Virology Journal

RESEARCH

Open Access

# Characterization and genomic analysis of Sharanji: a jumbo bacteriophage of *Escherichia coli*

Sharayu Magar[1], Sivaraj Barath[1], Debmitra Sen[1], Ranjith Kumar Singari[1], T. Nagarajan[1], Anjali Parmar[1] and Sutharsan Govindarajan[1*]

## Abstract

**Background** Bacteriophages are the most genetically diverse biological entities in nature. Our current understanding of phage biology primarily stems from studies on a limited number of model bacteriophages. Jumbo phages, characterized by their exceptionally large genomes, are less frequently isolated and studied. Some jumbo phages exhibit remarkable genetic diversity, unique infection mechanisms, and therapeutic potential.

**Methods** In this study, we describe the isolation of Sharanji, a novel *Escherichia coli* jumbo phage, isolated from chicken feces. The phage genome was sequenced and analyzed extensively through gene annotation and phylogenetic analysis. The jumbo phage was phenotypically characterized through electron microscopy, host range analysis, and survival at different pH and temperatures, and one-step growth curve assay. Finally, Sharanji mediated infection of *E. coli* is studied through fluorescence microscopy, to analyze its mechanism of infection compared to well-studied nucleus-forming jumbo phages.

**Results** Whole genome sequencing reveals that Sharanji has a genome size of 350,079 bp and is a phage encompassing 593 ORFs. Genomic analysis indicates that the phage belongs to the *Asteriusvirus* genus and is related to *E. coli* jumbo phages PBECO4 and 121Q. Phenotypic analysis of isolated phage Sharanji, indicates that the phage size is 245.3 nm, and it is a narrow-spectrum phage infecting *E. coli* K12 strains, but not other bacteria including avian pathogenic *E. coli*. Infection analysis using microscopy shows that Sharanji infection causes cell filamentation. Furthermore, intracellular phage nucleus-like structures were not observed in Sharanji-infected cells, in contrast to infection by ΦKZ-like jumbo phages.

**Conclusions** Our study reports the isolation and characterization of Sharanji, one of the large *E. coli* jumbo phages. Both genotypic and phenotypic analyses suggest that Sharanji serves as a unique model system for studying phage-bacteria interactions, particularly within the context of non-nucleus-forming jumbo phages. Further exploration of jumbo phages holds promise for uncovering new paradigms in the study of microbial viruses.

**Keywords** Jumbo phage, *E. coli*, Bacteriophage genomics

## Background

Jumbo phages are giant bacterial viruses with exceptionally larger genomes greater than 200 kb [1]. However, more recent studies based on comparative genomic analysis suggested that phages having genome size greater than 180 kb can be potentially considered as jumbo

*Correspondence:
Sutharsan Govindarajan
sutharsan.g@srmap.edu.in; kamalsuthan@gmail.com
[1] Department of Biological Sciences, SRM University – AP, Amaravati, Andhra Pradesh 522 240, India

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

phages [2], while phages with genome size more than 500 kb are regarded as mega phages [3]. Jumbo phages are more commonly isolated from aquatic environments [4–6]. The current model suggests that jumbo phages are polyphyletic in nature, having evolved from smaller phages multiple times throughout evolution. As a result, jumbo phages represent a genetically diverse group characterized by their larger genome and capsid size as a common feature [2]. The majority of the isolated jumbo phages infect gram-negative bacteria, but a few jumbo phages infect gram-positive bacteria, especially *Bacillus* [1, 6, 7].

Jumbo phages have been investigated for various applications, including the treatment of multi-drug resistant pathogens [8, 9], wastewater management [4], and the control of pathogens in plants and aquaculture [5, 10–12]. The fundamental biology of most jumbo phages are poorly characterized. However, recent studies on Chimalliviridae family of jumbo phages have highlighted their distinct genome organization, evolution, structural organization and mechanism of infection. These phages display a remarkable feature in their ability to protect their replicating DNA from host defense systems like the DNA-targeting CRISPR-Cas systems and the restriction modification systems through a protein-based organelle called as the phage nucleus [13, 14]. Additionally, some of these jumbo phages exhibit novel features including the formation of phage-encoded tubulin-like cytoskeletal system for intracellular transport [15, 16], and the organization of virion DNA through a set of DNA organization proteins known as Inner Body proteins [17–20]. These extraordinary features of jumbo phages make them an intriguing model system for fundamental understanding of host–pathogen interaction [5, 21].

Despite their significance in biotechnological applications and fundamental phage biology, jumbo phages have been inadequately isolated, resulting in limited investigation [1]. This is especially true for large jumbo phages with a genome size exceeding 300 or 350 kb. Interestingly, metagenomic analysis of the Earth's microbiome reveals a high abundance of jumbo phages [6, 22, 23]. Furthermore, a large number of diverse jumbo phages were also detected in animal and human microbiomes [23–25]. However, the majority of these jumbo phages have not been isolated. The challenges in isolating jumbo phages may stem from the complexity of the host but also due to limitations of commonly employed procedures for phage isolation. For instance, the isolation of phages from sample filtrates, which are typically filtered using a 0.22-µm filter, and the propagation of phages forming distinct plaques pose difficulties. Due to their large capsid size, some jumbo phages may not effectively pass through the 0.22-µm filter, leading to their exclusion from isolation [1, 12]. This is also observed through metagenomic investigations in which the fraction of jumbo phages is more frequently present above the 0.22-µm filter [6]. Additionally, jumbo phages exhibit host bias, meaning they are commonly isolated from some hosts but rarely or never isolated from others [26]. For example, Mycobacteriophage are widely isolated throughout the world [27], yet none of the isolated phages belong to the category of jumbo phages. The reason for this host bias is currently unclear but it could be due to the ecological distribution or intracellular life style. Furthermore, jumbo phages often produce very tiny plaques in standard double-layered agar methods, which can often be overlooked by researchers involved in the isolation process [1, 28]. In this study, we address some of the challenges in isolating jumbo phages by describing the isolation and characterization of a novel *Escherichia coli* jumbo phage, Sharanji, through a modified phage isolation approach.

## Materials and methods

### Bacterial strains, phages, and plasmids

For phage isolation and propagation, the bacteria *Escherichia coli* strain MG1655 was used. *E. coli* MG1655 was cultured in Luria–Bertani (LB) broth at 37 °C with 120 rotations per minute (rpm) agitation. LB media was also used for the growth of *E. coli W3110, E. coli* MCC2552, *E. coli* ECN1, *E. coli* MCC2246, *P. aeruginosa* PA01, *K. pneumonia* ATCC 33495, *A. baumannii* MTCC1425, *B. subtilis PY79, S. aureus (clinical).* When appropriate, antibiotic ampicillin was added at 100 µg/ml concentration (Sisco Research Laboratories). *Escherichia* phage vB_EcoM_Goslar (DSM 104658) (Goslar), purchased from Leibniz Institute DSMZ—German Collection of Microorganisms and Cell Cultures GmbH, was propagated as previously described [13]. Plasmids expressing mCherry (pBAD-mCherry) were constructed by Gibson assembly in the pBAD18 plasmid. mCherry was amplified from pSG30T-sfCherry-csy1(IF) [13] using the primers F-pBAD18-rbs-cherry and R-pBAD18-mcherry primers. The resulting amplicons were Gibson assembled in the pBAD18 vector backbone which was amplified using F-pBAD18-gib and R-pBAD18 gib primers. Primer sequences are given below.

F-pBAD18-rbs-cherry—ccataccgttttttgggctagcgaattcTTAACTTTAAGAAGGAGATATAATGGAGGAGGACAACATGGC.

R-pBAD18-mcherry—cttctctcatccgccaaaacagccaagcttTCAGCCGCCGGTGCTGTGTC.

F-pBAD18-gib—aagcttggctgttttggcggatgagagaag.

R-pBAD18 gib—gaattcgctagcccaaaaaaacgggtatgg.

### Phage isolation through filter trapping

Chicken fecal samples were collected from Poultry farms located in Mangalagiri mandal, Andhra Pradesh, India. The samples were homogenized by vortexing with sterile SM buffer (50 mM Tris–HCl, 100 mM NaCl, 8.5 mM MgSO4, pH 7.5) in a 50 mL falcon tube. After homogenization, the samples were briefly centrifuged at 3000 rpm for 10 min to remove large particles and debris. To eliminate bacteria from the supernatant, a few drops of chloroform were added, and the mixture was vigorously vortexed before being centrifuged at 8000 rpm for 10 min. The resulting supernatant, free from bacteria and debris, was then filtered through a 0.22-µm syringe filter. The filter was washed twice with a 10 ml sterile SM buffer to ensure complete elution of small phages that are not trapped. Subsequently, the filter was inverted, and phages trapped on the filter surface, which may include jumbo phages, were eluted with 1 ml SM buffer using a syringe.

To test the presence of phages in the eluted sample, 100 µl of the elution were mixed with 5 ml LB and a 1:100 dilution of overnight culture of *E. coli* MG1655, followed by overnight (16 h) incubation at 30 °C with aeration. The absence of bacterial growth in selected tubes indicated the presence of phages. These samples were then further processed for phage isolation using the double-layer agar method. In this method, 3µL of phage lysate and overnight grown 100 µL of host bacteria (*E. coli* MG1655) added into 3 mL of 0.35% top agar. This mixture was then spread evenly on the top of 2% LB bottom agar and allowed to solidify completely. This plate was subsequently incubated at 30 °C for overnight for visualization of plaques.

