Data in Brief 57 (2024) 110897

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Genome sequence data of *Caudoviricetes* bacteriophage MK21 infecting *Xanthomonas citri*, the causative agent of citrus canker

Kira Moon[a], Seung Hui Song[a], Dae-Ju Oh[b], Ji-Gweon Park[b], Aslan Hwanwhi Lee[a], Miye Kwon[b,*]

[a] *Division of Environmental Materials, Honam National Institute of Biological Resources (HNIBR), Mokpo 58762, Republic of Korea*
[b] *Biodiversity Research Institute, Jeju Technopark (JTP), Jeju 63608, Republic of Korea*

ARTICLE INFO

Article history:
Received 1 July 2024
Revised 26 August 2024
Accepted 27 August 2024
Available online 5 September 2024

Dataset link: Novel bacteriophage mk21 infecting xanthomonas citri (Original data)
Dataset link: Novel bacteriophage mk21 infecting xanthomonas citri (Original data)

Keywords:
Phage
Phage therapy
Siphoviridae
Xanthomonas

ABSTRACT

This dataset reports the isolation and genomic characterization of the *Caudoviricetes* bacteriophage MK21, a novel bacteriophage infecting *Xanthomonas citri* subsp. *citri* (XCC), collected from soil samples on Jeju Island, South Korea. The phage was isolated and enriched using double agar layer plaque assays on nutrient media. Genomic analysis revealed that the phage MK21 is a double-stranded circular DNA genome of 43,495 bp, comprising 61 genes with high coding density. The dataset includes detailed genomic information, highlighting genes related to structural components, lysis mechanisms, and DNA/RNA metabolism. Phylogenetic analysis shows a close relationship with *Xanthomonas* phage CP1, supporting its potential use in comparative genomic studies and the development of antibacterial agents against citrus canker. This dataset offers valuable insights for the advancement of phage therapy and sustainable agricultural practices.

* Corresponding author.
E-mail address: miyekwon@jejutp.or.kr (M. Kwon).

https://doi.org/10.1016/j.dib.2024.110897

Specifications Table

Subject	Genomics
Specific subject area	Whole genome sequencing and genome annotation of a bacteriophage
Type of data	Tables, Image, Figures
Data collection	The *Xanthomonas*-infecting phage, *Caudoviricetes* bacteriophage MK21 was isolated from soil samples collected from Jeju Island, Korea, using *Xanthomonas citri* enriched suspension. Phage DNA was extracted by phage enrichment method and sequenced using Illumina technology. Raw whole-genome sequence was assembled using SPAdes v.3.13.1 and coding sequences were predicted by the RAST web service, Prokka, and NCBI BLASTp. Annotation of the genome was performed using BLASTP provided by the NCBI.
Data source location	Institution: Jeju Biodiversity Institute, Jeju Technopark, Korea Region: Jeju Island Country: Republic of Korea Latitude and Longitude (GPS coordinates) for collected samples/data: 33.31N, 126.63W
Data accessibility	Repository name: GenBank, NCBI Data identification number: PP780497 Direct URL to data: data not available until this article is published
Related research article	None.

1. Value of the Data

- *Xanthomonas citri* causes citrus canker on mandarin farms, leading to significant economic losses. Environmentally friendly preventive measures against this pathogenic bacterium are crucial for healthier fruit harvest. Therefore, whole-genome data of a novel bacteriophage that effectively infects *X. citri* are highly valuable for the development of phage therapy systems for citrus canker.
- This dataset includes detailed genomic information on the bacteriophage, featuring a lysis cassette containing holin, endolysin, and o-spanin genes. These components are essential for understanding the mechanisms of bacteriophage lysis, especially in Gram-negative bacterial strains. These lysis genes can be utilized by other researchers to further study bacteriophage genomes. These, genes have potential applications in the development of phage-derived antibacterial agents.
- The data provided are the whole-genome sequences of the novel bacteriophage that can be utilized for comparative analysis with other bacteriophage genomes. Moreover, this genome can facilitate phylogenetic studies and enhance our understanding of the evolution and relationships among bacteriophages.

