

# Complete genome sequence of a *Pseudomonas fluorescens* bacteriophage UNO-G1W1 isolated from freshwater ice in Nebraska

Thomas T. Schulze,<sup>1,2</sup> Andrew J. Neville,<sup>1</sup> Gabrielle F. Watson,<sup>1,2</sup> Austin G. Sanford,<sup>1</sup> Harim I. Won,<sup>1</sup> Mackenzie E. Conrin,<sup>1</sup> Connor G. Eastman,<sup>1</sup> LeeAnna M. Lui,<sup>1</sup> M. Yunos Alizai,<sup>1</sup> Matthias J. Walters,<sup>1</sup> Paul H. Davis,<sup>1</sup> William E. Tappich<sup>1</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 2.

**ABSTRACT** We provide the complete genome sequence for a novel *Pseudomonas fluorescens* bacteriophage named UNO-G1W1. This phage was isolated from a single ice cover sampling. The genome was sequenced on the Nanopore MinION, generated with the direct terminal repeat-phage-pipeline and polished with Illumina short reads. Sequence identity classifies the phage as an *otagovirus*.

**KEYWORDS** bacteriophages, *Pseudomonas fluorescens*, *Pseudomonas*, genomics, genomes, long-read sequencing, Oxford Nanopore, DTR, hybrid assembly

*Pseudomonas fluorescens* is a gram-negative bacterium commonly found in water and soil that is known for its versatile metabolome. While primarily described as a plant-promotive, nonpathogenic microbe, it has been associated with rare cases of bacteremia in humans (1). This effort sought to examine the feasibility of recovering viable bacteriophages in local freshwater resources with ice cover sampling, using *P. fluorescens* as the target host organism. The bacteriophage UNO-G1W1 was extracted by auger drill ice sampling from Lake Wanhoo at global positioning system (GPS) coordinates 41.251597 N, -96.612549 W (Wahoo, Nebraska, USA) on 20 December 2016. The thawed sample was filtered (0.22 µm), diluted, and added to log phase host bacteria (*P. fluorescens* Migula strain; ATCC 27663) at room temperature for 20 minutes for viral invasion. These bacteria were subsequently added to 0.7% agarose (45°C), plated and incubated overnight at 26°C. A single plaque was isolated and purified twice, and morphology was stable. A high titer stock was used to check purity via transmission electron microscopy (TEM) and to isolate genomic DNA by phenol-chloroform extraction as described previously (2).

A short-read library was prepared with the Illumina Nextera XT kit and sequenced on a Hi-Seq 2500 with 151 bp paired-end reads. FastQC 0.11.9 was used to verify read quality (<http://www.bioinformatics.bsrc.ac.uk/projects/fastqc/>). A long-read library was prepared with the Oxford Nanopore Technologies ligation kit (SQK-LSK109). Briefly, to enrich for and improve recovery of long fragments following adapter ligation, a Long Fragment Buffer was used for washes, and incubation with Elution Buffer was carried out at 37°C. Sequencing was performed with a MinION Mk1B device using a FLO-MIN106D (R9.4.1) flow cell and MinKNOW software (v23.11.7). Live super-accurate basecalling was performed by Dorado (v7.2.13). A sequencing summary is provided in Table 1.

To construct a complete genome, long reads were first filtered with a Q-score threshold of 10, and adapters were trimmed using PoreChop\_ABI v0.5.0 (7). The direct terminal repeat (DTR)-phage-pipeline (<https://github.com/nanoporetech/DTR-phage-pipeline>, original release acc. 15 March 2024; default parameters, except "MEDAKA:model" as "r941\_min\_sup\_g507") was used to generate a polished long-read

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

Address correspondence to Thomas T. Schulze, [thomasschulze@unomaha.edu](mailto:thomasschulze@unomaha.edu).

Thomas T. Schulze and Andrew J. Neville contributed equally to this article. The author order was determined on the basis of seniority.

The authors declare no conflict of interest.

See the funding table on p. 3.

**Received** 12 April 2024

**Accepted** 10 May 2024

**Published** 7 June 2024

Copyright © 2024 Schulze 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 Sequencing summary

Read library	Raw/total reads	Mean Q-score <sup>a</sup>	Total bases	<i>N</i> <sub>50</sub> (bases) <sup>d</sup>	Mean/median read length (bases) <sup>d</sup>	Estimated seq. depth <sup>b</sup>	Mean depth of coverage <sup>c</sup>
Illumina R1	543,186	37.4	70.9 Mb	--	--	719×	1,349×
Illumina R2	543,186	36.7	70.9 Mb	--	--	719×	
Nanopore trimmed	451,635	Mean: 7.9	7.44 Gb	31,888	Mean: 16,467	75,448×	64,869×
All reads		Median: 10.0			Median: 9,726		
Nanopore assembly	230,572	Mean: 12.3	4.08 Gb	33,491	Mean: 17,714	41,435×	40,604×
Reads (trimmed Q ≥ 10)		Median: 13.2			Median: 10,888		

<sup>a</sup>Illumina Q-scores calculated with BioPython and NumPy, Nanopore Q-score by NanoPlot v1.42.0 (3).

<sup>b</sup>Total bases divided by complete genome size (98,572 bp).

<sup>c</sup>Average depth by position calculated using SAMtools (v1.19.2, 4) Burrow-Wheeler Aligner (BWA, v0.7.17-r1188, 5) used to map Illumina reads. Minimap2 (v2.28-r1209, 6) used to map Nanopore reads.

