Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Comparative genomics and super-resolution microscopy of novel cyanophages reveals genomic isomerism, auxiliary genes, and a non-synchronous infection process

Yue Shi
Chinese Academy of Sciences
Marcus Ziemann
University of Freiburg
Yang Zhao
Chinese Academy of Sciences
Viktoria Reimann
University of Freiburg
Tao Zhu
Chinese Academy of Sciences
Wolfgang R. Hess
wolfgang.hess@biologie.uni-freiburg.de
University of Freiburg
Xuefeng Lu
Chinese Academy of Sciences

Research Article

Keywords: Cyanophages, Anabaena sp. PCC 7120, Comparative genome analysis, Genomic isomerism, phoD, tnpB

Posted Date: June 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9610408/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Microbial communities of multicellular, heterocyst-forming cyanobacteria assimilate organic nitrogen from dinitrogen gas (N_2) through nitrogen fixation and are important photosynthetic primary producers. This makes them ecologically significant, particularly in natural freshwater bodies, wetlands, and agricultural ecosystems. Cyanophages, viruses that specifically infect cyanobacteria play a key role in lysing them, terminating cyanobacterial blooms, and turning over the nutrients accumulated by these communities, with tremendous relevance for the biogeochemical cycles in these environments. However, only a limited number of cyanophages infecting N_2-fixing cyanobacteria have been genome sequenced to date, and it has remained unknown how their interaction proceeds with the individual cells within the filaments forming the backbone of these communities.

Results

We isolated and characterized three *Caudoviricetes* cyanophages (A-Lf14, A-Alj1 and A-Hlh1) specifically infecting the N_2-fixing cyanobacterium *Anabaena* sp. PCC 7120. Super-resolution microscopy revealed heterogeneous infection outcomes among adjacent cells within individual host filaments, demonstrating cell-level specificity in host defense. Comparative genomics revealed a conserved ~ 30 kb invertible genomic region flanked by inverted repeats, a structural configuration not previously recognized in cyanophages. Using diagnostic PCR, Sanger sequencing, and read-mapping, we validated that two isomers stably coexist within phage populations. In contrast to the previously characterized cyanophage A-1(L), the three new phages encode a non-cyanobacterial alkaline phosphatase (PhoD). The expression of *phoD* was highly upregulated during infection, suggesting a phage-driven mechanism for phosphorus acquisition. Conversely, these phages lack the early-expressed *tnpB* endonuclease gene carried by A-1(L). Transcription of *tnpB* was strongly upregulated for both strands ($tnpB_{fwd}$ and $tnpB_{rev}$) during A-1(L) infection, while infection by the three novel phages modestly induced the transcription of five identical host-encoded *tnpB* copies. In a highly unusual setting, both *tnpB* strands are protein-coding over their entire length, leading to the expression of a 347-residue protein from the $tnpB_{rev}$ strand.

Conclusions

This study reveals that genomic isomerism is a source of structural plasticity in freshwater cyanophages and uncovers functionally distinct auxiliary gene repertoires that link phage infection strategies to nutrient status. Our findings provide mechanistic insight into how cyanophages diversify their genomes, interact with their hosts and show that a phage gene can be protein-coding on both strands.

Background

Cyanophages play a crucial ecological role by influencing the abundance, diversity, and metabolic output of cyanobacterial communities in freshwater ecosystems [1, 2]. By lysing bloom-forming species, they affect nutrient cycling and microbial succession, yet the mechanisms underlying their adaptation to complex, dynamic environments remain poorly understood [2–4]. Compared to cyanophages that target unicellular hosts, this knowledge gap is even more pronounced for those infecting multicellular, heterocyst-forming cyanobacteria like *Anabaena*, which are central for the assimilation of organic nitrogen from dinitrogen gas through nitrogen (N_2) fixation and certain bloom events [5–8].

Prior to this study, seven cyanophages infecting hosts from the *Anabaena*/*Nostoc* clade had been fully sequenced: A-1(L) [9–13], N1 [11, 12], A-4(L) [11, 14, 15], Elbi [16], YongM [17], Dfl-JY23 [18], and YongM2. With the exception of A-4(L), these belong to the class *Caudoviricetes*, which encompasses tailed phages with linear, double-stranded DNA genomes [19–21]. The genome of A-1(L) was recently re-sequenced and reannotated, revealing sequence variations compared to the originally deposited version [12, 13]. This reannotation also led to the identification of previously unannotated genes, such as those encoding the neck fiber proteins Gp80–Gp82, which form a distinctive bead-chain-like structure [13]. These findings underscore the ongoing need for accurate genomic reference sequences to underpin more detailed mechanistic and molecular studies.

Complementing genomic analysis, high-resolution microscopic and structural techniques are crucial for linking the genetic architecture to functional morphology and infection dynamics [13, 22, 23]. While advances in cryo-electron microscopy and tomography have revealed conserved viral assembly pathways within unicellular cyanobacteria [22], these often provide static snapshots [22, 24]. For filamentous, heterocyst-forming hosts like *Anabaena*, infection was predicted to involve intricate spatial and temporal coordination across cellular networks [13, 25]. However, the application of *in situ* imaging approaches, which are capable of resolving virion transmission, heterogeneous cellular outcomes, and lysis progression, has remained limited. Overcoming this methodological barrier is therefore essential to uncover how cyanophages navigate, exploit, and ultimately reshape structured multicellular hosts.

Genomic plasticity is a hallmark of viral evolution, enabling rapid adaptation to host defenses and environmental fluctuations. In bacteriophages, site-specific DNA inversion, mediated by dedicated recombinases flipping segments (typically several kilobases) between inverted repeats (IRs), is a mechanism for regulating host tropism and gene expression. The classic model for this regulation is the invertible G segment of phage Mu, whose orientation controls the expression of tail-fiber genes and thereby determines host range. Lytic infection strongly selects for one orientation (termed the “+” orientation), whereas induction from lysogeny typically produces a near-balanced mixture of both orientations, directly linking inversion to viral infectivity and lifestyle [26–32]. Although such recombinase-mediated inversion systems have been extensively documented for other bacteriophages (e.g., P1 [32–35], P7 [32, 34] and HK022 [36]), bacterial pathogens (e.g., *Salmonella* [30–32]), and for the *Saccharomyces cerevisiae* 2-micron plasmid [32, 37, 38], they remain scarcely reported in cyanophages. A notable exception is cyanophage Elbi, which carries a ~ 30 kb invertible region flanked by a 504 bp IR [16]. Similar IRs of ~ 500–700 bp were also noticed in analogous genomic locations in the previously

studied cyanophages A-1(L) and N-1. However, it is unclear whether they form part of a functional inversion system [16]. Furthermore, the potential role of other genetic elements such as transposases in shaping phage or host genome architecture and defense remains an open question. This includes TnpB nuclease genes [39–42], which are found both in the cyanophage A-1(L) and the *Anabaena* host genome [13, 43]. Intriguingly, a recent transcriptomic study of the *Microcystis* cyanophage Mic1 identified a TnpB-like transposase (*gp49*) as part of a "pioneer cluster" of early genes that are highly expressed immediately upon infection [44], suggesting a conserved and potentially pivotal role for such elements in reprogramming host metabolism. Thus, the extent and mechanisms of structural genomic variation in freshwater cyanophages are intriguing yet understudied aspects of virus-host coevolution.

The infection cycle of cyanophages is intimately linked to host physiology, which in turn is profoundly influenced by environmental variables such as light and nutrient availability [7, 45–47]. Light fuels host photosynthesis, providing essential energy and carbon for phage replication [47], while nitrogen availability regulates processes like heterocyst differentiation, N_2-fixation, and overall metabolic activity in filamentous hosts [7, 45]. Phosphorus limitation represents another major constraint, particularly in freshwater ecosystems, and triggers sophisticated acclimation responses in both cyanobacterial hosts and their phages [48]. In marine systems, cyanophage infection dynamics are modulated by the host's phosphorus status and some phages carry auxiliary metabolic genes, including *pstS* encoding the phosphate-binding protein PstS, and *phoA*, encoding alkaline phosphatase whose expression is regulated by the host's phosphate-sensing machinery [49, 50]. In freshwater cyanobacteria, the alkaline phosphatase/phosphodiesterase gene *phoD* is widely distributed and plays an important role in scavenging phosphate from organic sources under phosphorus-limited conditions [48–52]. Accordingly, the two *phoD* homologs in filamentous *Anabaena* are upregulated during phosphate starvation [48]. However, the presence, role and regulation of similar phosphorus-acquisition genes remains unclear for freshwater cyanophages that infect heterocyst-forming hosts. In fact, it has remained unknown to what extent phosphorus availability shapes phage-host interactions in these environments at all.

To extend the set of cyanophages available for studying their interactions with filamentous cyanobacteria, we searched for new isolates from freshwater habitats in China and obtained three new *Caudoviricetes* cyanophages specific to *Anabaena* sp. PCC 7120. Using comparative genomics, super-resolution imaging, and physiological assays, we investigated the occurrence and biological relevance of genomic isomerism in these phages, examined how environmental factors modulate infection dynamics, and explored the functional roles of accessory genes *phoD* and *tnpB*. By integrating our results with recent advances in the Mic1 system [44], we highlight a potentially widespread strategy among freshwater cyanophages that involves both large-scale genomic inversion and the deployment of TnpB-family proteins. Our findings provide insights into the multi-layered adaptive strategies of cyanophages, underscoring the dynamic interplay between genomic plasticity, environmental cues, and phage-host coevolution in freshwater ecosystems.

Methods

Cultivation of cyanobacterial strains

Cyanobacterial strains were maintained on BG-11 agar plates. A single colony was inoculated into 20 mL of liquid BG-11 medium and grown under continuous white light at an intensity of 50 μmol photons m^{-2} s^{-1}, at 30°C with shaking at 150 rpm. Cultures were incubated for 4–5 days until they reached the exponential growth phase (OD_{730} = 0.5–0.8).

Cyanophage isolation and purification

Planktonic and benthic samples were collected from 36 sampling sites across freshwater ecosystems, including lakes, ponds, and wetlands, in Heilongjiang Province, northeastern China, during 2022–2023. Planktonic material was harvested via horizontal towing at a depth of 0.5 m using 20-μm mesh nets. Benthic specimens were collected from substrates by scraping with sterile brushes or forceps. All samples were immediately transported on ice and stored at -80°C until processing. For sample preparation, solid specimens were suspended in 500 μL of sterile water. The samples from certain site were homogenized, clarified by gravitational sedimentation, and then filtered through 0.22-μm membranes (Merck Millipore) to obtain cell-free environmental filtrates.

For primary infection assays, thirty-six aliquots of exponentially growing *Anabaena* sp. PCC 7120 were prepared by resuspending cultures in 36 mL BG-11 medium to an OD_{730} of 0.2. Each aliquot was inoculated with 100 μL of environmental filtrate from one of the sampling sites in 24-well plates (900 μL/well). Cultures exhibiting complete lysis within 24–48 h, evidenced by chlorosis and culture clarification, were centrifuged at 12,000 × g for 10 min. The supernatants were then filtered through 0.22-μm membranes to obtain primary cyanophage suspensions. This process of centrifugation, filtration, and reinfection was repeated for three successive rounds to enrich the cyanophage population. Ultimately, three independent cyanophage strains were isolated and designated A-Lf14, A-Alj1, and A-Hlh1.

Cyanophage genome sequencing and annotation

Genomic DNA was extracted from purified cyanophage particles using the UNIQ-10 Column Virus DNA Isolation Kit (Sangon Biotech). DNA quality was assessed by spectrophotometry (with an $OD_{260/280}$ ratio of 1.8–2.0), yielding a minimum of 6 μg per sample. Fragment libraries were constructed from high-quality DNA and sequenced on the Illumina NovaSeq 6000 platform (2 × 150 bp; Biozeron Biotechnology).

Raw sequencing reads were processed with Trimmomatic v0.36, which involved quality trimming by removing bases with Q-scores < 20 and adapter sequences. The optimal results for the assembly were obtained by splicing the optimized sequence with multiple Kmer parameters using ABySS (http://www.bcgsc.ca/platform/bioinfo/software/abyss). GapCloser

(https://sourceforge.net/projects/soapdenovo2/files/GapCloser/) was applied to fill up the remaining local inner gaps and correct single-base polymorphisms, ultimately enhancing the accuracy and completeness of the final assembly.

Open reading frames (ORFs) were predicted using two annotation pipelines. The first pipeline used GeneMarkS v4.32 (http://exon.gatech.edu/GeneMark/genemarks.cgi) with heuristic model selection, and the second employed RAST (http://rast.nmpdr.org/rast.cgi) via the PATRIC platform. Consensus ORF calls were generated by merging overlapping predictions from both pipelines. All primers used for validation of genome sequences and orientation isomers are listed in Additional file 2 Table S1.

Transmission electron microscope (TEM)

Purified cyanophages (10 μL) were applied to carbon-coated copper grids and incubated for 10 min. Excess liquid was removed using filter paper. The grids were negatively stained with 3% uranyl acetate for 1–3 min, followed by removal of the residual stain with filter paper.

Virions were imaged using a JEM-1200EX transmission electron microscope (JEOL, Tokyo, Japan) operating at an acceleration voltage of 80 kV. Morphometric analysis focused on the characteristic icosahedral capsids, contractile tails and neck fibers.

Laser scanning confocal microscope (LSCM)

Anabaena sp. PCC 7120 cultures were infected with cyanophages and incubated for 12 h under standard growth conditions to allow phage adsorption. Host-phage mixtures were stained with DAPI to a final concentration of 1 ng/μL and incubated for 10 min at room temperature in the dark. Samples were then mounted on glass slides. Images were captured using a Zeiss laser scanning confocal microscope (Carl Zeiss AG, Germany) that featured a ×63 oil-immersion objective lens with a numerical aperture of 1.4. Images were recorded in 16-bit, 2,186 × 2,186-pixel format and each pixel was 24 × 24 nm. Phycobilisome autofluorescence was detected with excitation at 548 nm and emission at 561 nm, while DAPI was imaged with excitation at 401 nm and emission at 422 nm. Microscope control and image acquisition used Zeiss software ZEN 3.12.

