Genome Sequences of 17 Diverse *Pseudomonas aeruginosa* Phages

Ross A. Campbell, Jason Farlow, Helen R. Freyberger, Yunxiu He, Amanda M. Ward, Damon W. Ellison, Derese Getnet, Brett E. Swierczewski, Mikeljon P. Nikolich, Andrey A. Filippov

*The Geneva Foundation, Fort Detrick, Maryland, USA*

Wound Infections Department, Bacterial Diseases Branch, Walter Reed Army Institute of Research, Silver Spring, Maryland, USA

Farlow Scientific Consulting, LLC, Lewiston, Utah, USA

**ABSTRACT** Here, we describe genome sequences of 17 *Pseudomonas aeruginosa* phages, including therapeutic candidates. They belong to the families *Myoviridae*, *Podoviridae*, and *Siphoviridae* and six different genera. The genomes ranged in size from 42,788 to 88,805 bp, with G+C contents of 52.5% to 64.3% and numbers of coding sequences from 58 to 179.

Phages are attracting increasing attention as alternative antibacterial agents due to the wide spread of multidrug-resistant (MDR) infections. Phages have been successfully used against *Pseudomonas aeruginosa* infections in humans as expanded-access treatment and even in controlled clinical trials but are preferable to use as phage cocktails to cover multiple clinical isolates (1). To develop broad host range therapeutic cocktails against MDR *P. aeruginosa*, we have recently isolated 10 lytic phages and reported their whole genomes (2). Here, we describe the complete genome sequences of 17 additional diverse *P. aeruginosa* phages (Table 1), of which many also have potential for use in durable fixed therapeutic cocktails.

The main source of these novel phages was raw sewage collected in Washington, DC, except for EPa38 and EPa39 (from lake water in Frederick County, MD), EPa40 (from soil in Montgomery County, MD), and EPa41 (from chicken feces collected in Montgomery County). Several diverse *P. aeruginosa* strains were used for enrichment (Table 1). Each phage was purified by three rounds of growth from individual plaques, propagated on the enrichment strain in broth, and concentrated by high-speed centrifugation (3). After the removal of host RNA and DNA from lysates using RNase A and DNase, phage DNA was purified by proteinase K and SDS treatment, phenol-chloroform extraction, and precipitation with salt and ethanol (3).

Sequencing libraries were prepared using a Nextera XT DNA library preparation kit (Illumina, San Diego, CA). Validation and quantification of sequencing libraries were done with a TapeStation D5000 kit (Agilent Technologies, Inc., Santa Clara, CA) and an Invitrogen Qubit double-stranded DNA (dsDNA) broad-range (BR) assay kit (Thermo Fisher Scientific, Waltham, MA). The libraries were purified using AMPure XP beads (Beckman Coulter Diagnostics, Brea, CA) and sequenced with a 600-cycle MiSeq reagent kit v3 on an Illumina MiSeq instrument that produced 300-bp paired-end reads. FastQC 0.11.5 ([https://www.bioinformatics.babraham.ac.uk/projects/fastqc/](https://www.bioinformatics.babraham.ac.uk/projects/fastqc/)) was used for read quality control. Raw reads (Table 1) were trimmed using Geneious Prime 2019.2.3 with default parameters, with the exception of EPa18 reads which were trimmed with fastp using default parameters (4), and phage genomes were *de novo* assembled using PATRIC genome assembly service (5), also with default parameters. Phage genomes were annotated on the RAST server (6), and nucleic acid sequence similarity searches were carried out using default parameters in BLASTn (7).