### Phage propagation and purification

Phages obtained from individual plaques were collected and mixed with 200 µL of SM buffer. This phage lysate was subsequently used for a whole-plate assay to propagate and purify the phage. The whole-plate assay is similar to the double-layer agar method described earlier; however, at the end of the process, 4 mL of SM buffer was added on top and left at room temperature for 4 h. Subsequently, the buffer containing the phage was transferred into a new Falcon tube, a few drops of chloroform were added, and the mixture was vortexed vigorously for 2 min. The content was then centrifuged at 10,000 rpm for 10 min, and the supernatant containing the phage lysate was carefully collected. Each phage was processed for three consecutive purifications, and the final lysates were stored at 4 °C.

### Transmission electron microscopy

For TEM imaging, 10 µL of high-titer purified phage lysate was spotted onto a 200-mesh carbon film grid and allowed to set at room temperature for 5 min. Excess phage lysate was then removed using Kimwipes. The adsorbed phages were negatively stained with Uranyless (Electron Microscopy Sciences) stain. Subsequently, the phage-coated carbon grids were imaged using a Joel JEM 2100 plus electron microscope at 200 kV voltage. Head length and widths, tail lengths were measured as previously described [12].

### Phage DNA purification and whole genome sequencing

High titer ($10^8$ PFU/ml) lysates of phage were produced from plate lysates as described above, and genomic DNA (gDNA) was extracted from 5 ml of lysates. The genomic DNA of Sharanji was isolated from the lysate using the Promega Wizard DNA Clean Up Kit (Promega catalogue #A7280) according to the manufacturer's instructions and as described previously [13]. The extracted DNA samples were analyzed on a 1% agarose gel by agarose gel electrophoresis sand imaged using Bio-Rad Gel Doc. Libraries for whole genome sequencing were prepared using the KAPA DNA HyperPrep kit, followed by sequencing on an Illumina NovaseqX plus instrument to obtain 2×151 bp reads and around 3.7–4.3 GB data was generated for the sample. We checked the following parameters from the fastq file—Base quality score distribution, sequence quality score distribution, average base content per read, GC distribution in the reads, PCR amplification issue, overrepresented sequences and adapters. Based on the quality report of fastq files reads were trimmed, to only retain high-quality sequences for further analysis. In addition, the low-quality sequence reads were excluded from the analysis. The adapter trimming was performed using fastq mcfv1.04.803 (https://expressionanalysis.github.io/ea-utils/). De-novo whole genome assembly was carried out using spades assembler (v3.11.1) [29] with parameters -k 21,33,55, --phred-offset 64 and –cov-cutoff 100 and careful correction.

### Genome annotation and comparative genome analysis

BLASTN search was performed to find out closest relatives of these sequenced phage genomes. ORFs were predicted by using Rapid Annotation using Subsystem Technology Pipeline (RASTtk, https://rast.nmpdr.org/rast.cgi) [30]. This annotation was further complemented using HMMSCAN software (http://hmmer.org) against Pfam Database (shifted to InterPro database; earlier at ftp://ftp.ebi.ac.uk/pub/databases/Pfam/current_release/Pfam-A.hmm.dat.gz) with E-value cut-off of 1e−5 and further parsed using hmmscanparser.sh to identify

confirmatory non-overlapping domains of known functions and motifs within all genes. Additionally, NCBI CDD database (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi) and Prokka (https://github.com/tseemann/prokka) with default parameters were used to annotate Sharanji genome. PhageScope v1.3 [31] and PhageTerm v1.0.11 [32] tools were utilized for predicting phage life cycle and genomic DNA nature. The taxonomic classification of Sharanji was carried out using taxmyPHAGE, a tool that compares the query phage genome against viruses classified by the ICTV. It provides insights into whether the phage represents a new species or genus based on ANI (Average Nucleotide Identity) scores of the genomes [33]. The comparative analysis began by aligning the sequenced phage genome with related genomes using the Viptree web server. Following this alignment, a proteomic tree was constructed based on the genome-wide sequence similarities which were calculated by tBLASTx. Circular and rectangular phylogenetic tree was then created to visualize the results. ANI of closely related phages of Sharanji was calculated using VIRIDIC web server with 95 and 75 percentage of threshold for species and genus [34]. Circular map of Sharanji phage genome was visualized using SnapGene Viewer 7.0.3 (https://aur.archlinux.org/snapgene-viewer.git). The Linear Genome Plot tool from Galaxy version 1.0 (https://cpt.tamu.edu/new-tool-linear-genome-plot-2/) was employed to create a linear map of Sharanji [35]. Proteinortho6 was used to identify a pan-genome of Sharanji and its closely related species (Parameters used; -identity=35 -cov=50), revealing unique genes present in the Sharanji genome [36]. DiGAlign (Dynamic Genomic Alignment server) (https://www.genome.jp/digalign/) server was used to align (tBLASTx function) and visualize the genome alignments. ABRicate v 1.0.1 software (https://github.com/tseemann/abricate) was used to identify antimicrobial resistance and virulence genes against NCBI, CARD, ARG-ANNOT, Resfinder, *E. coli*_VF and Virulence factors databases with 30% and 50% identity and coverage respectively.

#### Host tropism assay

To analyze the host range of the phage, plaque assay was performed by double layer agar method as described above. Tested strains include *E. coli* MG1655 *E. coli W3110, E. coli* MCC2552, *E. coli* ECN1, *E. coli* MCC2246, *P. aeruginosa* PA01, *K. pneumonia* ATCC 33495, *A. baumannii* MTCC 1425, *B. subtilis* PY79, *and S. aureus (clinical isolate).*

#### pH stability analysis

To assess the pH stability of phage, SM buffer of various pH (values of 2, 4, 6, 7, 8, 10 and 12) were prepared using 1 M HCl and 5 N NaOH. 50 µl of phages ($10^{10}$ PFU/ml) were added to 450 µl of SM buffer (of different pH) and incubated at 37 °C for 1 and 2 h followed by serial dilution. Spot tests were performed as described previously and incubated at 37 °C overnight. The assay was done in triplicates independently. The mean of PFU/ml were plotted along with standard errors of the means (SEM) and unpaired student t-test was performed using GraphPad Prism 10.

#### Thermostability analysis

The temperature stability of the phage was analyzed by incubating 50 µl of phage ($10^{10}$ PFU/ml) at different temperatures (4 °C, 20 °C, 30 °C, 37 °C, 40 °C, 50 °C, 60 °C) for 1 and 2 h. Serial dilution followed by spot tests were performed. The plates were incubated overnight at 37 °C. The assay was done in triplicates independently. The mean of PFU/ml were plotted along with standard errors of the means and unpaired student t-test was performed using GraphPad Prism 10.

#### One step growth curve

The one-step growth curve was performed to calculate the phage growth kinetics and the latent period. The method was performed as previously suggested in [37]. A multiplicity of infection (MOI) of 1 was maintained. 100 µL of *E. coli* MG1655 was grown in 10 mL of fresh LB until an OD of 0.4 was reached. Then, 100 µL of Sharanji phage lysate (corresponding to MOI of 1) was added to the 10 mL culture of OD 0.4 *E. coli* MG1655. Phages were allowed to adsorb for 10 min at 37 °C in static condition. Following the adsorption, unadsorbed phages were removed by centrifugation at 5000 rpm for 8 min, and the resulting pellet was resuspended in 10 mL of fresh LB. The culture was maintained at 37 °C in static condition and was sampled at 10 min intervals to determine the phage titer by spot tests. The values were plotted with SEMs, and burst size was calculated based on the average of latent and post latent values. Each assay was performed in triplicate.

#### Fluorescence microscopy for monitoring phage infection

Fluorescence microscopy for monitoring phage infection was performed as previously described [13]. *E. coli* MG1655 containing plasmid that expresses mCherry (pBAD18-mCherry) was grown on an LB agar plate supplemented with ampicillin at 37℃ overnight. From the agar plate, a single colony was picked, washed, and resuspended in 20% LB broth. 5 µl of this culture was spotted on the prepared 1% agarose imaging pads made with 20% LB supplemented with 0.05% arabinose. The culture was evenly spread with the bottom of a 0.5 ml Eppendorf tube. The imaging pads with the culture were then incubated in a humidifier at 37℃ for 3 h. After 3 h, 5 µl

of the phage lysates was spotted to the agarose pads and evenly spread. The pads were further incubated at 37℃. After 60 min of phage infection, the pads were taken and 5 µl DAPI (5 µg/ml) was spotted and spread. After drying, a coverslip was placed on the agarose pad and samples were imaged by Nikon Eclipse Ti2-E equipped with a 100X CFI Plan Apochromat oil objective and a DSQi-2 Monochrome Camera (Nikon). Images were processed using the NIS Elements AR software (Nikon). For capturing cell bursting, infection was carried out as described in *E. coli* MG1655. The video was captured in 70 min post infection using Nikon Eclipse Ti2-E microscope and processed with NIS-Elements AR and ImageJ-win64.

