Complete Genome Sequences of a Diverse Group of 13 Propionibacterium acnes Bacteriophages Isolated from Urban Raw Sewage

Gustavo Ybazeta,a Jenna Graham,a Jelena Trifkovic,b Lyne Giroux,c Mazen Saleh,e Syed A. Sattar,b,d Reza Nokhbeh,c

aHealth Sciences North Research Institute, Sudbury, Ontario, Canada
bDepartment of Biochemistry, Microbiology and Immunology, Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada
cNorthern Ontario School of Medicine, Sudbury, Ontario, Canada
dCREM Co. Laboratories, Mississauga, Ontario, Canada
eDepartment of Biology, Faculty of Science, Laurentian University, Sudbury, Ontario, Canada

ABSTRACT We present complete genome sequences of 13 Propionibacterium acnes phages isolated from urban raw sewage. They belong to the family Siphoviridae, have genome sizes of 29,450.6 ± 256.5 nucleotides and G+C contents of 54.14% ± 0.22% and contain 42 to 45 coding DNA sequences (CDS). Genomic sequences of 9 of 13 phages were divergent by 6 to 10%, distinguishing them as species.

Restricted diversity among phages isolated from limited sources cannot represent the actual global diversity (1). Raw sewage from metropolitan areas with multiethnic communities is a rich and unique source of phages, reflecting the global diversity of their human populations. Such a source is crucial for obtaining a diverse collection of phages against a given pathogen for use in cocktails effective in clinical practice.

For this study, we sampled urban raw sewage every 1 to 2 months over a 6-month period from two sewage treatment facilities (Ottawa and Gatineau, Canada) and successfully obtained a genetically diverse population of phages against eight Propionibacterium acnes indicator strains.

The phage plaques, isolated by the agar overlay method, were purified and used for genomic DNA isolation. Our initial BamHI restriction analysis of the phages’ genomic DNA revealed multiple restriction banding patterns indicative of genetic diversity (not shown), which prompted further analysis by full-genome sequencing. The genomes of 13 representative Propionibacterium acnes phages from a larger collection were sequenced using Illumina HiSeq 2500 (with a paired-end 125-bp read length) and Nanopore MinION MK1 platforms. De novo hybrid assembly of genomes was performed using SPAdes v3.20.1 (2) and Unicycler v0.4.3 (3) as previously described (4). SeqKit v0.4.5, Sickle v1.33, and FastQC v0.11.5 were used for downsampling the Illumina reads to ~100×, trimming reads, and quality controlling, respectively (5–7). The MinION reads were demultiplexed by Metrichor, and Porechop v0.2.1 was used to remove the adapters (8). The MinION long reads were used to produce bridges in hybrid assembly using Illumina output to generate single contigs (4, 9).

The assembled single contigs for each phage were further analyzed for accuracy and quality using Bandage v0.8.1 (10) and annotated using Rapid Annotations using Subsystems Technology (RAST) v2.0 (11). The phages are cos type with 11-nucleotide 3’-overhang sequences at either end. They have an average genome size of 29,450.6 ± 256.5 nucleotides (Table 1) with an average G+C content of 54.14% ± 0.22% and 42 to 45 CDS.

Received 21 February 2018 Accepted 1 June 2018 Published 28 June 2018

Citation Ybazeta G, Graham J, Trifkovic J, Giroux L, Saleh M, Sattar SA, Nokhbeh R. 2018. Complete genome sequences of a diverse group of 13 Propionibacterium acnes bacteriophages isolated from urban raw sewage. Genome Announc 6:e00224-18. https://doi.org/10.1128/genomeA.00224-18.
BLASTN analysis of these genomes against the previously published genome sequences of similar phages in the GenBank nucleotide database showed 6 to 10% variation for 9 of the 13 phages. The majority of these sequence differences were mapped to the right arm coding strand for nonstructural proteins, highlighting significant diversity among the isolates (12).

The International Committee on Taxonomy of Viruses (ICTV) defines “species” on the basis of > 5% differences between two genomes (13; see also http://www.ictv.global/proposals-16/2016.034a-dB.A.v1.Pa6virus.pdf). Based on this criterion, 9 of the 13 phage genomes described here are considered distinct species. These include pa310 and pa59, pa6919-4 and pa9-6919-4, pa15 and pa615, pa33 and pa35, pa27, pa29399-1-D_1, pa29399-1-D_2, pa63, and pa28. In conclusion, successful isolation of a diverse group of phages using raw sewage from multiethnic urban centers, together with the use of multiple bacterial indicator strains, highlights the importance of the classical methods of constructing phage collections. This has important implications for strategizing development of phage cocktails for clinical use.

