

Complete genome sequence of bacteriophages infecting *Escherichia*, *Enterobacter*, and *Pseudomonas* isolates

Aubrey Joy P. Tejada,¹ Michael Angelou L. Nada,¹ Ruth Antoinette D. Chin,¹ Janna Ysabelle O. Casidsid,¹ Joseph B. Ancla,¹ Marel Jan G. Joloro,¹ Mark Christian C. Reterta,¹ Anton Roi G. Collado,¹ Sharmen C. Berlin,¹ Arra B. Asejo,² Nikka Mae R. Yadao,¹ Virgilio P. De Paz, Jr.,¹ Ursela G. Bigol,¹ Rommel J. Gestuveo^{1,3}

AUTHOR AFFILIATIONS See affiliation list on p. 5.

ABSTRACT We reported six new species of bacteriophages belonging to the genus *Drulivirus*, *Mosigvirus*, *Kayfunavirus*, and *Kagunavirus*, with a shared genome similarity of 77.2% to 94.5%. Seven isolates are suitable candidates for phage therapy, thereby expanding our knowledge about biocontrol alternatives against infections caused by multidrug-resistant bacteria.

KEYWORDS phage therapy, genome sequencing, antimicrobial resistant bacteria, clinical isolates, bacteriophages, Philippines

The spread of antimicrobial resistance (AMR) raises public health concerns, rendering bacteria immune to antibiotics. Phage therapy is being explored as an alternative approach to combat AMR. Thus, phage identification and characterization through laboratory assays and genome sequencing are imperative to determine suitable candidates for phage therapy (1). Here, we report the complete genome of eight bacteriophages infecting selected bacterial isolates.

Wastewater samples (1 L) from various locations (Table 1) were collected, centrifuged, and filtered sequentially using 0.45 µm and 0.22 µm membrane filters. Direct isolation was used to isolate the *Escherichia* phage, VIPECOMC06, and the *Pseudomonas* phage, VIPPAEUMC02; otherwise, enriched isolation was utilized. Briefly, 20 mL of filtrate in 20 mL double-strength Luria-Bertani broth (HiMedia #M1151) containing 1 mM CaCl₂ was inoculated with target bacterial host (1 mL) (Table 1) for 24 hr (37°C, 160 rpm). Both methods employed double-agar overlay; single plaques were picked after 24 hr, purified for at least three rounds, and propagated to obtain high-titer phage lysates (Table 1) (2, 3).

DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen) incubating samples at 75°C for 5–10 min to inactivate DNase 1/RNase A (4) and further purified using AMPure XP beads (Beckman Coulter). To prepare the sequencing libraries, 100–500 ng of DNA were tagged and amplified for 5 cycles using the Illumina DNA prep kit and sequenced (loading concentration: 15 pM) on the Illumina MiSeq platform following the manufacturer's instructions. On average, 1,612,925 paired-end (251 bp) reads were generated per sample.

Read quality was assessed using FastQC v0.11.9 (5) and trimmed with Trimmomatic v0.39 (6). Subsampled reads were *de novo* assembled using SPAdes v3.15.5 (parameters: --only-assembler -k 21,33,55,77,99,127) (7). Genomes were reoriented based on terminase (terL) genes of the closest relative using SeqFu (8) followed by error correction using Pilon v1.24 (9). Genome coverage was determined using BBDMap v39.06 (10) and SAMtools v1.19.2 (11). Finally, phage genomes were annotated using Prokka v1.14.6 (12) and the PHROGs (13) database. Default parameters were used for all software unless otherwise specified.

Editor John J. Dennehy, Department of Biology, Queens College, Queens, New York, USA

Address correspondence to Aubrey Joy P. Tejada, ajptejada@gmail.com, or Michael Angelou L. Nada, mikeangelounada@gmail.com.

The authors declare no conflict of interest.

See the funding table on p. 5.

