AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Sulfitobacter* Phage ϕGT1, Isolated from a Tidal Flat

Chung Y. Hwang,[a,b] Yirang Cho,[c] Gwang I. Jang,[d] Byung C. Cho,[b,e] Stephen C. Hardies[f]

[a]School of Earth and Environmental Sciences, Seoul National University, Seoul, South Korea
[b]Research Institute of Oceanography, Seoul National University, Seoul, South Korea
[c]Bioinformatics, Johns Hopkins University, Baltimore, Maryland, USA
[d]West Sea Fisheries Research Institute, National Institute of Fisheries Science, Incheon, South Korea
[e]Saemangeum Environmental Research Center, Kunsan National University, Kunsan, South Korea
[f]Department of Biochemistry and Structural Biology, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA

ABSTRACT The *Sulfitobacter* bacteria are ubiquitous and important players in organic sulfur cycling in marine environments. Here, we report the complete genome sequence of ϕGT1 infecting *Sulfitobacter* sp. HGT1, both of which were isolated from coastal sediment. ϕGT1 has a 40,019-bp genome containing 69 predicted protein-encoding genes.

The genus *Sulfitobacter* belongs to the family *Rhodobacteraceae*, which is one of the major alphaproteobacterial groups in aquatic environments (1). *Sulfitobacter* bacteria have been isolated from diverse marine habitats and are known to affect the biogeochemical sulfur cycle in the ocean (1–5). Five *Sulfitobacter* phages have been isolated (Table 1), all of which are marine podoviruses (6–8).

The host bacterium *Sulfitobacter* sp. HGT1 and phage ϕGT1 were isolated from a tidal flat (37.62°N, 126.37°E) in the Yellow Sea, South Korea, in August 2011. A surface sediment sample was resuspended with a filtrate of ambient seawater that had been prefiltered with a 0.22-μm-pore membrane filter (Millipore). An aliquot of the slurry was spread on marine agar (Difco), and the plate was incubated aerobically at 30°C. Strain HGT1 was purified and identified as described previously (3). Phages in the filtrate (0.22-μm pore size) of ambient seawater were concentrated using Amicon Ultra-4 centrifugal filter units (Millipore). The phage concentrate was inoculated onto a host bacterial culture according to the double-layer plaque assay protocol (9). A single plaque, designated ϕGT1, was selected and purified by four rounds of double-layer plaque assays.

To extract genomic DNA, ϕGT1 concentrates were precipitated with polyethylene glycol 8000 (Sigma) from fully lysed plates as described previously (10). DNA was extracted using the QIAamp MinElute virus spin kit (Qiagen). A sequencing library was constructed with the Accel-NGS 2S PCR-free DNA library kit (insert size, ~300 bp; Swift Biosciences) and sequenced by Macrogen, Inc. (Seoul, South Korea), using the Illumina MiSeq platform with a 2 × 300-bp paired-end format. Raw reads were processed with Trimmomatic version 0.39 (11) with default settings to trim the adapter and low-quality sequences. PhiX control reads were removed using the bbduk.sh script in the BBMap package version 38.84 (12). The trimmed reads (705,348 reads in total) were assembled using *de novo* assembly with default settings in CLC Genomics Workbench version 8.51 (Qiagen, Aarhus, Denmark), resulting in a circular contig with an average coverage of 3,795×. Phage termini and packaging mode were identified using PhageTerm with default settings (13). Open reading frames (ORFs) were predicted using Phage Search Tool Enhanced Release (PHASTER), GeneMark, GeneMarkS, and

Citation Hwang CY, Cho Y, Jang GI, Cho BC, Hardies SC. 2020. Complete genome sequence of *Sulfitobacter* phage ϕGT1, isolated from a tidal flat. Microbiol Resour Announc 9:e00779-20. https://doi.org/10.1128/MRA.00779-20.

Editor Simon Roux, DOE Joint Genome Institute

Address correspondence to Chung Y. Hwang, chung.hwang@snu.ac.kr.

Received 5 July 2020
Accepted 15 July 2020
Published 13 August 2020

TABLE 1 Genomic characteristics of ϕGT1 and other *Sulfitobacter* phages

Phage	Life cycle	GenBank accession no.	Genome size (bp)	GC content (%)	No. of ORFs	Coverage with ϕGT1 (%)	Identity with ϕGT1 (%)	Source or reference
ϕGT1	Lytic	MT584811	40,019	56.4	69	100	100	This study
NYA-2014a	Temperate	NC_027299	42,092	58.5	71	25	74.8	GenBank
ϕCB2047-A	Temperate	NC_020858	40,929	58.8	73	21	80.5	6
ϕCB2047-C	Temperate	NC_020856	40,931	59.0	71	21	78.2	6
ϕCB2047-B	Lytic	NC_020862	74,480	43.0	92	NS[a]	NS	7
EE36ϕ1	Lytic	NC_012696	73,325	47.0	79	NS	NS	8

[a] NS, no significant similarity.

