



# Genomic Analysis of Mic1 Reveals a Novel Freshwater Long-Tailed Cyanophage

Feng Yang, Hua Jin, Xiao-Qian Wang, Qiong Li, Jun-Tao Zhang, Ning Cui, Yong-Liang Jiang, Yuxing Chen, Qing-Fa Wu\*, Cong-Zhao Zhou\* and Wei-Fang Li\*

Hefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, China

## OPEN ACCESS

### Edited by:

Krishna Khaimar,  
National Environmental Engineering  
Research Institute (CSIR), India

### Reviewed by:

George S. Bullerjahn,  
Bowling Green State University,  
United States  
Daichi Morimoto,  
Kyoto University, Japan

### \*Correspondence:

Qing-Fa Wu  
wuqf@ustc.edu.cn  
Cong-Zhao Zhou  
zcz@ustc.edu.cn  
Wei-Fang Li  
liwf@ustc.edu.cn

### Specialty section:

This article was submitted to  
Virology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 12 October 2019

**Accepted:** 05 March 2020

**Published:** 08 April 2020

### Citation:

Yang F, Jin H, Wang X-Q, Li Q, Zhang J-T, Cui N, Jiang Y-L, Chen Y, Wu Q-F, Zhou C-Z and Li W-F (2020) Genomic Analysis of Mic1 Reveals a Novel Freshwater Long-Tailed Cyanophage. *Front. Microbiol.* 11:484. doi: 10.3389/fmicb.2020.00484

Lake Chaohu, one of the five largest freshwater lakes in China, has been suffering from severe cyanobacterial blooms in the summer for many years. Cyanophages, the viruses that specifically infect cyanobacteria, play a key role in modulating cyanobacterial population, and thus regulate the emergence and decline of cyanobacterial blooms. Here we report a long-tailed cyanophage isolated from Lake Chaohu, termed Mic1, which specifically infects the cyanobacterium *Microcystis aeruginosa*. Mic1 has an icosahedral head of 88 nm in diameter and a long flexible tail of 400 nm. It possesses a circular genome of 92,627 bp, which contains 98 putative open reading frames. Genome sequence analysis enabled us to define a novel terminase large subunit that consists of two types of intein, indicating that the genome packaging of Mic1 is under fine control via posttranslational maturation of the terminase. Moreover, phylogenetic analysis suggested Mic1 and mitochondria share a common evolutionary origin of DNA polymerase  $\gamma$  gene. All together, these findings provided a start-point for investigating the co-evolution of cyanophages and its cyanobacterial hosts.

**Keywords:** freshwater cyanosiphophage, Lake Chaohu, genome sequence, terminase large subunit, DNA polymerase  $\gamma$ , evolution

## INTRODUCTION

Cyanobacteria, which have existed on the earth for more than 3.5 billion years, are widely distributed in aquatic environments (Zhang and Gui, 2018). As a group of photosynthetic bacteria, cyanobacteria provide the source of primary production of oxygen, nitrogen, and carbon, and act as a model organism for studying the coordination of carbon and nitrogen metabolisms (Jiang et al., 2018). Beyond contributing to biogeochemical cycle, blooms also yield toxicity and hypoxia of waterbodies, which was first investigated in the Lake Alexandrina of Australia in 1878 (Francis, 1878). From the ecological perspective, blooms are increasing in frequency, magnitude, and duration in recent years, and cause the death of fish and risk to human diet (Paerl et al., 2001; Huisman et al., 2018). Therefore, it becomes imperative to devise an effective strategy to mitigate and control the water blooms.

Cyanophage, the classic companion virus of cyanobacteria, is a key factor that mediates the host communities, food web, carbon cycling, and nutrient recycling. It also has a potential impact on the regulation of cyanobacterial bloom through lysis-induced mortality, metabolic outputs as well as altering diversity and community structures (Zhang and Gui, 2018). Since the first complete genome sequence of cyanophage P60 that infects the marine *Synechococcus* was reported in 2002 (Chen and Lu, 2002), the genomes of 112 cyanophages have been sequenced and deposited in Virus-Host Database.<sup>1</sup> However, only 14 genomes of freshwater cyanophages have been sequenced. The first one came from the myophage Ma-LMM01 of *Microcystis aeruginosa* in 2008 (Yoshida et al., 2008). Afterward, four contractile-tailed phages, MaMV-DC (Ou et al., 2015a), S-CRM01 (Dreher et al., 2011), A-1 and N-1 (Chenard et al., 2016), five short-tailed phages, PP (Zhou et al., 2013), Pf-WMP3 (Liu et al., 2008), Pf-WMP4 (Liu et al., 2007), A-4L (Ou et al., 2015b) and S-EIV1 (Chenard et al., 2015) and one so-called tailless phage PaV-LD (Gao et al., 2012) were sequentially isolated from different sources of freshwater. Nevertheless, only three complete genomes of freshwater siphophages, S-2L (Marliere, 2006), S-LBS1 (Zhong et al., 2018), and CrV-01T (Martin et al., 2019) have been reported to date.

Lake Chaohu (117°16'54"E-117°51'46"E, 30°25'28"N-31°43'28"N), located at the south of the capital city Hefei of Anhui province, is one of the five largest freshwater lakes in China. It annually suffers from massive water blooms due to the fast growth of cyanobacteria, accompanied with the rapid industrialization and urbanization of the surrounding areas in the past decades. We successfully isolated a freshwater cyanosiphophage Mic1 from Lake Chaohu. Genome sequencing showed that Mic1, which depicts *Microcystis*-specific, possesses a genome of 92,627 bp (GenBank accession No. MN013189), consisting of 98 putative open reading frames (ORFs). Genome sequence analysis indicated that Mic1 contains a hypothetical ParABS plasmid partition system and a prophage antirepressor, suggesting that the prophage of Mic1 might also exist in the host. Moreover, Mic1 encodes a novel terminase large subunit that consists of two types of intein, which might be involved in the fine control of DNA packaging and phage maturation. In addition, Mic1 encodes a mitochondrion-like DNA polymerase gene, which might be transferred from a common ancestor of mitochondrion or its companion phage.

## MATERIALS AND METHODS

### Isolation and Purification of Cyanophage

Water samples were collected semimonthly from the estuaries of 11 rivers toward Lake Chaohu since 2016 (Figure 1A). After being concentrated to about 100-fold by ultrafiltration, the water samples were applied to infect 11 cyanobacterial strains isolated from Lake Chaohu (data not shown), then the lytic cyanophages infecting previously named *Microcystis wesenbergii* FACHB 1339 were further isolated by the serial dilution method

(Yoshida et al., 2006). Notably, the host *Microcystis wesenbergii* FACHB 1339, which was bought from the Freshwater Algae Culture Collection at the Institute of Hydrobiology, Wuhan, should be classified into *Microcystis aeruginosa* (Harke et al., 2016). Then the crude lysate after phage infection was treated with 1 µg/mL DNase I and RNase at 37°C for 1 h. Afterward, NaCl was added to a final concentration of 0.5 M, followed by incubation at 4°C for 1 h. After centrifugation at 8,000 g for 20 min, the cyanophage particles in the supernatant were pooled and incubated with 10% polyethylene glycol 6,000 at 4°C for 10 h. The pellets after centrifugation were resuspended in SM buffer (50 mM Tris, pH 7.5, 10 mM MgSO<sub>4</sub>, 100 mM NaCl), and further purified by ultracentrifugation at 100,000 g for 4 h with a CsCl density gradient. The cyanophage band with the highest opalescence was collected by a syringe and dialyzed against SM buffer. The isolated cyanophage was named Mic1, denoting the first strain of cyanosiphophage toward its host *M. aeruginosa*.

### Transmission Electron Microscopy

The 4 µL suspension of freshly purified cyanophage Mic1 was layered onto a hydrophilized carbon-coated copper grid, and Mic1 particles were negatively stained with 2% uranyl acetate for 1 min. The particles were examined with a Tecnai G2 Spirit BioTWIN 120 kV transmission electron microscope (FEI Company).

### Genomic DNA Extraction

The phage sample in 2 × lysis buffer (20 mM EDTA and 0.5% SDS) was incubated with 50 µg/mL proteinase K at 56°C for 1 h. The phage suspension was sequentially treated with an equal volume of phenol, phenol-chloroform-isoamyl alcohol (25:24:1) and chloroform, respectively. Subsequently, the DNA was precipitated with 0.3 mol/L CH<sub>3</sub>COONa, pH 7.5 and 3-fold volume of ethanol at -80°C for 1 h. The precipitate of phage genomic DNA was washed twice with 70% ethanol and resuspended with sterile water.