Cryo-scanning electron microscopy (Cryo-SEM)

For cryo-SEM analysis, infected cells collected 4–8 h post-infection (initial lysis events were observed) were prepared by placing 5 μL aliquots between gold-plated copper planchettes (3 mm diameter), followed by plunge-freezing in liquid ethane. Samples were then transferred to a pre-cooled stage (-145°C) in a Zeiss Ultra Plus HR-SEM (Carl Zeiss AG, Germany) equipped with Gemini optics. Imaging was performed at a beam acceleration voltage (BAV) of 2 kV and a working distance (WD) of 4.9 to 7 mm, using in-lens secondary electron detectors for signal collection.

Plaque assay for cyanophage quantification

Fresh cyanophage lysates were prepared from the cyanobacterium *Anabaena* sp. PCC 7120. The host was grown to exponential phase, harvested by centrifugation (5,000 × g, 5 min), washed twice with BG-11 medium, and resuspended in fresh BG-11 to an OD_{730} of approximately 0.6. One milliliter of cyanophage stock was added to 100 mL of the host culture, and the mixture was incubated at 30°C under continuous illumination (50 μmol photons m^{-2} s^{-1}) for 2–3 days, until clearing of the culture was observed. The resulting lysate was centrifuged (12,000 × g, 5 min), and the supernatant was filtered through a 0.22 μm sterile filter to remove cell debris.

To determine the cyanophage titer, ten-fold serial dilutions of the lysate were prepared. For each dilution, 70 μL of the dilution was mixed with 7 mL of host cell suspension (*Anabaena* sp. PCC 7120, OD_{730} = 0.6) prepared in BG-11 soft agar (0.6% agar). Following adsorption at 30°C for 15 min, the mixture was poured directly onto BG-11 hard agar plates (1.5% agar). Plates were incubated under the same light and temperature conditions for 3–4 days, after which plaques were counted. Each dilution was tested in triplicate. The titer (plaque-forming units per mL, PFU mL^{-1}) was calculated as (plaque count × dilution factor) / inoculum volume (mL). All four cyanophage stock solution were prepared fresh for each experiment, and their titers were adjusted to 1 × 10^8 PFU mL^{-1} prior to use.

Host specificity analysis

Ten phylogenetically diverse cyanobacterial strains, including *Anabaena variabilis* FACHB-319, *Nostoc punctiforme* ATCC 29133, *Nostoc* sp. NIES-2110, *Nostoc spongiaeforme* FACHB-130, *Tolypothrix tenuis* FACHB-128, *Scytonema javanicum* FACHB-887, *Trichormus* sp. ZY3-43, *Limnospira platensis* MMECC-01, *Synechococcus elongatus* PCC 7942, *and Synechocystis* sp. PCC 6803, together with *Anabaena* sp. PCC 7120, were cultured under standard conditions. Cells were then harvested and resuspended in fresh BG-11 medium to an OD_{730} of 0.6.

Infection assays were performed in 24-well plates, with each well containing 1 mL of cyanobacterial suspension inoculated with 10 μL of fresh cyanophage stock solution (1 × 10^8 PFU mL^{-1}). Plates were incubated at 30°C with orbital shaking at 100 rpm under identical light conditions.

Lytic activity was assessed after 24–48 h based on visual observation of lysis severity, supplemented by microscopic examination of filament morphology and autofluorescence. Strains exhibiting visible chlorosis and culture clarification relative to uninfected controls were classified as susceptible.

Infection kinetics under environmental stress

Anabaena sp. PCC 7120 cultures were adjusted to an OD_{730} of 0.6 in BG-11 medium. For light/dark assays, 200 mL cultures were infected with 2 mL fresh cyanophage stock solution (1 × 10^8 PFU mL^{-1})

and split into duplicate flasks: one kept under standard illumination (50 μmol photons m^{-2} s^{-1}) and the other in complete darkness, both at 30°C. Samples were taken every 12 h to measure OD_{730} and generate growth curves. For nitrogen response assays, exponentially growing cells from BG-11 and BG-11(-N) cultures were harvested by centrifugation. Cells from BG-11 medium were resuspended either in BG-11 (+ N group) or in BG-11(-N) (-N group), respectively. Cells from the BG-11(-N) pre-culture were resuspended in BG-11(-N) to form the -N-N group. All three groups were then adjusted to an OD_{730} of 0.6 in 20 mL cultures and infected with 200 μL fresh cyanophage stock solution (1×10^8 PFU mL^{-1}). The infected cultures were incubated at 30°C under 50 μmol photons m^{-2} s^{-1} with shaking at 150 rpm. Samples were taken every 12 h to measure OD_{730} and generate growth curves. For phosphorus response assays, cells were collected by centrifugation from exponentially growing cultures after a 12-hours pre-treatment in BG-11 or BG-11(-P) medium. Cells were resuspended in the same media, i.e., BG-11 (+ P group), or BG-11(-P) (-P group). Both groups were adjusted to an OD_{730} of 0.6, after which 2.5 mL aliquots were dispensed into 12-well plates. The cultures were infected with 25 μL fresh cyanophage stock solution (1×10^8 PFU mL^{-1}) and incubated at 30°C under 50 μmol photons m^{-2} s^{-1} with shaking at 110 rpm. OD_{730} was measured and photographs were taken at 0, 6, 12 and 24 h post-infection. Across all conditions, uninfected controls were included as references.

Viptree analysis and library creation

The phage genome database was created by combining cyanophage genome sequences from the Virus-Host Database [53] and GenBank. To avoid duplications, phages with the same sequences were deleted, only one member per phage species was used and only complete genomes were allowed. For the tree figure, the Viptree algorithm was used [54] and the host information was taken from Virus-Host Database [53] and GenBank. The taxonomy information of the phages and hosts was taken from NCBI.

Quantitative reverse transcription real-time PCR (qRT-PCR)

For RNA extraction, 2 mL samples were collected at 6 h post-infection from both infected and uninfected *Anabaena* sp. PCC 7120 cultures and immediately mixed with a pre-chilled lysis buffer to preserve RNA integrity. Total RNA was purified using the Vazyme Bacterial RNA Extraction Kit. Residual genomic DNA was eliminated, and cDNA was synthesized with the iScript III RT Supermix for qPCR (+ gDNA wiper) kit (Vazyme) following the manufacturer's protocol. Quantitative PCR was performed using a SYBR Premix ExTaqTM (Takara, Dalian, China) on a Roche LightCycler® 480 II sequence detection system (Roche, Shanghai, China). The housekeeping gene *rnpB* (encoding RNase P subunit B) of *Anabaena* sp. PCC 7120 served as an internal reference. Primer sequences for target genes (*phoD*, neck fiber genes, and *tnpB*) are provided in Additional file 2: Table S1. All qRT-PCR assays were carried out in three biological replicates, which yielded consistent results.

Results

Isolation and host specificity of novel cyanophages

To isolate cyanophages specific to filamentous, heterocyst-forming cyanobacteria, we collected environmental samples from diverse freshwater habitats in Heilongjiang Province, northeastern China, a region with documented cyanophage diversity [55–57]. A total of 36 sampling sites spanning lakes, ponds, and wetlands were sampled in 2022 and 2023 (Additional file 2: Table S2). Water and sediment samples were processed by filtration through 0.22-μm membranes to obtain cell-free environmental filtrates. Each filtrate was screened for lytic activity against the nitrogen-fixing cyanobacterium *Anabaena* sp. PCC 7120, a model strain for heterocyst differentiation studies, at an initial host density of OD_{730} = 0.2 in 24-well plates.

From this screening, lysates from three sites produced complete lysis of the host cultures within 24–48 h post-infection, evidenced by visible chlorosis and culture clarification (Fig. 1a). Following three successive rounds of plaque purification and re-infection on *Anabaena* sp. PCC 7120, three cyanophage strains were obtained and designated A-Lf14, A-Alj1 and A-Hlh1. Each isolate produced clear and distinct lytic halos on host lawns after 72 h of incubation (Fig. 1b).

To determine the host range of these isolates, we tested their ability to infect a panel of 10 phylogenetically diverse cyanobacterial strains at an initial host density OD_{730} = 0.6. All three phages, as well as the reference phage A-1(L), caused complete lysis exclusively of *Anabaena* sp. PCC 7120, evidenced by visible chlorosis and culture clarification within 48 h (Fig. 1c). No lysis or loss of autofluorescence was observed for any of the other tested strains, including *Anabaena variabilis* FACHB-319, *Nostoc punctiforme* ATCC 29133, *Nostoc* sp. NIES-2110, *Nostoc spongiaeforme* FACHB-130, *Tolypothrix tenuis* FACHB-128, *Scytonema javanicum* FACHB-887, *Trichormus* sp. ZY3-43, *Limnospira platensis* MMECC-01, *Synechococcus elongatus* PCC 7942, *and Synechocystis* sp. PCC 6803 (Fig. 1c).

Multi-modal imaging reveals spatially asynchronous infection dynamics in filamentous cyanobacteria

Having established the lytic activity and host specificity of these phages, we next sought to characterize their virion morphology and visualize the infection process within the multicellular host. Transmission electron microscopy (TEM) of negatively stained particles confirmed that A-Lf14, A-Alj1, and A-Hlh1 all shared the canonical *Caudoviricetes* morphology of an icosahedral capsid (64 ± 2 nm in diameter) and a contractile tail (95 ± 5 nm in length) (Fig. 2a). Notably, all three phages displayed elaborate neck fibers (Fig. 2a, yellow arrows), a feature structurally reminiscent of the bead-chain-like neck fibers formed by Gp80–Gp82 in cyanophage A-1(L) [13]. However, these newly isolated cyanophages possess neck fibers that are substantially longer than those of A-1(L) [13], extending approximately 2–3 folds beyond the capsid surface compared to approximately 100 nm in A-1(L) [13].

To visualize the infection process *in situ*, we applied Airyscan super-resolution confocal microscopy to *Anabaena* sp. PCC 7120 filaments infected with A-Lf14, using 4’,6-diamidino-2-phenylindole (DAPI) to stain DNA (pseudocolored green) and detecting host phycobilisome autofluorescence (magenta) (Fig. 2b). Relative to uninfected control filaments, which displayed uniform autofluorescence and diffuse

DAPI staining of the nucleoplasm (Fig. 2b, panel I), infected filaments exhibited a progressive series of DAPI signal patterns. At early infection stages, punctate DAPI signals accumulated along the filament periphery, consistent with adsorbed phage particles (Fig. 2b, panels II-III). At later stages, intense, delocalized intracellular DAPI signals appeared within individual cells (Fig. 2b, panel IV). In the final stages, DAPI-stained particles were observed external to cells, associated with remnants of disrupted cellular sheaths (Fig. 2b, panels V–VIII).

Three-dimensional reconstruction of Z-stack images further revealed spatially heterogeneous infection within individual filaments. In some filament regions, peripheral DAPI signals were confined to discrete focal sites (Additional file 1: Fig. S1, Panel Infection 1), whereas in other regions, dispersed signals consistent with host chromosome degradation and phage progeny release were observed (Additional file 1: Fig. S1, Panel Infection 2). Volumetric rendering of a representative infected filament resolved multiple coexisting infection stages, including surface-associated phages, intracellular virion clusters, and presumptive cell-to-cell phage spread, within contiguous cells of a single filament (Additional file 3: Video S1).

To obtain ultrastructural corroboration at nanometer resolution, we performed cryo-scanning electron microscopy (cryo-SEM). Imaging of cell filaments infected with A-1(L) at 8 h post-infection provided direct evidence of mature virions assembled within the host cytoplasm (Fig. 2c). Critically, cryo-SEM captured the striking asynchrony of phage-induced lysis, where two neighboring cells within the same filament exhibited vastly different degrees of disruption (Additional file 1: Fig. S2a). A series of single-cell views detailed a continuum of structural damage, from early interior alterations to complete cellular disintegration (Additional file 1: Fig. S2b-l), directly visualizing the single-cell heterogeneity of infection outcomes in a multicellular context.

Genome sequencing and comparative overview of Anabaena-infecting cyanophages

To understand the genetic basis underlying the above morphological and infection-related phenotypes, we determined the complete phage genome sequences. In parallel, we re-sequenced the genome of A-1(L) to establish a high-confidence reference sequence. All three newly isolated phages possess double-stranded DNA genomes, with sizes of 70,561 bp (A-Lf14), 70,112 bp (A-Alj1), and 71,024 bp (A-Hlh1), and a GC content of approximately 37% (Table 1). These genome sizes are comparable to that of A-1(L) [12, 13] (67,883 bp; this study) and other previously reported phages infecting *Anabaena* or related cyanobacterial hosts, including Elbi (68,626 bp), YongM (65,429 bp), YongM2 (66,888 bp), Dfl-JY23 (69,209 bp), and N1 (64,690 bp) (Table 1). Gene prediction using GeneMarkS and RAST identified 109, 108, and 110 open reading frames (ORFs) in A-Lf14, A-Alj1 and A-Hlh1, respectively, compared to 105 ORFs in A-1(L) (Table 1).

Table 1
Genome sequence comparison of cyanophages infecting *Anabaena*.

Cyanophage	Genome length[a]	GC(%)	Genes[b]	Reference
A-Alj1	70,112	37	109	This work, PX432480
A-Hlh1	71,024	37	108	This work, PX432481
A-Lf14	70,561	37	110	This work, PX432479
A-1(L)	67,883	36.5	105	This work, PX432482
A-1(L) (2024)	67,884	36.5	95	OR360731 [13]
A-1(L) (2016)	68,304	36.5	96	KU234533 [12]
YongM2	66,888	36.2	107	OP834087
YongM	65,429	35.5	93	MT426122 [17]
Dfl-JY23	69,209	36	95	PP438415 [18]
Elbi	68,626	36.7	108	MZ078141 [16]
N1	64,690	35.4	89	KU234532 [12]

[a]The genome length is given in bp.

[b]The numbers for N1 and A-1(L) (2016) are as given in the GenBank files.

To resolve ambiguities in the A-1(L) reference genome, we compared two previously published datasets [12, 13] with our newly generated sequence and identified 12 discordant positions (Additional file 2: Table S3). Sanger sequencing confirmed the nucleotide identity at each of these sites, resolving 9 of the 12 variants to a single high-confidence base call. The remaining three sites were confirmed as genuine single-nucleotide polymorphisms (SNPs) within the A-1(L) population (Additional file 2: Table S3). The corrected A-1(L) genome comprises 67,883 bp and 105 predicted ORFs (Table 1).