2. Background

Xanthomonas citri subsp. *citri* (XCC) is a major phytopathogen that causes substantial economic damage to farmers by causing citrus canker [1]. Jeju Island, the predominant citrus cultivation region in South Korea, is particularly susceptible to this disease [2]. Although citrus canker is currently not the most prevalent phytopathological issue in South Korea, it still contributes to the reduced quality of citrus fruits, thereby diminishing their market value.

Copper (Cu)-based bactericides are commonly employed to combat XCC [3]. While effective, extensive use of these bactericides has led to the emergence of Cu-resistant phytopathogens [4]. Furthermore, the heavy reliance on Cu-based treatments has raised considerable environmental concerns, particularly regarding soil contamination and pollution because of Cu accumulation.

In light of these challenges, bacteriophages (phages) have been proposed as an alternative and eco-friendly method for the treatment and prevention of citrus canker [5]. Phage therapy offers a targeted approach, as phages are exclusively effective against specific host bacterial strains. This specificity ensures little risk of environmental pollution or adverse effects on the surround-

ing phytosphere. Therefore, culturing and revealing the whole-genome of a novel phage infecting *X. citri* is fundamental for the development of phage therapy.

3. Data Description

The current study reports a novel phage, *Caudoviricetes* bacteriophage MK21, classified within the kingdom *Heunggongvirae* and phylum *Uroviricota*. This phage is classified as siphoviridae [6], featuring an icosahedral head with a diameter of 48.93 ± 1.28 nm and a long contractile tail measuring 153.57 nm ± 15.43 nm (Fig. 1). The *Caudoviricetes* bacteriophage MK21 possesses a double-stranded circular DNA genome, 43,495 bp in length, with a G+C content of 53.08 % (Table 1). The closest related phage is *Xanthomonas* phage CP1 (NC_019933) [7], with a two-way Average Nucleotide Identity (ANI) value of 92.8 %.

Table 1 summarizes the genomic characteristics of *Caudoviricetes* bacteriophage MK21. The genome spans 43,506 bp and contains 61 genes with 36 assigned functions. It exhibits high coding density of 92.74 % and lacks tRNA genes. Genes with known functions are categorized into four groups: structural, recombination-associated, lytic, and DNA/RNA metabolic genes (Table 2). Most genes are associated with structural functions, including the terminase large subunit, head maturation protein, and phage tail protein (Supplementary Table). Nine genes, including endonu-

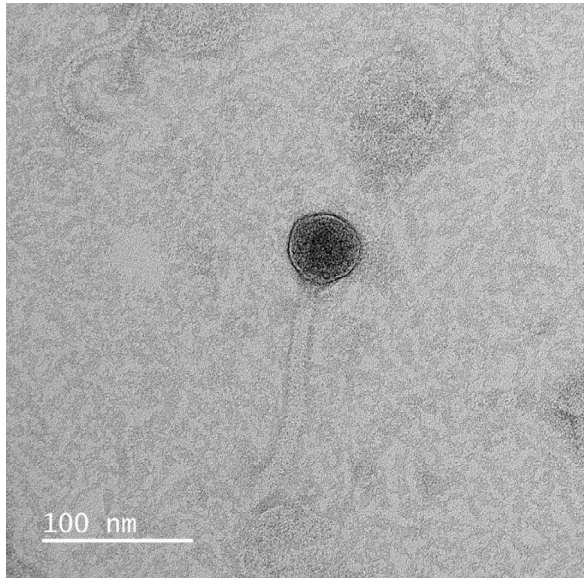

Fig. 1. Transmission electron micrographs of the *Caudoviricetes* bacteriophage MK21.

Table 1
Genome details of *Caudoviricetes* bacteriophage MK21.

Feature	*Caudoviricetes* bacteriophage MK21
Genome Size	43,506 bp
Number of genes	61
Genes assigned a function	36
G+C content	53.08 %
Coding density	92.74 %
tRNAs	0
Morphology	Siphoviridae
Closest relative	Xanthomonas phage CP1(NC_019933)

Table 2
Comparison of functional genes predicted in *Caudoviricetes* bacteriophage MK21.