### Genome Sequencing

The sequencing library of phage genomic DNA was constructed using TruePrep DNA library Prep Kit V2, and a total of 10 Gb raw data was generated using Illumina Hiseq 2000 platform. The software Velvet (Version 1.2.07) was applied for genome assembly and 12 contigs showing somewhat similarity with known cyanophages at the protein sequence level were chosen for further analysis. The order and orientation of the 12 contigs were determined by PCR and Sanger sequencing. Finally, the assembled genome was verified using DNAMAN (Lynnon Biosoft).

### Genome Annotation and Characterization

The ORFs were predicted using Glimmer (Delcher et al., 1999) and GeneMarkS (Lukashin and Borodovsky, 1998). The translated ORFs were compared with nr protein database in NCBI<sup>2</sup> using BLASTp program with *e*-values <10<sup>-3</sup> and

<sup>1</sup><https://www.genome.jp/virushostdb/index/virus/all>

<sup>2</sup><https://www.ncbi.nlm.nih.gov/>

the protein hits with the minimal *e*-value in each species were considered to be orthologs. HHpred analysis against the pfamA database was carried out using the default parameters.<sup>3</sup> Promoters of the genome were predicted by BPPROM (LDF > 5).<sup>4</sup> Multiple-sequence alignment was performed using the Multalin program.<sup>5</sup> The circular genome map was drawn using CGView.<sup>6</sup> The cyanophage proteomic tree and genome alignments were conducted with Viptree<sup>7</sup> (Nishimura et al., 2017). The genome-wide similarity score (SG) cutoff for clustering was set to  $\geq 0.15$  (viral genus-level cutoff, where the SG value of 1 stands for two identical genomes and 0 for no detectable high-scoring segment pairs by tBLASTx), according to the previous study (Morimoto et al., 2019). The heatmap was generated using Gegenees (tBLASTx method, accurate parameters-fragment length: 200 bp; step size: 100 bp), and the Splits Tree dendrogram was calculated using the Nexus file exported from Gegenees.

## Mass-Spectrometric Identification of Phage Proteins

The phage proteins separated by 4–12% gradient polyacrylamide gel were analyzed by liquid chromatography/mass spectrometry (LC-MS/MS) using ion-trap mass spectrometer (Thermo LUMOS) and identified by comparing to the protein/peptide sequences deduced from Mic1 genome as previously described (Jin et al., 2019).

## RESULTS AND DISCUSSION

### Isolation and Morphology of Mic1

Lake Chaohu has been severely polluted in recent years, owing to a large amount of industrial waste water and domestic sewage flowing from the rivers of Shiwuli, Nanfei, Shuangqiao and Xiapai, which are major rivers to the lake. Accordingly, we chose the estuaries of 11 major rivers entering Lake Chaohu (Figure 1A) to investigate the cyanophages and their hosts. Cyanophage Mic1 was isolated from a surficial water sample collected at Shuangqiao estuary on October 21st, 2017, which was enriched with *Microcystis*-dominated blooms. Host-range assays showed that Mic1 is only able to lyse a strain of *Microcystis aeruginosa* (previously named *M. wesenbergii* FACHB 1339) isolated from Lake Chaohu, indicating that the infectivity of Mic1 is limited to specific cyanobacterial strains.

Purification of Mic1 from the lysate was performed by CsCl density gradient centrifugation, and the fourth opalescence band containing cyanophages was collected (Figure 1B). The purified cyanophage particles were negatively stained for morphological analysis by electron microscopy, revealing that all particles have a siphovirid morphology. As shown in Figure 1C, Mic1 has an

icosahedral head with a diameter of 88 nm and a flexible and non-contractile tail of 400 nm in length.

### Genome Sequence of Mic1

Mic1 has a circular double-stranded DNA genome of 92,627 bp with a G + C content of 35% (Figure 2A). BLASTp and HHpred analyses indicated that Mic1 has 98 ORFs, which encode hypothetical proteins/peptides of 36–2,572 residues in length. Sequence comparison indicated that only 50 hypothetical proteins have homologs, 36 of which have known functions (Supplementary Table 1). These 36 proteins could be divided into five groups: structural proteins, nucleotide metabolism, DNA replication and packing, auxiliary metabolism and others (Figure 2A).

Bioinformatics analyses revealed that the majority of Mic1 structural proteins are similar to those found in other long-tailed phages, for example, the portal protein (ORF2), receptor binding protein (ORF16), phage minor tail (ORF27, ORF32, ORF34), phage tail (ORF28), phage tail L (ORF31), and tail tape measure (ORF35). Moreover, the cryoelectron microscopy (cryo-EM) structure of the icosahedral capsid of Mic1 demonstrated that the major capsid protein and the cement protein are encoded by ORF40 and ORF47, respectively (Jin et al., 2019). To further identify the structural proteins, the purified Mic1 particles were analyzed by mass spectrometry, which identified 23 proteins in total (Jin et al., 2019). As expected, the most abundant proteins are the major capsid protein ORF40 and tail tube protein ORF20. The genes encoding the 23 structural proteins are mainly distributed in three regions (Figure 2B): portal protein ORF2 followed by two proteins ORF3 and ORF4 of unknown function, tail-related proteins containing ORF20 to ORF35 in addition to a downstream receptor-binding protein ORF16, and major capsid ORF40 with accessory proteins (such as the head scaffold and cement). Notably, most of the 23 structural genes are transcribed in the opposite direction against the majority of non-structural genes. In addition, Mic1 harbors two auxiliary metabolic genes, *phoH* (ORF97) and *mazG* (ORF98), the products of which are involved in elevating the phage fitness by altering host metabolism during infection. PhoH belongs to the phosphate regulon that regulates phosphate uptake and metabolism under conditions of low-phosphate and phosphate limitation (Goldsmith et al., 2011), whereas MazG is responsible to reactivate the macromolecular synthesis pathways, via modulating the ppGpp pool, for the propagation of phage progenies in the nutrient-limited cyanobacteria (Bryan et al., 2008; Dreher et al., 2011). Besides, Mic1 encodes nine hypothetical proteins, which share no sequence similarity to any previously identified proteins (Supplementary Table 1 and Figure 2A).

Comparison with the genomes of three freshwater cyanosiphophages CrV-01T, S-LBS1 and S-2L, which are of 104,262, 34,236, and 42,000 bp, respectively, Mic1 possesses genome of 92,627 bp. Despite sharing a similar genome architecture, Mic1 encodes 22 unique components, including ParA (ORF5), ParB (ORF82), head scaffold (ORF42), cement (ORF47), and DNA polymerase  $\gamma$  (ORF85 and ORF86) (Figure 2C).

