AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of the Virulent *Klebsiella pneumoniae* Phage Geezett Infecting Multidrug-Resistant Clinical Strains

Belinda Loh,[a] Liwei Zhang,[b] Xiaoting Hua,[c] Yunsong Yu,[c] Long Ma,[b] Xiaoqing Wang,[b,d] Prasanth Manohar,[b] Ramesh Nachimuthu,[e] Willames M.B.S. Martins,[f] Mark A. Toleman,[f] Sebastian Leptihn[b,c,g]

[a]Department of Biological Sciences, Xi'an Jiaotong-Liverpool University, Suzhou, China
[b]Zhejiang University-University of Edinburgh (ZJU-UoE) Institute, Zhejiang University, International Campus, Haining, Zhejiang, China
[c]Department of Infectious Diseases, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China
[d]Medical School, Lishui University, Lishui, China
[e]Antibiotic Resistance and Phage Therapy Laboratory, School of Biosciences and Technology, Vellore Institute of Technology (VIT), Vellore, Tamil Nadu, India
[f]Department of Medical Microbiology, Division of Infection and Immunity, Cardiff University, Cardiff, United Kingdom
[g]College of Medicine & Veterinary Medicine, University of Edinburgh Medical School, Edinburgh, United Kingdom

ABSTRACT Geezett was isolated from hospital sewage in Hangzhou, China, and exhibits lytic activity against clinical isolates of the nosocomial pathogen *Klebsiella pneumoniae*. The bacteriophage is a myovirus and has a double-stranded DNA (dsDNA) genome 50,707 bp long, containing 79 open reading frames (ORFs).

In many countries, *Klebsiella pneumoniae* is a leading cause of hospital-acquired infections (1), which include skin and soft tissue infection, infections of the urinary tract, and also life-threatening bloodstream infections and pneumonia. According to the World Health Organization, the emergence of *K. pneumoniae* strains resistant to carbapenems and third-generation cephalosporins represents an urgent need for development of new antimicrobial agents such as therapeutic phages (2, 3). Phage Geezett was isolated from sewage water obtained from the Sir Run Run Shaw Hospital in Hangzhou, China, using an enrichment culture of the clinical multidrug-resistant *K. pneumoniae* strain GZ-1. Characterized primarily by its head-tail structure and a long contractile tail, the phage morphology indicates that it belongs to the *Myoviridae* family of the order *Caudovirales* (Fig. 1).

Phages were obtained from single plaques and amplified prior to DNA extraction, as described previously (4). Phages in the filtrate were used for extracting DNA using the Biomed virus rapid DNA/RNA kit (Beijing, China). Sequencing libraries were prepared using the NEBNext Ultra II DNA library prep kit for Illumina. The phage genome was then sequenced using the Illumina HiSeq platform. A total of 4,404,022 raw reads were obtained with read lengths of 150 bp (paired-end format). The genome coverage was 7,867×. The short-read sequence data were assembled using Unicycler v.0.4.8 (5). Genome annotation and analysis were conducted using default settings via the CPT Galaxy (6) and Web Apollo (7) interfaces. Open reading frames (ORFs) were identified using GeneMarkS v.4.28 (8), GLIMMER v.3 (9), and MetaGeneAnnotator v.1.0 (10) and were manually validated using NCBI BLAST v.2.9.0 searches (11) against the NCBI nonredundant database, the Swiss-Prot database (12), and the Bacterial Virulence Factor Database (VFDB) (13). Default parameters were used unless stated otherwise.

Geezett has a double-stranded DNA (dsDNA) genome of 50,707 bp with a GC content of 48%. It is predicted to encode 79 proteins, of which 23 align to phage genes of known functions. These include proteins involved in transcription regulation, replication, DNA packaging, host lysis, and structural proteins. No genes were found to encode toxins or antibiotic

Editor Catherine Putonti, Loyola University Chicago

Address correspondence to Belinda Loh, belinda.loh@gmail.com, or Sebastian Leptihn, sebastian.leptihn@ed.ac.uk.

Received 29 July 2021
Accepted 6 November 2021
Published 2 December 2021

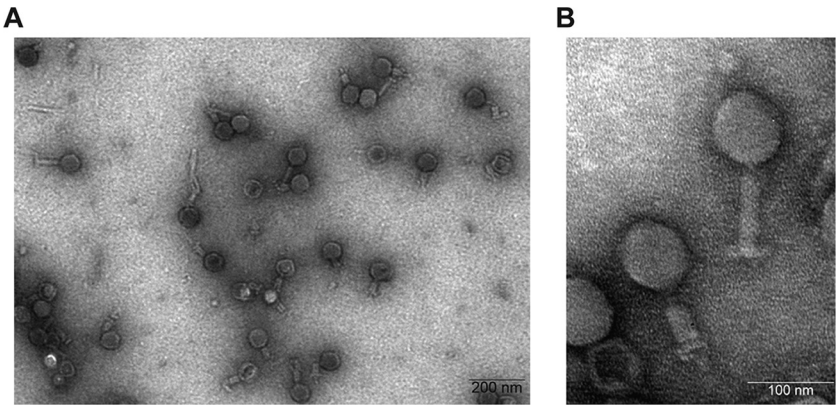

FIG 1 Transmission electron micrograph of *Klebsiella pneumoniae* phage Geezett. Phages were negative stained using 2% uranyl acetate. (A) Scale bar = 200 nm; (B) scale bar = 100 nm.

