



# Genome Sequences of Bacteriophages KaiHaiDragon and OneinaGillian, Isolated from *Microbacterium foliorum* in Riverside, California

Gillian Miller,<sup>a</sup> Steven Tran,<sup>a</sup> Rhiannon Abrahams,<sup>a</sup> Daniel Bazan,<sup>a</sup> Ethan Blaylock,<sup>a</sup> Jessica Choi,<sup>a</sup> Shannon Grewal,<sup>a</sup> Elmira Hernandez,<sup>a</sup> Daniel Kim,<sup>a</sup> Kay Kim,<sup>a</sup> Yumin Lee,<sup>a</sup> Meagan Lopez,<sup>a</sup> Rachel Specht,<sup>a</sup> Nancy Kalaj,<sup>a</sup> Arun Muthiah,<sup>a</sup> Natasha Dean,<sup>a</sup>  Arturo Diaz<sup>a</sup>

<sup>a</sup>Biology Department, La Sierra University, Riverside, California, USA

**ABSTRACT** KaiHaiDragon and OneinaGillian are two bacteriophages which have been recovered from soil samples using the bacterial host *Microbacterium foliorum*. Their genome lengths are 52,992 bp and 61,703 bp, with 91 and 104 predicted open reading frames, respectively. KaiHaiDragon belongs to cluster EC, and OneinaGillian is a member of cluster EG.

To increase knowledge regarding bacteriophage diversity, La Sierra University's chapter of the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) program (1) collected soil samples from Riverside, California, and isolated bacteriophages by directly plating the environmental samples on a bacterial lawn of *Microbacterium foliorum* NRRL B-24224. The genomes of two of the isolated phages, KaiHaiDragon and OneinaGillian, were selected by the class for sequencing and annotation.

DNA was isolated from high-titer phage lysates ( $>5 \times 10^9$  PFU/ml) generated after two rounds of purification using the Wizard DNA cleanup system (Promega, catalog number A7280), and libraries were prepared using an NEB Ultra II FS kit at the Pittsburgh Bacteriophage Institute. The libraries were run on an Illumina MiSeq instrument, yielding at least 700,000 250-base paired-end reads per sample. Raw reads were assembled by using Newbler (version 2.6) and Consed (version 29) (2) into a single contig for each phage, with 1,039-fold and 2,049-fold coverage for KaiHaiDragon and OneinaGillian, respectively. The genome sequences were manually annotated with the assistance of DNA Master (<http://cobamide2.bio.pitt.edu/computer.htm>), ARAGORN (3), BLASTP (4), HHpred (5), and Phamerator (6), as outlined in the SEA-PHAGES bioinformatics guide (<https://seaphagesbioinformatics.helpdocsonline.com/home>).

Phages were assigned to clusters based on genomic sequence similarity, genomic synteny, and phylogenetic analysis using the Actinobacteriophage Database (PhagesDB) and the Phamerator software with default settings (6, 7). KaiHaiDragon was assigned to cluster EC, and OneinaGillian was assigned to cluster EG. KaiHaiDragon's genome is circularly permuted, with a length of 52,992 base pairs. The genome has a 68.9% G+C content and encodes 91 genes. Interestingly, all of KaiHaiDragon's open reading frames (ORFs) are transcribed in the forward direction. Functions were assigned to 30 of the predicted genes, and no tRNA genes were identified. Of note, KaiHaiDragon gp83 was found to encode an RNA polymerase sigma factor.

The OneinaGillian genome has a G+C content of 67.1% and is 61,703 base pairs long, and each terminus has 203-base pair direct terminal repeats. Average nucleotide identity (ANI) was determined by OrthoANI (8). OneinaGillian is only 72.1% and 70.49% identical, respectively, to Squash and Hyperion, which are also phages assigned to cluster EG. The beginning and the end of the genome of OneinaGillian are highly

**Citation** Miller G, Tran S, Abrahams R, Bazan D, Blaylock E, Choi J, Grewal S, Hernandez E, Kim D, Kim K, Lee Y, Lopez M, Specht R, Kalaj N, Muthiah A, Dean N, Diaz A. 2019. Genome sequences of bacteriophages KaiHaiDragon and OneinaGillian, isolated from *Microbacterium foliorum* in Riverside, California. *Microbiol Resour Announc* 8:e00002-19. <https://doi.org/10.1128/MRA.00002-19>.

**Editor** Christina Cuomo, Broad Institute of MIT and Harvard University

**Copyright** © 2019 Miller 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/).

Address correspondence to Arturo Diaz, [adiaz@lasierra.edu](mailto:adiaz@lasierra.edu).

**Received** 2 January 2019

**Accepted** 10 March 2019

**Published** 11 April 2019

dissimilar to those of Squash and Hyperion due to the presence of 55 orphans, genes without detectable homologs in other sequenced phages, located in these regions. In comparison, Squash and Hyperion are 90.75% identical to each other. Out of a total of 104 genes in the OneinaGillian genome, only 29 genes were assigned a function, including a tRNA. OneinaGillian is organized with its virion structure and assembly genes in the leftmost 31 kbp of the genome, with a centrally located lysin A (gp45) and membrane protein genes. The nonstructural genes are present to the right of the lysin A gene, while the rightmost portion of the genome contains mostly orphan genes. Additional details on OneinaGillian and KaiHaiDragon can be found in the Actinobacteriophage Database (7).

**Data availability.** The GenBank and SRA accession numbers are [MH590600](#) and [SRX5352621](#) for KaiHaiDragon and [MH727556](#) and [SRX5352620](#) for OneinaGillian, respectively.

## ACKNOWLEDGMENTS

This work was made possible by the La Sierra University Biology Department and the Howard Hughes Medical Institute SEA-PHAGES program.

## REFERENCES

- Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, Dennehy JJ, Denver DR, Dunbar D, Elgin SCR, Findley AM, Gissendanner CR, Golebiewska UP, Guild N, Hartzog GA, Grillo WH, Hollowell GP, Hughes LE, Johnson A, King RA, Lewis LO, Li W, Rosenzweig F, Rubin MR, Saha MS, Sandoz J, Shaffer CD, Taylor B, Temple L, Vazquez E, Ware VC, Barker LP, Bradley KW, Jacobs-Sera D, Pope WH, Russell DA, Cresawn SG, Lopatto D, Bailey CP, Hatfull GF. 2014. A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. *mBio* 5:e01051-13. <https://doi.org/10.1128/mBio.01051-13>.
- Gordon D, Abajian C, Green P. 1998. Consed: a graphical tool for sequence finishing. *Genome Res* 8:195–202. <https://doi.org/10.1101/gr.8.3.195>.
- 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>.
- 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](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:W244–W248. <https://doi.org/10.1093/nar/gki408>.
- Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatics tool for comparative bacteriophage genomics. *BMC Bioinformatics* 12:395. <https://doi.org/10.1186/1471-2105-12-395>.
- Russell DA, Hatfull GF. 2017. PhagesDB: the Actinobacteriophage Database. *Bioinformatics* 33:784–786. <https://doi.org/10.1093/bioinformatics/btw711>.
- Lee I, Ouk Kim Y, Park S-C, Chun J. 2016. OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66:1100–1103. <https://doi.org/10.1099/ijsem.0.000760>.