<sup>3</sup><https://toolkit.tuebingen.mpg.de/hhpred>

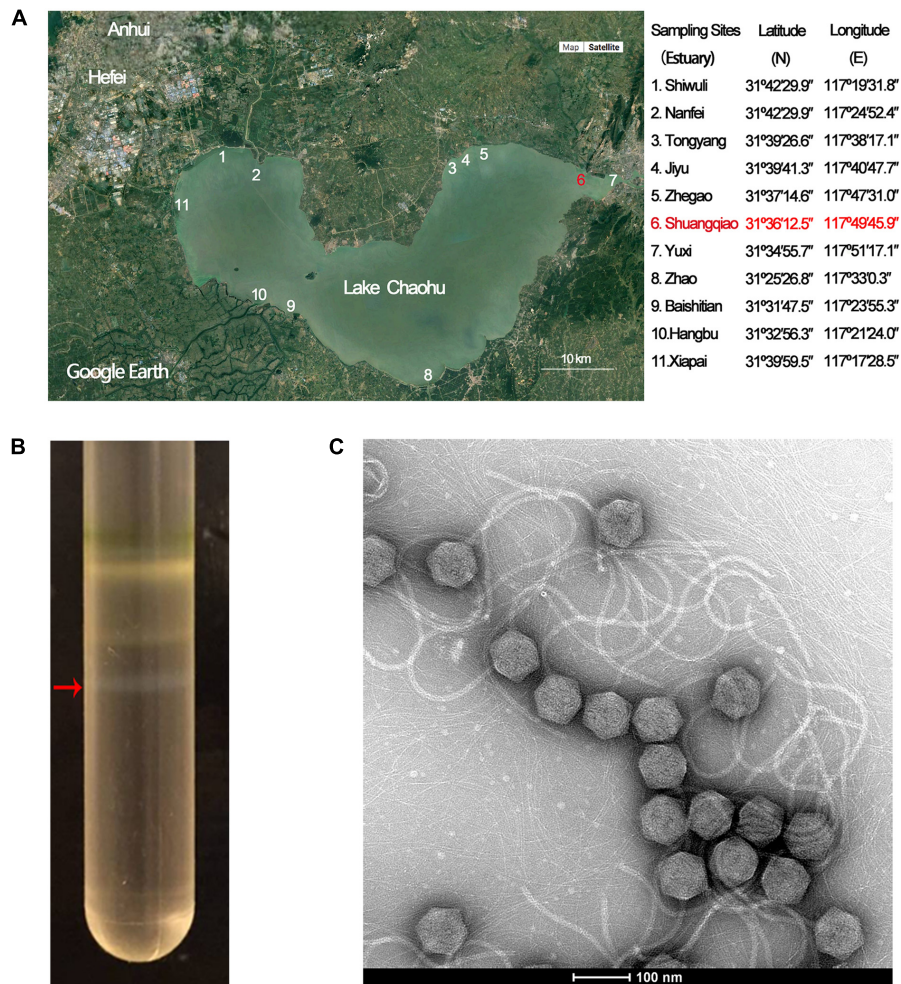
<sup>4</sup><http://www.softberry.com/berry.phtml?topic=bpprom&group=programs&subgroup=gfi>

<sup>5</sup><http://multalin.toulouse.inra.fr/multalin/>

<sup>6</sup><http://cgview.ca/>

<sup>7</sup><https://www.genome.jp/viptree/>





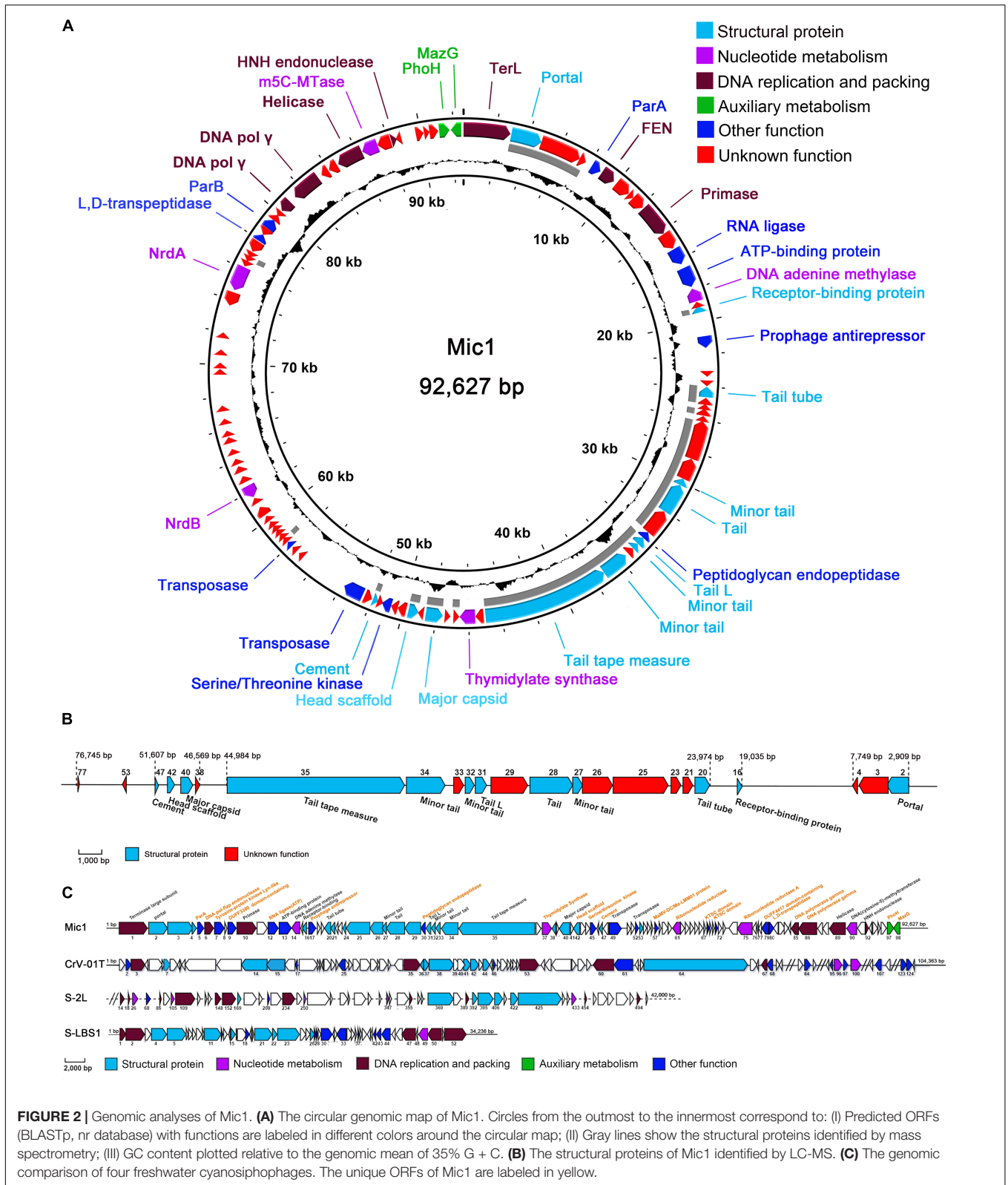
**FIGURE 1** | Isolation, purification, and morphological identification of Mic1. **(A)** Location of Lake Chaohu (from the Google Earth) and the coordinates of 11 major river estuaries to Lake Chaohu. The red mark shows the site where Mic1 was isolated. **(B)** CsCl density gradient ultracentrifugation of Mic1. Mic1 concentrated in the fourth layer gradient is indicated by a red arrow. **(C)** Electron micrograph negatively stained Mic1 particles. The cyanophage Mic1 were negatively stained with uranyl acetate and visualized by transmission electron microscopy at 120 kV accelerating voltage. The phages have an isometric head of ~88 nm in diameter and a long, non-contractile tail of ~400 nm in length.

Notably, we presented here the first report of a cyanophage encoding ParA and ParB, which are involved in DNA partition. ParA and ParB were reported to exist in the low-copy-number phage-plasmids, such as *Escherichia coli* bacteriophage P1 (Vecchiarelli et al., 2010) and N15 (Ravin, 2011). Moreover, a palindromic sequence (8828AAATCACCTAAGTTAGGTGATTT8850) was also found at the downstream of *parA* gene in Mic1 genome (Figure 3), similar to the previously reported *parS* partition site at the upstream or downstream of *parAB* operon (Brooks and Hwang, 2017). Identification of ParA/B homologs combined with the *parS*-like site indicated that Mic1 might also adopt the ParABS plasmid partition system in the lysogenic cycle, similar to that of phage P1 (Vecchiarelli et al., 2010) and N15 (Ravin, 2011). Moreover, we found that ORF17 of Mic1 most likely encodes a prophage antirepressor (Figure 3A), which was proposed to mediate the lytic induction among temperate

phages in the order Caudovirale (Kim and Ryu, 2013). In fact, the phenomena of the bull's-eyed plaques were found when amplifying Mic1 (Figure 3B). However, the defined condition that switches the lysis-lysogeny transition remains unknown, despite the lysogenic state is most likely favored at the conditions such as low productivity of host cells, poor nutrients, exposure to high temperature or UV (Howard-Varona et al., 2017). Altogether, identification of the ParABS plasmid partition system and the prophage antirepressor strongly indicated that Mic1 might also have a lysogenic cycle at some yet unknown conditions.