## Results

### Bacteriophage isolation, morphological characterization, and genome sequencing

Phage isolation was performed from chicken feces, as schematically represented in Fig. 1A. After processing the chicken feces sample for phage isolation, we identified two distinct plaques from the same sample (Fig. 1B)—one forming a tiny plaque and another forming a slightly bigger plaque through double agar layer method containing 0.35% top agar. Both phages were purified for analysis through transmission electron microscopy (TEM). The small plaque-forming phage, which we named Sharanji, appeared as a myophage with an average particle size of 247.2 nm. The bigger plaque-forming phage, which we named PCM001, appeared as a myophage with an average size of 224.2 nm (Fig. 1B). To further understand the nature of the isolated phages, genomic DNA was isolated and whole-genome sequencing was performed. For PCM001, we obtained a partial genome. The reason for obtaining partial genome could be due to various reasons ranging from extracted DNA sample quality, poor sequencing coverage, the presence of complex genomic areas such as high GC content or nucleotides modifications. However, the partial genome of PCM001 closely matches the genomes of *Escherichia* phage vB_Eco_NicPhage, *Escherichia* phage AlbertHofmann, and *Escherichia* phage 55, which have genome sizes ranging from 168 to 170 kb (Fig. S1). In the case of Sharanji, we were able to obtain the complete genome corresponding to 350,079 bp. Since Sharanji is a jumbo phage with complete genome, we decided to further characterize it.

### Genomic features of the jumbo phage Sharanji

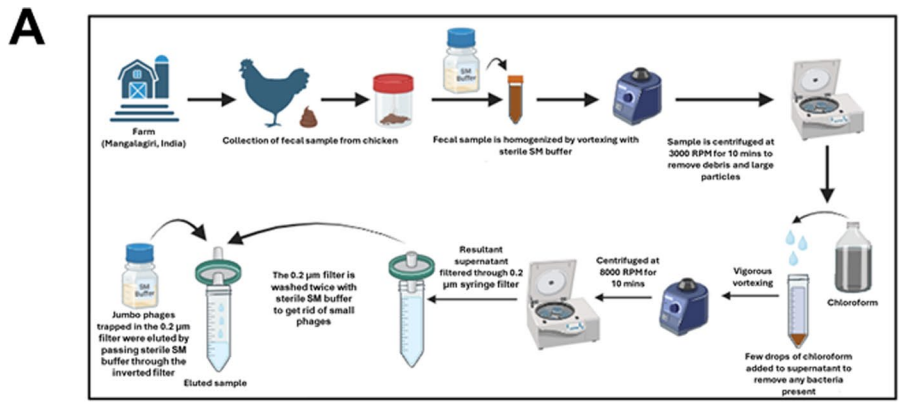
Sharanji has a linear double-stranded DNA with a genome size of 350,079 bp, and an average GC content of 34.1%, which is lower than the host *E. coli* GC content, which is 50.8%. Analysis of Sharanji genome through PhageTerm showed that Sharanji contains large direct terminal repeats (DTR) corresponding to 20,668 bp (Fig. S2). Figure 2A illustrates the circular representation of the Sharanji genome, featuring annotated predicted ORFs and their respective genomic locations. Figure 3 illustrates the linear representation of the Sharanji genome showing the location and function of some of the important annotated genes. In general, Sharanji harbors 593 protein-encoding open reading frames (ORFs) and 7 tRNAs. Genes associated with DNA replication-recombination-repair, translation, transcription, structural components, lysis, and hypothetical categories are distinguished by different colors. The lack of genes associated with integrase, excisionase, repressor, and an attachment site in the genome of Sharanji suggests that this phage is lytic, consistent with the lytic nature observed in all documented jumbo phages. The lytic character of Sharanji was further confirmed through the use of the PhageScope, a tool designed for predicting phage lifestyles [31].

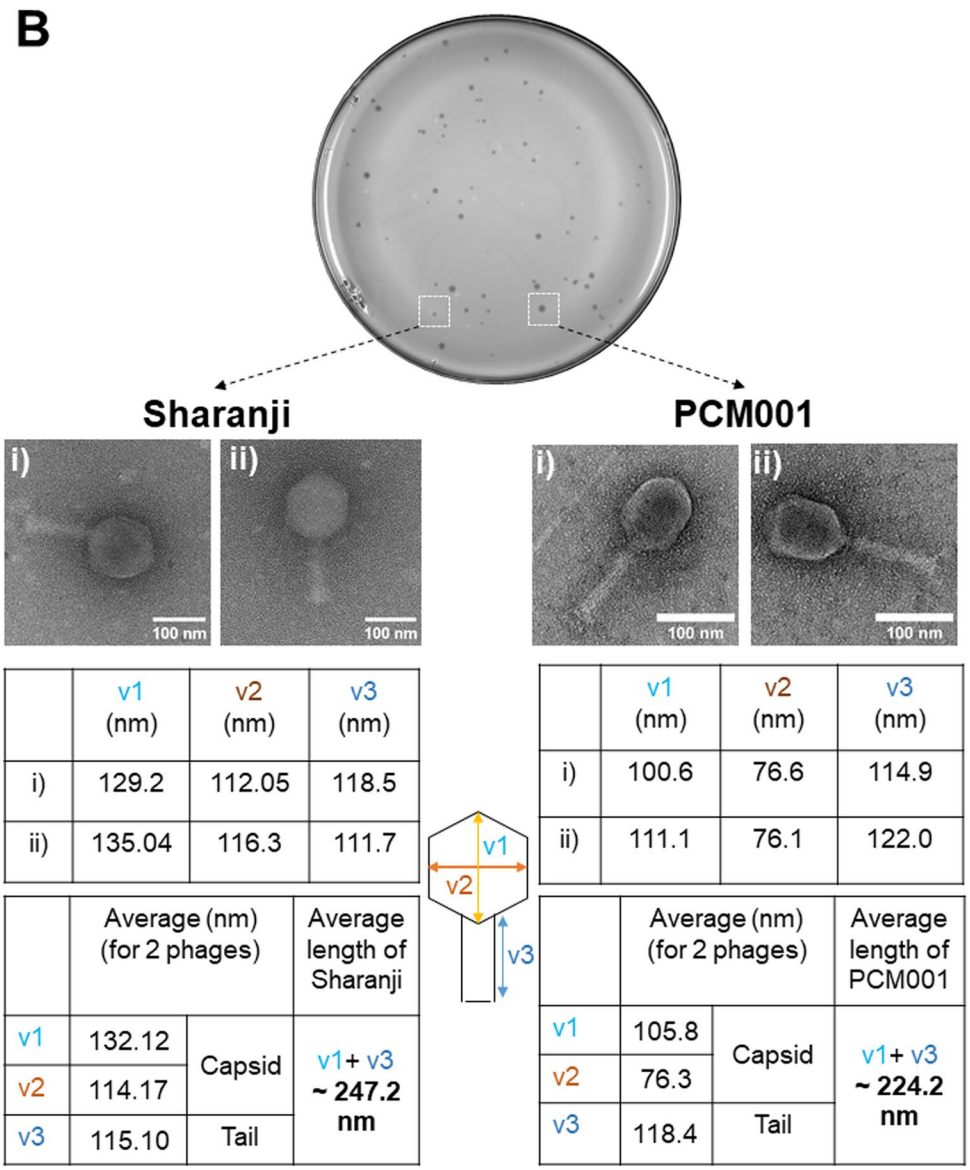
Out of a total of 593 protein encoding ORFs, we were able to predict the functions of 119 proteins using four distinct approaches, namely RASTtk, Pfam, CDD, and Prokka. The number of genes predicted by each approach and their overlap with other predictions are summarized in a Venn diagram shown in Fig. 2B. Among the approaches, the Conserved Domain Database (CDD) identified functions for the largest number of ORFs (103/119). Similarly, RASTtk, the Pfam database, and Prokka predicted functions for 29/119, 82/119, and 65/119 ORFs, respectively. Only 17/119 ORFs had functions consistently predicted by all four methods. The predicted functions of these annotated ORFs, along with their genomic loci, are detailed comprehensively in Table S1.

A description of genes in some of the important categories is mentioned below.