Received 20 December 2024

Accepted 6 March 2025

Published 1 April 2025

Copyright © 2025 Tejada et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

TABLE 1 Genome characterizations of the *Escherichia*, *Enterobacter*, and *Pseudomonas* phages

Phage name	Enterobacter phage vB_	Escherichia phage vB_	Escherichia phage vB_	Pseudomonas phage vB_	Pseudomonas phage vB_	Escherichia phage vB_	Escherichia phage vB_
	VIPECLMCO6	VIPECOMC04	VIPECOMC06	VIPECOTH07	VIPEAETPH1	VIPEAUMC02	VIPECOTPH05A°
Sampling location	Masalasa Creek	Masalasa Creek	Masalasa Creek	Tertiary Hospital	Tarlac Provincial Hospital	De La Salle Medical and Health Sciences Institute	Tarlac Provincial Hospital
Sampling coordinates	15°28'33" N 12°35'42" E	15°28'33" N 12°35'42" E	15°28'33" N 12°35'42" E	14°34'39" N 120°59'08" E	15°28'28" N 120°35'09" E	14°19'40" N 120°56'32" E	15°28'28" N 120°35'09" E
Propagation host	<i>Enterobacter cloacae</i>	<i>Escherichia coli</i>	<i>Escherichia coli</i> ATCC 11229	<i>Escherichia coli</i> OMCS-22-1338	<i>Pseudomonas aeruginosa</i> OMCS-21-1583	<i>Pseudomonas aeruginosa</i> ATCC 15442	<i>Escherichia coli</i> OMCS-21-1589B
Isolation method	Enriched	Enriched	Direct	Enriched	Enriched	Direct	Enriched
Phage titer (PFU/mL)	3.07 × 10 ¹⁰	2.6 × 10 ¹⁰	4.5 × 10 ⁹	6.2 × 10 ¹⁰	2.13 × 10 ⁹	1.0 × 10 ¹⁰	2.6 × 10 ¹⁰
No. of raw reads	1,292,346	1,713,847	1,405,721	1,322,377	1,389,816	2,388,688	1,377,052
No. of cleaned reads	431,893	1,371,138	1,171,304	1,013,834	867,728	751,771	183,193
No. of subsampled reads	150,000	5,000	200,000	30,000	300,000	300,000	100,000
Genome size (bp)	40,245	45,658	166,719	44,205	66,340	92,031	44,015
GC content (%)	51.2	50.7	37.5	50.8	55.6	49.3	53.8
Mean coverage (x)	1554	36	112	80	1726	1297	428
No. of CDS	50	76	264	76	93	172	61
No. of CDS with predicted functions	32 (64%)	40 (53%)	155 (59%)	38 (50%)	36 (39%)	53 (30%)	28 (46%)
No. of tRNAs	1	0	1	0	0	14	0
Lifestyle prediction	Lytic	Lytic	Lytic	Lytic	Lytic	Temperate	Lytic
Temperate marker genes	Absent	Absent	Absent	Absent	Absent	Cro gene	Absent
Presence of ARGs ^b	Absent	Absent	Absent	Absent	Absent	Absent	Absent
Presence of virulence genes	Absent	Absent	Absent	Absent	Absent	Absent	Absent
Closest relative (accession number)	<i>Enterobacteria</i> phage EcoDS1 (EU734172)	<i>Escherichia</i> phage vB_EcoS_fPoEco01 (NC_073317)	<i>Shigella</i> phage phi25-307 (MG589383)	<i>Escherichia</i> phage vB_EcoS_XY1 (NC_073325)	<i>Pseudomonas</i> phage vB_PaeM_LS1 (MG897799)	<i>Pseudomonas</i> phage vB_VIPPAEUMC01 (OQ721915)	<i>Klebsiella</i> phage Kp2 (KT367886)
Genome similarity (%)	77.2	79.5	94.5	89.2	95.6	100	77.9
Phage family ^c	Autographiviridae	Guernseyvirinae	Straboviridae	Guernseyvirinae	Lindbergviridae	Vandenendeviridae	Autographiviridae
Proposed binomial name	<i>Kayfunavirus jftwin</i>	<i>Kagunavirus ohbebero</i>	<i>Mosivirus lindsay</i>	<i>Kagunavirus ligaya</i>	<i>Pbunavirus LST^d</i>	<i>Pakpunavirus vippaeumc^d</i>	<i>Drulivirus working-ina</i>
BioProject accession no.	PRJNA1164958	PRJNA1164962	PRJNA1164964	PRJNA1164959	PRJNA1164966	PRJNA1164965	PRJNA1164961