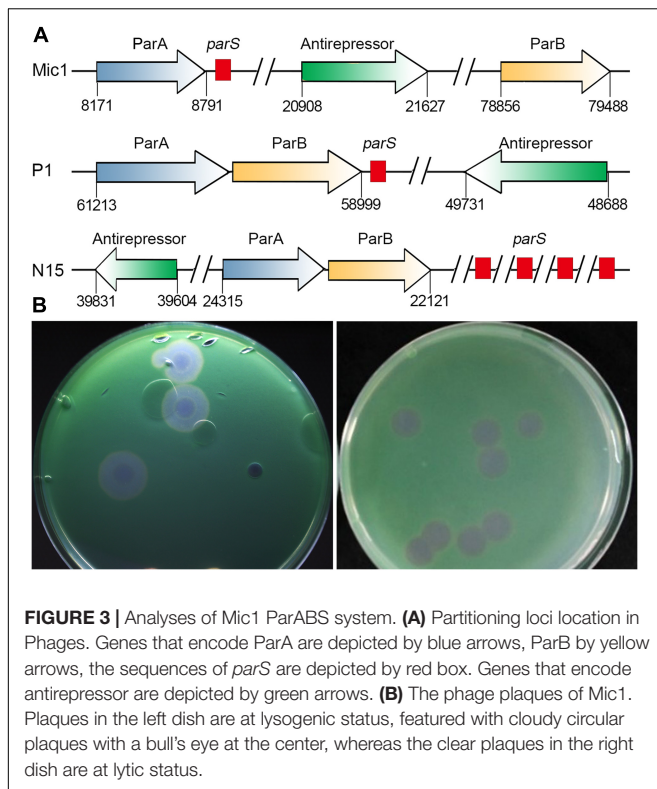
Phylogenetic analysis based on the large terminase subunit TerL showed that Mic1 falls into the "P22-like headful" cluster (Figure 4A), which has a terminally redundant and circularly permuted genome (Casjens and Gilcrease, 2009). Mic1 is closely related to siphovirus *Lactobacillus thermophilus* prophage Lj964 (Desiere et al., 2000) and *Staphylococcus aureus* phage phiETA





(Yamaguchi et al., 2000) based on the pac-site DNA packaging and long tail morphogenesis modules, but distinct from the  $\lambda$ -like freshwater cyanosiphophage S-LBS1 (Zhong et al., 2018).

Notably, TerL of Mic1 contains two different inteins, which are inserted in the ATPase domain and the nuclease domain, respectively; and moreover, the large intein has a homing



**FIGURE 3 |** Analyses of Mic1 ParABS system. **(A)** Partitioning loci location in Phages. Genes that encode ParA are depicted by blue arrows, ParB by yellow arrows, the sequences of *parS* are depicted by red box. Genes that encode antirepressor are depicted by green arrows. **(B)** The phage plaques of Mic1. Plaques in the left dish are at lysogenic status, featured with cloudy circular plaques with a bull's eye at the center, whereas the clear plaques in the right dish are at lytic status.

endonuclease domain (HED) (Figure 4B). Inteins are often mobile, with the ability to post-translationally excise themselves out of the precursor proteins (Perler et al., 1994), usually triggered by external stimuli, such as reactive oxygen species and reactive nitrogen species (Topilina et al., 2015a), reducing agent (Callahan et al., 2011), high temperature (Topilina et al., 2015b), low pH (Wood et al., 1999), high concentration of salt (Reitter et al., 2016), or DNA damage (Lennon et al., 2016). Thus inteins are thought to be biosensors that allow instantaneous splicing to produce the activated proteins. We proposed that the two inteins in the TerL precursor might act as biosensors to control the DNA packaging and maturation of Mic1.

Sequence analysis showed that Mic1 genome has some highly repetitive sequences in the non-coding region. For example, the motif.

GGRYATAT(A)ACA WWTCTAAAAAAGATGACATAATG  
GTAACATAAAGAAAAACACAGG that contains the predicted promoter (−35 box TATACA ~ −10 box TGACATAAT) repeats 10 times in the region from 66,031 to 71,646. Another motif GTGCCAGTTYATAAAGTGTCACAATTAGAYTTGACAAAA  
AGCAAAGTTTATGCAGGC(G)ATAAT containing the predicted promoter (−35 box TTGACA ~ −10 box AGGCAT AAT) repeats 8 times in the region from 67,047 to 69,344. However, more transcriptomic investigations are needed to elucidate the putative multiple promoters that drive the efficient transcription in the lytic cycle of Mic1. In addition, a short repetitive sequence ATCAGTT repeats 20 times in the region from 59,828 to 60,196. In fact, this 7 nt repetitive unit is also widespread in the repetitive regions (ranging from 2 to 40

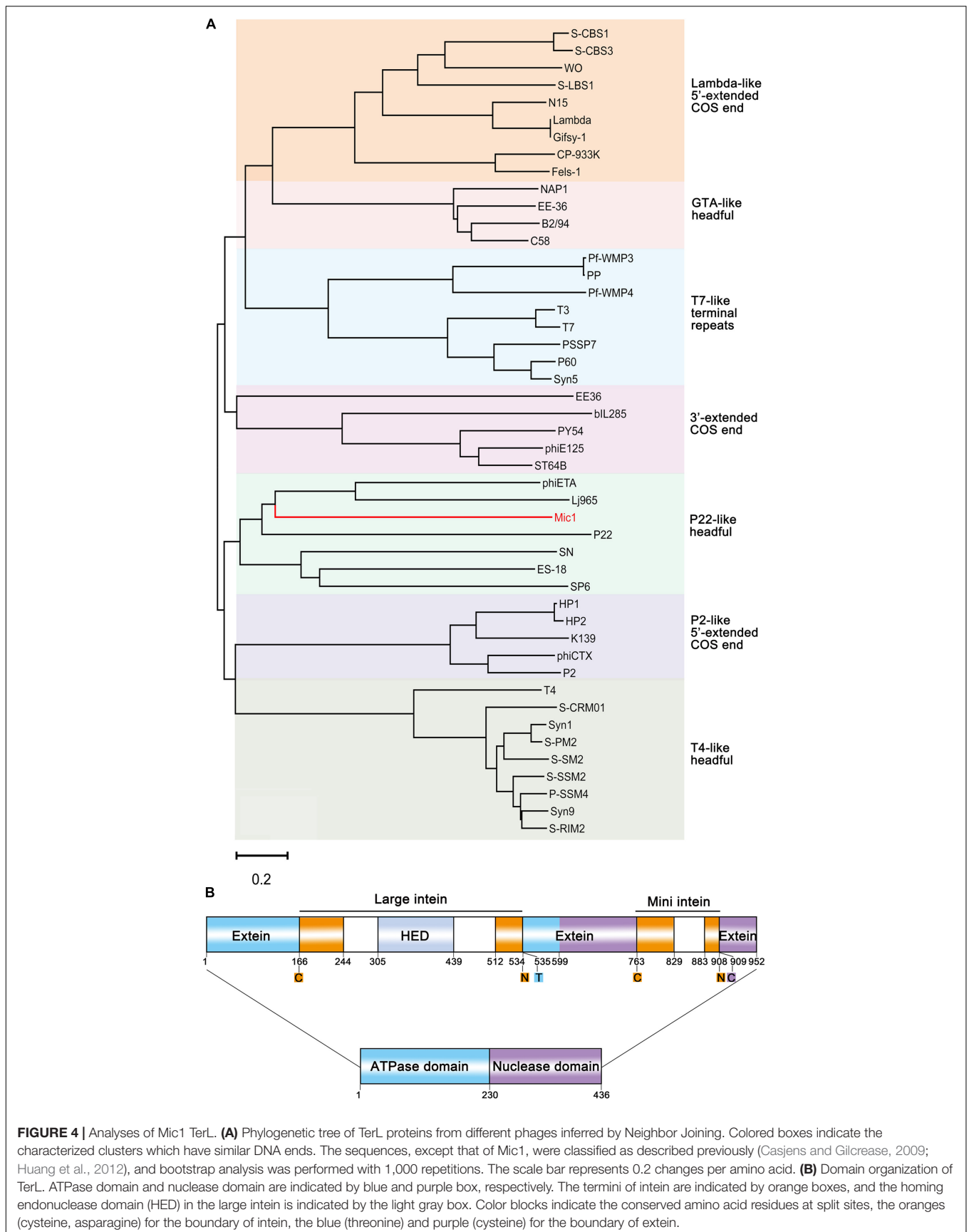
repetitive units) of *Microcystis aeruginosa* NIES-843 genome. It suggested that this short repetitive unit of Mic1 might come from the host genome through horizontal gene transfer, despite its function in the cyanophage or host remains unknown.