Gene Functional Group	Number and details of *Caudoviricetes* bacteriophage MK21 genes
Structural genes	**19** (1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 17, 19, 21, 28, 31)
Genes associated with recombination	**9** (2, 13, 18, 20, 27, 29, 38, 57, 59)
Lytic genes	**1** (26)
DNA/RNA metabolism genes	**7** (35, 37, 39, 41, 42, 44, 45)
Genes with hypothetical functions	**25** (22, 23, 24, 25, 30, 32, 33, 34, 36, 40, 43, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56 58, 60, 61)
Total	61

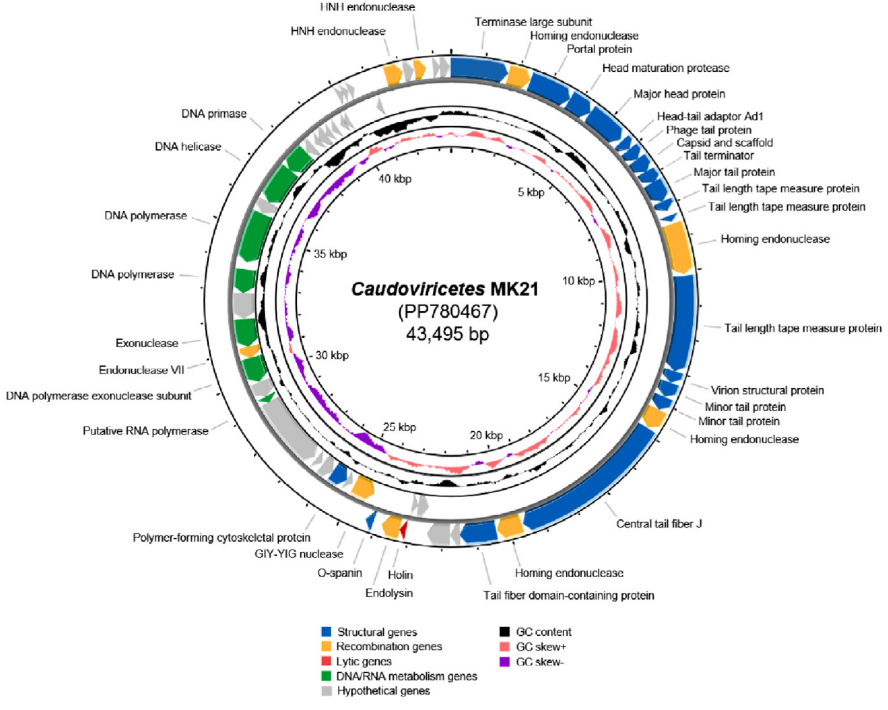

Fig. 2. Genome map of the *Caudoviricetes* bacteriophage MK21. The genes predicted to encode structural genes are colored blue, those predicted to encode recombination genes are colored yellow, the gene predicted to encode lytic gene is colored red, those predicted to encode DNA/RNA metabolism genes are colored green, and those predicted to encode hypothetical genes are colored grey.

cleases and endolysins, are involved in nucleic acid recombination, whereas seven genes are related to DNA/RNA metabolism, including DNA and RNA polymerases, DNA helicase, and DNA primase. Notably, MK21 carries a lysis gene cassette comprising holin, endolysin, and o-spanin genes which are known to disrupt peptidoglycan layer and lipopolysaccharide outer membrane of the host [8].

The phage MK21 genome displays a unique gene orientation (Fig. 2, Supplementary Table). Genes 57 (HNH endonuclease) and 28 (o-spanin protein) are predominantly located on the positive strand with a positive G+C skew. The genomic region is associated with various structural functions. Conversely, genes 29 (GIY-YIG nuclease family protein) to 52 (hypothetical protein)