Having established high-confidence genome sequences, we compared the gene content of the four phages to identify features that might explain their phenotypic differences. Comparative analysis revealed a suite of differentially distributed accessory genes, including genes encoding a NarI/FixJ family transcription factor and a putative antirepressor, both unique to A-1(L), a likely restriction enzyme shared by A-1(L) and A-Alj1, a antirepressor KilAC domain protein shared by A-1(L) and A-Lf14, and distinct neck-fiber variants (Additional file 2: Table S4; see also Fig. 3B in the phylogenomic section below). Among these variable loci, we focused on two genes whose functions are directly linked to environmental adaptation and genome plasticity. The *phoD* gene (cluster P124 in Additional file 2: Table S5), encoding a putative alkaline phosphatase/phosphodiesterase, was present in A-Lf14, A-Alj1, and A-Hlh1 but absent from A-1(L). Conversely, the *tnpB* gene (cluster P70 in Additional file 2: Table S5), encoding a transposase-associated RNA-guided DNA endonuclease, was present in A-1(L) but absent

from the three new isolates (Additional file 2: Table S4). The *tnpB* sequence in A-1(L) is identical to five gene copies present in the host *Anabaena* sp. PCC 7120 chromosome (*all3986*, *alr4104*) and its alpha and gamma plasmids *(alr7228*, *alr7231*, *all8010*), indicating recent acquisition or exchange with host mobile genetic elements [43]. These endogenous insertion sequences (ISs) are known to disrupt host genes: the IS elements carrying *all3986* and *alr4104* each split a two-component histidine kinase family gene, while *alr7231* interrupts a Ser/Thr kinase family gene [43], suggesting that IS-associated *tnpB* copies are active mobile elements capable of genome rearrangement.

Phylogenomic placement anchors the new phages within a freshwater cyanophage radiation

We performed phylogenomic analysis to position the newly sequenced genomes with publicly available cyanophage genomes, to reconstruct phylogenetic relationships and assess gene-content variation within this ecologically relevant viral lineage. Phylogenomic analysis was conducted using a whole-proteome approach implemented in the program Viptree [54]. A total of 280 complete cyanophage genomes retrieved from GenBank and the Virus-Host Database [53, 58] were included, alongside the three new isolates (Fig. 3a; see Additional file 2: Table S4 for accession list). This analysis placed A-Lf14, A-Alj1 and A-Hlh1 within a well-defined clade consisting of six previously described *Caudoviricetes* phages that infect *Anabaena* or related cyanobacterial hosts: *Nostoc*/*Anabaena* phages A-1(L), N1, YongM, YongM2, and Elbi, as well as *Dolichospermum* phage Dfl-JY23. This tight phylogenetic clustering is congruent with the conserved *Caudoviricetes* morphological features revealed by TEM imaging (Fig. 2a), confirming the evolutionary integrity of this phage lineage and establishing a solid phylogenetic foundation for subsequent comparative genomic analysis.

With this well-resolved phylogenetic framework in place, we further performed systematic comparative genomic analysis to characterize gene content variation across the above nine closely related cyanophages. A total of 181 gene clusters were identified, of which 82 could be assigned predicted functions or conserved domains by homology search against curated databases [12, 13]. These included 28 clusters encoding structural proteins, 13 encoding transmembrane-domain-containing proteins, 15 encoding regulatory factors, and 9 involved in DNA replication (Additional file 2: Table S6). Of the 181 clusters, 69 were conserved across all nine phages, forming a core gene set, whereas 56 clusters were unique to individual phages (Additional file 2: Table S6). Most of the 56 isolate-specific clusters corresponded to short sequences (< 300 nt) lacking detectable similarity to known proteins. We further screened the entire gene set for putative anti-CRISPR proteins using three predictive algorithms (LSTM, GRU, and LINEAR) [59], yielding three candidate homologs encoded by clusters P039, P121, and P160 (Additional file 2: Table S6). Detailed functional annotation for all 943 predicted proteins encoded by the nine phage genomes is summarized in Additional file 2: Table S7.

A conserved invertible genomic region reveals active isomerism

Beyond the gene content differences described above, comparative genome alignment of the nine phages in this clade revealed a higher-order structural feature: a large, syntenic region flanked by inverted repeats (IRs) ranging from 400 to 700 nt in length (Fig. 3b). Alignment of these IR sequences shows short conserved motifs, though these are not shared among all phages (Additional file 1: Fig. S3). Despite the sequence divergence, the consistent presence of these IRs in analogous genomic locations suggests that they serve as recognition sites for a conserved rearrangement mechanism.

Remarkably, whole-genome alignments revealed that the IR-bounded segment occurred in opposite orientations across different isolates: for example, A-1(L), A-Hlh1 and N1 shared one orientation, whereas A-Alj1, A-Lf14 and Elbi exhibited the reverse orientation. For clarity of visual comparison, all genomes were re-drawn with the start codon of cluster 94 positioned at the left terminus and the invertible region uniformly oriented (Fig. 3b).

To experimentally determine whether the observed IR-flanked genomic inversion represents a true biological polymorphism or a technical artifact, we first visualized the genomic organization of the four cyanophages. Comparative alignment using GBKviz clearly illustrated the syntenic region flanked by IRs, with the orientation of the intervening segment reversed in different phages (Fig. 4a). This inversion was observed between A-1(L) and A-Hlh1 (one orientation) and A-Alj1 and A-Lf14 (opposite orientation), supporting the bioinformatic prediction of genomic isomerism.

To directly test for the coexistence of both isomers within phage populations, we designed a diagnostic PCR assay, in which a common primer (Uni) anchored in the non-inverting flanking region was paired with outward-facing primers P1 or P2 located within the invertible segment (Fig. 4b). PCR amplification with each primer pair yielded products of the predicted sizes from all four phage lysates: Uni/P1 generated amplicons of 2,545 bp (A-1(L)), 2,358 bp (A-Lf14), 2,367 bp (A-Alj1), and 2,366 bp (A-Hlh1); Uni/P2 generated amplicons of 4,116 bp (A-1(L)), 2,541 bp (A-Lf14), 2,551 bp (A-Alj1), and 2,558 bp (A-Hlh1) (Fig. 4b, c). Sanger sequencing of these amplicons confirmed exact nucleotide identity to the sequences predicted for each isomer (Fig. 4b, c; Additional file 1: Fig. S4).

To independently verify that both isomers were present in the sequenced populations and not the result of assembly artifacts, we mapped the raw Illumina sequencing reads to both isomeric reference sequences for each phage. Read coverage was abundant and continuous across the IR-flanked region in both orientations for all four phages (Additional file 1: Fig. S5).

A broader search across publicly available cyanophage genomes identified structurally similar IR-flanked regions in five additional phages, which infect taxonomically distant hosts such as *Cylindrospermopsis*, *Synechococcus*, and *Microcystis* (Additional file 2: Table S8).

Divergent auxiliary metabolic gene repertoires shape infection strategies

To investigate the functional significance of differentially distributed accessory genes in the infection process under environmentally relevant conditions, we examined the phage-encoded *phoD*. Given the presence of *phoD* in the three new phages and its absence in A-1(L), we monitored host culture growth (OD_{730}) and appearance over 24 h under phosphate-replete (+ P) and phosphate-depleted (-P) conditions following infection with A-1(L), A-Lf14, A-Alj1 and A-Hlh1 to assess the potential functional impact of this gene on cyanophage infectivity (Fig. 5a, Additional file 1: Fig. S6). Uninfected control (CK) cultures grew better under + P than under -P, confirming phosphate as a growth-limiting nutrient for this host. Visual inspection of culture color (Additional file 1: Fig. S6) showed that phosphate repletion is also required for efficient infection: in + P cultures, yellowing of the cultures was evident as early as 6 h for A-Lf14, A-Alj1 and A-Hlh1, with severe clearing by 12 h, whereas -P cultures remained green throughout. Infection with A-1(L) resulted in a mild, delayed, and largely phosphate-independent lysis phenotype: host OD_{730} remained stable until 24 h, then dropped sharply in both + P and -P cultures with no significant difference between conditions. In contrast, A-Lf14, A-Alj1 and A-Hlh1 exhibited strong, phosphate-dependent lytic activity. Under + P, these phages caused a rapid drop in OD_{730} by 12 h, with host biomass nearly exhausted by 24 h. Under -P, host growth was inherently suppressed and phage-mediated lysis was significantly attenuated, with much higher residual OD_{730} at 12 h (Fig. 5a). Consistent with these growth data, yellowing was far more pronounced in + P cultures infected with A-Lf14, A-Alj1 and A-Hlh1 at 12 h than in A-1(L)-infected cultures (Fig. 5a, Additional file 1: Fig. S6), highlighting the distinct lytic strategies of these phage groups.

The phage-encoded PhoD proteins shared low amino acid identity (< 30%) with the two host alkaline phosphatases, PhoD1 (Alr2234) and PhoD2 (Alr4976), and higher similarity to PhoD sequences from non-cyanobacterial taxa, including Verrucomicrobia. In uninfected *Anabaena* sp. PCC 7120, *phoD2* transcript levels exceeded those of *phoD1* under both + P and -P conditions (Fig. 5b). Upon infection with A-Lf14, A-Alj1, or A-Hlh1, the phage-borne *phoD* genes were expressed at levels approximately two orders of magnitude higher than host *phoD2* transcripts, and their expression was further significantly upregulated under -P relative to + P (Fig. 5c).

In all three *phoD*-carrying phages, the *phoD* gene was located adjacent to neck fiber (*nf)* structural genes. qRT-PCR analysis showed that *phoD* transcript levels did not correlate with those of the adjacent *nf* genes, which were expressed at uniformly high levels irrespective of phosphate availability (Fig. 5d, e). The *nf2* gene, encoding the major neck fiber subunit, was among the most highly expressed genes in all three phages. Sequence alignment of the *nf2* homologs across the four phages revealed substantial divergence in the predicted neck fiber protein length, consistent with the morphological differences in neck fiber dimensions observed by TEM (Fig. 2a).

In parallel, we investigated the *tnpB* locus, which distinguishes A-1(L) from the three *phoD*-carrying phages (Additional file 2: Table S4). Reanalysis of differential RNA-seq data from our previous study [60] allowed us to examine the transcriptional organization of the *tnpB* locus in *Anabaena* sp. PCC 7120. Analysis of transcriptional start sites (TSS) confirmed that *tnpB* on the forward strand ($tnpB_{fwd}$) is actively transcribed (Fig. 6a). Unexpectedly, we identified a long open reading frame on the reverse

strand, potentially encoding a previously undescribed 347-amino acid protein, which we designated $TnpB_{rev}$ (Fig. 6a). Using the TSS information [60], we precisely determined the complete coding sequence of $TnpB_{rev}$. A BLASTP search with the $TnpB_{rev}$ sequence against the NCBI database revealed limited similarity only to a hypothetical protein (*NSP_9610*) from the Baltic Sea cyanobacterium *Nodularia spumigena* CCY9414 (52% identity, 67% similarity over 81% of the sequence). Reanalysis of the corresponding genomic region and TSS data from *Nodularia spumigena* [61] confirmed the presence of syntenic $tnpB_{fwd}$/$tnpB_{rev}$ gene pairs (*NSP_9620/NSP_9610*) in this evolutionarily distant cyanobacterium (Fig. 6a).

Multiple sequence alignment of $TnpB_{fwd}$ and $TnpB_{rev}$ from both species is presented in Fig. 6b. Importantly, reanalysis of publicly available proteomics datasets from *Anabaena* sp. PCC 7120 [62, 63] identified several peptides that matched both $TnpB_{fwd}$ and $TnpB_{rev}$ (highlighted in yellow in Fig. 6b), providing direct evidence that both proteins were produced *in vivo*. Furthermore, examination of previous microarray data [64] revealed that the expression of both $tnpB_{fwd}$ and $tnpB_{rev}$ was regulated. Transcriptomic profiles from a *hetR* (encoding the master regulator of heterocyst differentiation) inactivation mutant showed altered expression levels of both transcripts compared to the wild-type (Fig. 6c).

The qRT-PCR assays confirmed that both strands were transcribed at low levels in uninfected *Anabaena* sp. PCC 7120 (Fig. 6d). However, upon infection with A-1(L), which carries a phage-encoded *tnpB* copy, expression of both $tnpB_{fwd}$ and $tnpB_{rev}$ increased by approximately 12.7-fold and 10-fold, respectively. In contrast, infection with A-Lf14, A-Alj1, and A-Hlh1 resulted in only modest upregulation of host-encoded $tnpB_{fwd}$ and $tnpB_{rev}$ (Fig. 6d).

Environmental cues modulate cyanophage infection kinetics

The above results demonstrate that the differentially distributed *phoD* and *tnpB* genes profoundly influence infection phenotypes under specific nutrient conditions. To determine whether phage infection dynamics are more broadly responsive to host physiological state, we examined the effects of light and nitrogen availability. Light plays a central role in all stages of the cyanophage life cycle, from adsorption to replication, and release. Furthermore, in filamentous nitrogen-fixing cyanobacteria such as *Anabaena* sp. PCC 7120, nitrogen availability critically regulates heterocyst development and has been shown to shape the evolution of phage resistance [7, 8]. Therefore, we examined how variations in light intensity and nitrogen supply affect the infection dynamics of our isolated cyanophages.

Under standard light conditions (50 µmol photons m^{-2} s^{-1}), all four phages caused complete host lysis within 24 h (Fig. 7a). In continuous darkness, lysis was delayed but reached completion by 48 h for all four phages. Under high light (150 µmol photons m^{-2} s^{-1}), the initial decrease in OD_{730} was slightly decelerated relative to standard light for most phages, although complete lysis still occurred by 48 h. A-

1(L) showed a distinct pattern, with slower lysis under standard light but comparatively faster lysis in darkness after 36 h (Fig. 7a). Under nitrogen-replete conditions (+ N), all four phages caused rapid host lysis. When nitrogen was removed at the time of infection (+ N-N) or in cultures pre-starved of nitrogen (- N-N), lysis was markedly attenuated but was nevertheless completed by 48 h for all phages (Fig. 7b).