## Evolutionary Analyses at the Genome Level

Before Mic1, only 14 cyanosiphophages have been genome-sequenced, including 11 marine and 3 freshwater cyanophages. To determine the degree of genomic variability between the cyanosiphophages, the heatmap was clustered based on Bray-Curtis similarity (Zhong et al., 2018) and compared the gene similarities between two phages (Figure 5A). The cyanophages A-HIS1 and A-HIS2 of the unicellular cyanobacterium *Acaryochloris marina* (Chan et al., 2011), CrV-01T of the filamentous *Cylindrospermopsis (Raphidiopsis) raciborskii* (Martin et al., 2019), vB\_NpeS-2AV2 of the filamentous nitrogen fixing cyanobacterium *Nodularia* sp. (Coloma et al., 2017), PSS2 of *Prochlorococcus* (Sullivan et al., 2009) were clustered according to their hosts, respectively. In contrast, the two fresh water cyanophage S-LBS1, S-2L were clustered with six marine siphophages S-CBS (1,3,4), KBS-S-2A, S-ESS1, S-SKS1, most likely because they all infect similar hosts of unicellular cyanobacterium *Synechococcus* (Figure 5A). Notably, Mic1, which infects the abundant freshwater unicellular cyanobacteria in the genera *Microcystis*, is distinct from the 14 previously sequenced cyanosiphophages and classified in a unique category. These results implied that cyanosiphophages infecting different host cyanobacteria might possess different genetic lineages.

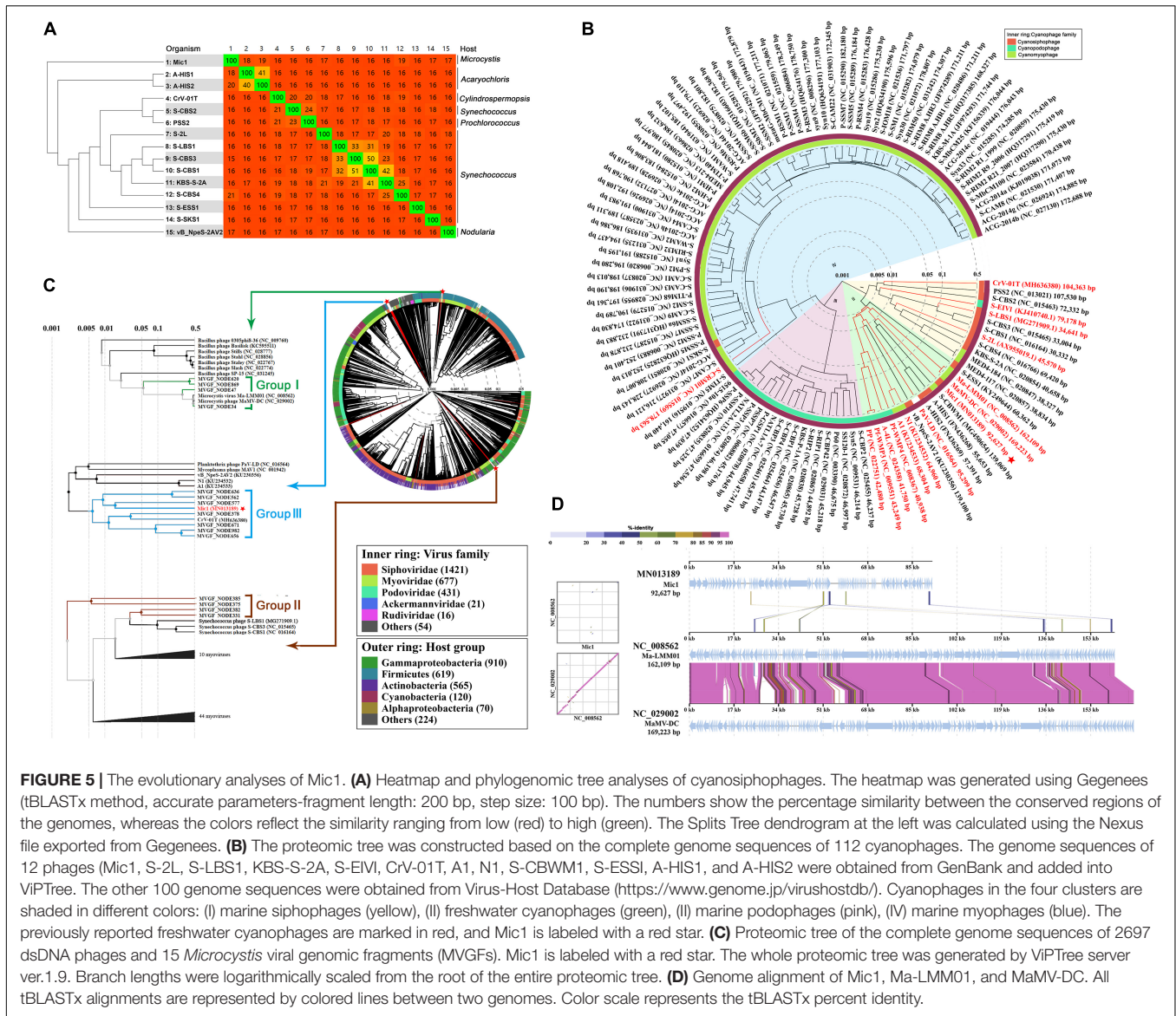
To further define the relationship between Mic1 and other known cyanophages, proteomic tree was constructed using Viptree based on the genomes of 112 completely sequenced cyanophages from Virus-Host Database (Figure 5B). The cyanophages were grouped into four clusters based on the genome similarity scores (SG): the cyanosiphophages, cyanopodophages and cyanomyophages from marine belong to the clusters I, III, and IV, respectively, whereas most of the freshwater cyanophages including Mic1 fall into cluster II. Despite there is a clear boundary between cluster I and II, some freshwater cyanophages and marine cyanosiphophages are chimerically distributed in clusters I and II, respectively. For example, four freshwater cyanophages CrV-01T, S-EIV1, S-LBS1, and S-2L were grouped into the cluster I, mainly due to homology in tail components or nucleotide synthesis enzymes with their neighboring marine cyanosiphophages. A fine proteomic comparison of cyanophages in cluster II revealed that Mic1 shares similar DNA replication components with A-HIS1, A-HIS2, and S-CBWM1, such as DNA polymerase  $\gamma$ , flap endonuclease (FEN, also known as 5' nuclease, ORF6), DNA helicase (ORF89) and primase (ORF10). Moreover, Mic1 encodes a DNA polymerase  $\gamma$  with a relatively high similarity to those of A-HIS1, A-HIS2, and S-CBWM1 up to 29–32% sequence identity.

To have a global view of the evolutionary position of Mic1 among the phages, we constructed the proteomic tree (Figure 5C) with the reference genomes of 2,697 phages, in



**FIGURE 4 |** Analyses of *Mic1* TerL. **(A)** Phylogenetic tree of TerL proteins from different phages inferred by Neighbor Joining. Colored boxes indicate the characterized clusters which have similar DNA ends. The sequences, except that of *Mic1*, were classified as described previously (Casjens and Gilcrease, 2009; Huang et al., 2012), and bootstrap analysis was performed with 1,000 repetitions. The scale bar represents 0.2 changes per amino acid. **(B)** Domain organization of TerL. ATPase domain and nuclease domain are indicated by blue and purple box, respectively. The termini of intein are indicated by orange boxes, and the homing endonuclease domain (HED) in the large intein is indicated by the light gray box. Color blocks indicate the conserved amino acid residues at split sites, the oranges (cysteine, asparagine) for the boundary of intein, the blue (threonine) and purple (cysteine) for the boundary of extein.