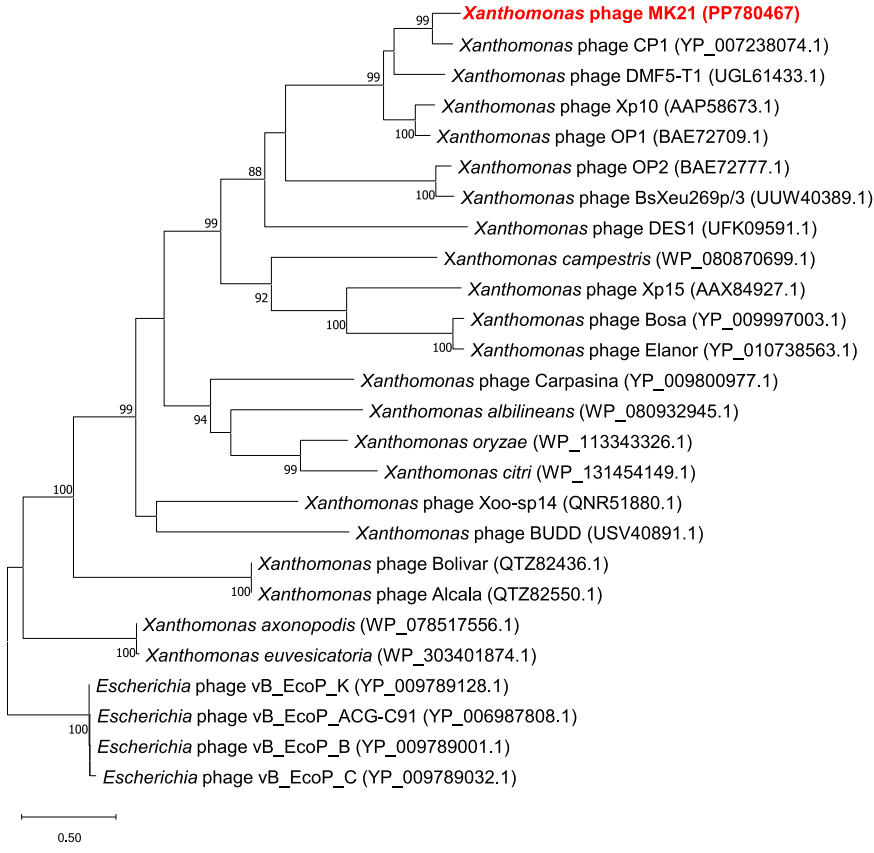

Fig. 3. Maximum-likely phylogenetic tree constructed from the terminase large subunit gene sequences of the *Caudoviricetes* bacteriophage MK21 and its related phages. Bootstrap values under 70 were removed. Accession numbers for the phages are written in parenthesis.

are primarily negative strand with a negative G+C skew, mostly encoding DNA/RNA metabolism and hypothetical proteins. For phylogenetic analysis, the terminase large subunit gene of phage MK21 was compared with that of closely related *Xanthomonas* phages. A maximum likelihood tree was constructed using the Poisson model and a bootstrap value of 500 after aligning the sequences with Clustal Omega. According to the phylogenetic tree (Fig. 3), phage MK21 grouped with *Xanthomonas* phages CP1, DMF5-T1, XP10, and OP1.

4. Experimental Design, Materials and Methods

The *Caudoviricetes* bacteriophage MK21, which infects XCC 27-12, was segregated from canker-infected leaf samples collected from a citrus farm on Jeju Island, South Korea. To isolate the phage, the collected citrus leaves were finely chopped and soaked in 0.85 % sterile saline solution at a 10 % ratio (v/v). The mixture was agitated at 25 °C and 180 rpm for 18 h. The solution of suspended citrus leaves was filtered through a on 0.22 μm pore size filter to remove bacterial debris. The filtrate was then used for a plaque assay to enrich and purify the phage particles. In total, 2 mL of filtrate was mixed and 4 mL of host bacterial strain inoculum with 14 mL of nutrient broth media (BD Difco, US); and then incubated at 28 °C for 25 min in a water bath. After incubation, 20 mL of 2 x Nutrient agar (BD Difco, US) was added to the mixture of host-phage followed by overnight cultivation on the petri dish for 24 h at 28 °C in the dark.

