CB0101 and their transcriptomic responses to environmental stressors. *Front. Microbiol.* **2017**, *8*, 1213. [CrossRef] [PubMed]
- 63. Fucich, D.; Chen, F. Presence of toxin-antitoxin systems in picocyanobacteria and their ecological implications. *ISME J.* **2020**, *14*, 2843–2850. [CrossRef] [PubMed]
- 64. Marsan, D.; Wommack, K.E.; Ravel, J.; Chen, F. Draft Genome Sequence of *Synechococcus* sp. Strain CB0101, Isolated from the Chesapeake Bay Estuary. *Genome Announc.* **2014**, *2*, e01111-13. [CrossRef] [PubMed]
- 65. Engelberg-Kulka, H.; Hazan, R.; Amitai, S. *mazEF*: A chromosomal toxin-antitoxin module that triggers programmed cell death in bacteria. *J. Cell Sci.* 2005, 118, 4327–4332. [CrossRef]
- Šulčius, S.; Šimoliūnas, E.; Alzbutas, G.; Gasiūnas, G.; Jauniškis, V.; Kuznecova, J.; Miettinen, S.; Nilsson, E.; Meškys, R.; Roine, E.; et al. Genomic Characterization of Cyanophage vB_AphaS-CL131 Infecting Filamentous Diazotrophic Cyanobacterium *Aphanizomenon flos-aquae* Reveals Novel Insights into Virus-Bacterium Interactions. *Appl. Environ. Microbiol.* 2019, 85, e01311-18. [CrossRef]
- 67. Iyer, L.M.; Leipe, D.D.; Koonin, E.V.; Aravind, L. Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* **2004**, *146*, 11–31. [CrossRef]
- 68. Snider, J.; Houry, W.A. MoxR AAA+ ATPases: A novel family of molecular chaperones? J. Struct. Biol. 2006, 156, 200–209. [CrossRef]
- 69. Bakkouri, M.E.; Gutsche, I.; Kanjee, U.; Zhao, B.; Yu, M.; Goret, G.; Schoehn, G.; Burmeister, W.P.; Houry, W.A. Structure of RavA MoxR AAA+ protein reveals the design principles of a molecular cage modulating the inducible lysine decarboxylase activity. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 22499–22504. [CrossRef]
- Dieppedale, J.; Sobral, D.; Dupuis, M.; Dubail, I.; Klimentova, J.; Stulik, J.; Postic, G.; Frapy, E.; Meibom, K.L.; Barel, M.; et al. Identification of a Putative Chaperone Involved in Stress Resistance and Virulence in *Francisella tularensis*. *Infect. Immun.* 2011, 79, 1428–1439. [CrossRef]
- Scheele, U.; Erdmann, S.; Ungewickell, E.J.; Felisberto-Rodrigues, C.; Ortiz-Lombardía, M.; Garrett, R.A. Chaperone role for proteins p618 and p892 in the extracellular tail development of *Acidianus* two-tailed virus. *J. Virol.* 2011, *85*, 4812–4821. [CrossRef]
- 72. Sancar, G.B. DNA repair enzymes. Annu. Rev. Biochem. 1988, 57, 29–67. [CrossRef]
- 73. Lee, J.; Park, H.S.; Lim, S.; Jo, K. Visualization of UV-induced damage on single DNA molecules. *Chem. Commun.* 2013, 49, 4740–4742. [CrossRef]
- 74. Grafstrom, R.H.; Park, L.; Grossman, L. Enzymatic repair of pyrimidine dimer-containing DNA. A 5' dimer DNA glycosylase: 3'-apyrimidinic endonuclease mechanism from *Micrococcus luteus*. J. Biol. Chem. **1982**, 257, 13465–13474. [CrossRef]
- 75. Walker, R.K.; McCullough, A.K.; Lloyd, R.S. Uncoupling of Nucleotide Flipping and DNA Bending by the T4 Pyrimidine Dimer DNA Glycosylase. *Biochemistry* **2006**, *45*, 14192–14200. [CrossRef]
- 76. Bailly-Bechet, M.; Vergassola, M.; Rocha, E. Causes for the intriguing presence of tRNAs in phages. *Genome Res.* 2007, 17, 1486–1495. [CrossRef]
- 77. Albers, S.; Czech, A. Exploiting tRNAs to Boost Virulence. Life 2016, 6, 4. [CrossRef]
- 78. Wilson, J.H. Function of the bacteriophage T4 transfer RNA's. J. Mol. Biol. 1973, 74, 753–757. [CrossRef]

- 79. Bobrovskyy, M.; Vanderpool, C.K. Regulation of Bacterial Metabolism by Small RNAs Using Diverse Mechanisms. *Annu. Rev. Genet.* **2013**, *47*, 209–232. [CrossRef]
- 80. Oglesby-Sherrouse, A.G.; Murphy, E.R. Iron-responsive bacterial small RNAs: Variations on a theme. *Metallomics* **2013**, *5*, 276–286. [CrossRef] [PubMed]
- 81. Mika, F.; Hengge, R. Small RNAs in the control of RpoS, CsgD, and biofilm architecture of *Escherichia coli*. *RNA Biol*. **2014**, *11*, 494–507. [CrossRef]
- 82. Ng, W.L.; Bassler, B.L. Bacterial quorum-sensing network architectures. Annu. Rev. Genet. 2009, 43, 197–222. [CrossRef] [PubMed]
- Klähn, S.; Bolay, P.; Wright, P.R.; Atilho, R.M.; Brewer, K.I.; Hagemann, M.; Breaker, R.R.; Hess, W.R. A glutamine riboswitch is a key element for the regulation of glutamine synthetase in cyanobacteria. *Nucleic Acids Res.* 2018, 46, 10082–10094. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.