Phage genomes varied in length from 42,788 (EPa40) to 88,805 nucleotides (EPa26), with G+C contents ranging between 52.5% (EPa4) and 64.3% (EPa38). The genomes contained...
| Phage name | Enrichment strain | Family | Genus | Genome length (bp) | G+C content (%) | No. of CDSs<sup>a</sup> | Genome coverage (x) | No. of raw reads | GenBank accession no. | SRA accession no. |
|------------|------------------|--------|-------|-------------------|-----------------|----------------|-------------------|----------------|-------------------|------------------|
| EPa4       | PAO1             | Podoviridae | Bruynoghevirus | 45,439           | 52.5            | 80               | 1,356.7           | 271,576         | MT118288          | SRR13222877      |
| EPa7       | PAO1             | Myoviridae | Phunavirus     | 65,629           | 55.5            | 96               | 2,628.8           | 773,481         | MT118289          | SRR13196079      |
| EPa10      | PAO1             | Myoviridae | Phunavirus     | 66,774           | 55.7            | 104              | 5,998.2           | 1,702,554       | MT118290          | SRR13196078      |
| EPa12      | MRSN 1680        | Myoviridae | Phunavirus     | 66,520           | 55.7            | 102              | 2,005.9           | 540,861         | MT118291          | SRR13196070      |
| EPa13      | MRSN 1680        | Myoviridae | Phunavirus     | 65,680           | 55.5            | 96               | 1,014.8           | 310,304         | MT118292          | SRR13196069      |
| EPa14      | MRSN 1680        | Myoviridae | Phunavirus     | 65,797           | 55.3            | 107              | 3,906.4           | 1,089,559       | MT118293          | SRR13196068      |
| EPa16      | MRSN 1680        | Myoviridae | Nankokuvirus   | 88,727           | 54.8            | 178              | 5,969.1           | 2,354,906       | MT118294          | SRR13196067      |
| EPa18      | MRSN 3705        | Myoviridae | Nankokuvirus   | 88,109           | 54.7            | 175              | 1,021.0           | 378,783         | MT118295          | SRR13196066      |
| EPa20      | MRSN 1680        | Myoviridae | Phunavirus     | 66,505           | 55.6            | 105              | 992.8            | 274,774         | MT118297          | SRR13196064      |
| EPa21      | MRSN 1680        | Myoviridae | Phunavirus     | 66,764           | 55.6            | 101              | 774.4            | 214,388         | MT118298          | SRR13196063      |
| EPa25      | MRSN 1680        | Myoviridae | Phunavirus     | 66,811           | 55.6            | 101              | 1,497.2           | 436,993         | MT118299          | SRR13196077      |
| EPa26      | PAO1             | Myoviridae | Nankokuvirus   | 88,805           | 54.8            | 179              | 3,106.9           | 1,245,653       | MT118300          | SRR13196076      |
| EPa33      | PAO1             | Podoviridae | Hollowayvirus | 64,021           | 63.5            | 80               | 2,991.0           | 800,353         | MT118301          | SRR13196075      |
| EPa38      | PAO1             | Siphoviridae | Yuavirus      | 61,775           | 64.3            | 96               | 1,338.4           | 328,359         | MT118302          | SRR13196074      |
| EPa39      | PAO1             | Myoviridae | Phunavirus     | 66,708           | 54.9            | 102              | 2,130.1           | 619,285         | MT118303          | SRR13196073      |
| EPa40      | ATCC 10145       | Siphoviridae | Septimatrevirus | 42,788           | 53.2            | 58               | 1,987.6           | 357,786         | MT118304          | SRR13196072      |
| EPa41      | ATCC 10145       | Siphoviridae | Septimatrevirus | 43,258           | 53.2            | 60               | 3,517.7           | 629,579         | MT118305          | SRR13196071      |

<sup>a</sup>CDSs, protein-coding sequences.
58 to 179 coding sequences (Table 1). The phages were classified into the families Myoviridae (genera Pbunavirus and Nankokuvirus), Podoviridae (genera Bruynoghevirus and Hollowayvirus), and Siphoviridae (genera Septimatrevirus, and Yuavirus) based on DNA sequence identity to characterize phages using a threshold of >50% for placement in the same genus (8). Pbunavirus phages comprised the most numerous group, including nine representatives, namely, EPa7, EPa10, EPa12, EPa13, EPa14, EPa20, EPa21, EPa25, and EPa39. BLASTn and BLASTp analyses showed no significant similarity to genes and proteins related to the lysogenic life style or gene transfer, including integrases, recombinases, transposases, excisionases, and repressors of the lytic cycle, or any bacterial genes or proteins. A similar pattern was found for Nankokuvirus phages EPa16, EPa18, and EPa26. Such a strictly lytic nature is typical of myophages from the genera Pbunavirus (2, 9) and Nankokuvirus (2, 10) that makes them safe and potent therapeutic phages.

Only two phages were the members of the family Podoviridae, namely, EPa4 and EPa33. BLASTn sequence comparisons showed that phage EPa4, like EPa1 and EPa2 isolated in our laboratory earlier (2), belongs to the genus Bruynoghevirus and shows 96.5% identity to lytic phage LUZ24 (GenBank accession number AM910650) (11). Genomic analysis showed that EPa4, similar to EPa1, EPa2, and LUZ24, lacks genes typical for temperate phages, suggesting that they are strictly virulent and potential therapeutic candidates. As opposed to EPa4, podophage EPa33 belonged to the genus Hollowayvirus, which includes a large number of temperate phages similar to F116, the generalized transducing phage (12). BLASTn analysis revealed multiple extensive regions of EPa33 genome identity to P. aeruginosa chromosomal DNA (e.g., GenBank accession numbers CP030075, CP039988, and CP015377, and many others), suggesting that EPa33 is also a temperate phage and potential transducer and cannot be used for therapy.

Three Siphoviridae phages included the members of two different genera. EPa40 and EPa41 (genus Septimatrevirus) showed no signs of temperate phages and thus appear to be obligately lytic phages and suitable candidates for phage therapy, as previously shown for this group by other authors (13). Phage EPa38 (genus Yuavirus), like EPa5 and EPa43 (genus Abidjanvirus) isolated by our team earlier (2), encoded putative proteins designated by others as an integrase and a repressor (ORF22 and ORF21 in the Ab18 genome, GenBank accession number LN610577) (14). Our previous analysis identified only primase-related domains and no integrase-associated domains in the ORF22 product in EPa5, EPa43, and other Abidjanvirus phages (2), which also applies to EPa38 and other Yuavirus phages.