<sup>d</sup>Fields with "--" indicate a value that was not applicable or calculated.

consensus genome (8). This sequence was further polished with Polypolish [v0.6.0 (9)], using Illumina short reads which revised a single single nucleotide polymorphism (SNP) in a low complexity region. The provided genome has a total length of 98,572 bp and guanine-cytosine (GC) content of 48.3%. Annotation was performed using Pharokka v1.7.1 and corresponding v1.4.0 database (10), which predicted 209 genomic features. These included 18 tRNAs and 132 predicted as hypothetical or proteins of unknown function (11, 12).

The DTR-phage-pipeline predicted the presence of a 544 bp DTR, but no clear evidence of circular permutation was detected. Whole genome BLASTn (22 March 2024) identified *P. fluorescens* bacteriophage phiPsa374 (accession: [NC\\_023601.2](https://ncbi.nlm.nih.gov/nucl/NC_023601.2)) as the closest relative with 84.62% sequence identity and 63% coverage. Based on identity and recent taxonomic revisions (13), this novel bacteriophage belongs to the *Caudoviricetes* class (NCBI: txid2731619) and *Otagovirus* genus (NCBI: txid2560197).

## ACKNOWLEDGMENTS

Illumina next-generation sequencing was performed at the DNA Sequencing Core Facility located at the University of Nebraska Medical Center in Omaha, NE. The UNMC DNA Sequencing Core Facility receives partial support from the Nebraska Research Network In Functional Genomics, NE-INBRE P20GM103427, The Molecular Biology of Neurosensory Systems CoBRE P30GM110768, The Fred and Pamela Buffett Cancer Center - P30CA036727, and The Center for Root and Rhizobiome Innovation (CRR) 36-5150-2085-20. This work utilized the Holland Computing Center of the University of Nebraska, which receives support from the Nebraska Research Initiative. We would like to thank John M. Eppley, University of Hawai'i at Mānoa, for providing technical expertise with the DTR-phage-pipeline.

## AUTHOR AFFILIATIONS

<sup>1</sup>Department of Biology, University of Nebraska at Omaha, Omaha, Nebraska, USA

<sup>2</sup>Department of Pathology, Microbiology, and Immunology, University of Nebraska Medical Center, Omaha, Nebraska, USA

## PRESENT ADDRESS

Mackenzie E. Conrin, Department of Environmental Health and Safety, University of Nebraska Medical Center, Omaha, Nebraska, USA

## AUTHOR ORCIDs

Thomas T. Schulze  <http://orcid.org/0000-0002-3597-5727>

Andrew J. Neville  <http://orcid.org/0000-0001-6237-9640>

Harim I. Won  <http://orcid.org/0000-0002-7182-8599>

Paul H. Davis  <http://orcid.org/0000-0002-0526-3841>

## FUNDING

Funder	Grant(s)	Author(s)
UNMC   Fred and Pamela Buffett Cancer Center (Fred & Pamela Buffett Cancer Center)	P30CA036727	Paul H. Davis
NU   Center for Root and Rhizobiome Innovation, University of Nebraska-Lincoln (CRR)	36-5150-2085-20	Paul H. Davis
HHS   NIH   National Institute of General Medical Sciences (NIGMS)	P30GM110768	Paul H. Davis
HHS   NIH   National Institute of General Medical Sciences (NIGMS)	P20GM103427	Paul H. Davis

## DATA AVAILABILITY

The complete genome sequence is available on NCBI GenBank as accession [PP551948](#). The version described in this paper is the 2nd version, [PP551948.2](#). Illumina and Nanopore raw reads are available from SRA accessions [SRR28523409](#) and [SRR28523408](#), respectively.

## REFERENCES

1. Scales BS, Dickson RP, LiPuma JJ, Huffnagle GB. 2014. Microbiology, genomics, and clinical significance of the *Pseudomonas fluorescens* species complex, an unappreciated colonizer of humans. *Clin Microbiol Rev* 27:927–948. <https://doi.org/10.1128/CMR.00044-14>
2. Conrin M. 2020. Characterization of caudovirales isolated from freshwater samples against *Pseudomonas fluorescens* and host range testing against *Pseudomonas aeruginosa* Thesis, University of Nebraska at Omaha
3. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
4. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, Li H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* 10:giab008. <https://doi.org/10.1093/gigascience/giab008>
5. Li H, Durbin R. 2009. Fast and accurate short read alignment with burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
6. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
7. Bonenfant Q, Noé L, Touzet H. 2023. Porechop\_ABI: discovering unknown adapters in Oxford Nanopore technology sequencing reads for downstream trimming. *Bioinform Adv* 3:vbac085. <https://doi.org/10.1093/bioadv/vbac085>
8. Beaulaurier J, Luo E, Eppley JM, Uyl PD, Dai X, Burger A, Turner DJ, Pendelton M, Juul S, Harrington E, DeLong EF. 2020. Assembly-free single-molecule sequencing recovers complete virus genomes from natural microbial communities. *Genome Res* 30:437–446. <https://doi.org/10.1101/gr.251686.119>
9. Wick RR, Holt KE. 2022. Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 18:e1009802. <https://doi.org/10.1371/journal.pcbi.1009802>
10. Bouras G, Nepal R, Houtak G, Psaltis AJ, Wormald PJ, Vreugde S. 2023. Pharokka: a fast scalable bacteriophage annotation tool. *Bioinformatics* 39:btac776. <https://doi.org/10.1093/bioinformatics/btac776>
11. Shen A, Millard A. 2021. Phage genome annotation: where to begin and end. *PHAGE (New Rochelle)* 2:183–193. <https://doi.org/10.1089/phage.2021.0015>
12. Turner D, Adriaenssens EM, Tolstoy I, Kropinski AM. 2021. Phage annotation guide: guidelines for assembly and high-quality annotation. *PHAGE (New Rochelle)* 2:170–182. <https://doi.org/10.1089/phage.2021.0013>
13. Turner Dann, 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>