Individual plaques were picked with a sterile toothpick and stored in 1 mL of sodium-magnesium (SM) buffer at 4 °C; and then sterilized with 20 µL of chloroform. Further enrichment was achieved by generating confluent double-agar layer plates. To induce diffusion, 10 mL of SM buffer was added to the confluent phage agar plate and incubated at 4 °C overnight [9]. After incubation, phage particles that diffused into the SM buffer were collected. To concentrate phage particles, PEG8000 was added at 10 % w/v to the phage inoculate and incubated at 0 °C for 24 h to precipitate phage particles. The mixture was centrifuged at 11,000 × g for 40 min, the supernatant was discarded, and the pellet was resuspended in 3 mL SM buffer. The resuspension was treated with chloroform (1:1 ratio), vortexed, and centrifuged at 3000 rpm at 4 °C for 30 min. The supernatant containing phage particles was collected and further concentrated via ultracentrifugation at 48,000 × g for 2 h at 4 °C [10]. The resulting pellet was resuspended in 0.1 mL SM buffer.

A 10 µL aliquot of the phage concentrate was examined using transmission electron microscopy at Gangwon National University upon negative staining with 2 % v/v uranyl acetate [11]. The microscopic images were analyzed using the Gatan Microscope Suite (GMS3).

Phage genomic DNA was extracted using a QIAGEN Blood & Tissue Kit (QIAGEN, Germany) following the manufacturer's instruction. DNA concentration and quality were assessed using a NanoDrop spectrophotometer (Spectrophotometer AND-1000, South Korea). DNA was sequenced using an Illumina HiSeq 4000 (Macrogen, South Korea), yielding 150-bp, paired-end reads and resulting in approximately 5 Gb dataset. Sequenced reads were assembled using SPAdes v. 3.13.1 [12].

Genome annotation was conducted using the RAST web server [13], Prokka 1.13 [14] and NCBI BLASTP against the nr database. Genetically related reference phage genomes were selected from a protein sequence phylogenetic tree constructed using the ViPTree webserver [15]. The ANI value with the related phages were calculated using the ANI calculator (enveomics.ce.gatech.edu/ani/index) [16]. Terminase large subunit amino acid sequences of the related reference phage genomes were obtained from the NCBI database. After alignment using Clustal Omega [17], a maximum-likelihood phylogenetic tree with a bootstrap of 500 replicates was constructed using MEGA 11 program [18]. The phage genome tree was visualized using Proksee [19].

Limitations

None.

Ethics Statement

The authors have read and followed the ethical requirements for publication in Data in Brief and confirm that the current work does not involve human subjects, animal experiments, or any data collected from social media platforms.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the authors used chatGPT4.0o in order to improve language. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data Availability

Novel bacteriophage mk21 infecting xanthomonas citri (Original data) (NCBI)
Novel bacteriophage mk21 infecting xanthomonas citri (Original data) (NCBI)

CRediT Author Statement

Kira Moon: Methodology, Investigation, Data curation, Software, Writing – original draft, Visualization; **Seung Hui Song:** Methodology; **Dae-Ju Oh:** Project administration, Supervision; **Ji-Gweon Park:** Project administration, Supervision; **Aslan Hwanwhi Lee:** Project administration, Supervision; **Miye Kwon:** Conceptualization, Methodology, Resources, Investigation, Data curation, Validation, Visualization, Writing – review & editing, Supervision.

Acknowledgments

The bacterial host strain, *Xanthomonas citri* subsp. *citri* (XCC) 27-12 used in this study was provided by the Citrus Research Institute of National Institute of Horticultural and Herbal Science. This work was supported by the grant from National Research Foundation of Korea (NRF), funded by the Korean government (MSIT) (2019R1C1C1007725) and Honam National Institute of Biological Resources (HNIBR), funded by the Ministry of Environment (MOE) of the Republic of Korea (HNIBR202402114).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2024.110897.