Accession number(s). The GenBank accession numbers of these phages are listed in Table 1.

**TABLE 1** Summary of the phage genomes deposited in GenBank and their accession numbers

| Accession no. | Phage name | Length (bp) | Avg coverage (x) | SD    |
|---------------|------------|-------------|------------------|-------|
| MG820632      | pa310      | 29,508      | 276.29           | 16.62 |
| MG820633      | pa59       | 29,507      | 267.26           | 16.35 |
| MG820634      | pa27       | 29,593      | 607.95           | 24.66 |
| MG820635      | pa29399-1-D_1 | 29,370   | 379.04           | 19.47 |
| MG820636      | pa29399-1-D_2 | 28,892   | 567.77           | 20.99 |
| MG820637      | pa63       | 29,446      | 691.46           | 26.30 |
| MG820638      | pa6919-4   | 29,784      | 241.80           | 15.55 |
| MG820639      | pa9-6919-4 | 29,784      | 686.45           | 26.20 |
| MG820640      | pa15       | 29,309      | 738.33           | 27.17 |
| MG820641      | pa615      | 29,307      | 619.74           | 24.89 |
| MG820642      | pa28       | 29,733      | 196.05           | 14.00 |
| MG820643      | pa35       | 29,491      | 762.90           | 27.62 |
| MG820644      | pa33       | 29,135      | 773.60           | 27.81 |

ACKNOWLEDGMENTS

We thank Susan Springthorpe, formerly of the Faculty of Medicine, University of Ottawa, for her initial input and support for this project.

This work was supported by Ontario AHSC AFP Innovation Fund–Northern Ontario Academic Medicine Association (NOAMA).

Illumina sequencing was performed at the McGill University and Génome Québec Innovation Centre, Montréal, QC, Canada. We acknowledge Alexandre Montpetit, the director of the Illumina sequencing platform at the Génome Québec Innovation Centre, and Annie Verville and Janick St-Cyr for their technical assistance. MinION sequencing and genome analysis were performed at the Bioinformatics & Genomics Office–Health Sciences North Research Institute.

REFERENCES

1. Marinelli LJ, Fitz-Gibbon S, Hayes C, Bowman C, Inkeles M, Loncaric A, Russell DA, Jacobs-Sera D, Cokus S, Pellegrini M, Kim J, Miller JF, Hatfull GF, Modlin RL. 2012. *Propionibacterium acnes* bacteriophages display limited genetic diversity and broad killing activity against bacterial skin isolates. mBio 3:e00279-12. https://doi.org/10.1128/mBio.00279-12.

2. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pyshkin AV, Pyzyk SN, 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.

3. Wick RR, Louise MJ, 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.

4. Ybazeta G, Douglas L, Graham J, Fraleigh NL, Yanal M, Perez J, Diaz-Mitoma F, Tille K, Nokhbeh R. 2017. Complete genome sequence of *Enterococcus thailandicus* strain a523 isolated from urban raw sewage. Genome Announc 5:e01298-17. https://doi.org/10.1128/genomeA.01298-17.

5. Wei S, Le S, Li Y, Hu F. 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS One 11:e0163910. https://doi.org/10.1371/journal.pone.0163910.

6. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
7. Joshi NA, Fass JN. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for fastQ files. https://github.com/najoshi/sickle.

8. Ryan W. 2017. Porechop. https://github.com/rrwick/Porechop.

9. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Completing bacterial genome assemblies with multiplex MinION sequencing. Microb Genom 3. https://doi.org/10.1099/mgen.0.000132.

10. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 31:3350–3352. https://doi.org/10.1093/bioinformatics/btv383.

11. Ross O, Olson R, Gordon D, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206–D214. https://doi.org/10.1093/nar/gkt1226.

12. Jared L, Yan R, Zhong Q, Ngo S, Bangayan NJ, Nguyen L, Lui T, Liu M, Erfe MC, Craft N, Tomida S, Li H. 2015. The diversity and host interactions of Propionibacterium Acnes bacteriophages on human skin. ISME J 9:2078–2093. https://doi.org/10.1038/ismej.2015.47.

13. Adams MJ, Lefkowitz EJ, King AMQ, Harrach B, Harrison RL, Knowles NJ, Kropinski AM, Krupovic M, Kuhn JH, Mushegian AR, Nibert M, Sabanadzovic S, Sanfaçon H, Siddell SG, Simmonds P, Varsani A, Zerbini FM, Gorbalenya AE, Davison AJ. 2017. Changes to taxonomy and the international code of virus classification and nomenclature ratified by the International Committee on Taxonomy of Viruses (2017). Arch Virol 162:2505–2538. https://doi.org/10.1007/s00705-017-3358-5.