resistance factors. A search for related phages using nBLAST showed that Geezett is novel, with its closest relative being *Klebsiella* phage vB_KpnM_FZ14 (GenBank accession number MK521906.1), with a sequence coverage of only 66% (at 91.64% nucleotide sequence identity) (14). Several genes are dissimilar, such as the tail spike protein, which has only 68% amino acid sequence coverage with the corresponding protein in phage vB_KpnM_FZ14. The tail fiber protein of Geezett has no similarity with phage vB_KpnM_FZ14; it does, however, have 98% amino acid sequence coverage with that of *Klebsiella* phage vB_KpnP_KpV48 but with only a 45% amino acid sequence identity, indicating that Geezett might have a different host range compared to other *Klebsiella* phages.

Lysogeny-related genes and virulence factors were not found in Geezett during genome annotation. The phage is categorized as lytic using the program PhageAI (15), which might allow the deployment of Geezett as a therapeutic phage.

Data availability. The complete genome of Geezett has been deposited at GenBank under the accession number MZ504995.1 and the SRA accession number SRR15367659.

ACKNOWLEDGMENTS

We thank Andrew Millard (University of Leicester, UK) for his advice on phage genomics.

This work was supported by the National Natural Science Foundation of China (32011530116).

REFERENCES

1. De Oliveira DMP, Forde BM, Kidd TJ, Harris PNA, Schembri MA, Beatson SA, Paterson DL, Walker MJ. 13 May 2020. Antimicrobial resistance in ESKAPE pathogens. Clin Microbiol Rev 33:e00181-19. https://doi.org/10.1128/CMR.00181-19.
2. World Health Organization. 2020. Antimicrobial resistance. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance. Accessed 5 July 2021.
3. Leptihn S. 2019. Welcome back to the pre-penicillin era: why we desperately need new strategies in the battle against bacterial pathogens. Infect Microbes Dis 1:33. https://doi.org/10.1097/IM9.0000000000000009.
4. Loh B, Wang X, Hua X, Chook HW, Ma L, Zhang L, Manohar P, Jin Y, Leptihn S. 2021. Complete genome sequence of the lytic bacteriophage Phab24, which infects clinical strains of the nosocomial pathogen Acinetobacter baumannii. Microbiol Resour Announc 10:e00669-21. https://doi.org/10.1128/MRA.00669-21.
5. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
6. Cock PJ, Grüning BA, Paszkiewicz K, Pritchard L. 2013. Galaxy tools and workflows for sequence analysis with applications in molecular plant pathology. PeerJ 1:e167. https://doi.org/10.7717/peerj.167.
7. Lee E, Helt GA, Reese JT, Munoz-Torres MC, Childers CP, Buels RM, Stein L, Holmes IH, Elsik CG, Lewis SE. 2013. Web Apollo: a Web-based genomic annotation editing platform. Genome Biol 14:R93. https://doi.org/10.1186/gb-2013-14-8-r93.
8. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 29:2607–2618. https://doi.org/10.1093/nar/29.12.2607.
9. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res 27:4636–4641. https://doi.org/10.1093/nar/27.23.4636.
10. Noguchi H, Taniguchi T, Itoh T. 2008. MetaGeneAnnotator: detecting species-specific patterns of ribosomal binding site for precise gene prediction in anonymous prokaryotic and phage genomes. DNA Res 15:387–396. https://doi.org/10.1093/dnares/dsn027.
11. 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/S0022-2836(05)80360-2.

12. Bairoch A, Boeckmann B. 1994. The Swiss-Prot protein sequence data bank: current status. Nucleic Acids Res 22:3578–3580.
13. Liu B, Zheng D, Jin Q, Chen L, Yang J. 2019. VFDB 2019: a comparative pathogenomic platform with an interactive Web interface. Nucleic Acids Res 47:D687–D692. https://doi.org/10.1093/nar/gky1080.
14. Zurabov F, Zhilenkov E. 2019. Complete genome sequences of lytic polysaccharide-degrading Klebsiella pneumoniae bacteriophages vB_KpnS_FZ10, vB_KpnP_FZ12, vB_KpnM_FZ14, and vB_KpnS_FZ41. Microbiol Resour Announc 8:e00914-19. https://doi.org/10.1128/MRA.00914-19.
15. Tynecki P, Guziński A, Kazimierczak J, Jadczuk M, Dastych J, Onisko A. 2020. PhageAI—bacteriophage life cycle recognition with machine learning and natural language processing. bioRxiv https://doi.org/10.1101/2020.07.11.198606.