References

[1] S. Ali, A. Hameed, G. Muhae-Ud-Din, M. Ikhlaq, M. Ashfaq, M. Atiq, F. Ali, Z.U. Zia, S.A.H. Naqvi, Y. Wang, Citrus canker: a persistent threat to the worldwide citrus industry—an analysis, Agronomy 13 (4) (2023) 1112.
[2] J.W. Hyun, P.H. Yi, R.Y. Hwang, K.H. Moon, Aspect of incidence of the major citrus diseases recently, Res. Plant Dis. 19 (2) (2013) 102–107.
[3] C.M. Ference, A.M. Gochez, F. Behlau, N. Wang, J.H. Graham, J.B. Jones, Recent advances in the understanding of *Xanthomonas citri* ssp. *citri* pathogenesis and citrus canker disease management, Mol. Plant Pathol. 19 (6) (2018) 1302–1318, doi:10.1111/mpp.12638.
[4] D. Richard, N. Tribot, C. Boyer, M. Terville, K. Boyer, S. Javegny, M. Roux-Cuvelier, O. Pruvost, A. Moreau, A. Chabirand, First report of copper-resistant *Xanthomonas citri* pv. *citri* pathotype A causing Asiatic citrus canker in Réunion, France, Plant Dis. 101 (3) (2017) 503.
[5] R. Nakayinga, A. Makumi, V. Tumuhaise, W. Tinzaara, *Xanthomonas* bacteriophages: a review of their biology and biocontrol applications in agriculture, BMC Microbiol. 21 (1) (2021) 291, doi:10.1186/s12866-021-02351-7.
[6] C.M. Chibani, A. Farr, S. Klama, S. Dietrich, H. Liesegang, Classifying the unclassified: a phage classification method, Viruses 11 (2) (2019) 195, doi:10.3390/v11020195.
[7] A.A. Ahmad, M. Ogawa, T. Kawasaki, M. Fujie, T. Yamada, Characterization of bacteriophages Cp1 and Cp2, the strain-typing agents for Xanthomonas axonopodis pv. citri, Appl. Environ. Microbiol. 80 (1) (2014) 77–85, doi:10.1128/aem.02310-13.
[8] R. Young, Phage lysis: three steps, three choices, one outcome, J. Microbiol. 52 (3) (2014) 243–258, doi:10.1007/s12275-014-4087-z.
[9] P. Serwer, S.J. Hayes, J.A. Thomas, S.C. Hardies, Propagating the missing bacteriophages: a large bacteriophage in a new class, Virol. J. 4 (2007) 21, doi:10.1186/1743-422x-4-21.
[10] M.R. Green, J. Sambrook, Molecular Cloning: a Laboratory Manual, 4th ed., Cold Spring Harbor, New York, 2012.
[11] H.W. Ackermann, Bacteriophage electron microscopy, Adv. Virus Res. 82 (2012) 1–32, doi:10.1016/b978-0-12-394621-8.00017-0.
[12] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. 19 (5) (2012) 455–477.

[13] T. Seemann, Prokka: rapid prokaryotic genome annotation, Bioinformatics 30 (14) (2014) 2068–2069.
[14] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, O. Zagnitko, The RAST Server: rapid annotations using subsystems technology, BMC Genom. 9 (75) (2008) 1–15.
[15] Y. Nishimura, T. Yoshida, M. Kuronishi, H. Uehara, H. Ogata, S. Goto, ViPTree: the viral proteomic tree server, Bioinformatics 33 (15) (2017) 2379–2380.
[16] L.M. Rodriguez-R, K.T. Konstantinidis, Bypassing cultivation to identify bacterial species, Microbe 9 (3) (2014) 111–118.
[17] F. Sievers, D.G. Higgins, Clustal Omega, accurate alignment of very large numbers of sequences, Methods Mol. Biol. 1079 (2014) 105–116, doi:10.1007/978-1-62703-646-7_6.
[18] K. Tamura, G. Stecher, S. Kumar, MEGA11: molecular evolutionary genetics analysis version 11, Mol. Biol. Evol. 38 (7) (2021) 3022–3027, doi:10.1093/molbev/msab120.
[19] J.R. Grant, E. Enns, E. Marinier, A. Mandal, E.K. Herman, C.Y. Chen, M. Graham, G. Van Domselaar, P. Stothard, Proksee: in-depth characterization and visualization of bacterial genomes, Nucleic Acids Res. 51 (W1) (2023) W484–w492, doi:10.1093/nar/gkad326.