#### *DNA replication, repair and recombination*

Jumbo phages are well known to encode multiple genes involved in DNA replication, repair and recombination [2]. In the case of Sharanji, nearly 19 known genes exhibit potential functions in these processes. Notably, Sharanji encompasses genes such as DNA polymerase I (WNN14621.1), DNA polymerase II (WNN14630.1), DNA polymerase III subunit tau (WNN14562.1), DNA polymerase III DnaE (WNN14566.1), DnaX (WNN14562.1), and DnaC (WNN14602.1), replicative DNA helicase (WNN14602.1) indicative of a comprehensive suite of genes facilitating DNA replication. Sharanji also features thymidylate synthase (WNN14671.1) and anaerobic ribonucleoside-triphosphate reductase (WNN14743.1), enabling the biosynthesis of DNA from precursors like thymidylate and ribonucleotides, respectively. Furthermore, the identification of genes encoding DNA gyrase subunits A (WNN14569.1) and




| | v1 (nm) | v2 (nm) | v3 (nm) |
|---|---|---|---|
| i) | 129.2 | 112.05 | 118.5 |
| ii) | 135.04 | 116.3 | 111.7 |

| | Average (nm) (for 2 phages) | | Average length of Sharanji |
|---|---|---|---|
| v1 | 132.12 | Capsid | v1+ v3 ~ 247.2 nm |
| v2 | 114.17 | | |
| v3 | 115.10 | Tail | |

| | v1 (nm) | v2 (nm) | v3 (nm) |
|---|---|---|---|
| i) | 100.6 | 76.6 | 114.9 |
| ii) | 111.1 | 76.1 | 122.0 |

| | Average (nm) (for 2 phages) | | Average length of PCM001 |
|---|---|---|---|
| v1 | 105.8 | Capsid | v1+ v3 ~ 224.2 nm |
| v2 | 76.3 | | |
| v3 | 118.4 | Tail | |

**Fig. 1** **A** Schematic illustrating phage isolation procedure by the filter trapping approach. **B** Images displaying the plaques formed by Sharanji and PCM001 on *E. coli* MG1655 lawn and transmission electron micrograph of phage Sharanji and PCM001 with a scale bar of 100 nm. A table with details of phage size is also given

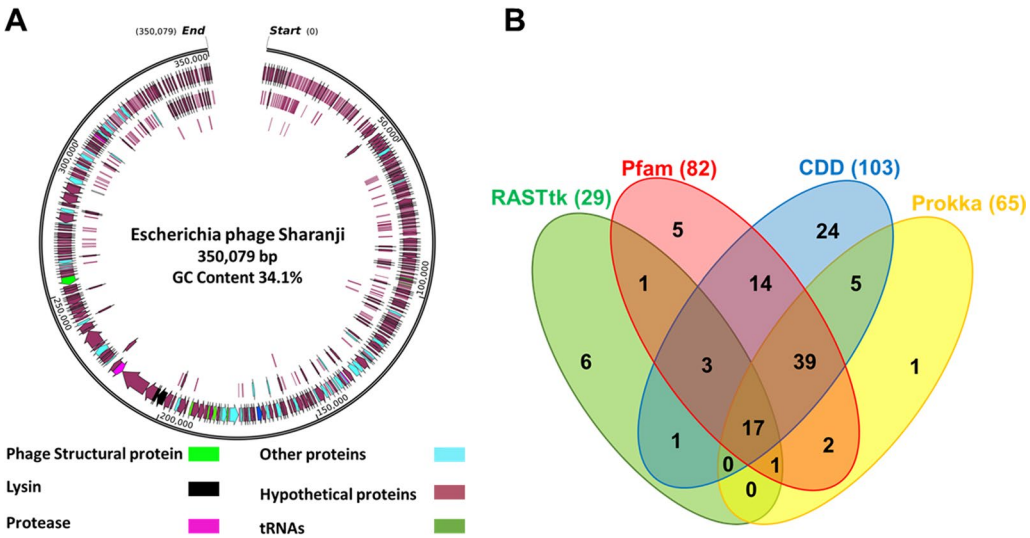



**Fig. 2** **A** Circular genome map of *E. coli* jumbo phage Sharanji, indicating various functional categories and all ORFs, with transcription direction indicated. **B** Venn diagram depicting the fraction of ORFs annotated by RASTtk, Pfam, CDD, and Prokka

B (WNN14570.1) suggests a potential role in advancing replication fork progression during phage DNA replication. Additionally, other essential genes identified include those encoding DNA ligase (WNN14781.1) and a replicative DNA helicase (WNN14602.1). Proteins crucial for DNA repair, such as RecBCD enzyme subunit RecD (WNN14578.1), exonuclease subunit (WNN14648.1), putative nucleotidyltransferase (WNN14733.1), 5' nucleotidase deoxy (pyrimidine) cytosolic type C protein (WNN14561.1), pyrimidine dimer DNA glycosylase (WNN14420.1), exodeoxyribonuclease (WNN14613.1), and recombinase A (WNN14615.1), were also identified. Collectively, the data suggests that Sharanji encodes an extensive array of genes involved in DNA replication, repair, and recombination, which likely play a pivotal role in the efficient replication, repair, and recombination of its large genome.

### *Transcription*

Sharanji encodes a phage-encoded RNA polymerase sigma factor RpoD (WNN14470.1). It is likely that WNN14470.1 associates with the host RNA polymerase and mediates transcription of the phage genes. Recently, jumbo phages have been categorized into 3 major categories based on a set of genes involved in transcription and replication. The presence of phage-encoded RNA polymerase sigma factor RpoD is a hallmark of Group 2 jumbo phages [2].

### *Translation*

Sharanji encodes seven tRNAs covering amino acids, including Thr (Sharanji_gp208), Asn (Sharanji_gp213), Met (Sharanji_gp214), Arg (Sharanji_gp221), Gly (Sharanji_gp225), and two tRNAs for Ser (Sharanji_gp215, Sharanji_gp222). Additionally, it encodes factors that regulate translation, including tRNA-specific adenosine deaminase (WNN14772.1), sigma 54 modulation protein/S30EA ribosomal protein (WNN14706.1), and multifunctional CCA protein (WNN14700.1). A gene encoding translation initiation factor IF-3 (WNN14575.1) was also identified. In bacteria, IF3 is well-known for regulating the precision and speed of bacterial mRNA translation initiation [38]. However, its exact function in bacteriophages cannot be definitively stated. In the case of *P. aeruginosa* jumbo phage ΦKZ, host ribosome takeover has recently been demonstrated as one of the earliest steps in the jumbo phage infection process [39]. Although a homolog of host ribosome takeover factor of ΦKZ (NP_803580.1) is absent in the genome of Sharanji, it is plausible that IF3 and some of the other factors may function as ribosome-associated phage factors, mediating host takeover and modulation of the translational apparatus.

### *Genome protection*

Phages employ diverse mechanisms to protect their genomes against host immune enzymes, such as CRISPR-Cas and restriction nucleases. For instance, jumbo phages like ΦKZ shield their genome within a proteinaceous phage nucleus [13, 14]. Similarly, T4 phages protect their genome from specific Cas nucleases by modifying cytosine to glucosyl-hydroxymethylcytosine (glc-HMC) [40–42]. Examination of the Sharanji genome revealed the absence of homologs for *chmA*, indicating that it

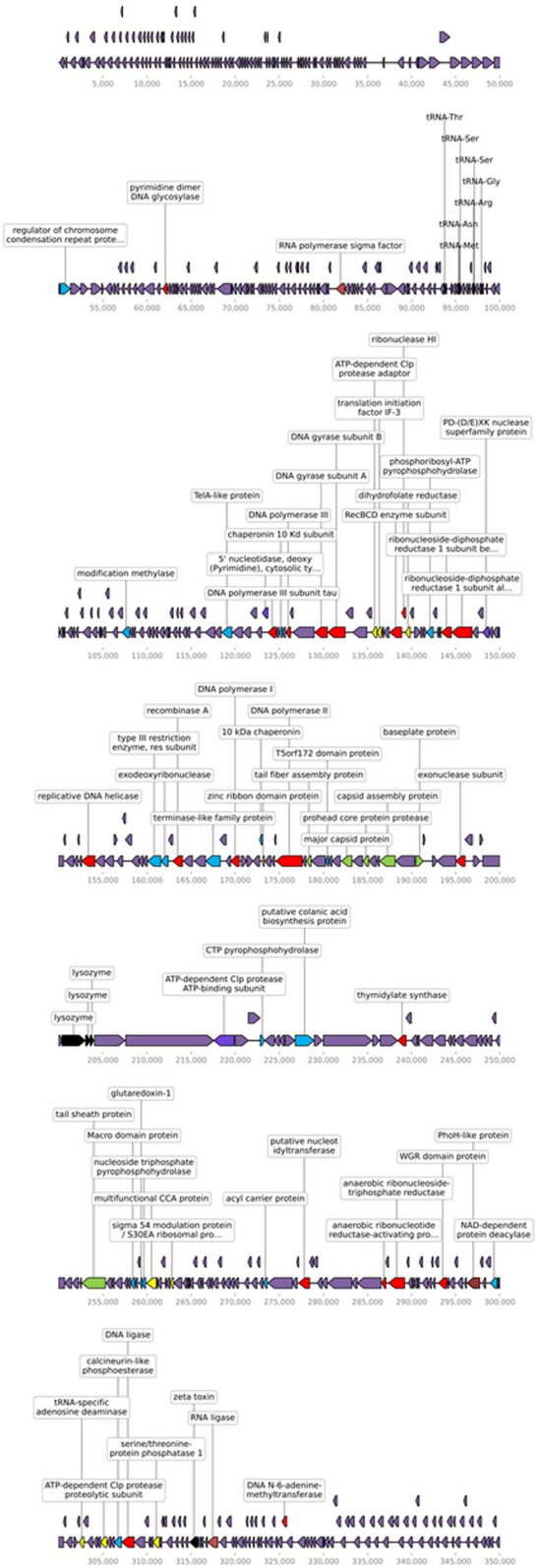



**Fig. 3** Linear genome map of *E. coli* jumbo phage Sharanji, depicting the positions and predicted functions of ORFs

does not fall within the category of Chimalliviridae and may not form a phage nucleus. However, Sharanji does encode other genes potentially involved in DNA protection. This includes DNA N-6-adenine-methyltransferase (WNN14823.1), methyltransferase (WNN14829.1) which could prevent degradation of viral DNA by host restriction-modification defense enzymes. It also encodes a type I restriction endonucleases (WNN14711.1), Type III restriction enzyme res subunit (WNN14612.1) and a modification methylase HaeIII (WNN14532.1), generally known to protect DNA cleavage by the HaeIII endonuclease. Analysis of Sharanji genome through AcrFinder [43] did not identify known anti-CRISPRs. However, through our genomic analysis, we identified a solo Cas4 (vCas4) (WNN14593.1) in Sharanji, which may associate with bacterial CRISPR-Cas system and induce autoimmunity, thereby facilitating successful phage infection, as described for other phages [44, 45]. Furthermore, we identified a homolog of the T7 bacteriophage protein kinase gene in Sharanji (WNN14432.1). This protein kinase is suggested to phosphorylate host defense proteins Retron-Eco9 and DarTG1, thereby weakening the host's immunity against the phage and facilitating successful infection [46]. It is possible that the Sharanji protein kinase serves a similar function in compromising host immunity. Additionally, Sharanji possesses another protein kinase (WNN14351.1), which function is unknown. The exact functions of these proteins can only be confirmed through experimental studies.