(Continued on next page)

TABLE 1 Genome characterizations of the *Escherichia*, *Enterobacter*, and *Pseudomonas* phages (Continued)

Phage name	<i>Enterobacter</i> phage vB_	<i>Escherichia</i> phage vB_	<i>Escherichia</i> phage vB_	<i>Pseudomonas</i> phage vB_	<i>Pseudomonas</i> phage vB_	<i>Escherichia</i> phage vB_	<i>Escherichia</i> phage vB_
	VIPECLMC06	VIPECOMC04	VIPECOMC06	VIPEATH07	VIPEATH01	VIPEATH02	VIPEATH05A ^a
SRA accession no.	SRR30791123	SRR30791285	SRR30791201	SRR30791124	SRR30791126	SRR30791202	SRR30791125
GenBank accession	PQ411323	PQ423993	PQ416618	PQ417132	PQ417133	PQ429081	PQ429079

^aGenomes were assembled from the same sequencing library indicating that two phages were present in the sample. Further purification using Cesium chloride (CsCl) and ultracentrifugation is needed to fully separate the two phages.
^bAntibiotic resistance genes.
^cBased on current ICTV Taxonomy (MSL v40).
^dICTV-registered species.

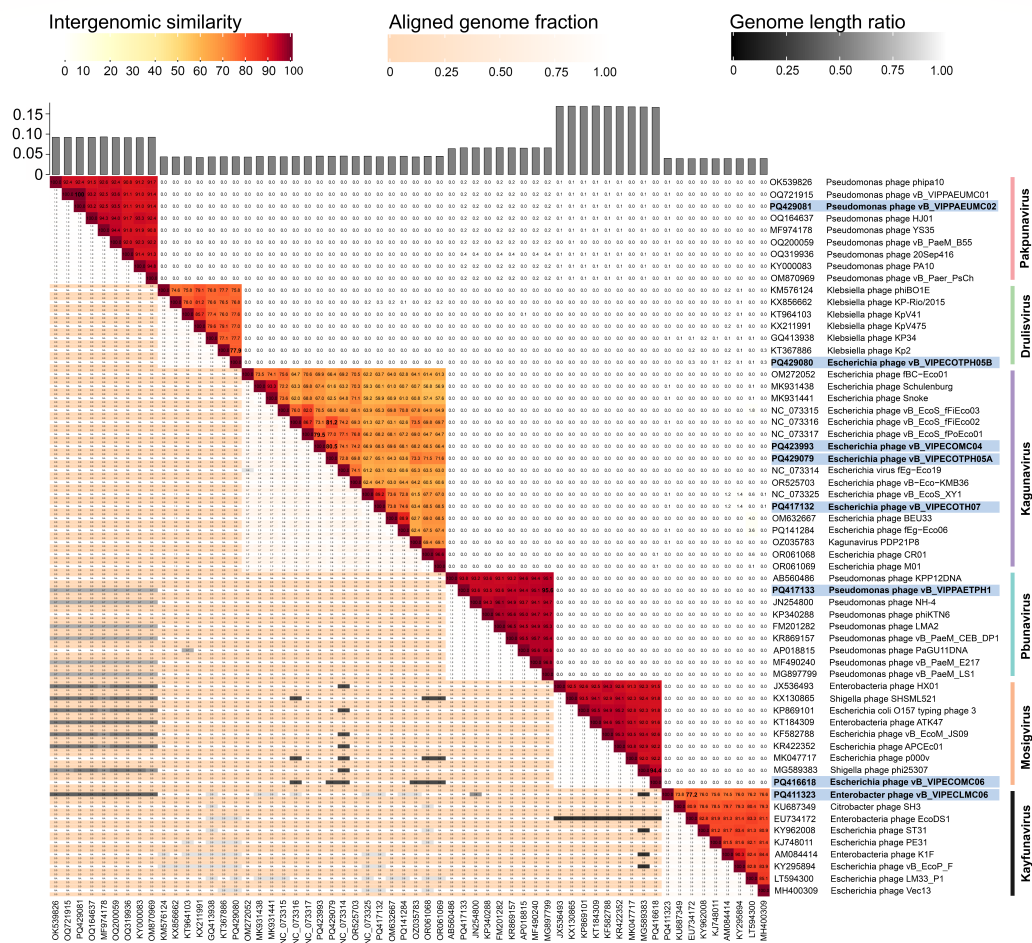


FIG 1 Whole-genome comparison using Virus Intergenomic Distance Calculator (VIRIDIC). The closest relatives (distance cut-off: ≤ 0.09) were determined with mash distance estimation using INPHARED (v1Sep2024) database, and the reference genomes of the top eight hits were downloaded from NCBI. Dark red color indicates high genome similarities, with percent similarities of the closest relative in bold. Phages with genome similarity of $\geq 95\%$ are identical species. In closely related phages, a high fraction (orange to white) of the genome is aligned and is expected to have similar genome length (black to white). The phage isolates and their accession numbers are highlighted in blue.

Genome size ranges from 40,245 to 166,719 bp with GC content of 37.5% to 55.6% and mean coverage of 37 $\times$ to 1,726 $\times$ (Table 1). All genomes were identified as “complete with direct terminal repeats” using checkV v1.4 (14). Taxonomic assignment was performed using taxmyPHAGE (15). For samples that cannot be classified using ICTV-registered genomes, the closest relatives were identified using Mash v2.3 (distance cut-off: 0.09) (16) and INPHARED (v1Sep2024) database (<https://github.com/RyanCook94/inphared>) (17). Reference genomes were retrieved from the NCBI nucleotide database, and a phylogenetic tree was constructed based on Genome-BLAST Distance Phylogeny using VICTOR (<https://ggdc.dsmz.de/victor.php#>) (18).

We identified six new phage species with genome similarity of 77.2% to 94.5% using VIRIDIC v1.1 (19) (Fig. 1). These isolates belonging to the genus *Drulisvirus* (1), *Mosigivirus* (1), *Kayfunavirus* (1), and *Kagunavirus* (3) were named following the ICTV guidelines (Table 1) (20, 21). Meanwhile, the *Pseudomonas* phages VIPPAETPH1 and VIPPAEUMC02 were identical to vB_PaeM_LS1 (95.6%) and vB_VIPPAEUMC01 (100%), respectively. Finally, PhageLeads (22) predicted seven isolates with lytic lifestyle, no antibiotic resistance, and virulence genes, making them suitable for phage therapy.