GeneMark.hmm (14–16). Translated ORFs were examined using PsiBLAST version 2.2.28+ and HHpred version 3.2.0 (17, 18) searches against the Protein Data Bank (PDB), Pfam, and NCBI conserved domain databases (19–21). Phylogenetic analysis of the large terminase sequence of ϕGT1 (phiGT1_2) was performed using the UCSC Sequence Alignment and Modeling system version 3.5 (22) and MrBayes version 3.2 (23), as described by Casjens et al. (24). The complete genome sequence of ϕGT1 was used as a query for MegaBLAST searching with default settings against the NCBI nonredundant/nucleotide database (release 237.0) to find highly similar sequences.

The host bacterium *Sulfitobacter* sp. HGT1 was most closely related to the type strain of *Sulfitobacter marinus* (25), based on the 16S rRNA gene sequence similarity (99.7%), with a GC content of 57.9%. The complete genome of ϕGT1 is 40,019 bp, with a GC content of 56.4% and a total of 69 predicted ORFs (Table 1). The ϕGT1 genome was concluded to be circularly permuted by a headful packaging system based on the results from the analysis of reads using PhageTerm and the similarity of the large terminase sequence to other headful packaging terminases. For reporting, the first base of the ϕGT1 genome was intentionally set at a start position of the gene encoding a putative small terminase subunit (phiGT1_1), the coding sequence of which typically contains the recognition and cleavage sites for the first end produced on the concatemer (26). The genome-wide comparison showed that the best hit for ϕGT1 was the temperate *Sulfitobacter* phage NYA-2014a in the *Podoviridae* family (Table 1). In terms of genome size and GC content, ϕGT1 was more similar to the three known temperate *Sulfitobacter* phages (40,929 to 42,092 bp, with GC contents of 58.5 to 59.0%) than the two known lytic *Sulfitobacter* phages (73,325 to 74,480 bp, with GC contents of 43.0 to 47.0%) (Table 1). Although ϕGT1 was distantly related to the three known temperate *Sulfitobacter* phages, with 21 to 25% coverage and 74.8 to 80.5% identity (Table 1), those three temperate phages were aligned among themselves over substantial fractions of their genomes (74 to 95%), with high identity values (99.5 to 99.9%). No significant match was found between ϕGT1 and the two known lytic *Sulfitobacter* phages (Table 1). ϕGT1 displayed a lytic life cycle based on plaque characterization (i.e., clear plaque formation) and gene content, including the presence of lysis genes (endolysin [phiGT1_23] and holin [phiGT1_24]) and the absence of an integrase gene. Neither terminal repeats nor RNA polymerase genes were found in ϕGT1, but ORFs homologous to podoviral core tail proteins P22 gp10 (phiGT1_15) and T7 tubular tail A (phiGT1_14) were found, indicating that ϕGT1 is likely to be a new marine lytic *Sulfitobacter* phage in the family *Podoviridae*.

Data availability. The annotated complete sequence of ϕGT1 has been deposited in DDBJ/ENA/GenBank under the accession number MT584811 (BioProject number PRJDB9847 and BioSample number SAMD00233255). The version of the phage genome described in this paper is the first version. The raw sequence reads are available in the DDBJ Sequence Read Archive with the accession number DRA010407. The genome sequence of the host bacterium *Sulfitobacter* sp. HGT1 is also available in DDBJ/ENA/GenBank under the accession number BLWI01000000 (BioSample number SAMD00228206).

ACKNOWLEDGMENT

This work was supported by a National Research Foundation of South Korea grant funded by the government of South Korea (grant 2019R1A6A1A10073437).