### *Structural proteins*

Sharanji encodes six proteins with predicted functions related to phage morphogenesis and structure. These include genes for the major capsid protein (WNN14637.1), capsid assembly (WNN14642.1), prohead core protein protease (WNN14639.1), tail sheath protein (WNN14689.1), tail fiber protein (WNN14632.1), baseplate protein J-like (WNN14656.1), flagellar hook protein (WNN14670.1) and base plate protein (WNN14644.1). Additionally, Sharanji encodes a phage head completion protein (WNN14689.1), homologous to the transposase A gene of Tn7 [47]. The phage head completion protein is suggested to cleave packaged DNA, enabling the joining of heads to tails [48].

### *Lysis proteins*

Sharanji encodes a lysozyme (WNN14653.1), well-known for its action on peptidoglycan in bacterial cell walls, facilitating phage release [49]. Additionally, it harbors a gene for zeta toxin (WNN14797.1), whose primary role involves inhibiting cell wall biosynthesis, potentially serving as a bactericidal agent [50].

### Phylogenetic and comparative genome analysis of the jumbo phage Sharanji

Taxonomy analysis of the jumbo phage Sharanji through taxmyPHAGE [33] classified Sharanji within the genus *Asteriusvirus* and class of *Caudoviricetes,* and representing it as a new species. Summary of the taxonomy analysis is provided in Table S2. Furthermore, a phylogenetic analysis using Viptree was conducted to assess the relationship between Sharanji and other phages infecting the genus *Escherichia* (Fig. 4A). Among 177 related phages from the *Pseudomonadota* group that infect *Escherichia* bacteria, it was found that Sharanji is closely related to *Escherichia* phages 121Q and PBECO4. These phages belong to the same clade, consistent with taxonomy classification (denoted in red box). The Average Nucleotide Identity (ANI) scores were determined using the VIRIDIC web server, yielding scores of 94% and 94.2% with 121Q and PBECO4, respectively (Fig. 4B). Since Sharanji has an ANI score of less than 95%, it is regarded as a new species, as per the guidelines of Genome-Based Phage Taxonomy [51].

To further explore the genetic similarity between Sharanji and the two other jumbo phages, 121Q and PBECO4, we performed a pan-genome analysis (Fig. 4C and Table S3). Of the 593 genes encoded by Sharanji, 528 (89%) represent core genes conserved across all three phages. Additionally, Sharanji encodes 55 accessory genes (9.3%) which are shared with either 121Q or PBECO4, and 10 unique genes (1.7%) found exclusively in the Sharanji genome. The list of unique genes present in Sharanji is given in Table S4. Furthermore, the genome alignment of Sharanji with 121Q and PBECO4 reveals genomic regions that are shared but arranged differently among these phages, highlighting their modular genome (Fig. 4D).

### Phenotypic characterization of the jumbo phage Sharanji

We next decided to characterize the various biological properties of Sharanji including its host range, the capacity of the phage to tolerate various stress including pH and temperature, and one-step growth curve to understand the phage infection efficiency. Figure 5A shows that Sharanji specifically infects *E. coli* K12-based strains, but not other *E. coli* or other bacteria including *P. aeruginosa*, *K. pneumoniae, A. baumannii, B. subtilis, and S. aureus*. This suggests that Sharanji is a narrow spectrum

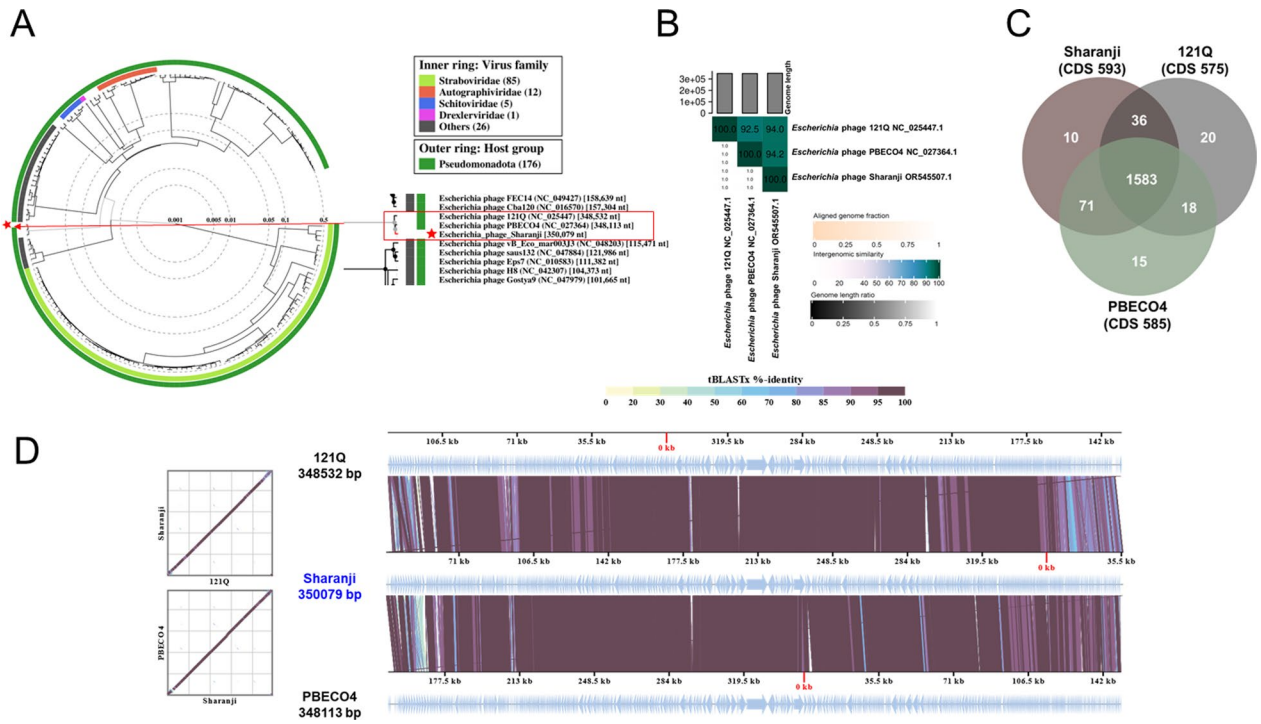



**Fig. 4** **A** Phylogenetic tree showing the genome relationships of Sharanji, closely associated with jumbo phages *Escherichia* phage 121Q and *Escherichia* phage PBECO4 in a distinct clade (highlighted in red box). **B** ANI score matrix of Sharanji with *Escherichia* phages 121Q and PBECO4, indicating 94% and 94.2% similarity, respectively. **C** Venn diagram illustrates core, accessory and unique genes among Sharanji, 121Q and PBECO4. **D** The genome alignment of Sharanji (indicated in blue) with its closest relatives, 121Q and PBECO4, reveals a high degree of genomic similarity, with substantial conservation across some regions. The dot plots (left panel) show a syntenic pattern, with linear alignments highlighting conserved core regions