Therefore, we report the whole-genome sequences of 17 P. aeruginosa phages that belong to 3 families and 6 genera. Fifteen of them (12 myophages, as well as Bruynoghevirus phage EPa4 and Septimatrevirus phages EPa40 and EPa41) appear to be strictly virulent phages and safe therapeutic candidates, while more research is needed to clear a siphophage EPa38 for therapeutic use, and EPa33 is a temperate and potentially transducing phage unsuitable as a therapeutic agent.

**Data availability.** The 17 complete phage genome sequences were deposited in GenBank and the NCBI Sequence Read Archive (SRA) under the accession numbers listed in Table 1.

**ACKNOWLEDGMENTS**

This material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The opinions or assertions contained herein are the private views of the authors, and are not to be construed as official, or as reflecting true views of the Department of the Army or the Department of Defense.

This study was supported by the Military Infectious Diseases Research Program, grant W0252_18_WR, and Peer Reviewed Medical Research Program, Focused Program Award PR182667.

The Multidrug Resistant Organism Repository and Surveillance Network (MRSN) at the Walter Reed Army Institute of Research provided P. aeruginosa clinical isolates used for phage isolation and host range testing. Wiriya Rutvisuttinunt, Tao Li, and Richard Jarman (Viral Diseases Branch, Walter Reed Army Institute of Research) are greatly appreciated for...
conducting DNA sequencing. We are grateful to Nathan Brown (University of Leicester, Leicester, United Kingdom) for his help in phage genome analysis.

REFERENCES

1. Nikolich MP, Filippov AA. 2020. Bacteriophage therapy: developments and directions. Antibiotics (Basel) 9:135. https://doi.org/10.3390/antibiotics9030135.

2. Farlow J, Freyberger HR, He Y, Ward AM, Rutvisutinunt W, Li T, Campbell R, Jacobs AC, Nikolich MP, Filippov AA. 2020. Complete genome sequences of 10 phages lytic against multidrug-resistant Pseudomonas aeruginosa. Microbiol Resour Announc 9:e00503-20. https://doi.org/10.1128/MRA.00503-20.

3. Sergueev KV, Filippov AA, Farlow J, Su W, Kvachadze L, Balarjishvili N, Kutateladze M, Nikolich MP. 2019. Correlation of host range expansion of therapeutic bacteriophage Sb-1 with allele state at a hypervariable repeat locus. Appl Environ Microbiol 85:e0109-19. https://doi.org/10.1128/AEM.01209-19.

4. Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

5. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Dietrich EM, Dovernick JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg KE, Olsen GJ, Murphy-olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Stevens RL. 2017. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. Nucleic Acids Res 45: D535–D542. https://doi.org/10.1093/nar/gkw1017.

6. Aziz RK, Bartels D, Best AA, Dejongh M, Disz T, Edwards RA, Formsmo K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil KK, Paarmann D, Paczian T, Parrillo B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.

7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/0022-2836(90)80360-2.

8. Adriamssens EM, Brister JR. 2017. How to name and classify your phage: an informal guide. Viruses 9:70. https://doi.org/10.3390/v9040070.

9. Lavigne R, Darius P, Summer EJ, Seto D, Mahadevan P, Nilsson AS, Ackermann HW, Kropinski AM. 2009. Classification of Myoviridae bacteriophages using protein sequence similarity. BMC Microbiol 9:224. https://doi.org/10.1186/1471-2180-9-224.

10. Morello E, Saussereau E, Maura D, Huerre M, Touqui L, Debarbieux L. 2011. Pulmonary bacteriophage therapy on Pseudomonas aeruginosa cystic fibrosis strains: first steps towards treatment and prevention. PLoS One 6:e16963. https://doi.org/10.1371/journal.pone.0016963.

11. Ceyssens P-J, Hertveldt K, Ackermann H-W, Noben J-P, Demeke M, Volkaert G, Lavigne R. 2008. The intron-containing genome of the lytic Pseudomonas phage LUZ24 resembles the temperate phage PaP3. Virology 377:233–238. https://doi.org/10.1016/j.virol.2008.04.038.

12. Byrne M, Kropinski AM. 2005. The genome of the Pseudomonas aeruginosa generalised transducing bacteriophage F116. Gene 346:187–194. https://doi.org/10.1016/j.gene.2004.11.001.

13. Karumidze N, Thomas JA, Kvadadze N, Goderdzhishvili M, Hakala KW, Weintraub ST, Alavizde Z, Hardies SC. 2012. Characterization of lytic Pseudomonas aeruginosa bacteriophages via biological properties and genomic sequences. Appl Microbiol Biotechnol 94:1609–1617. https://doi.org/10.1007/s00253-012-4119-8.

14. Essoh C, Latino L, Midoux C, Blouin Y, Loukou G, Nguetta S-PA, Lathr0 S, Cablanmian A, Kouassi AK, Vergnaud G, Pourcel C. 2015. Investigation of a large collection of Pseudomonas aeruginosa bacteriophages collected from a single environmental source in Abidjan, Cote d’Ivoire. PLoS One 10:e0130548. https://doi.org/10.1371/journal.pone.0130548.