Discussion

In this study, we isolated and characterized three novel freshwater cyanophages that specifically infect the filamentous, heterocyst-forming cyanobacterium *Anabaena* sp. PCC 7120. By combining multi-modal imaging, high-resolution comparative genomics, experimental validation of genomic isomerism, and functional analysis of auxiliary metabolic genes, we uncovered multiple layers of adaptive strategy in this ecologically important but understudied phage group.

All four phages examined here displayed an exceptionally narrow host range, lysing only *Anabaena* sp. PCC 7120 among a panel of 10 phylogenetically diverse cyanobacterial strains (Fig. 1c). This strict specificity suggests that infection depends on highly specific host factors present in *Anabaena* sp. PCC 7120 but absent even from closely related genera, including *Anabaena variabilis* and multiple *Nostoc* species (Fig. 1c). The narrow host range further implies specialized adaptation to *Anabaena* sp. PCC 7120-like hosts, which appear to exist in Chinese freshwater ecosystems despite the original isolation of this strain in North America, originally described as *Nostoc muscorum* [9, 67]. Consistent with this, related cyanophages have been isolated from geographically distant locations: A-1(L) and A-4(L) from the former Soviet Union [10, 14], Elbi from Illinois, U.S. [16], N-1 from Wisconsin, U.S. [9], and YongM from lake Dian Chi, China [17]. The successful isolation of A-Lf14, A-Alj1, and A-Hlh1 by direct host screening (Fig. 1a, b) demonstrates their compatibility with laboratory cultivation, establishing them together with A-1(L) as a tractable comparative model for molecular dissection of phage-host interactions in filamentous cyanobacteria, a group for which only a handful of phage isolates are available.

Beyond establishing these phages as a model system, a key methodological contribution of this study is the integrated imaging framework that combined negative-stain TEM, Airyscan super-resolution confocal microscopy, three-dimensional Z-stack reconstruction, and cryo-SEM (Fig. 2; Additional file 1: Fig. S1 and S2; Additional file 3: Video S1). Using this approach, we visualized a highly productive and asynchronous lytic cycle, in which efficient intracellular virion assembly, consistent with the high burst sizes reported for related phages [11, 15, 68], coexisted with striking cell-to-cell heterogeneity within individual filaments. Neighboring cells in the same filament displayed vastly different degrees of infection progression, from nascent phage adsorption to complete cellular disintegration (Additional file 1: Fig. S2). This comprehensive visualization strategy provides a framework to dissect infection dynamics and host-level variability in complex, structured multicellular systems. Furthermore, the substantial divergence in neck fiber *nf2* gene sequences across the four phages (Fig. 5d, e), which correlates with the marked differences in neck fiber length observed by TEM (Fig. 2a), provides a genetic basis for these structural differences and may influence adsorption efficiency or receptor specificity [13].

Moving from imaging to genomics, a central finding of our comparative genomic analysis is the identification and experimental validation of a conserved invertible genomic region in *Anabaena* infecting cyanophages. The inversion was confirmed to be a dynamic, reversible process: both orientations coexist within individual propagated populations, as demonstrated by diagnostic PCR, Sanger sequencing, and independent read mapping (Fig. 4; Additional file 1: Fig. S4 and S5). The mosaic distribution of orientations across the nine-phage clade (Fig. 3b; Fig. 4a), with closely related phages exhibiting opposite orientations, rules out a single ancestral inversion event and instead indicates that isomerism occurs recurrently within lineages. Analogous IR-flanked regions were also identified in cyanophages infecting taxonomically distant hosts (Additional file 2: Table S8) [16], suggesting that this genomic architecture represents a conserved feature within a broader subset of tailed phages. The molecular mechanism driving isomer switching remains unresolved. One possibility is that a site-specific recombinase mediates the inversion, analogous to the Gin/Cin/Mu systems in enterobacterial phages [26–32], although no canonical recombinase gene was identified within or adjacent to the invertible region. Alternatively, the headful packaging mechanism employed by these phages, which proceeds via circular intermediates and concatemers [69–73] and generates circularly permuted, terminally redundant ends [21, 71–73], may itself facilitate large-scale genomic rearrangements including the maintenance of segmental isomers [74]. Regardless of the mechanism, the coexistence of two stable genomic isomers has implications for phage biology. In other systems, DNA inversions regulate host tropism by alternating the expression of tail fiber genes [26–32]. Although the invertible region in these cyanophages does not contain recognizable tail fiber loci, it encompasses genes of unknown function whose orientation-dependent expression could contribute to phenotypic diversification or adaptation in infection dynamics.

Beyond this shared structural feature, the four phages exhibited a striking divergence in their auxiliary metabolic gene repertoires (Additional file 2: Table S4; Fig. 3b). The *phoD* gene, present in A-Lf14, A-Alj1, and A-Hlh1 but absent from A-1(L), encodes an alkaline phosphatase whose functional relevance was demonstrated by the pronounced phosphate-dependent lysis phenotype of the *phoD*-carrying phages (Fig. 5a; Additional file 1: Fig. S6). Under phosphate-replete conditions, these phages lysed the host rapidly; under phosphate depletion, their lytic activity was significantly attenuated. Phage-borne *phoD* transcripts accumulated to levels approximately two orders of magnitude higher than the dominant host alkaline phosphatase gene (*phoD2*) and were further upregulated under phosphate limitation (Fig. 5b, c), indicating active deployment of this module precisely when phosphorus becomes limiting. The low amino acid identity between phage and host PhoD proteins, together with higher similarity to sequences from non-cyanobacterial taxa, suggests acquisition via horizontal gene transfer from an external source. The independent transcriptional regulation of *phoD* from the adjacent neck fiber genes (Fig. 5d, e), despite their conserved physical linkage, indicates that *phoD* is controlled by separate, condition-dependent signals rather than being co-transcribed with structural genes.

In contrast to the phosphate-scavenging function of PhoD, the *tnpB* gene, carried by A-1(L) and present in five identical copies in the host genome [43], exhibited a distinct expression pattern with potential implications for genome plasticity. Reanalysis of the *tnpB* locus in both *Anabaena* sp. PCC 7120 and the

distantly related *Nodularia spumigena* CCY9414 revealed a conserved dual-strand coding architecture (Fig. 6a, b), with both the forward strand TnpB and a previously unannotated reverse-strand protein, $TnpB_{rev}$, producing detectable transcripts and proteins *in vivo* [60, 61]. During A-1(L) infection, both strands were strongly induced (Fig. 6d), whereas infection with the *tnpB*-lacking phages induced only modest upregulation of the host copies. These observations, together with the early induction of a TnpB-like transposase in the *Microcystis* cyanophage Mic1 [44], suggest that phage-encoded TnpB-family proteins play a conserved role in the early stages of freshwater cyanophage infection. We hypothesize that the TnpB system may contribute to the genomic isomerism observed in this phage group. As an RNA-guided DNA endonuclease capable of programmable DNA cleavage [39–41], TnpB could facilitate targeted rearrangements at or near the inverted repeats, potentially directed by a phage-encoded guide RNA. Alternatively, TnpB could function in transposon homing, cleaving at donor sites to promote recombination that retains the phage element within host replicons. The co-expression of $TnpB_{rev}$ adds an additional regulatory dimension; this novel protein could act as an inhibitor, guide RNA chaperone, or specificity determinant that modulates TnpB activity. Direct biochemical characterization, including *in vitro* cleavage assays with the invertible element's inverted repeats as substrates, will be required to test these hypotheses.

Finally, the infection kinetic assays under varying light and nitrogen regimes (Fig. 7) further demonstrated that cyanophage lytic progression is tightly coupled to host physiological state. Darkness delayed but did not abolish lysis (Fig. 7a), indicating that phages can sustain replication using stored host energy reserves. High light slightly decelerated lysis, possibly reflecting transient photoinhibition of host metabolism. Notably, A-1(L) exhibited faster lysis in darkness after 36 h compared to standard light, suggesting genetic adaptation for resource utilization under energy-limited conditions that may be linked to its unique gene complement, including *tnpB*. Under nitrogen-depleted conditions, lysis was markedly attenuated but ultimately completed (Fig. 7b), consistent with the capacity to scavenge nitrogen from intracellular macromolecular pools. Taken together, our findings suggest a model in which distinct auxiliary gene repertoires equip different phages with differential fitness under specific environmental scenarios. Phages carrying *phoD* possess an inducible phosphorus-scavenging module of clear ecological relevance in freshwater ecosystems where phosphorus availability often constrains cyanobacterial growth [48–52]. Phages carrying *tnpB* harbor a genetic element with the potential to drive genomic rearrangements, possibly facilitating adaptation to host defenses or enabling persistence through genome association. The reciprocal distribution of these two modules among phages sharing a highly conserved core genome and infecting the same host suggests that they represent alternative adaptive strategies shaped by distinct selective pressures in their respective environments.

Several limitations should be acknowledged. We have not identified the molecular machinery responsible for genome inversion, nor determined whether isomer ratios shift in response to environmental conditions or host state. The functional significance of phage-encoded PhoD was inferred from transcriptional upregulation and phosphate-dependent infection kinetics; direct biochemical evidence that the phage PhoD protein actively hydrolyzes organic phosphates and contributes to the host

phosphate pool is still needed. The hypothesis that TnpB participates in genome isomerization remains speculative in the absence of genetic or biochemical validation. Future work combining genetic manipulation with single-molecule assays to dissect isomer recombination mechanisms, and integrating field sampling with metagenomics to validate these adaptive traits in native habitats, will be essential to move from correlation to causation and to clarify the roles of cyanophage evolutionary strategies in microbial community structure and biogeochemical cycling across dynamic aquatic ecosystems.

Conclusion

This study reveals the layered adaptive mechanisms of *Anabaena*-infecting *Caudoviricetes* cyanophages, establishing genomic isomerism as a previously underappreciated plasticity trait and demonstrating how environmental cues and accessory genes shape infection dynamics and ecological niche differentiation. Our integrated imaging and genomic analyses link structural and genetic features to functional adaptation, expand the experimental framework for studying virus-host interactions in heterocyst-forming cyanobacteria, and advance understanding of phage-driven plasticity in freshwater ecosystems. The discovery of active genomic isomerism, the functional validation of phage-encoded PhoD as a phosphorus-responsive fitness determinant, and the identification of a conserved dual-strand TnpB locus with potential roles in genome dynamics collectively illustrate the multi-layered strategies by which cyanophages diversify their genomes and interact with their hosts.

Abbreviations

BLAST: Basic local alignment search tool; CRISPR: Clustered regularly interspaced short palindromic repeats; Cryo-SEM: Cryo-scanning electron microscopy; DAPI: 4’,6-diamidino-2-phenylindole; dsDNA: Double-stranded DNA; GC: Guanine-cytosine; IR: Inverted repeat; IS: Insertion sequence; LSCM: Laser scanning confocal microscopy; MOI: Multiplicity of infection; *nf*: Neck fiber; OD: Optical density; ORF: Open reading frames; PFU: Plaque-forming unit; qRT-PCR: Quantitative reverse transcription PCR; RNA-seq: RNA sequencing; RNase P: Ribonuclease P; SNP: Single-nucleotide polymorphism; TEM: Transmission electron microscopy; TSS: Transcriptional start site; WT: Wild-type.

Declarations

Acknowledgements

We thank Chunpeng Yang, Xiaofei Yan, and Fei Li for their assistance with confocal laser scanning microscopy (CLSM), cryo-scanning electron microscopy (Cryo-SEM), and transmission electron microscopy (TEM), respectively. We are grateful to Prof. Xudong Xu for providing the cyanophage A-1(L) and to Prof. Yan Liu for providing part of the environmental samples. We also thank Alexander Mitrofanov and Sven Hauns for the algorithmic search for anti-CRISPR genes and Dr. Zhuo Chen for his help with the proteomics data analysis.

Authors’ Contributions

WRH, XL and TZ designed the study. YS and TZ isolated and sequenced the cyanophages and conducted the morphological and physiological experiments. MZ did the majority of bioinformatic analyses. VR maintained phage and cyanobacterial stocks and performed expression analyses. YZ and TZ carried out the environmental sampling. YS, MZ, TZ and WRH wrote the manuscript with contributions from all authors. All authors reviewed the manuscript.

Funding

Financial support for this work was provided by the National Key Research and Development Program of China (2021YFA0909700 and 2024YFB4106501), the Joint Sino-German Research Program (M-0214 to WRH and XL), the DFG (HE 2544/14-2 and HE 2544/20-2 to WRH), the Strategic Priority Research Program of Chinese Academy of Sciences (XDB1290100), the National Natural Science Foundation of China (32570172), the QIBEBT International Cooperation Project (QIBEBT ICP202301), the Key Research and Development Program of Shandong Province (2025CXPT029), the Qingdao Science and Technology Benefit People Demonstration Guide Special Project (24-1-8-xdny-17-nsh) and the Shandong Taishan Scholarship (to XL).

Data Availability

The genomic sequences of A-Lf14, A-Alj1, A-Hlh1 and A-1(L) have been deposited in GenBank under accession numbers PX432479 to PX432482, respectively. The direct links to these data are https://www.ncbi.nlm.nih.gov/nuccore/PX432479, https://www.ncbi.nlm.nih.gov/nuccore/PX432480, https://www.ncbi.nlm.nih.gov/nuccore/PX432481, and https://www.ncbi.nlm.nih.gov/nuccore/PX432482. Source data are provided with this paper.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing Interests

The authors declare no competing interests.

References

1. Zhou Z, Tran PQ, Martin C, Rohwer RR, Baker BJ, McMahon KD, et al. Unravelling viral ecology and evolution over 20 years in a freshwater lake. Nat Microbiol. 2025;10(1):231–45. https://doi.org/10.1038/s41564-024-01876-7.