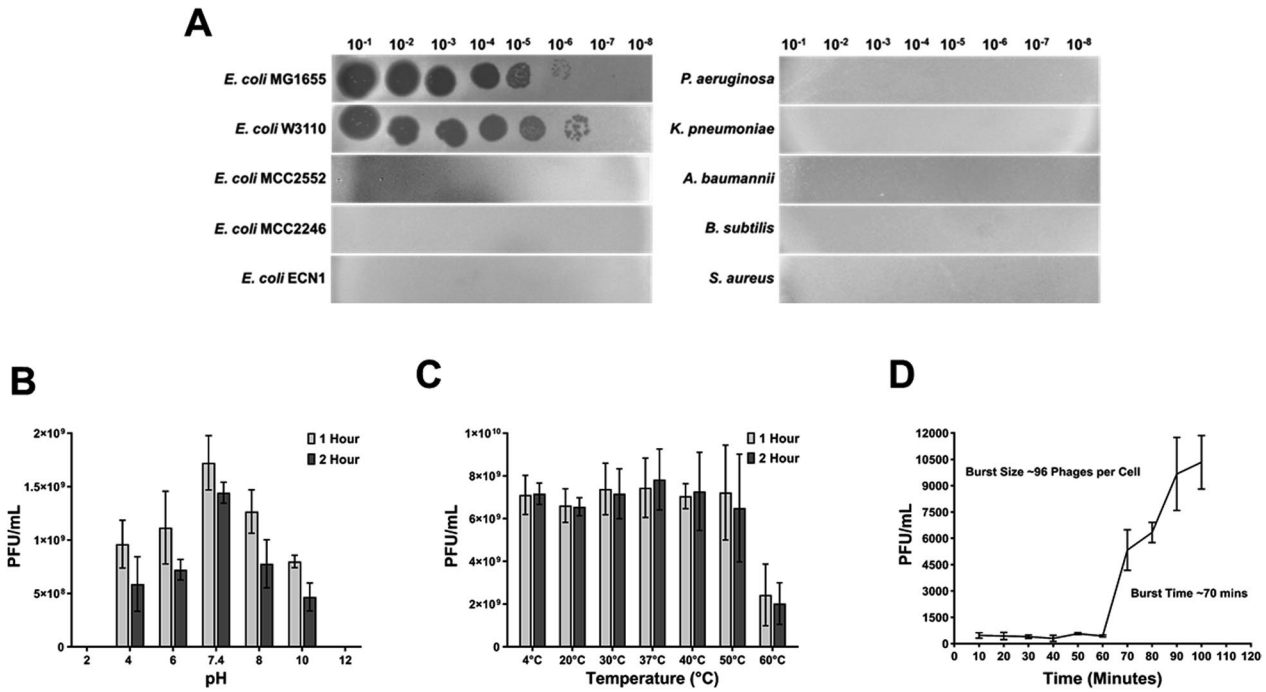



**Fig. 5** **A** Host tropism evaluation, showing clear plaques formed by Sharanji on bacterial lawns of *E. coli* K12 strains namely MG1655 and W3110 but not on other bacterial strains indicated. **B** Analysis of pH stability of Sharanji. **C** Analysis of the thermostability of Sharanji. **D** One-step growth curve analysis. Data presented in B, C, and D represent averages from three independent experiments

phage capable of infecting only specific strains of *E. coli*. Next, we assessed the stability of Sharanji across various pH levels (2, 4, 6, 7, 8, 10 and 12) and temperatures (4 °C, 20 °C, 30 °C, 37 °C, 40 °C, 50 °C, 60 °C) after incubating it for 1 or 2 h. Results presented in Fig. 5B show that Sharanji is more stable at pH 7.4, and significant reduction in stability was observed at pH 4, 6, 8 and 10. However, its stability is majorly affected at pH 2 and 12 (Fig. 5B and Table S5). In the case of temperature, Sharanji was stable when incubated at temperature range of 4 to 50 °C. However, its stability significantly declined at 60 °C (Fig. 5C and Table S5).

Next, we performed a one-step growth curve for Sharanji to understand the dynamics of phage replication and the timing of infection cycles. The results presented in Fig. 5D indicate that the latent period of the Sharanji is 60 min, with a burst size of about 96 virions per infected cell. Taken together, these phenotypic characterizations indicate that Sharanji is a robust phage with a specific host range, considerable stability across different pH levels and temperatures, and efficient replication dynamics.

### Cytological profiling of Sharanji infection through fluorescence microscopy

Jumbo phages like ΦKZ are well-known to induce distinct morphological changes in host cells due to the formation of replication compartments [15]. Other morphological changes have been observed for non-nucleus-forming jumbo phages as well [52]. To determine whether Sharanji infection causes intracellular morphological transitions in the host cell, we performed single-cell infection microscopy and compared it with the infection of the nucleus-forming *E. coli* jumbo phage, Goslar [53]. Cells were expressing fluorescent protein mCherry, as a marker for cytoplasm, and stained with the DNA-binding dye DAPI. In uninfected cells, DAPI stained DNA and mCherry markers appeared smooth and spread out throughout the cell. Cells infected by Goslar formed a ball-like DAPI-stained replication compartment, which is the phage nucleus. The cytoplasmic marker mCherry was clearly excluded from the phage nucleus as expected (Fig. 6A). In contrast, Sharanji-infected cells did not show a localized replication compartment; instead, the DAPI staining appeared distributed throughout the cell (Fig. 6A). Notably, we observed that Sharanji infected cells were filamented. To probe this further, we performed phase contrast microscopy following Sharanji's late stage of infection. Figure 6B illustrates representative images at the late stage of Sharanji infection (70 min' post infection (mpi)). The upper panel shows the infected bacterial cells just prior to lysis, and the lower panel indicates the lysed cell, indicated by red arrows. Sharanji infected cells are not only filamented, but formed bulged regions prior to lysis. Time lapse imaging of Sharanji

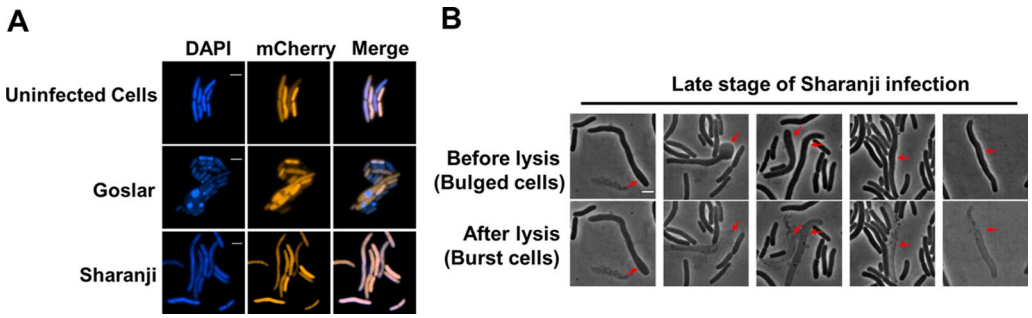



**Fig. 6** **A** Fluorescence microscopy images of uninfected cells (top) and *E. coli* cells infected with nucleus-forming *E. coli* jumbo phage Goslar (middle), and Sharanji (bottom). mCherry is the reporter for cytoplasm and DAPI is the reporter for DNA. **B** Phase contrast images depicting the late stage of *E. coli* cells infected by Sharanji. The upper panel shows the infected cells just prior to lysis (70 min' post infection) and the lower panel displays the lysed host cells, indicated by red arrows. Scale bar represents 2 µm

infected cells shows bursting of cells occur at bulged regions (Video S1). These observations provide insights in the mechanism of Sharanji infection.

## Discussion

Jumbo phages have emerged as unique model systems for understanding phage-bacteria interactions. In recent years, some of the remarkable findings based on jumbo phages including exquisite intracellular organization during infection via formation of phage nucleus, cytoskeletal structures, and their ability to evade several immune systems, highlights yet to be explored biology from these systems. However, the number of jumbo phages isolated and available for fundamental studies are far less compared to other types of phages. Moreover, while jumbo phages have been isolated in recent years, isolation and characterization of very large jumbo phages exceeding 350 kb are rare. The major finding from our study is the isolation and characterization of such a large phage, Sharanji, capable of infecting *E. coli*.

Sharanji was isolated using a modified phage isolation technique, focusing on phages retained by a 0.22-µm filter and capable of forming very small plaques. We propose that this method preferentially enriches jumbo phages, increasing their likelihood of isolation. However, non-jumbo phages may also be isolated, occasionally due to their elongated morphology. In our study, both a jumbo phage, Sharanji, and a non-jumbo phage, PCM001, were isolated from the same sample. Thus, further optimization is required. Previous research has utilized differential centrifugation and other approaches to enrich jumbo phages, ensuring their preferential isolation [54–56]. It is possible that a combination of approaches could be employed for targeted isolation of jumbo phages from diverse sources.

The genomic analysis of Sharanji identifies it as a jumbo phage with a linear double-stranded DNA genome of 350,079 bp, encoding 593 protein-coding ORFs and 7 tRNAs. Similar to other known jumbo phages, Sharanji is lytic in nature, as evidenced by the absence of genes associated with lysogeny. The presence of a large number of genes involved in DNA replication, repair, transcription, and translation suggests robust replication capabilities. Our results indicate that nearly 96 phages are produced from a single infected *E. coli* cell. Given the size of Sharanji's genome, this corresponds to the production of 33,607 kbp of DNA per cell, which is over 7 times the size of the *E. coli* genome. We speculate that such massive DNA synthesis could be mediated by DNA metabolism-related genes encoded by Sharanji, the degradation of host DNA for recycling deoxyribonucleotides, or a combination of both. These processes are likely activated during the early to mid-infection period, after the phage's takeover of the host's transcription and translation machinery.