ACKNOWLEDGMENTS

The authors thank Dr. Maita Theresa P. Rigor of the Tarlac Provincial Hospital (TPH), the management of De La Salle Medical and Health Sciences Institute (DLSMHSI), the local government of Tarlac for their assistance during sampling collection, and the DOST-ASTI Computing and Archiving Research Environment (COARE) for computing resources.

This study was funded by the Department of Science and Technology Grants-In-Aid (GIA) Program Project number LFP-EBD-2021-02.

AUTHOR AFFILIATIONS

¹Department of Science and Technology, Industrial Technology Development Institute, Taguig City, Philippines

²Department of Science and Technology, Science Education Institute, Taguig City, Philippines

³Department of Science and Technology, S&T Fellows Program, Taguig City, Philippines

AUTHOR ORCID*s*

Aubrey Joy P. Tejada  <http://orcid.org/0009-0003-9940-2914>

Michael Angelou L. Nada  <http://orcid.org/0000-0002-3466-3465>

FUNDING

Funder	Grant(s)	Author(s)
Department of Science and Technology Grants-In-Aid Program	LFP-EBD-2021-02	Rommel J. Gestuveo

AUTHOR CONTRIBUTIONS

Aubrey Joy P. Tejada, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review and editing | Michael Angelou L. Nada, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Visualization, Writing – original draft | Ruth Antoinette D. Chin, Data curation, Methodology, Project administration, Validation, Writing – review and editing | Janna Ysabelle O. Casidsid, Data curation, Methodology, Project administration, Validation, Writing – review and editing | Joseph B. Ancla, Conceptualization, Data curation, Methodology, Validation, Writing – review and editing | Marel Jan G. Joloro, Conceptualization, Data curation, Methodology, Validation, Writing – review and editing | Mark Christian C. Reterta, Data curation, Methodology, Project administration, Validation, Writing – review and editing | Anton Roi G. Collado, Data curation, Methodology, Project administration, Validation, Writing – review and editing | Sharmen C. Berlin, Data curation, Methodology, Validation, Writing – review and editing | Arra B. Asejo, Data curation, Validation, Writing – review and editing | Nikka Mae R. Yadao, Project administration, Writing – review and editing | Virgilio P. De Paz, Jr., Project administration, Writing – review and editing | Ursela G. Bigol, Funding acquisition, Project administration | Rommel J. Gestuveo, Funding acquisition, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

The raw sequences (SRA) and genome assemblies were deposited in DDBJ/ENA/GenBank under the accession numbers in Table 1.