2. Zimmerman AE, Howard-Varona C, Needham DM, John SG, Worden AZ, Sullivan MB, et al. Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. Nat Rev Microbiol. 2020;18(1):21–34. https://doi.org/10.1038/s41579-019-0270-x.
3. Li Q, Yang F, Zhou CZ. Cyanophages: Billions of years of coevolution with cyanobacteria. Annu Rev Microbiol. 2025;79(1):639–61. https://doi.org/10.1146/annurev-micro-042924-095145.
4. Nair S, Zhang Z, Wang X, Zhang B, Jiao N, Zhang Y. Engineering microbiomes to enhance macroalgal health, biomass yield, and carbon sequestration. Green Carbon. 2025;3(1):63–73. https://doi.org/10.1016/j.greenca.2024.11.001.
5. Golden JW, Yoon H-S. Heterocyst development in *Anabaena*. Curr Opin Microbiol. 2003;6(6):557–63. https://doi.org/10.1016/j.mib.2003.10.004.
6. Nieves-Morión M, Flores E, Whitehouse MJ, Thomen A, Foster RA, Giovannoni SJ. Single-cell measurements of fixation and intercellular exchange of C and N in the filaments of the heterocyst-forming cyanobacterium *Anabaena* sp. strain PCC 7120. mBio. 2021;12(4):e01314–21. https://doi.org/10.1128/mBio.01314-21.
7. Kolan D, Cattan-Tsaushu E, Enav H, Freiman Z, Malinsky-Rushansky N, Ninio S, et al. Tradeoffs between phage resistance and nitrogen fixation drive the evolution of genes essential for cyanobacterial heterocyst functionality. ISME J. 2024;18(1). https://doi.org/10.1093/ismejo/wrad008.
8. Higazi M, Kolan D, Cattan-Tsaushu E, Freiman Z, Ninio S, Avrani S. Nitrogen availability shapes evolution of phage resistance in cyanobacteria. ISME J. 2025;19(1). https://doi.org/10.1093/ismejo/wraf180.
9. Adolph KW, Haselkorn R. Isolation and characterization of a virus infecting the blue-green alga *Nostoc muscorum*. Virology. 1971;46(2):200–8. https://doi.org/10.1016/0042-6822(71)90023-7.
10. Koz'iakov S, Gromov BV, Khudiakov I. A-1(L)-cyanophage of the blue-green alga *Anabaena variabilis*. Mikrobiologiia. 1972;41(3):555–9.
11. Hu NT, Thiel T, Giddings TH, Jr., Wolk CP. New *Anabaena* and *Nostoc* cyanophages from sewage settling ponds. Virology. 1981;114(1):236–46. https://doi.org/10.1016/0042-6822(81)90269-5.
12. Chénard C, Wirth JF, Suttle CA, Hendrix R, Hatfull G, Clokie M. Viruses infecting a freshwater filamentous cyanobacterium (*Nostoc* sp.) encode a functional CRISPR array and a proteobacterial DNA polymerase B. mBio. 2016;7(3):e00667–16. https://doi.org/10.1128/mBio.00667-16.
13. Yu RC, Yang F, Zhang HY, Hou P, Du K, Zhu J, et al. Structure of the intact tail machine of *Anabaena* myophage A-1(L). Nat Commun. 2024;15(1):2654. https://doi.org/10.1038/s41467-024-47006-z.
14. Khudiakov I, Gromov BV. Temperate cyanophage A-4 (L) of the blue-green alga *Anabaena variabilis*. Mikrobiologiia. 1973;42(5):904–7.
15. Padan E, Shilo M. Cyanophages-viruses attacking blue-green algae. Bacteriol Rev. 1973;37(3):343–70. https://doi.org/10.1128/br.37.3.343-370.1973.
16. Patel R, Chen J, Xu J, Erdmann E, Laughlin Z, Alvey RM. Complete genome sequence of the *Anabaena* myophage Elbi. Microbiol Resour Announc. 2021;10(29):e0055221.

https://doi.org/10.1128/MRA.00552-21.

17. Zhang S, Zhao B, Li J, Song X, Tong Y, An W. Host cyanobacteria killing by novel lytic cyanophage YongM: A protein profiling analysis. Microorganisms. 2022;10(2):257. https://doi.org/10.3390/microorganisms10020257.
18. Lin W, Li D, Pan L, Li M, Tong Y. Cyanobacteria-cyanophage interactions between freshwater and marine ecosystems based on large-scale cyanophage genomic analysis. Sci Total Environ. 2024;950:175201. https://doi.org/10.1016/j.scitotenv.2024.175201.
19. Koonin EV, Dolja VV, Krupovic M, Varsani A, Wolf YI, Yutin N, et al. Global organization and proposed megataxonomy of the virus world. Microbiol Mol Biol Rev. 2020;84(2):e00061–19. https://doi.org/10.1128/MMBR.00061-19.
20. Turner D, Kropinski AM, Adriaenssens EM. A roadmap for genome-based phage taxonomy. Viruses. 2021;13(3):506. https://doi.org/10.3390/v13030506.
21. Gulyaeva A, Garmaeva S, Kurilshikov A, Vich Vila A, Riksen NP, Netea MG, et al. Diversity and ecology of *Caudoviricetes* phages with genome terminal repeats in fecal metagenomes from four dutch cohorts. Viruses. 2022;14(10):2305. https://doi.org/10.3390/v14102305.
22. Luque D, Castón JR. Cryo-electron microscopy for the study of virus assembly. Nat Chem Biol. 2020;16(3):231–9. https://doi.org/10.1038/s41589-020-0477-1.
23. Ortiz de Ora L, Wiles ET, Zund M, Banuelos MS, Haro-Ramirez N, Suder DS, et al. Phollow reveals in situ phage transmission dynamics in the zebrafish gut microbiome at single-virion resolution. Nat Microbiol. 2025;10(5):1067–83. https://doi.org/10.1038/s41564-025-01981-1.
24. Dai W, Fu C, Raytcheva D, Flanagan J, Khant HA, Liu X, et al. Visualizing virus assembly intermediates inside marine cyanobacteria. Nature. 2013;502(7473):707–10. https://doi.org/10.1038/nature12604.
25. Xiong Z, Wang Y, Dong Y, Zhang Q, Xu X. Cyanophage A-1(L) adsorbs to lipopolysaccharides of *Anabaena* sp. strain PCC 7120 via the tail protein lipopolysaccharide-interacting protein (ORF36). J Bacteriol. 2019;201(3):10.1128/jb.00516–18. https://doi.org/doi:10.1128/jb.00516-18.
26. Howe MM. The invertible G segment of phage Mu. Cell. 1980;21(3):605–6. https://doi.org/10.1016/0092-8674(80)90423-7.
27. van de Putte P, Cramer S, Giphart-Gassler M. Invertible DNA determines host specificity of bacteriophage Mu. Nature. 1980;286(5770):218–22. https://doi.org/10.1038/286218a0.
28. Giphart-Gassler M, Plasterk RHA, van de Putte P. G inversion in bacteriophage Mu: a novel way of gene splicing. Nature. 1982;297(5864):339–42. https://doi.org/10.1038/297339a0.
29. Plasterk RH, Brinkman A, van de Putte P. DNA inversions in the chromosome of *Escherichia coli* and in bacteriophage Mu: Relationship to other site-specific recombination systems. Proc Natl Acad Sci U S A. 1983;80(17):5355–8. https://doi.org/10.1073/pnas.80.17.5355.
30. Kamp D, Kardas E, Ritthaler W, Sandulache R, Schmucker R, Stern B. Comparative analysis of invertible DNA in phage genomes. Cold Spring Harb Symp Quant Biol. 1984;49:301–11. https://doi.org/10.1101/sqb.1984.049.01.036.

31. Van de Putte P, Goosen N. DNA inversions in phages and bacteria. Trends Genet. 1992;8(12):457–62. https://doi.org/10.1016/0168-9525(92)90331-w.
32. Johnson RC, Rice P, Craig N. Site-specific DNA inversion by serine recombinases. Microbiol Spectrum. 2015;3(1):1–36. https://doi.org/10.1128/microbiolspec.MDNA3-0047-2014.
33. Kennedy KE, Iida S, Meyer J, Stålhammar-Carlemalm M, Hiestand-Nauer R, Arber W. Genome fusion mediated by the site specific DNA inversion system of bacteriophage P1. Mol Genet Genomics. 1983;189(3):413–21. https://doi.org/10.1007/bf00325903.
34. Iida S, Hiestand-Nauer R, Meyer J, Arber W. Crossover sites cix for inversion of the invertible DNA segment C on the bacteriophage P7 genome. Virology. 1985;143(1):347–51. https://doi.org/10.1016/0042-6822(85)90123-0.
35. Sandmeier H, Iida S, Arber W. DNA inversion regions Min of plasmid p15B and Cin of bacteriophage P1: evolution of bacteriophage tail fiber genes. J Bacteriol. 1992;174(12):3936–44. https://doi.org/10.1128/jb.174.12.3936-3944.1992.
36. Dorgai L, Oberto J, Weisberg RA. Xis and Fis proteins prevent site-specific DNA inversion in lysogens of phage HK022. J Bacteriol. 1993;175(3):693–700. https://doi.org/10.1128/jb.175.3.693-700.1993.
37. Volkert FC, Broach JR. Site-specific recombination promotes plasmid amplification in yeast. Cell. 1986;46(4):541–50. https://doi.org/10.1016/0092-8674(86)90879-2.
38. Chan K-M, Liu Y-T, Ma C-H, Jayaram M, Sau S. The 2 micron plasmid of *Saccharomyces cerevisiae*: A miniaturized selfish genome with optimized functional competence. Plasmid. 2013;70(1):2–17. https://doi.org/10.1016/j.plasmid.2013.03.001.
39. Karvelis T, Druteika G, Bigelyte G, Budre K, Zedaveinyte R, Silanskas A, et al. Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. Nature. 2021;599(7886):692–6. https://doi.org/10.1038/s41586-021-04058-1.
40. Nakagawa R, Hirano H, Omura SN, Nety S, Kannan S, Altae-Tran H, et al. Cryo-EM structure of the transposon-associated TnpB enzyme. Nature. 2023;616(7956):390–7. https://doi.org/10.1038/s41586-023-05933-9.
41. Sasnauskas G, Tamulaitiene G, Druteika G, Carabias A, Silanskas A, Kazlauskas D, et al. TnpB structure reveals minimal functional core of Cas12 nuclease family. Nature. 2023;616(7956):384–9. https://doi.org/10.1038/s41586-023-05826-x.
42. Faure G, Shmakov SA, Yan WX, Cheng DR, Scott DA, Peters JE, et al. CRISPR-Cas in mobile genetic elements: counter-defence and beyond. Nat Rev Microbiol. 2019;17(8):513–25. https://doi.org/10.1038/s41579-019-0204-7.
43. Wolk CP, Lechno-Yossef S, Jäger KM. The insertion sequences of *Anabaena* sp. strain PCC 7120 and their effects on its open reading frames. J Bacteriol. 2010;192(20):5289–303. https://doi.org/10.1128/jb.00460-10.
44. Wang XQ, Du K, Chen C, Hou P, Li WF, Chen Y, et al. Profiling the interplay and coevolution of *Microcystis aeruginosa* and cyanosiphophage Mic1. Microbiol Spectr. 2024;12(6):e0029824. https://doi.org/10.1128/spectrum.00298-24.

45. Padan E, Ginzburg D, Shilo M. The reproductive cycle of cyanophage LPP1-G in *Plectonema boryanum* and its dependence on photosynthetic and respiratory systems. Virology. 1970;40(3):514–21. https://doi.org/10.1016/0042-6822(70)90194-7.
46. Lin X, Ding H, Zeng Q. Transcriptomic response during phage infection of a marine cyanobacterium under phosphorus-limited conditions. Environ Microbiol. 2016;18(2):450–60. https://doi.org/10.1111/1462-2920.13104.
47. Waldbauer JR, Coleman ML, Rizzo AI, Campbell KL, Lotus J, Zhang L. Nitrogen sourcing during viral infection of marine cyanobacteria. Proc Natl Acad Sci U S A. 2019;116(31):15590–5. https://doi.org/10.1073/pnas.1901856116.
48. Liu Z, Wu C. Response of alkaline phosphatases in the cyanobacterium *Anabaena* sp. FACHB 709 to inorganic phosphate starvation. Curr Microbiol. 2012;64(6):524–9. https://doi.org/10.1007/s00284-012-0101-z.
49. Lin W, Zhao D, Luo J. Distribution of alkaline phosphatase genes in cyanobacteria and the role of alkaline phosphatase on the acquisition of phosphorus from dissolved organic phosphorus for cyanobacterial growth. J Appl Phycol. 2018;30(2):839–50. https://doi.org/10.1007/s10811-017-1267-3.
50. Keroack CD, Elsworth B, Tennessen JA, Paul AS, Hua R, Ramirez-Ramirez L, et al. Comparative chemical genomics in Babesia species identifies the alkaline phosphatase PhoD as a determinant of antiparasitic resistance. Proc Natl Acad Sci U S A. 2024;121(9):e2312987121. https://doi.org/10.1073/pnas.2312987121.
51. Kageyama H, Tripathi K, Rai AK, Cha-Um S, Waditee-Sirisattha R, Takabe T. An alkaline phosphatase/phosphodiesterase, PhoD, induced by salt stress and secreted out of the cells of *Aphanothece halophytica*, a halotolerant cyanobacterium. Appl Environ Microbiol. 2011;77(15):5178–83. https://doi.org/10.1128/AEM.00667-11.
52. Luo H, Benner R, Long RA, Hu J. Subcellular localization of marine bacterial alkaline phosphatases. Proc Natl Acad Sci U S A. 2009;106(50):21219–23. https://doi.org/10.1073/pnas.0907586106.
53. Mihara T, Nishimura Y, Shimizu Y, Nishiyama H, Yoshikawa G, Uehara H, et al. Linking virus genomes with host taxonomy. Viruses. 2016;8(3):66. https://doi.org/10.3390/v8030066.
54. Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S, et al. ViPTree: the viral proteomic tree server. Bioinformatics. 2017;33(15):2379–80. https://doi.org/10.1093/bioinformatics/btx157.
55. Wang X, Jing R, Liu J, Yu Z, Jin J, Liu X, et al. Narrow distribution of cyanophage *psbA* genes observed in two paddy waters of Northeast China by an incubation experiment. Virol Sin. 2016;31(2):188–91. https://doi.org/10.1007/s12250-015-3673-5.
56. Wang X, Liu J, Yu Z, Jin J, Liu X, Wang G. Novel groups and unique distribution of phage *phoH* genes in paddy waters in Northeast China. Sci Rep. 2016;6:38428. https://doi.org/10.1038/srep38428.
57. Li X, Sun Y, Liu J, Yao Q, Wang G. Molecular diversity of Cyanopodoviruses in two coastal wetlands in Northeast China. Curr Microbiol. 2019;76(7):863–71. https://doi.org/10.1007/s00284-019-01700-0.

