Complete Genome Sequence of *Escherichia coli* O157:H7 Phage PhiG17

Muyiwa A. Akindolire,a Bukola R. Aremu,a,b ©Collins N. Atebaa,b

a Department of Microbiology, School of Biological Sciences, Faculty of Natural and Agricultural Sciences, North-West University, Mmabatho, South Africa
b Food Security and Safety Niche Area, North-West University, Mmabatho, South Africa

**ABSTRACT** Here, we announce the complete genome sequence of bacteriophage PhiG17, which is virulent to *Escherichia coli* O157:H7 strains and was isolated from cattle feces in the North West province of South Africa. This report presents the major genetic features of the phage PhiG17 based on its whole-genome sequence.

*Escherichia coli* O157:H7 is an important foodborne pathogen of significant public health concern worldwide (1, 2). In order to prevent human exposure to this pathogen, virulent bacteriophages have been suggested as a promising tool that can be used in reducing *E. coli* O157:H7 in both farm animals and food products (3–7). In the application of these promising biocontrol and therapeutic agents, it is important to screen the genomes of potential candidates to ensure their safety and efficacy (8). In this report, we present for the first time the complete genome sequence of a PhiG17 phage isolated from cattle feces in South Africa. The phage was capable of infecting virulent environmental Shiga toxin-producing *E. coli* (STEC) O157:H7 strains that were characterized in a previous report (9). Transmission electron micrographs of phage particles revealed an icosahedral head (75 ± 3.6 nm) and a short tail (23 ± 3.1 nm), as shown in Fig. 1, which are typical of the Podoviridae family (10).

Phage G17 genomic DNA was extracted and purified using a Norgen phage DNA isolation kit (Norgen Biotek Corp., Ontario, Canada) according to the manufacturer’s instructions. Purified phage DNA was sequenced at Inqaba Biotechnical Industries (Pty) Ltd. in South Africa, where the phage DNA was fragmented using an ultrasonication approach (Covaris). DNA fragments were size-selected (300 to 800 bp) using AMPure XP beads, the fragments were end-repaired, and Illumina-specific adapter sequences were ligated to each fragment. The sample was indexed, and a second size selection was performed. The sample was then quantified using a fluorometric method, diluted to a standard concentration (4 nM), and then sequenced on an Illumina MiSeq v3 platform (600-cycle kit) following a standard protocol as described by the manufacturer. A total of 50 Mb of data (2 × 300-bp long paired-end reads) were produced for the sample, and a total of 154,330 reads with an average read length of 274.9 bp were obtained, filtered, and trimmed using Trimmmomatic v0.36 with default parameters (11). Further quality checking was done with FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc), and de novo assembly of raw data was carried out using SPAdes v3.11.1 genome assembler software; the result was evaluated with QUAST (12).

The complete genome sequence of phage G17 revealed a genome size of 68,270 bp assembled in a single contig with an *N₅₀* value of 68,270 bp and an average G+C content of 43.5%. Genes and coding sequences in the genome were predicted and annotated using the PHAge Search Tool (PHAST) (13) and Rapid Annotations using Subsystems Technology (RAST) v2.0 server (14) online analysis tools. Predicted proteins in the genome were further annotated by BLAST against the NCBI nonredundant GenBank database (15). The PhiG17
The genome contained a total of 78 coding sequences ranging in length from 114 bp to 10,353 bp and 1 isoleucine-tRNA as predicted by ARAGORN (16) and tRNAscan-SE (17). The genome lacks genes encoding known antibiotic resistance as determined by ResFinder v3.0 (18). Genes encoding toxins and those implicated in transduction were absent from the genome of phage PhiG17. Nucleotide BLAST search of the genome revealed 94% nucleotide similarity with APEC7 (GenBank accession number KF562340) and 95% with APEC5 (GenBank accession number KF192075), both of which are members of the newly assigned G7cvirus. Further CoreGenes analysis (18) of the genome revealed a significant similarity to the same recently assigned G7cvirus, notably Escherichia phage vB EcoP PhAPEC7 (GenBank accession number KF562340). Members of this recent group are classified as belonging to the Podoviridae family and the G7cvirus genus and are considered to be safe biocontrol candidates (19, 20). Thus, Escherichia phage G17 is suggested to be a member of the G7cviruses.

**Data availability.** The genome sequence of *Escherichia* phage PhiG17 has been deposited in the NCBI GenBank database under the accession number MH358458. The raw sequences are available in the NCBI SRA database under the accession number SRP159624.