Phenotypic characterization of Sharanji revealed that it exhibited physicochemical stability across a range of temperatures and pH levels. It was affected only under extreme acidic conditions (pH 2) and extreme basic conditions (pH 12). Similarly, its stability significantly declined only at temperatures above 60 °C. These observations indicate that purified Sharanji can be maintained and stored under optimal conditions for long-term use, which is highly beneficial for phage-based applications in agriculture, clinical settings, and other fields. Further phenotypic characterization of Sharanji revealed a very narrow host range infecting K12-like *E. coli* strains, which was used originally as the host for isolating it. Although Sharanji was originally isolated from chicken feces, it was found to be incapable of infecting avian pathogenic *E. coli*. Several jumbo phages are known to have a broad host range, but the reason for the narrow host range of Sharanji is currently unknown. Notably, a recent study involving a large-scale analysis of *E. coli* phages and

their hosts indicate that the presence of phage receptors, rather than host defense mechanisms, largely determines phage infectivity [57]. Thus, identification of host receptor for Sharanji might provide an explanation for its narrow host range. A recent study involving an *E. coli* jumbo phage SHEFM2K, which belongs to the Asteriusvirus family, suggested the use of flagellar structures as potential receptors [58]. Since Sharanji also belongs to the Asteriusvirus family, it is likely to use similar structures, although this hypothesis needs to be experimentally verified in future studies.

Being a very large jumbo phage, Sharanji holds potential as a model system for studying fundamental aspects of phage biology. These include the mechanism of large DNA injection, novel host modulation factors synthesized during the early stages of infection, subcellular organization processes from DNA synthesis to virion assembly without phage nucleus-like compartmentalization, and the unique ability to pack large DNA into its capsid without relying on known jumbo phage factors like the inner body. Notably, cytological profiling reveals that the mechanism of host cell infection by Sharanji differs from well-studied nucleus-forming jumbo phages such as the *P. aeruginosa*-infecting ΦKZ or the *E. coli*-infecting Goslar. Cell filamentation was observed during Sharanji infection, a phenomenon also reported in other phages. For example, the Kil peptide of bacteriophage λ [59], KilR of Rac cryptic prophage [60], and Gp0.4 of T7 bacteriophage [61] are known inhibitors of FtsZ, the central bacterial cell division protein. However, such cell division inhibitors have not been identified and characterized from jumbo phage genomes, to the best of our knowledge. We suspect that Sharanji encodes novel cell division inhibition factors, as it lacks previously known division-inhibition elements mentioned above. Interestingly, during the late stages of infection, cell bulging was observed, and time-lapse imaging revealed that cells burst at these bulged regions. Thus, Sharanji encodes various cell envelope-disrupting factors, which could have potential biotechnological applications. Future studies on Sharanji and related jumbo phages could provide valuable insights into intracellular organization, genome replication, protection mechanisms, and the evolution of nucleus-lacking jumbo phages.

## Conclusions

We report the isolation and characterization of a novel *E. coli* jumbo phage Sharanji which was isolated through a modified phage isolation approach. Sharanji's distinct infection mechanisms and cytological features offer an opportunity to explore the fundamental aspects of jumbo phage biology particularly in the context of non-nucleus-forming jumbo phages. The findings from this study emphasize the need for further exploration of these underrepresented phages, which could reveal novel biological processes and contribute to our understanding of phage evolution, DNA protection, and host interaction strategies. Continued exploration of very large jumbo phages like Sharanji can reveal new paradigms in virology and microbial ecology and contribute to our understanding of giant viruses in general.

### Abbreviations

| | |
|---|---|
| LB | Luria–Bertani |
| TEM | Transmission electron microscopy |
| PCR | Polymerase chain reaction |
| ORF | Open reading frame |
| ICTV | International Committee on Taxonomy of Viruses |
| ANI | Average Nucleotide Identity |
| NCBI | National Center for Biotechnology Information |
| CARD | Comprehensive antibiotic resistance database |
| MOI | Multiplicity of infection |
| PFU | Plaque forming units |
| OD | Optical density |
| DTR | Direct terminal repeats |
| CRISPR | Clustered regularly interspaced short pallindromic repeats |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12985-025-02646-5.

Additional file 1.

### Acknowledgements

The authors acknowledge SRM Central Instrumentation Facility from SRM Institute of Science and Technology (SCIF, SRMIST) for providing HR TEM facility and MedGenome Labs Ltd (Bengaluru) for support with whole genome sequencing. We thank Leibniz Institute DSMZ—German Collection of Microorganisms and Cell Cultures GmbH for providing *Escherichia* phage vB_EcoM_Goslar (DSM 104658). We thank National Centre for Microbial Resource at National Centre for Cell Science (NCCS), Pune, India, for providing *E. coli* MCC2552 and MCC2246. We appreciate helpful discussions with Dr. Gaurav Sharma, IIT Hyderabad, and members of the Department of Biological Sciences, SRM University – AP.

### Author contributions

SG, and SM formulated the study design and plans. SM conducted bioinformatics analyses independently, and experiments were performed with support from BS, DS, NT, and AP. Manuscript was written by SG and SM. All authors read and approved the manuscript.

### Funding

Sutharsan Govindarajan is supported by the DBT-Wellcome Trust Early Career Fellowship (IA/E/19/1/504958) and DST-SERB (CRG/2020/003295). Sharayu Magar acknowledges the Senior Research Fellowship from DST-SERB (CRG/2020/003295). All authors appreciate the infrastructure support received from SRM University AP (SRMAP/URG/GENERAL/2023-24/010), (SRMAP/URG/E&PP/2022-23/018) and (SRMAP/URG/GENERAL/2024-25/039).

### Availability of data and materials

Nucleotide sequence accession number: The complete genome sequence of Sharanji and PCM001 is available in GenBank under Accession Number OR545507 and OR529306, respectively.

## Declarations

**Ethics approval and consent to participate**
This research does not involve human participants or animals in any experiments conducted by the authors. Views: The views expressed in this study are those of the authors and not necessarily those of either the funding agency or any other institution.

**Competing interests**
The authors declare that they have no competing interests.

Received: 20 August 2024 Accepted: 30 January 2025
Published online: 10 March 2025