58. Zhu J, Yang F, Du K, Wei ZL, Wu QF, Chen Y, et al. Phylogenomics of five *Pseudanabaena* cyanophages and evolutionary traces of horizontal gene transfer. Environ Microbiome. 2023;18(1):3. https://doi.org/10.1186/s40793-023-00461-5.
59. Wandera KG, Alkhnbashi OS, Bassett HvI, Mitrofanov A, Hauns S, Migur A, et al. Anti-CRISPR prediction using deep learning reveals an inhibitor of Cas13b nucleases. Mol Cell. 2022;82(14):2714–26.e4. https://doi.org/10.1016/j.molcel.2022.05.003.
60. Mitschke J, Vioque A, Haas F, Hess WR, Muro-Pastor AM. Dynamics of transcriptional start site selection during nitrogen stress-induced cell differentiation in *Anabaena* sp. PCC7120. Proc Natl Acad Sci U S A. 2011;108(50):20130–5. https://doi.org/doi:10.1073/pnas.1112724108.
61. Kopf M, Möke F, Bauwe H, Hess WR, Hagemann M. Expression profiling of the bloom-forming cyanobacterium *Nodularia* CCY9414 under light and oxidative stress conditions. The ISME Journal. 2015;9(10):2139–52. https://doi.org/10.1038/ismej.2015.16.
62. Yu S, Yang M, Xiong J, Zhang Q, Gao X, Miao W, et al. Proteogenomic analysis provides novel insight into genome annotation and nitrogen metabolism in *Nostoc* sp. PCC 7120. Microbiol Spectrum. 2021;9(2):e0049021. https://doi.org/10.1128/Spectrum.00490-21.
63. Zhang Q, Yu S, Wang Q, Yang M, Ge F. Quantitative proteomics reveals the protein regulatory network of *Anabaena* sp. PCC 7120 under nitrogen deficiency. J Proteome Res. 2021;20(8):3963–76. https://doi.org/10.1021/acs.jproteome.1c00302.
64. Brenes-Álvarez M, Mitschke J, Olmedo-Verd E, Georg J, Hess WR, Vioque A, et al. Elements of the heterocyst-specific transcriptome unravelled by co-expression analysis in *Nostoc* sp. PCC 7120. Environ Microbiol. 2019;21(7):2544–58. https://doi.org/10.1111/1462-2920.14647.
65. Gouet P, Robert X, Courcelle E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res. 2003;31(13):3320–3. https://doi.org/10.1093/nar/gkg556.
66. Buikema WJ, Haselkorn R. Characterization of a gene controlling heterocyst differentiation in the cyanobacterium *Anabaena* 7120. Genes Dev. 1991;5(2):321–30. https://doi.org/10.1101/gad.5.2.321.
67. Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. J Gen Microbiol. 1979;111(Mar):1–61. https://doi.org/10.1099/00221287-111-1-1.
68. Ou T, Liao X-Y, Gao X-C, Xu X-D, Zhang Q-Y. Unraveling the genome structure of cyanobacterial podovirus A-4L with long direct terminal repeats. Virus Res. 2015;203:4–9. https://doi.org/10.1016/j.virusres.2015.03.012.
69. Bancroft I, Smith RJ. Restriction mapping of genomic DNA from five cyanophages infecting the heterocystous cyanobacteria *Nostoc* and *Anabaena*. New Phytol. 1989;113(2):161–6. https://doi.org/10.1111/j.1469-8137.1989.tb04702.x.
70. Khudyakov I, Wolk CP. Evidence that the *hanA* gene coding for HU protein is essential for heterocyst differentiation in, and cyanophage A-4(L) sensitivity of, *Anabaena* sp. strain PCC 7120. J Bacteriol.

1996;178(12):3572–7. https://doi.org/10.1128/jb.178.12.3572-3577.1996.

71. Casjens SR, Gilcrease EB. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. Methods Mol Biol. 2009;502:91–111. https://doi.org/10.1007/978-1-60327-565-1_7.
72. Li S, Fan H, An X, Fan H, Jiang H, Chen Y, et al. Scrutinizing virus genome termini by high-throughput sequencing. PLoS One. 2014;9(1):e85806. https://doi.org/10.1371/journal.pone.0085806.
73. Garneau JR, Depardieu F, Fortier L-C, Bikard D, Monot M. PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. Sci Rep. 2017;7(1):8292. https://doi.org/10.1038/s41598-017-07910-5.
74. Nanamiya H, Tanaka D, Hiyama G, Isogai T, Watanabe S. Detection of four isomers of the human cytomegalovirus genome using nanopore long-read sequencing. Virus Genes. 2024;60(4):377–84. https://doi.org/10.1007/s11262-024-02083-6.

Figures

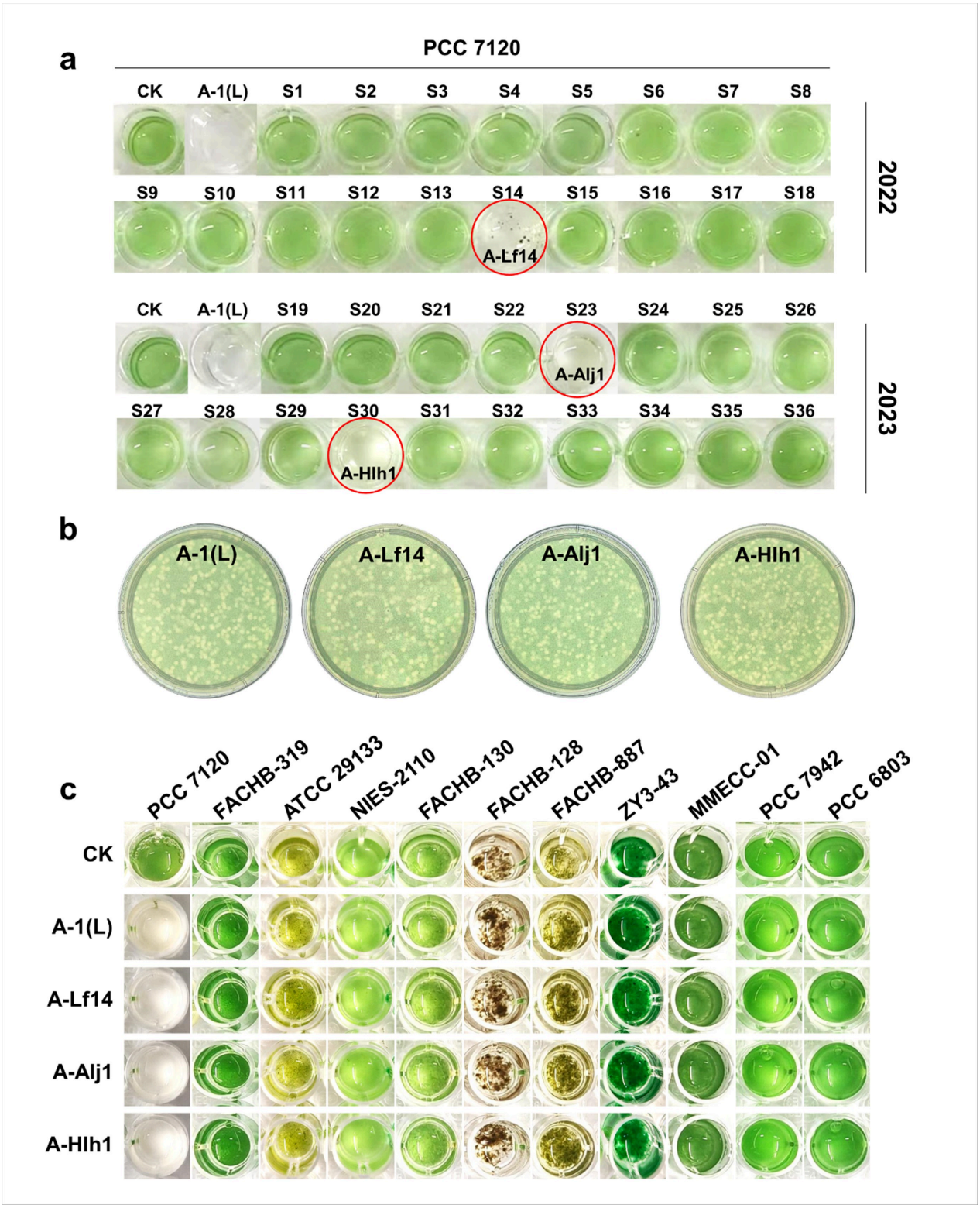


Figure 1

Isolation and host range analysis of novel cyanophages infecting *Anabaena* sp. PCC 7120. (**a**) Primary screening of cyanophage-induced lysis. Environmental filtrates from 36 freshwater sampling sites (S1–S18 collected in 2022; S19–S36 collected in 2023, Additional file 2: Table S2) were incubated with *Anabaena* sp. PCC 7120 in 24-well plates. The uninfected control (CK) and a positive control infected with the reference cyanophage A-1(L) are shown. Cultures exhibiting complete lysis (chlorosis and

clarification) within 24–48 h were selected for further purification. (**b**) Clear and distinct lytic halos formed by cyanophages on lawns of *Anabaena* sp. PCC 7120, confirming lytic activity in the initial liquid screening. Assays were performed using mid-exponential phase host cultures (OD_{730}=0.6) infected with fresh cyanophage stock solution (1×10^8 PFU mL^{-1}). Plates were incubated for 72 h at 30 °C under continuous illumination (50 µmol photons m^{-2} s^{-1}) using BG-11 medium with a hard agar base (1.5% w/v) and a soft agar overlay (0.6% w/v). (**c**) Host range profiling. Infection assays were performed using four *Caudoviricetes* cyanophages against 10 phylogenetically distinct cyanobacterial strains plus *Anabaena* sp. PCC 7120. The tested strains included Anabaena variabilis FACHB-319, Nostoc punctiforme ATCC 29133, Nostoc sp. NIES-2110, Nostoc spongiaeforme FACHB 130, Tolypothrix tenuis FACHB-128, Scytonema javanicum FACHB-887, Trichormus sp. ZY3-43, Limnospira platensis MMECC-01, Synechococcus elongatus PCC 7942, *Synechocystis* sp. PCC 6803 using four *Caudoviricetes* cyanophages [A-1(L), A-Lf14, A-Alj1, A-Hlh1]. All infection assays were conducted with biological triplicate replicates.

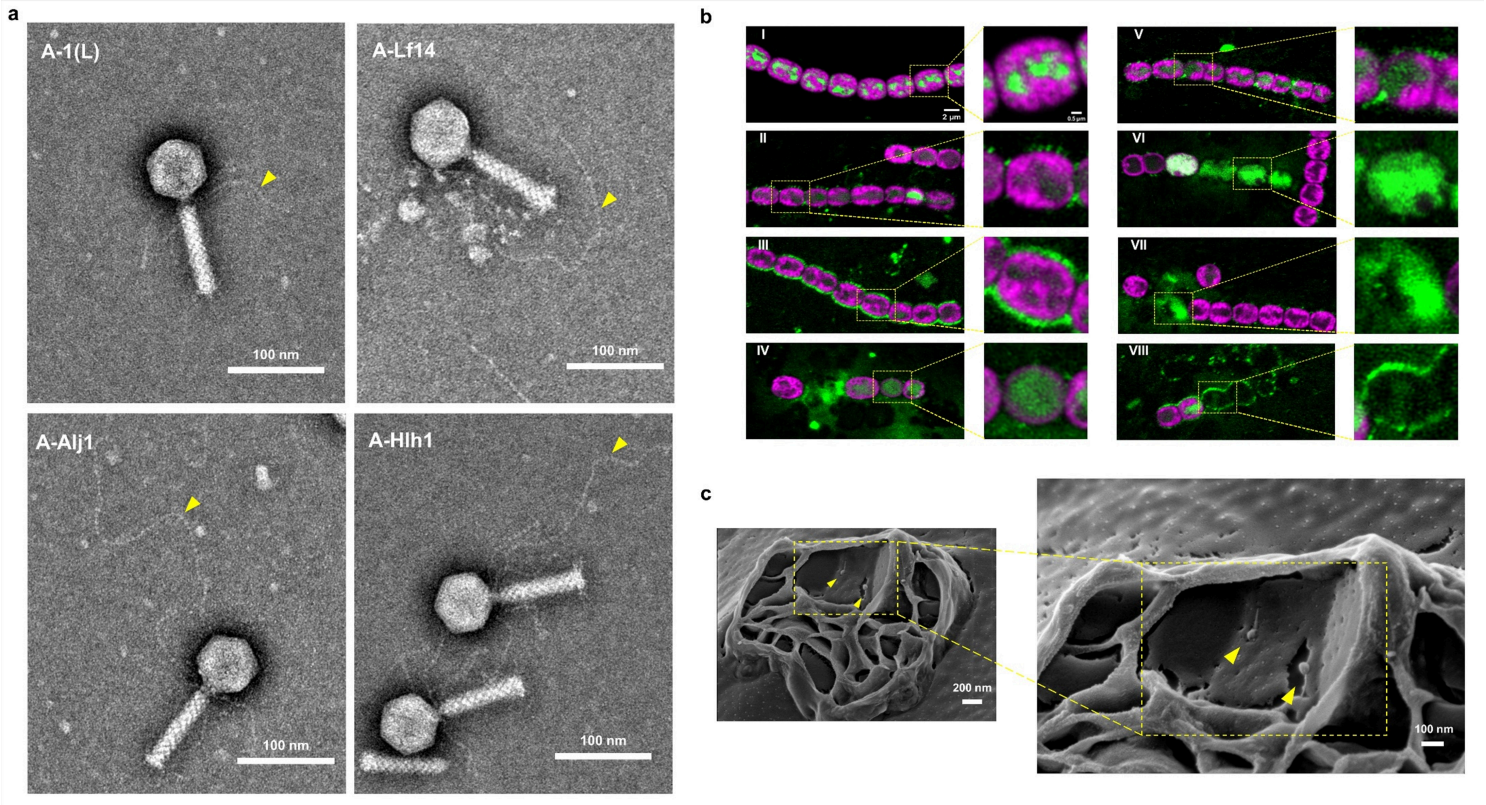


Figure 2

Morphological characterization of cyanophage virions and infection progression in *Anabaena* sp. PCC 7120. (**a**) TEM of negatively stained cyanophage particles. Virions of A-1(L), A-Lf14, A-Alj1, and A-Hlh1, stained with 3% uranyl acetate, display characteristic *Caudoviricetes* (Myovirus) morphology, with icosahedral capsids (64 ± 2 nm in diameter, n = 20), contractile tails (95 ± 5 nm in length, n = 20). Notably, the neck fibers (yellow arrows) in A-1(L) are substantially shorter than those observed in the three new isolates. (**b**) Visualization of cyanophage A-Lf14 infection by confocal fluorescence

microscopy. Samples containing uninfected or phage-infected *Anabaena* filaments were stained with DAPI (pseudocolored green; ex/em 401/422 nm) and imaged using a Zeiss LSM 880 microscope with Airyscan detection. Host phycobilisome autofluorescence is shown in magenta (ex/em 548/561 nm). The overlay images depict key stages of the infection cycle: (I) uninfected control filament; (II–III) phage adsorption and accumulation on the cell surface; (IV) intracellular virion assembly; (V–VII) release and propagation; (VIII) phage particles associated with empty cellular sheaths. See also Additional file 1: Fig. S1 and Additional file 3: Video S1. (**c**) Intracellular virion assembly observed by cryo-scanning electron microscopy (cryo-SEM). *Anabaena* sp. PCC 7120 cells infected with A-1(L) were plunge-frozen at 8 h post-infection. Arrows indicate mature phage particles within the host cytoplasm, exhibiting typical capsids and tails consistent with the *Caudoviricetes*. The corresponding enlarged view of the boxed area is presented on the right, highlighting the assembled virions. Imaging was performed on a Zeiss Ultra Plus HR-SEM at 2 kV with a stage temperature of -145 °C. See also Additional file 1: Fig. S2.