REFERENCES

- Nada MAL, Ancla JB, Yadao NMR, de Paz VP, Manalaysay JG, Samante FLD, Bigol UG. 2023. Isolation and genome sequencing of five lytic bacteriophages from hospital wastewater in the Philippines. *Microbiol Resour Announc* 12:e0031123. <https://doi.org/10.1128/MRA.00311-23>
- Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage discovery guide. Howard Hughes Medical Institute.

3. Clokie MR, Kropinski A. 2009. Methods and protocols, volume 1: Isolation, characterization, and interactions. *Methods in molecular biology*, p 69–81. Humana press.
4. Jakočiūnė D, Moodley A. 2018. A rapid bacteriophage DNA extraction method. *Methods Protoc* 1:27. <https://doi.org/10.3390/mps1030027>
5. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data
6. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
7. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>
8. Telatin A, Fariselli P, Birolo G. 2021. SeqFu: a suite of utilities for the robust and reproducible manipulation of sequence files. *Bioengineering (Basel)* 8:59. <https://doi.org/10.3390/bioengineering8050059>
9. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>
10. Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner
11. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
12. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
13. Terzian P, Olo Ndela E, Galiez C, Lossouarn J, Pérez Bucio RE, Mom R, Toussaint A, Petit M-A, Enault F. 2021. PHROG: families of prokaryotic virus proteins clustered using remote homology. *NAR Genom Bioinform* 3:lqab067. <https://doi.org/10.1093/nargab/lqab067>
14. Nayfach S, Camargo AP, Schulz F, Eloë-Fadrosch E, Roux S, Kyrpides NC. 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat Biotechnol* 39:578–585. <https://doi.org/10.1038/s41587-020-00774-7>
15. Millard A, Denise R, Lestido M, Thomas MT, Webster D, Turner D, Sicheritz-Pontén T. 2025. taxMyPhage: automated taxonomy of dsDNA phage genomes at the genus and species level. *Phage* 17:1–14. <https://doi.org/10.1089/phage.2024.0050>
16. Cook R, Brown N, Redgwell T, Rihtman B, Barnes M, Clokie M, Stekel DJ, Hobman J, Jones MA, Millard A. 2021. INfrastructure for a PHAge REference database: identification of large-scale biases in the current collection of cultured phage genomes. *Phage (New Rochelle)* 2:214–223. <https://doi.org/10.1089/phage.2021.0007>
17. Meier-Kolthoff JP, Göker M. 2017. VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics* 33:3396–3404. <https://doi.org/10.1093/bioinformatics/btx440>
18. Moraru C, Varsani A, Kropinski AM. 2020. VIRIDIC-A novel tool to calculate the intergenomic similarities of prokaryote-infecting viruses. *Viruses* 12:1268. <https://doi.org/10.3390/v12111268>
19. Turner D, Kropinski AM, Adriaenssens EM. 2021. A roadmap for genome-based phage taxonomy. *Viruses* 13:506. <https://doi.org/10.3390/v13030506>
20. Turner D, Shkoporov AN, Lood C, Millard AD, Dutilh BE, Alfenas-Zerbini P, van Zyl LJ, Aziz RK, Oksanen HM, Poranen MM, et al. 2023. Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Arch Virol* 168:74. <https://doi.org/10.1007/s00705-022-05694-2>
21. Yukgehaish K, Rajandas H, Parimannan S, Manickam R, Marimuthu K, Petersen B, Clokie MRJ, Millard A, Sicheritz-Pontén T. 2022. PhageLeads: rapid assessment of phage therapeutic suitability using an ensemble machine learning approach. *Viruses* 14:342. <https://doi.org/10.3390/v14020342>