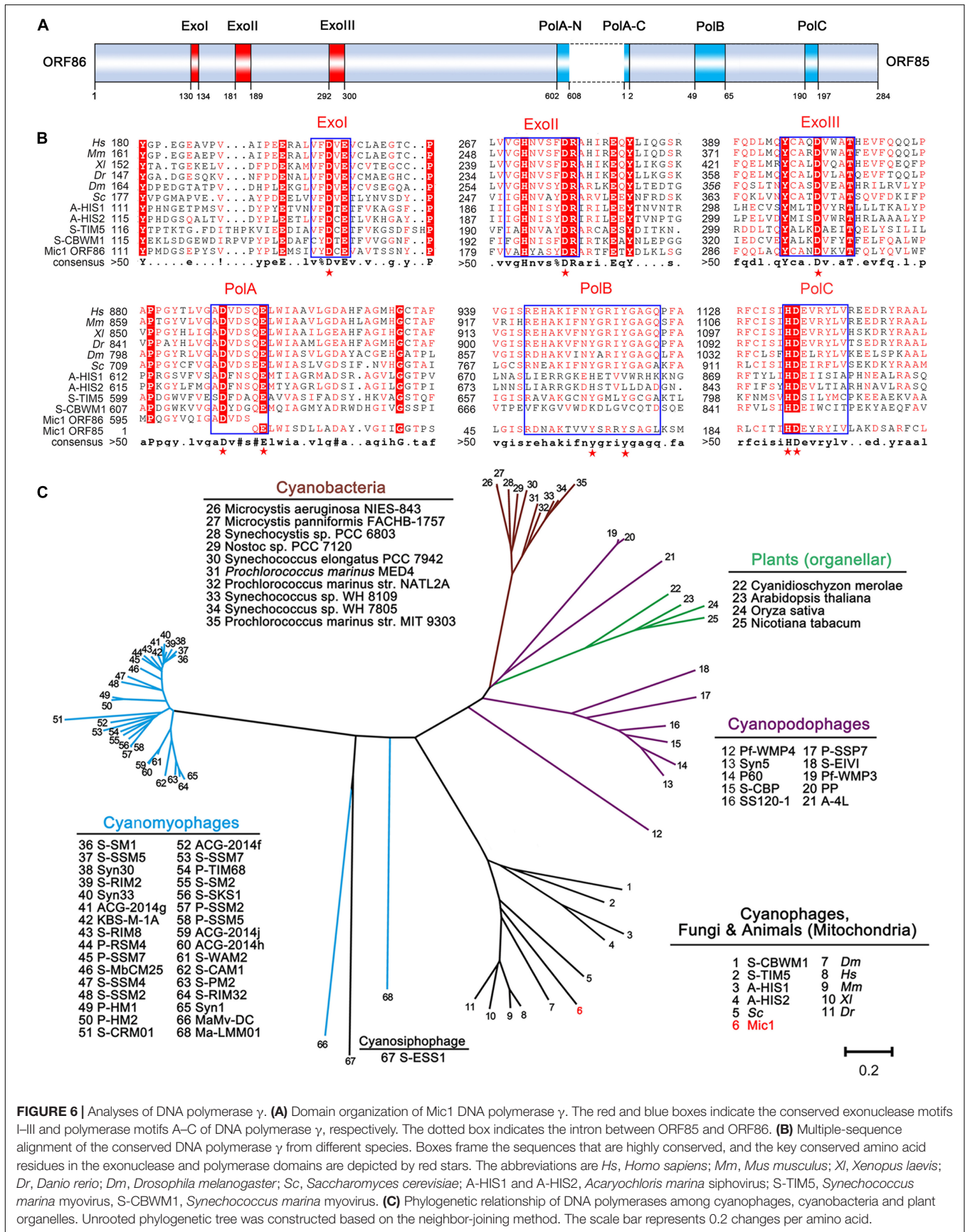
addition to 15 *Microcystis* viral genomic fragments (MVGFs) from Hirosawaniike pond in Japan recently sequenced by metagenomic approaches (Morimoto et al., 2019), using Viptree as previously described (Nishimura et al., 2017). It also indicated that Mic1 possesses a genome distinct from other phages. Moreover, the *Microcystis* cyanophage genomes and the 15 MVGFs could be classified into three groups; and Mic1 falls into group III, the cyanophages of which were reported to have a narrow-host-range that infect highly abundant hosts (Morimoto et al., 2019). It is also consistent with the specificity of Mic1 toward *M. aeruginosa* FACHB 1339, which is one of the most dominant bloom-forming cyanobacterial species in Lake Chaohu.

Despite falling into different groups as shown in Figure 5C, the three genome-sequenced *M. aeruginosa* cyanophages Mic1, Ma-LMM01, and MaMV-DC showed a relatively high similarity at the genome level (Figure 5B). Sequence alignment

of the genomes of these three phages also revealed high homology in several coding regions (Figure 5D), especially those corresponding to four Mic1 proteins: ORF45 (Serine/Threonine

**TABLE 1 |** Homologs in cyanophages that infect *M. aeruginosa*.

Mic1 ORF	Function	Homologs in Ma-LMM01 (Yoshida et al., 2008)		Homologs in MaMV-DC (Ou et al., 2015a)	
		ORF	Identity (%)	ORF	Identity (%)
45	Serine/Threonine kinase	25	58.24	40	59.34
49	Transposase	31	39.44	127	32.47
97	PhoH	183	37.32	169	37.32
57	Hypothetical protein	158	64.22	145	73.39



**FIGURE 6 |** Analyses of DNA polymerase  $\gamma$ . **(A)** Domain organization of Mic1 DNA polymerase  $\gamma$ . The red and blue boxes indicate the conserved exonuclease motifs I–III and polymerase motifs A–C of DNA polymerase  $\gamma$ , respectively. The dotted box indicates the intron between ORF85 and ORF86. **(B)** Multiple-sequence alignment of the conserved DNA polymerase  $\gamma$  from different species. Boxes frame the sequences that are highly conserved, and the key conserved amino acid residues in the exonuclease and polymerase domains are depicted by red stars. The abbreviations are *Hs*, *Homo sapiens*; *Mm*, *Mus musculus*; *Xl*, *Xenopus laevis*; *Dr*, *Danio rerio*; *Dm*, *Drosophila melanogaster*; *Sc*, *Saccharomyces cerevisiae*; A-HIS1 and A-HIS2, *Acarlyochloris marina* siphovirus; S-TIM5, *Synechococcus marina* myovirus; S-CBWM1, *Synechococcus marina* myovirus. **(C)** Phylogenetic relationship of DNA polymerases among cyanophages, cyanobacteria and plant organelles. Unrooted phylogenetic tree was constructed based on the neighbor-joining method. The scale bar represents 0.2 changes per amino acid.

kinase), ORF49 (transposase), ORF97 (PhoH), and ORF57 (hypothetical protein) (Table 1).

## Mic1 Encodes a DNA Polymerase $\gamma$

DNA polymerase  $\gamma$ , which is specifically localized in the mitochondria, is responsible for the replication of mitochondrial DNA in eukaryotes (Bolden et al., 1977). Despite mitochondria were proposed to be derived from  $\alpha$ -proteobacteria, to date no DNA polymerase  $\gamma$  gene has ever been found in any bacteria (Chan et al., 2011). Since 2011, the DNA polymerase  $\gamma$  gene was found in a couple of marine cyanophages, including cyanosiphophages A-HIS1 and A-HIS2 of *Acaryochloris* (Chan et al., 2011), cyanomyophages S-TIM5 (Sabehi et al., 2012), and S-CBWM1 (Xu et al., 2018) of *Synechococcus*. Here we reported a DNA polymerase  $\gamma$  gene in the freshwater cyanophage. The DNA polymerase  $\gamma$  of Mic1 is encoded by ORF85 and ORF86, with an intron of 298 bp between the two ORFs (Figure 6A). Similar to the motif organization of DNA polymerase  $\gamma$  in the model eukaryotes and previously reported cyanophages, ORF86 encodes ExoI~III motifs and N-terminal moiety of PolA, and ORF85 encodes C-terminal moiety of PolA, PolB, and PolC motifs (Figure 6B). Notably, like the mitochondrial DNA polymerase  $\gamma$ , Mic1 and S-TIM5 possess a PolB motif with a conserved tyrosine, which is absent in A-HIS1/A-HIS2 and S-CBWM1.

To date, the evolutionary origin of mitochondrial DNA polymerase remains unclear. Filee et al. (2002) hypothesized that the replication apparatus of mitochondrial DNA polymerase  $\gamma$  gene was originated from phage, and replaced that of bacterial ancestor. However, the homologs of DNA polymerase  $\gamma$  have been only found in cyanophages, but not present in heterotrophic bacteria or their phages by metagenomic recruitment assays (Xu et al., 2018). As shown in the phylogenetic tree (Figure 6C), DNA polymerases  $\gamma$  of Mic1, together with A-HIS1/A-HIS2, S-TIM5, and S-CBWM1, share a common origin with mitochondrial DNA polymerases in opisthokonts (fungi and animals), but distinct from those in photosynthetic eukaryotes, which belong to origin-unknown plant organellar DNA polymerases (Moriyama et al., 2011; Cupp and Nielsen, 2014). Thus we can predict that, with

more cyanophage genome sequences and metagenomics data available, more homologs of mitochondrial DNA polymerase  $\gamma$  genes will be identified. All together, we propose that the DNA polymerase  $\gamma$  genes in mitochondria and cyanophages share a common ancestor.