**References**

1. Yuan Y, Gao M. Jumbo bacteriophages: an overview. Front Microbiol. 2017;8:403.
2. Iyer LM, Anantharaman V, Krishnan A, Burroughs AM, Aravind L. Jumbo Phages: a comparative genomic overview of core functions and adaptions for biological conflicts. Viruses. 2021;13(1):63.
3. Devoto AE, Santini JM, Olm MR, Anantharaman K, Munk P, Tung J, et al. Megaphages infect Prevotella and variants are widespread in gut microbiomes. Nat Microbiol. 2019;4(4):693–700.
4. Hu M, Xing B, Yang M, Han R, Pan H, Guo H, et al. Characterization of a novel genus of jumbo phages and their application in wastewater treatment. Iscience. 2023;26(6):106947.
5. Nazir A, Ali A, Qing H, Tong Y. Emerging aspects of jumbo bacteriophages. Infect Drug Resist. 2021;14:5041–55.
6. Weinheimer AR, Aylward FO. Infection strategy and biogeography distinguish cosmopolitan groups of marine jumbo bacteriophages. ISME J. 2022;16(6):1657–67.
7. Hendrix RW. Jumbo bacteriophages. In: Lesser known large dsDNA viruses. Springer; 2009. pp. 229–240.
8. Cobián Güemes AG, Ghatbale P, Blanc AN, Morgan CJ, Garcia A, Leonard J, et al. Jumbo phages are active against extensively drug-resistant eyedrop-associated *Pseudomonas aeruginosa* infections. Antimicrob Agents Chemother. 2023;67(12):e00654-23.
9. Rai P, Shetty SS, Prabell S, Kuntar A, Pinto D, Kumar BK, et al. Characterisation of broad-spectrum phiKZ like jumbo phage and its utilisation in controlling multidrug-resistant *Pseudomonas aeruginosa* isolates. Microb Pathog. 2022;172:105767.
10. Thammatinna K, Sinprasertporn A, Naknaen A, Samernate T, Nuanpirom J, Chanwong P, et al. Nucleus-forming vibriophage cocktail reduces shrimp mortality in the presence of pathogenic bacteria. Sci Rep. 2023;13(1):17844.
11. Kanaizuka A, Sasaki R, Miyashita S, Ando S, Ito K, Fukuhara T, et al. Isolation of *Burkholderia* jumbo phages and their utilization as biocontrol agents to suppress rice seedling rot disease. J Gen Plant Pathol. 2023;89(1):24–34.
12. Attai H, Boon M, Phillips K, Noben J-P, Lavigne R, Brown PJB. Larger than life: isolation and genomic characterization of a jumbo phage that infects the bacterial plant pathogen, *Agrobacterium tumefaciens*. Front Microbiol. 2018;9:1861.
13. Mendoza SD, Nieweglowska ES, Govindarajan S, Leon LM, Berry JD, Tiwari A, et al. A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. Nature. 2020;577(7789):244–8.
14. Malone LM, Warring SL, Jackson SA, Warnecke C, Gardner PP, Gumy LF, et al. A jumbo phage that forms a nucleus-like structure evades CRISPR–Cas DNA targeting but is vulnerable to type III RNA-based immunity. Nat Microbiol. 2020;5(1):48–55.
15. Chaikeeratisak V, Nguyen K, Khanna K, Brilot AF, Erb ML, Coker JKC, et al. Assembly of a nucleus-like structure during viral replication in bacteria. Science. 2017;355(6321):194–7.
16. Kraemer JA, Erb ML, Waddling CA, Montabana EA, Zehr EA, Wang H, et al. A phage tubulin assembles dynamic filaments by an atypical mechanism to center viral DNA within the host cell. Cell. 2012;149(7):1488–99.
17. Li Y, Guan J, Hareendranath S, Crawford E, Agard DA, Makarova KS, et al. A family of novel immune systems targets early infection of nucleus-forming jumbo phages. bioRxiv. 2022;2009–22.
18. Krylov VN, Smirnova TA, Minenkova IB, Plotnikova TG, Zhazikov IZ, Khrenova EA. Pseudomonas bacteriophage contains an inner body in its capsid. Can J Microbiol. 1984;30(6):758–62.
19. Thomas JA, Weintraub ST, Wu W, Winkler DC, Cheng N, Steven AC, et al. Extensive proteolysis of head and inner body proteins by a morphogenetic protease in the giant *Pseudomonas aeruginosa* phage φKZ. Mol Microbiol. 2012;84(2):324–39.
20. Wu W, Thomas JA, Cheng N, Black LW, Steven AC. Bubblegrams reveal the inner body of bacteriophage φKZ. Science. 2012;335(6065):182.
21. Guan J, Bondy-Denomy J. Intracellular organization by jumbo bacteriophages. J Bacteriol. 2020;203(2):e00362-e420.
22. Al-Shayeb B, Sachdeva R, Chen L-X, Ward F, Munk P, Devoto A, et al. Clades of huge phages from across Earth's ecosystems. Nature. 2020;578(7795):425–31.
23. Carr VR, Shkoporov A, Gomez-Cabrero D, Mullany P, Hill C, Moyes DL. The human oral phageome is highly diverse and rich in jumbo phages. BioRxiv. 2020;2007–20.
24. Li C, Liu K, Gu C, Li M, Zhou P, Chen L, et al. Gastrointestinal jumbo phages possess independent synthesis and utilization systems of NAD+. Microbiome. 2024;12:268.
25. Banar M, Rokaya D, Azizian R, Khurshid Z, Banakar M. Oral bacteriophages: metagenomic clues to interpret microbiomes. PeerJ. 2024;12:e16947.
26. Cook R, Brown N, Redgwell T, Rihtman B, Barnes M, Clokie M, et al. INfrastructure for a PHAge REference Database: identification of large-scale biases in the current collection of phage genomes. bioRxiv. 2021.
27. Jacobs-Sera D, Marinelli LJ, Bowman C, Broussard GW, Bustamante CG, Boyle MM, et al. On the nature of mycobacteriophage diversity and host preference. Virology. 2012;434(2):187–201.
28. Lood C, Danis-Wlodarczyk K, Blasdel BG, Jang HB, Vandenheuvel D, Briers Y, et al. Integrative omics analysis of *Pseudomonas aeruginosa* virus PA5oct highlights the molecular complexity of jumbo phages. Environ Microbiol. 2020;22(6):2165–81.
29. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–77.
30. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST server: rapid annotations using subsystems technology. BMC Genomics. 2008;9:75.
31. Wang RH, Yang S, Liu Z, Zhang Y, Wang X, Xu Z, et al. PhageScope: a well-annotated bacteriophage database with automatic analyses and visualizations. Nucleic Acids Res. 2023;gkad979.
32. 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.
33. Millard AD, Denise R, Lestido M, Thomas M, Webster D, Turner D, et al. taxmyPHAGE: automated taxonomy of dsDNA phage genomes at the genus and species. bioRxiv. 2024;2008–24.
34. Moraru C, Varsani A, Kropinski AM. VIRIDIC—a novel tool to calculate the intergenomic similarities of prokaryote-infecting viruses. Viruses. 2020;12(11):1268.
35. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. Nucleic Acids Res. 2024;gkae410.
36. Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinform. 2011;12:124.
37. Wang Y, Li X, Dance DAB, Xia H, Chen C, Luo N, et al. A novel lytic phage potentially effective for phage therapy against Burkholderia pseudomallei in the tropics. Infect Dis poverty. 2022;11(1):87.
38. Nakamoto JA, Evangelista W, Vinogradova DS, Konevega AL, Spurio R, Fabbretti A, et al. The dynamic cycle of bacterial translation initiation factor IF3. Nucleic Acids Res. 2021;49(12):6958–70.
39. Gerovac M, Chihara K, Wicke L, Bottcher B, Lavigne R, Vogel J. Immediate targeting of host ribosomes by jumbo phage encoded proteins. bioRxiv. 2023;2002–23.

40. Bryson AL, Hwang Y, Sherrill-Mix S, Wu GD, Lewis JD, Black L, et al. Covalent modification of bacteriophage T4 DNA inhibits CRISPR-Cas9. MBio. 2015;6(3):10–1128.
41. Vlot M, Houkes J, Lochs SJA, Swarts DC, Zheng P, Kunne T, et al. Bacteriophage DNA glucosylation impairs target DNA binding by type I and II but not by type V CRISPR–Cas effector complexes. Nucleic Acids Res. 2018;46(2):873–85.
42. Liu Y, Dai L, Dong J, Chen C, Zhu J, Rao VB, et al. Covalent modifications of the bacteriophage genome confer a degree of resistance to bacterial CRISPR systems. J Virol. 2020;94(23):10–1128.
43. Yi H, Huang L, Yang B, Gomez J, Zhang H, Yin Y. AcrFinder: genome mining anti-CRISPR operons in prokaryotes and their viruses. Nucleic Acids Res. 2020;48(W1):W358–65.
44. Hooton SPT, Connerton IF. Campylobacter jejuni acquire new host-derived CRISPR spacers when in association with bacteriophages harboring a CRISPR-like Cas4 protein. Front Microbiol. 2015;5:744.
45. Zhang Z, Pan S, Liu T, Li Y, Peng N. Cas4 nucleases can effect specific integration of CRISPR spacers. J Bacteriol. 2019;201(12):10–1128.
46. Bartolec T, Mitosch K, Potel C, Corona F, Yang ALJ, Burtscher ML, et al. Pervasive phosphorylation by phage T7 kinase disarms bacterial defenses. bioRxiv. 2024;2012–24.
47. Benler S, Hung S-H, Vander Griend JA, Peters GA, Rohwer F, Segall AM. Gp4 is a nuclease required for morphogenesis of T4-like bacteriophages. Virology. 2020;543:7–12.
48. Fujisawa H, Morita M. Phage DNA packaging. Genes Cells. 1997;2(9):537–45.
49. Young RY. Bacteriophage lysis: mechanism and regulation. Microbiol Rev. 1992;56(3):430–81.
50. Zhang S-P, Wang Q, Quan S-W, Yu X-Q, Wang Y, Guo D-D, et al. Type II toxin–antitoxin system in bacteria: activation, function, and mode of action. Biophys Rep. 2020;6:68–79.
51. Turner D, Kropinski AM, Adriaenssens EM. A roadmap for genome-based phage taxonomy. Viruses. 2021;13(3):506.
52. Naknaen A, Samernate T, Wannasrichan W, Surachat K, Nonejuie P, Chaikeeratisak V. Combination of genetically diverse *Pseudomonas* phages enhances the cocktail efficiency against bacteria. Sci Rep. 2023;13(1):8921.
53. Birkholz EA, Laughlin TG, Armbruster E, Suslov S, Lee J, Wittmann J, et al. A cytoskeletal vortex drives phage nucleus rotation during jumbo phage replication in *E. coli*. Cell Rep. 2022;40(7):111179.
54. Wafa A, Mohamed-Saad A, Addy HS, Kawasaki T. Isolation protocol of jumbo phage from winter grass soil. In: IOP conference series: earth and environmental science. IOP Publishing; 2021. p. 12071.
55. Serwer P, Wright ET. In-gel isolation and characterization of large (and other) phages. Viruses. 2020;12(4):410.
56. Saad AM, Soliman AM, Kawasaki T, Fujie M, Nariya H, Shimamoto T, et al. Systemic method to isolate large bacteriophages for use in biocontrol of a wide-range of pathogenic bacteria. J Biosci Bioeng. 2019;127(1):73–8.
57. Gaborieau B, Vaysset H, Tesson F, Charachon I, Dib N, Bernier J, et al. Prediction of strain level phage–host interactions across the *Escherichia* genus using only genomic information. Nat Microbiol. 2024;9(11):2847–61.
58. Alsharif G, Pham TK, Connolly AN, Pyrzanowska K, Smith EM, Alrafaie A, et al. Isolation of a novel *E. Coli* flagellotrophic jumbophage Shefm2k that replicates in cytoplasmic assembly areas. Microbiol Res. 2025;293:128082.
59. Haeusser DP, Hoashi M, Weaver A, Brown N, Pan J, Sawitzke JA, et al. The Kil peptide of bacteriophage λ blocks *Escherichia coli* cytokinesis via ZipA-dependent inhibition of FtsZ assembly. PLoS Genet. 2014;10(3):e1004217.
60. Conter A, Bouche J-P, Dassain M. Identification of a new inhibitor of essential division gene ftsZ as the kil gene of defective prophage Rac. J Bacteriol. 1996;178(17):5100–4.
61. Kiro R, Molshanski-Mor S, Yosef I, Milam SL, Erickson HP, Qimron U. Gene product 0.4 increases bacteriophage T7 competitiveness by inhibiting host cell division. Proc Natl Acad Sci. 2013;110(48):19549–54.


## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.