**ACKNOWLEDGMENTS**

This work is based on research supported in part by the National Research Foundation of South Africa (NRF) (grant UID numbers 89228 and 98983). The study was also supported by the North-West University, South Africa.

Our profound gratitude goes to A. Jordan of the Laboratory for Electron Microscopy, Chemical Resource Beneficiation, North-West University, Potchefstroom Campus, for assisting with electron microscopy. We also acknowledge L. Chidamba of the University of Pretoria, South Africa, for assistance with the genomic analysis of phage PhiG17.

**REFERENCES**

1. Maruzumi M, Morita M, Matsuoka Y, Uekawa A, Nakamura T, Fuji K. 2005. Mass food poisoning caused by beef offal contaminated by *Escherichia coli* O157. Jpn J Infect Dis 58:397.

2. Sugiyama A, Iwade Y, Akachi S, Nakano Y, Matsuno Y, Yano T, Yamauchi A, Nakayama O, Sakai H, Yamamoto K, Nagasaka Y, Nakano T, Ihara T, Kamiya H. 2005. An outbreak of Shiga toxin-producing *Escherichia coli* O157:H7 in a nursery school in Mie Prefecture. Jpn J Infect Dis 58:398.

3. O’Flynn G, Ross RP, Fitzgerald GF, Coffey A. 2004. Evaluation of a cocktail of three bacteriophages for biocontrol of *Escherichia coli* O157:H7. Appl Environ Microbiol 70:3417–3424. https://doi.org/10.1128/AEM.70.6.3417-3424.2004.

4. Sheng H, Knecht HJ, Kudva IT, Hovde CJ. 2006. Application of bacteriophages to control intestinal *Escherichia coli* O157:H7 levels in ruminants. Appl Environ Microbiol 72:3359–3366. https://doi.org/10.1128/AEM.00099-06.

5. Hagens S, Loessner M. 2007. Application of bacteriophages for detection
and control of foodborne pathogens. Appl Microbiol Biotechnol 76: 513–519. https://doi.org/10.1007/s00253-007-1031-8.
6. Niu YD, Stanford K, Kropinski AM, Ackermann HW, Johnson RP, She YM, Ahmed R, Villegas A, McAllister TA. 2012. Genomic, proteomic and physiological characterization of a T5-like bacteriophage for control of Shiga toxin-producing Escherichia coli O157:H7. PLoS One 7:e34585. https://doi.org/10.1371/journal.pone.0034585.
7. Zinno P, Devirgiliis C, Ercolini D, Ongeng D, Mauriello G. 2014. Bacteriophage P22 to challenge Salmonella in foods. Int J Food Microbiol 191:69 –74. https://doi.org/10.1016/j.ijfoodmicro.2014.08.037.
8. Abedon ST. 2017. Information phage therapy research should report. Pharmaceuticals (Basel) 10:2. https://doi.org/10.3390/ph10020043.
9. Ackermann HW. 2003. Bacteriophage observations and evolution. Res Microbiol 154:245–251. https://doi.org/10.1016/S0923-2508(03)00067-6.
10. Bolger AM, Lohse M, Usadel B. 2014. Trimmmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114 –2120. https://doi.org/10.1093/bioinformatics/btu170.
11. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, 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.
12. Zhou Y, Liang Y, Lynch KH, Dennis J, Wishart DS. 2011. PHAST: a fast phage search tool. Nucleic Acids Res 39:W347–W352. https://doi.org/10.1093/nar/gkr485.
13. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Forma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LX, Paarmann D, Paczian T, Parrello 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.
14. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402. https://doi.org/10.1093/nar/25.17.3389.
15. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16. https://doi.org/10.1093/nar/gkh152.
16. Lowe TM, Chan PP. 2016. tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44: W54–W57. https://doi.org/10.1093/nar/gkw413.
17. Kulikov E, Kropinski AM, Golomidova A, Lingohr E, Govorun V, Serebryakova M, Prokhorov N, Letarov M, Manykin A, Strotskaya A, Letarov A. 2012. Isolation and characterization of a novel indigenous intestinal N4-related coliphage vB_EcoP_G7C. Virology 426:93–99. https://doi.org/10.1016/j.virol.2012.01.027.
18. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640–2644. https://doi.org/10.1093/jac/dks261.
19. Tsonos J, Oosterik LH, Tuntufye HN, Klumpp J, Butaye P, De Greve H, Hernalsteens JP, Lavigne R, Goddeears BM. 2014. A cocktail of in vitro efficient phages is not a guarantee for in vivo therapeutic results against avian colibacillosis. Vet Microbiol 171:470–479. https://doi.org/10.1016/j.vetmic.2013.10.021.