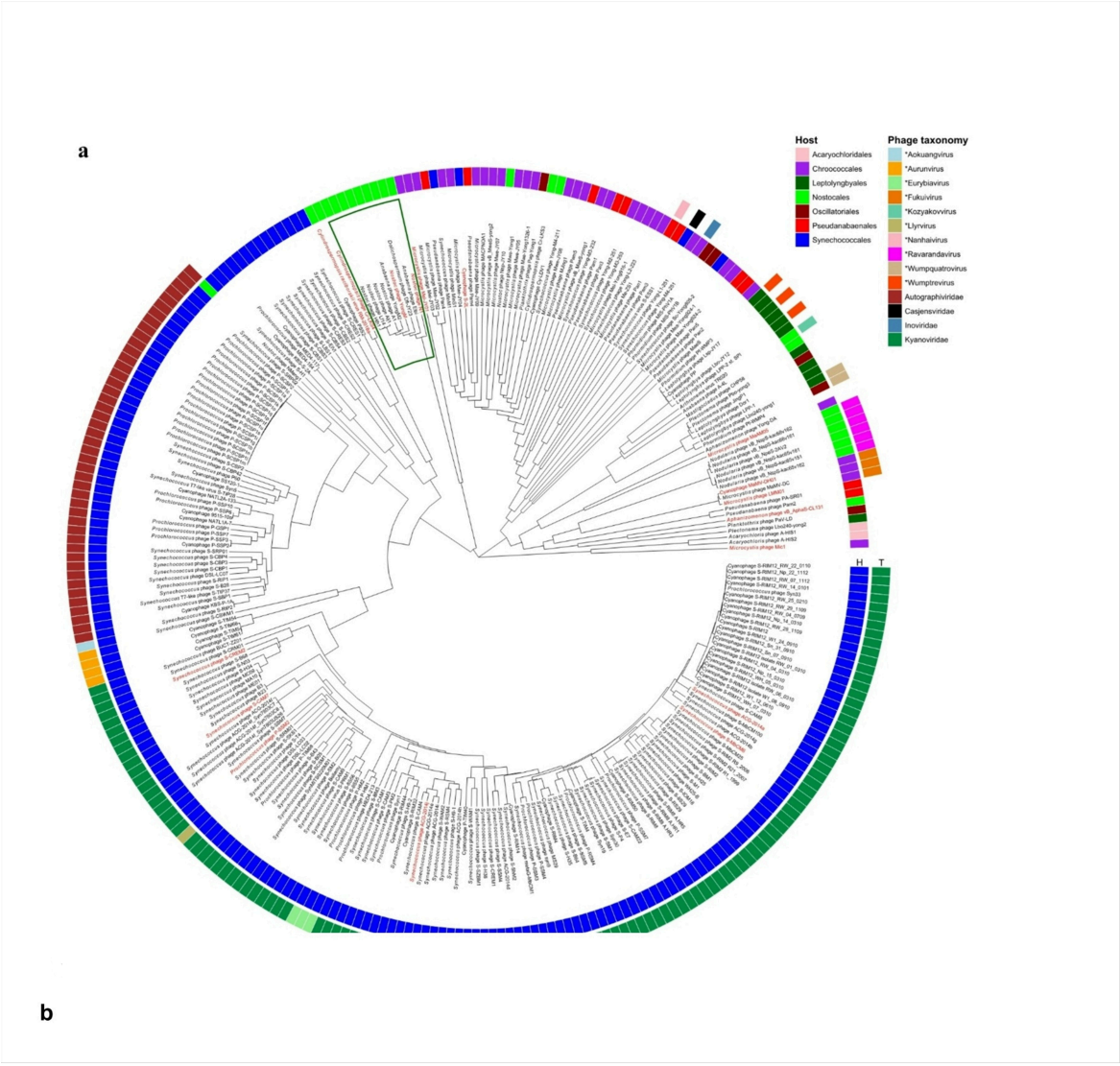


Figure 3

Phylogenomic and comparative genomic analysis of IR-containing cyanophages. (**a**) Phylogenomic analysis of 280 cyanophage whole-proteomes using Viptree [54]. The tree is presented at logarithmic scale, host (H) and phage taxonomy (T) were added to the picture. Phages encoding CRISPR arrays are highlighted in red. The clade containing *Anabaena* phage A-1(L) and related isolates is boxed in green. (**b**) Comparative genomic analysis of nine cyanophages containing a conserved invertible genomic

element (blue rectangle). Linear genome maps are shown, with genes colored according to predicted function and labeled with their corresponding gene cluster numbers. For better comparison, all genomes were re-drawn, with the start codon of the gene cluster 94 at the left terminus and all invertible regions oriented consistently. See Additional file 1: Fig. S3for an alignment of the IR sequences flanking the invertible regions.

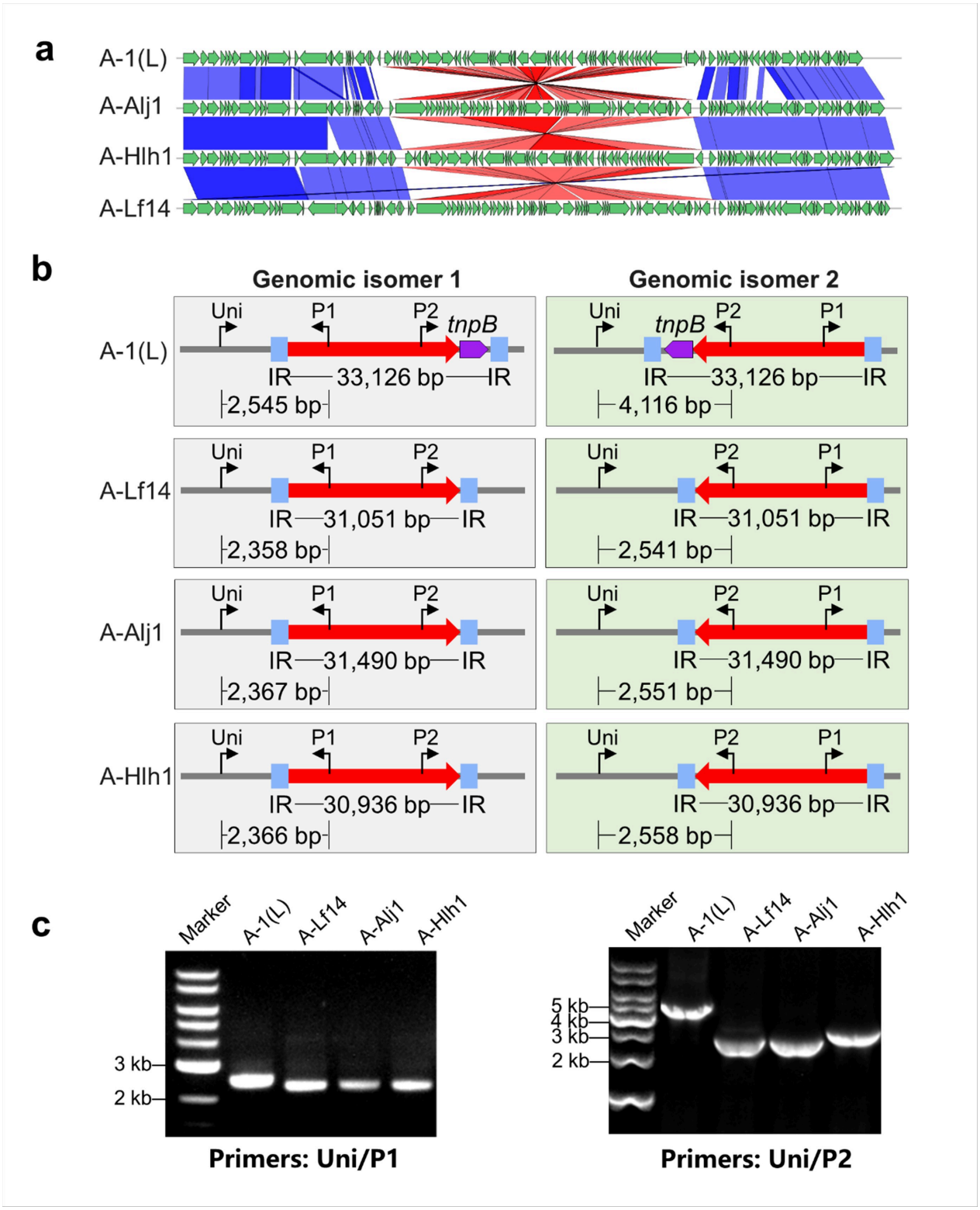


Figure 4

Reversible orientation isomers of a flexible genetic module in *Anabaena*-infecting *Caudoviricetes* cyanophages. (**a**) Comparative genomic visualization of the four *Anabaena* sp. PCC 7120-infecting cyanophages. Genomic comparisons were performed using GBKviz. The alignment relationships between the conserved flanking regions across the four phages are connected and shown in blue, while those between the internal invertible modules are connected and highlighted in red. (**b**) Schematic of the two genomic isomers and diagnostic PCR strategy. The diagram illustrates the two orientation isomers (Isomer 1 and 2) and shows the specific structures of phages A-Lf14, A-Alj1, and A-Hlh1. Inverted repeats (IRs) that flank the module are marked. To distinguish the isomers, a unique primer 'Uni' anchored in the flanking stable region was paired with outward-facing primers P1 or P2 located within the conserved invertible module, enabling specific amplification of each isomer. The predicted sizes (in bp) of the resulting Uni/P1 and Uni/P2 amplicons are indicated for each phage, and the length of the invertible fragment itself is annotated. For reference in later comparative analysis, the transposase gene *tnpB*, which located merely inside the invertible module and adjacent to the IR in phage A-1(L), is also labeled in the schematic. (**c**) Experimental validation of both isomers by PCR. Agarose gel electrophoresis of PCR products from the four cyanophage isolates using diagnostic primer pairs Uni/P1 (left panel) and Uni/P2 (right panel). Molecular size markers are indicated in lane M. The observed band sizes correspond to the predictions in panel (**b**), confirming the coexistence of both isomeric configurations in each phage culture.