## DATA AVAILABILITY STATEMENT

The raw datasets for this study can be found in GeneBank, accession No. MN013189.

## AUTHOR CONTRIBUTIONS

W-FL, C-ZZ, and Q-FW conceived, designed, and supervised the project. FY, HJ, X-QW, QL, and J-TZ performed the cyanophage isolation and characterization experiments. FY and Q-FW analyzed the genome data. NC, Y-LJ, and YC provided some good suggestions for the subject. W-FL, C-ZZ, and FY analyzed data and wrote the manuscript. W-FL and C-ZZ revised the manuscript. All authors discussed the data and read the manuscript.

## FUNDING

This work was supported by the Ministry of Science and Technology of China (<http://www.most.gov.cn>, Grant No. 2018YFA0903100) and the National Natural Science Foundation of China (<http://www.nsf.gov.cn>, Grant No. 31630001).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.00484/full#supplementary-material>

## REFERENCES

- Bolden, A., Noy, G. P., and Weissbach, A. (1977). DNA-polymerase of Mitochondria is a gamma-polymerase. *J. Biol. Chem.* 252, 3351–3356.
- Brooks, A. C., and Hwang, L. C. (2017). Reconstitutions of plasmid partition systems and their mechanisms. *Plasmid* 91, 37–41. doi: 10.1016/j.plasmid.2017.03.004
- Bryan, M. J., Burroughs, N. J., Spence, E. M., Clokie, M. R., Mann, N. H., and Bryan, S. J. (2008). Evidence for the intense exchange of MazG in marine cyanophages by horizontal gene transfer. *PLoS One* 3:e2048. doi: 10.1371/journal.pone.0002048
- Callahan, B. P., Topilina, N. I., Stanger, M. J., Van Roey, P., and Belfort, M. (2011). Structure of catalytically competent intein caught in a redox trap with functional and evolutionary implications. *Nat. Struct. Mol. Biol.* 18, 630–633. doi: 10.1038/nsmb.2041
- Casjens, S. R., and Gilcrease, E. B. (2009). Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. *Methods Mol. Biol.* 502, 91–111. doi: 10.1007/978-1-60327-565-1\_7
- Chan, Y. W., Mohr, R., Millard, A. D., Holmes, A. B., Larkum, A. W., Whitworth, A. L., et al. (2011). Discovery of cyanophage genomes which contain mitochondrial DNA polymerase. *Mol. Biol. Evol.* 28, 2269–2274. doi: 10.1093/molbev/msr041
- Chen, F., and Lu, J. (2002). Genomic sequence and evolution of marine cyanophage P60: a new insight on lytic and lysogenic phages. *Appl. Environ. Microbiol.* 68, 2589–2594. doi: 10.1128/AEM.68.5.2589-2594.2002
- Chenard, C., Chan, A. M., Vincent, W. F., and Suttle, C. A. (2015). Polar freshwater cyanophage S-EIV1 represents a new widespread evolutionary lineage of phages. *ISME J.* 9, 2046–2058. doi: 10.1038/ismej.2015.24
- Chenard, C., Wirth, J. F., and Suttle, C. A. (2016). Viruses infecting a freshwater filamentous cyanobacterium (*Nostoc* sp.) encode a functional CRISPR array and a proteobacterial DNA polymerase B. *mBio* 7:e667-16. doi: 10.1128/mBio.00667-16
- Coloma, S. E., Dienstbier, A., Bamford, D. H., Sivonen, K., Roine, E., and Hiltunen, T. (2017). Newly isolated *Nodularia* phage influences cyanobacterial community dynamics. *Environ. Microbiol.* 19, 273–286. doi: 10.1111/1462-2920.13601
- Cupp, J. D., and Nielsen, B. L. (2014). Minireview: DNA replication in plant mitochondria. *Mitochondrion* 19(Pt B), 231–237. doi: 10.1016/j.mito.2014.03.008



- Delcher, A. L., Harmon, D., Kasif, S., White, O., and Salzberg, S. L. (1999). Improved microbial gene identification with GLIMMER. *Nucleic Acids Res.* 27, 4636–4641. doi: 10.1093/nar/27.23.4636
- Desiere, F., Pridmore, R. D., and Brossow, H. (2000). Comparative genomics of the late gene cluster from *Lactobacillus* phages. *Virology* 275, 294–305. doi: 10.1006/viro.2000.0498
- Dreher, T. W., Brown, N., Bozarth, C. S., Schwartz, A. D., Riscoe, E., Thrash, C., et al. (2011). A freshwater cyanophage whose genome indicates close relationships to photosynthetic marine cyanomyophages. *Environ. Microbiol.* 13, 1858–1874. doi: 10.1111/j.1462-2920.2011.02502.x
- Filee, J., Forterre, P., Sen-Lin, T., and Laurent, J. (2002). Evolution of DNA polymerase families: evidences for multiple gene exchange between cellular and viral proteins. *J. Mol. Evol.* 54, 763–773. doi: 10.1007/s00239-001-0078-x
- Francis, G. (1878). Poisonous Australian lake. *Nature* 18, 11–12. doi: 10.1038/018011d0
- Gao, E. B., Gui, J. F., and Zhang, Q. Y. (2012). A novel cyanophage with a cyanobacterial nonbleaching protein A gene in the genome. *J. Virol.* 86, 236–245. doi: 10.1128/JVI.06282-11
- Goldsmith, D. B., Crosti, G., Dwivedi, B., McDaniel, L. D., Varsani, A., Suttle, C. A., et al. (2011). Development of phoH as a novel signature gene for assessing marine phage diversity. *Appl. Environ. Microbiol.* 77, 7730–7739. doi: 10.1128/AEM.05531-11
- Harke, M. J., Steffen, M. M., Gobler, C. J., Otten, T. G., Wilhelm, S. W., Wood, S. A., et al. (2016). A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. *Harmful Algae* 54, 4–20. doi: 10.1016/j.hal.2015.12.007
- Howard-Varona, C., Hargreaves, K. R., Abedon, S. T., and Sullivan, M. B. (2017). Lysogeny in nature: mechanisms, impact and ecology of temperate phages. *ISME J.* 11, 1511–1520. doi: 10.1038/ismej.2017.16
- Huang, S. J., Wang, K., Jiao, N. Z., and Chen, F. (2012). Genome sequences of siphoviruses infecting marine *Synechococcus* unveil a diverse cyanophage group and extensive phage-host genetic exchanges. *Environ. Microbiol.* 14, 540–558. doi: 10.1111/j.1462-2920.2011.02667.x
- Huisman, J., Codd, G. A., Paerl, H. W., Ibelings, B. W., Verspagen, J. M. H., and Visser, P. M. (2018). Cyanobacterial blooms. *Nat. Rev. Microbiol.* 16, 471–483. doi: 10.1038/s41579-018-0040-41
- Jiang, Y. L., Wang, X. P., Sun, H., Han, S. J., Li, W. F., Cui, N., et al. (2018). Coordinating carbon and nitrogen metabolic signaling through the cyanobacterial global repressor NdhR. *Proc. Natl. Acad. Sci. U.S.A.* 115, 403–408. doi: 10.1073/pnas.1716062115
- Jin, H., Jiang, Y. L., Yang, F., Zhang, J. T., Li, W. F., Zhou, K., et al. (2019). Capsid structure of a freshwater cyanophage siphoviridae Mic1. *Structure* 27, 1–9. doi: 10.1016/j.str.2019.07.003
- Kim, M., and Ryu, S. (2013). Antirepression system associated with the life cycle switch in the temperate Podoviridae phage SPC32H. *J. Virol.* 87, 11775–11786. doi: 10.1128/JVI.02173-2113
- Lennon, C. W., Stanger, M., and Belfort, M. (2016). Protein splicing of a recombinase intein induced by ssDNA and DNA damage. *Genes Dev.* 30, 2663–2668. doi: 10.1101/gad.289280.116
- Liu, X., Kong, S., Shi, M., Fu, L., Gao, Y., and An, C. (2008). Genomic analysis of freshwater cyanophage Pf-WMP3 infecting cyanobacterium *Phormidium foveolarum*: the conserved elements for a phage. *Microb. Ecol.* 56, 671–680. doi: 10.1007/s00248-008-9386-9387
- Liu, X., Shi, M., Kong, S., Gao, Y., and An, C. (2007). Cyanophage Pf-WMP4, a T7-like phage infecting the freshwater cyanobacterium *Phormidium foveolarum*: complete genome sequence and DNA translocation. *Virology* 366, 28–39. doi: 10.1016/j.viro.2007.04.019
- Lukashin, A. V., and Borodovsky, M. (1998). GeneMark.hmm: new solutions for gene finding. *Nucleic Acids Res.* 26, 1107–1115. doi: 10.1093/nar/26.4.1107
- Marliere, P. (2006). *Genomic Library of Cyanophage s-21 and Functional Analysis*. U.S. Patent No 2006027.0005A1. Washington, DC: U.S. Patent and Trademark Office.
- Martin, R. M., Moniruzzaman, M., Mucci, N. C., Willis, A., Woodhouse, J. N., Xian, Y., et al. (2019). *Cylindrospermopsis raciborskii* virus and host: genomic characterization and ecological relevance. *Environ. Microbiol.* 21, 1942–1956. doi: 10.1111/1462-2920.14425
- Morimoto, D., Tominaga, K., Nishimura, Y., Yoshida, N., Kimura, S., Sako, Y., et al. (2019). Cooccurrence of broad- and narrow-host-range viruses infecting the bloom-forming toxic cyanobacterium *Microcystis aeruginosa*. *Appl. Environ. Microbiol.* 85:e1170-19. doi: 10.1128/aem.01170-19
- Moriyama, T., Terasawa, K., and Sato, N. (2011). Conservation of POPs, the plant organellar DNA polymerases, in eukaryotes. *Protist* 162, 177–187. doi: 10.1016/j.protis.2010.06.001
- Nishimura, Y., Yoshida, T., Kuronishi, M., Uehara, H., Ogata, H., and Goto, S. (2017). ViPTree: the viral proteomic tree server. *Bioinformatics* 33, 2379–2380. doi: 10.1093/bioinformatics/btx157
- Ou, T., Gao, X. C., Li, S. H., and Zhang, Q. Y. (2015a). Genome analysis and gene nblA identification of *Microcystis aeruginosa* myovirus (MaMV-DC) reveal the evidence for horizontal gene transfer events between cyanomyovirus and host. *J. Gen. Virol.* 96, 3681–3697. doi: 10.1099/jgv.0.000290
- Ou, T., Liao, X. Y., Gao, X. C., Xu, X. D., and Zhang, Q. Y. (2015b). Unraveling the genome structure of cyanobacterial podovirus A-4L with long direct terminal repeats. *Virus Res.* 203, 4–9. doi: 10.1016/j.virusres.2015.03.012
- Paerl, H. W., Fulton, R. S. III, Moisaner, P. H., and Dyble, J. (2001). Harmful freshwater algal blooms, with an emphasis on cyanobacteria. *Sci. World J.* 1, 76–113. doi: 10.1100/tsw.2001.16
- Perler, F. B., Davis, E. O., Dean, G. E., Gimble, F. S., Jack, W. E., Neff, N., et al. (1994). Protein splicing elements - inteins and exteins - a definition of terms and recommended nomenclature. *Nucleic Acids Res.* 22, 1125–1127. doi: 10.1093/nar/22.7.1125
- Ravin, N. V. (2011). N15: the linear phage-plasmid. *Plasmid* 65, 102–109. doi: 10.1016/j.plasmid.2010.12.004
- Reitter, J. N., Cousin, C. E., Nicastrì, M. C., Jaramillo, M. V., and Mills, K. V. (2016). Salt-dependent conditional protein splicing of an intein from *Halobacterium salinarum*. *Biochemistry* 55, 1279–1282. doi: 10.1021/acs.biochem.6b01028
- Sabehi, G., Shaulov, L., Silver, D. H., Yanai, I., Harel, A., and Lindell, D. (2012). A novel lineage of myoviruses infecting cyanobacteria is widespread in the oceans. *Proc. Natl. Acad. Sci. U.S.A.* 109, 2037–2042. doi: 10.1073/pnas.1115467109
- Sullivan, M. B., Krastins, B., Hughes, J. L., Kelly, L., Chase, M., Sarracino, D., et al. (2009). The genome and structural proteome of an ocean siphovirus: a new window into the cyanobacterial ‘mobilome’. *Environ. Microbiol.* 11, 2935–2951. doi: 10.1111/j.1462-2920.2009.02081.x
- Topilina, N. I., Green, C. M., Jayachandran, P., Kelley, D. S., Stanger, M. J., Piazza, C. L., et al. (2015a). SufB intein of *Mycobacterium tuberculosis* as a sensor for oxidative and nitrosative stresses. *Proc. Natl. Acad. Sci. U.S.A.* 112, 10348–10353. doi: 10.1073/pnas.1512777112
- Topilina, N. I., Novikova, O., Stanger, M., Banavali, N. K., and Belfort, M. (2015b). Post-translational environmental switch of RadA activity by extein-intein interactions in protein splicing. *Nucleic Acids Res.* 43, 6631–6648. doi: 10.1093/nar/gkv612
- Vecchiarelli, A. G., Han, Y. W., Tan, X., Mizuuchi, M., Ghirlando, R., Biertumpfel, C., et al. (2010). ATP control of dynamic P1 ParA-DNA interactions: a key role for the nucleoid in plasmid partition. *Mol. Microbiol.* 78, 78–91. doi: 10.1111/j.1365-2958.2010.07314.x
- Wood, D. W., Wu, W., Belfort, G., Derbyshire, V., and Belfort, M. (1999). A genetic system yields self-cleaving inteins for bioseparations. *Nat. Biotechnol.* 17, 889–892. doi: 10.1038/12879
- Xu, Y., Zhang, R., Wang, N., Cai, L., Tong, Y., Sun, Q., et al. (2018). Novel phage-host interactions and evolution as revealed by a cyanomyovirus isolated from an estuarine environment. *Environ. Microbiol.* 20, 2974–2989. doi: 10.1111/1462-2920.14326
- Yamaguchi, T., Hayashi, T., Takami, H., Nakasone, K., Ohnishi, M., Nakayama, K., et al. (2000). Phage conversion of exfoliative toxin A production in *Staphylococcus aureus*. *Mol. Microbiol.* 38, 694–705. doi: 10.1046/j.1365-2958.2000.02169.x
- Yoshida, T., Nagasaki, K., Takashima, Y., Shirai, Y., Tomaru, Y., Takao, Y., et al. (2008). Ma-LMM01 infecting toxic *Microcystis aeruginosa* illuminates diverse cyanophage genome strategies. *J. Bacteriol.* 190, 1762–1772. doi: 10.1128/JB.01534-1537
- Yoshida, T., Takashima, Y., Tomaru, Y., Shirai, Y., Takao, Y., Hiroishi, S., et al. (2006). Isolation and characterization of a cyanophage infecting the toxic

- cyanobacterium *Microcystis aeruginosa*. *Appl. Environ. Microbiol.* 72, 1239–1247. doi: 10.1128/AEM.72.2.1239-1247.2006
- Zhang, Q. Y., and Gui, J. F. (2018). Diversity, evolutionary contribution and ecological roles of aquatic viruses. *Sci. China Life Sci.* 61, 1486–1502. doi: 10.1007/s11427-018-9414-7
- Zhong, K. X., Suttle, C. A., Baudoux, A. C., Derelle, E., Colombet, J., Cho, A., et al. (2018). A new freshwater cyanosiphovirus harboring integrase. *Front. Microbiol.* 9:2204. doi: 10.3389/fmicb.2018.02204
- Zhou, Q., Wei, N., Zheng, L., and Song, L. (2013). Host re-identification of cyanophage PP and its implications for host range and specificity. *Virol. Sin.* 28, 306–308. doi: 10.1007/s12250-013-3356-z
- 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.

Copyright © 2020 Yang, Jin, Wang, Li, Zhang, Cui, Jiang, Chen, Wu, Zhou and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.