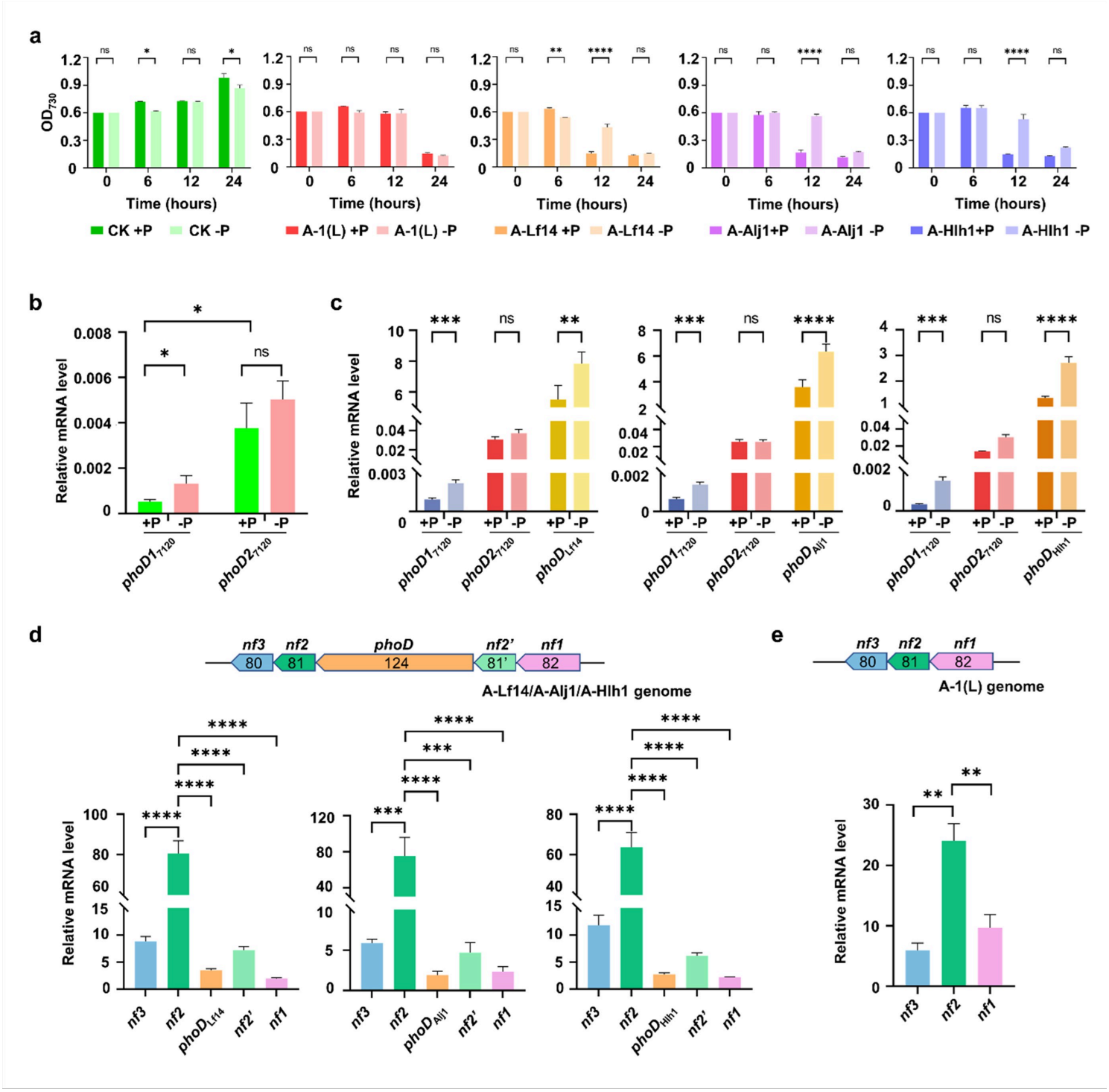


Figure 5

Phosphate availability modulates cyanophage infectivity in *Anabaena* sp. PCC 7120 via phage-encoded *phoD* homologs. (**a**) Growth curves of *Anabaena* sp. PCC 7120 under phosphate-replete (+P) and phosphate-deplete (-P) conditions, with and without infections by cyanophages A-1(L), A-Lf14, A-Alj1 and A-Hlh1. Cell density was measured as OD_{730} at 0, 6, 12, and 24 h post-infection. CK, uninfected control. (**b**) Transcript levels of host *phoD1* (*alr2234*) and *phoD2* (*alr4976*) in uninfected *Anabaena* sp. PCC 7120 under +P and -P conditions. (**c**) Relative transcript abundance of the host $phoD1_{7120}$ and $phoD2_{7120}$ and

cyanophage-encoded *phoD* homologs ($phoD_{Lf14}$, $phoD_{Alj1}$ and $phoD_{Hlh1}$) under +P and -P conditions at 6 h post-infection with A-Lf14, A-Alj1 and A-Hlh1, respectively. (**d**) Expression profiles of phage neck fiber (*nf*) genes located immediately adjacent to the *phoD* locus in A-Lf14, A-Alj1 and A-Hlh1 at 6 h post-infection (the A-Lf14/A-Alj1/A-Hlh1 genomes are collectively represented). (**e**) Expression of the neck fiber (*nf*) genes in the *phoD*-lacking phage A-1(L) at 6 h post-infection. The numbers shown on the gene maps in panels (**d**) and (**e**) indicate the respective gene cluster numbers (Additional file 2: Table S6). For all panels, data are presented as mean ± SD (n ≥ 3). Statistical significance was determined by unpaired Student's t-test; ns, not significant (P > 0.05); *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001.

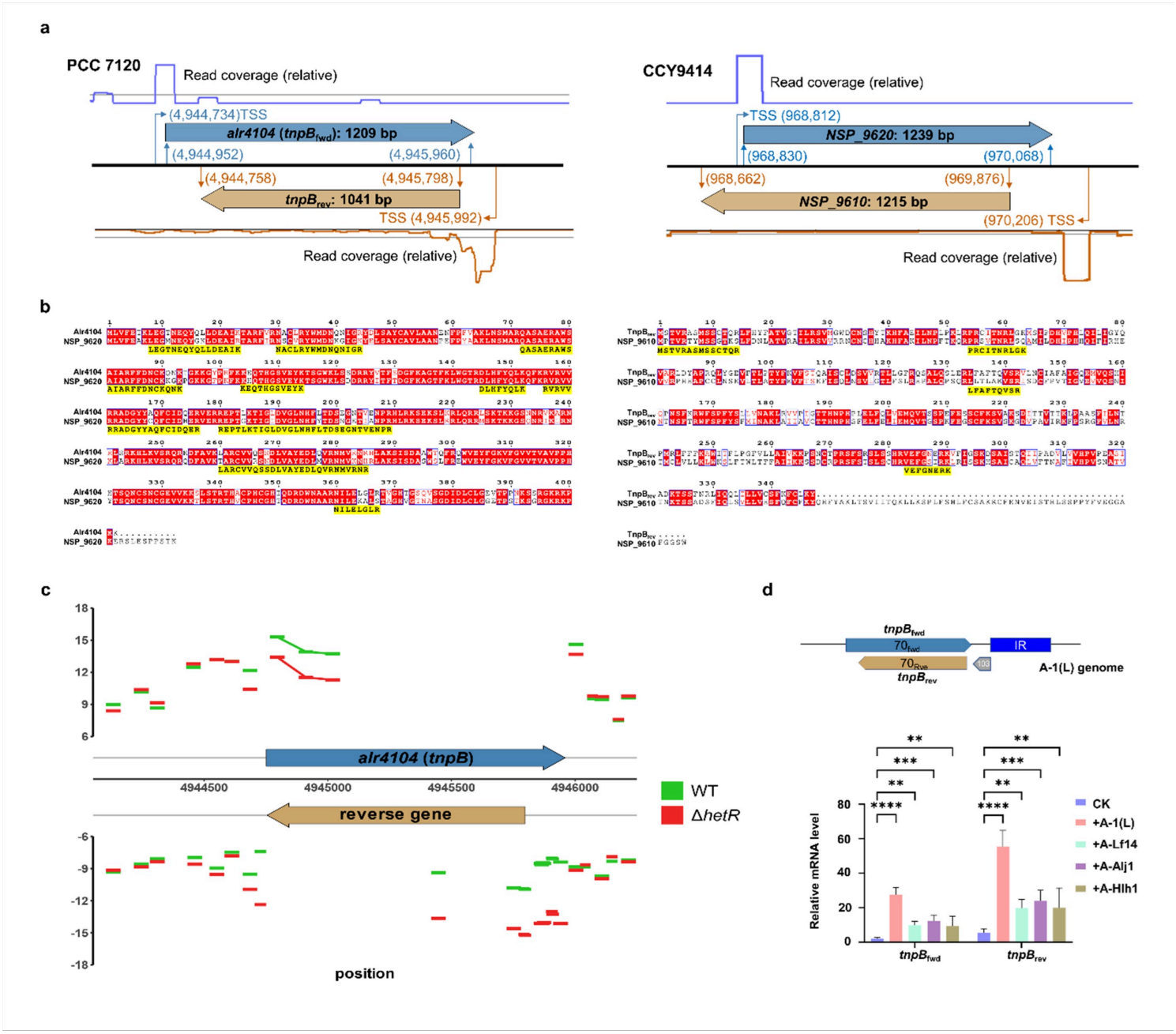


Figure 6

Dual-strand architecture and expression of the *tnpB* locus across species. (**a**) Transcriptional organization of the *tnpB* locus in *Anabaena* sp. PCC 7120 (left) and *Nodularia spumigena* CCY9414 (right). The forward-strand gene ($tnpB_{fwd}$, annotated as *alr4104* and *NSP_9620*, respectively) and the reverse-strand gene ($tnpB_{rev}$, NSP_9610) are shown. Arrows indicate transcriptional start sites (TSS) identified from differential RNA-seq data [60, 61]. Arrowheads mark start and stop codons. Gene models for *Nodularia spumigena* were refined based on TSS positions. (**b**) Sequence alignment of $TnpB_{fwd}$ and $TnpB_{rev}$ from *Anabaena* sp. PCC 7120 and the predicted homologs from *Nodularia spumigena* CCY9414 using program ESPript [65]. Yellow-highlighted peptides were identified in published proteomic datasets [62, 63] confirming translation of both proteins *in vivo*. (**c**) Transcriptomic data from WT (green) and a *hetR* (*alr2339*) inactivation mutant [66] (red) for the genomic region of $tnpB_{fwd}$ (*alr4104*) and $tnpB_{rev}$. The vertical tabs indicate the positions of individual oligonucleotide microarray probes. The RNA expression levels (Y axis) are shown on a $\log_2$ scale. The data originated from a publication by Brenes-Álvarez M, *et al*. (2019) [64]. WT expression values represent average expression levels derived from biological triplicates (GEO datasets GSM3399496, GSM3399497 and GSM3399512) and values for the Δ*hetR* mutant were averaged from duplicate samples GSM3399504 and GSM3399505. (**d**) qRT-PCR quantification of $tnpB_{fwd}$ and $tnpB_{rev}$ transcripts during infection with A-1(L) (which carries *tnpB*) and with A-Lf14, A-Alj1 and A-Hlh1 (which lack *tnpB*). Both strands were strongly induced specifically during A-1(L) infection. Numbers on the gene map correspond to gene cluster annotations in Additional file 2: Table S6. Data are presented as mean ± SD (n ≥ 3). Statistical significance was determined by unpaired Student's t-test; **P < 0.01, ***P < 0.001, ****P < 0.0001.

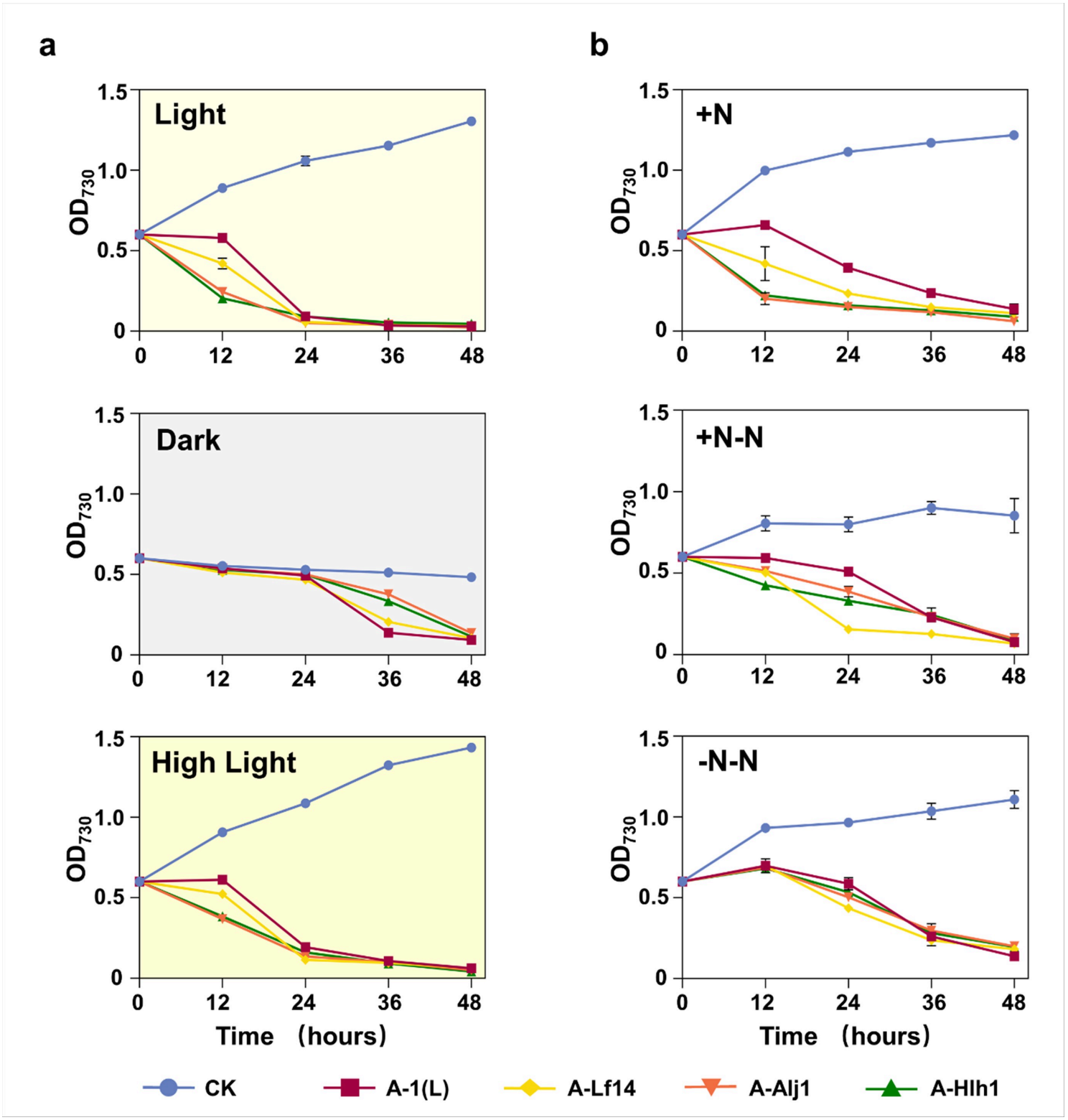


Figure 7

Environmental regulation of cyanophage infection dynamics in *Anabaena* sp. PCC 7120. (**a**) Light intensity modulates lytic kinetics. Growth curves for cyanophages A-1(L), A-Lf14, A-Alj1, and A-Hlh1 were determined under varying light conditions: standard light (50 µmol photons m^{-2} s^{-1}), darkness, and high light (150 µmol photons m^{-2} s^{-1}). Uninfected controls are included for each condition (CK). (**b**) Nitrogen availability constrains infection progression. Infection profiles were assessed under different nitrogen

regimes: nitrogen-replete conditions (+N), nitrogen-deprived infection in nitrogen-replete cultured cells (+N-N), and nitrogen-deprived infection in nitrogen-starved cells (-N-N). Uninfected controls for each nitrogen treatment are presented (CK). Data represent means ± SD of three biological replicates. Cultures were grown in BG-11 medium under standard conditions unless otherwise specified, and infection was monitored by measuring culture turbidity at OD_{730} over a 48-hour period.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.SupplementaryFig.S1S6.pdf
- Additionalfile2.SupplementaryTablesS1S8.xlsx
- Additionalfile3.LegendforSupplementaryVideo1.pdf
- Additionalfile3.SupplementaryVideo1.avi