REFERENCES

1. Sorokin DY, Rainey FA, Webb RI, Fuerst JA. 2015. *Sulfitobacter*. *In* Whitman WB, Rainey F, Kämpfer P, Trujillo M, Chun J, DeVos P, Hedlund B, Dedysh S (ed), Bergey's manual of systematics of archaea and bacteria. John Wiley & Sons, Hoboken, NJ. https://doi.org/10.1002/9781118960608.gbm00874.
2. Parte AC. 2018. LPSN—List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. Int J Syst Evol Microbiol 68:1825–1829. https://doi.org/10.1099/ijsem.0.002786.
3. Hwang CY, Cho BC. 2008. *Ponticoccus litoralis* gen. nov., sp. nov., a marine bacterium in the family *Rhodobacteraceae*. Int J Syst Evol Microbiol 58:1332–1338. https://doi.org/10.1099/ijs.0.65612-0.
4. Song J, Jang HJ, Joung Y, Kang I, Cho JC. 2019. *Sulfitobacter profundi* sp. nov., isolated from deep seawater. J Microbiol 57:661–667. https://doi.org/10.1007/s12275-019-9150-3.
5. Curson ARJ, Rogers R, Todd JD, Brearley CA, Johnston AWB. 2008. Molecular genetic analysis of a dimethylsulfoniopropionate lyase that liberates the climate-changing gas dimethylsulfide in several marine α-proteobacteria and *Rhodobacter sphaeroides*. Environ Microbiol 10:757–767. https://doi.org/10.1111/j.1462-2920.2007.01499.x.
6. Ankrah NYD, Budinoff CR, Wilson WH, Wilhelm SW, Buchan A. 2014. Genome sequences of two temperate phages, ϕCB2047-A and ϕCB2047-C, infecting *Sulfitobacter* sp. strain 2047. Genome Announc 2:e00108-14. https://doi.org/10.1128/genomeA.00108-14.
7. Ankrah NYD, Budinoff CR, Wilson WH, Wilhelm SW, Buchan A. 2014. Genome sequence of the *Sulfitobacter* sp. strain 2047-infecting lytic phage ϕCB2047-B. Genome Announc 2:e00945-13. https://doi.org/10.1128/genomeA.00945-13.
8. Zhao Y, Wang K, Jiao N, Chen F. 2009. Genome sequences of two novel phages infecting marine roseobacters. Environ Microbiol 11:2055–2064. https://doi.org/10.1111/j.1462-2920.2009.01927.x.
9. Adams MK. 1959. Bacteriophages. Interscience Publishers, Inc., New York, NY.
10. Holmfeldt K, Solonenko N, Shah M, Corrier K, Riemann L, VerBerkmoes NC, Sullivan MB. 2013. Twelve previously unknown phage genera are ubiquitous in global oceans. Proc Natl Acad Sci U S A 110:12798–12803. https://doi.org/10.1073/pnas.1305956110.
11. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
12. Bushnell B. 2014. BBMap: a fast, accurate, splice-aware aligner. Ernest Orlando Lawrence Berkeley National Laboratory, Berkeley, CA. https://www.osti.gov/biblio/1241166-bbmap-fast-accurate-splice-aware-aligner.
13. Garneau JR, Depardieu F, Fortier LC, Bikard D, Monot M. 2017. PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. Sci Rep 7:8292. https://doi.org/10.1038/s41598-017-07910-5.
14. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21. https://doi.org/10.1093/nar/gkw387.
15. 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.
16. Lukashin A, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res 26:1107–1115. https://doi.org/10.1093/nar/26.4.1107.
17. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421. https://doi.org/10.1186/1471-2105-10-421.
18. Zimmerman L, Stephens A, Nam S-Z, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J Mol Biol 430:2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007.
19. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. Nucleic Acids Res 28:235–242. https://doi.org/10.1093/nar/28.1.235.
20. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, Sonnhammer ELL, Hirsh L, Paladin L, Piovesan D, Tosatto SCE, Finn RD. 2019. The Pfam protein families database in 2019. Nucleic Acids Res 47:D427–D432. https://doi.org/10.1093/nar/gky995.
21. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH. 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222–D226. https://doi.org/10.1093/nar/gku1221.
22. Hughey R, Krogh A. 1996. Hidden Markov models for sequence analysis: extension and analysis of the basic method. Comput Appl Biosci 12:95–107. https://doi.org/10.1093/bioinformatics/12.2.95.
23. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542. https://doi.org/10.1093/sysbio/sys029.
24. Casjens SR, Gilcrease EB, Winn-Stapley DA, Schicklmaier P, Schmieger H, Pedulla ML, Ford ME, Houtz JM, Hatfull GF, Hendrix RW. 2005. The generalized transducing *Salmonella* bacteriophage ES18: complete genome sequence and DNA packaging strategy. J Bacteriol 187:1091–1104. https://doi.org/10.1128/JB.187.3.1091-1104.2005.
25. Yoon JH, Kang SJ, Oh TK. 2007. *Sulfitobacter marinus* sp. nov., isolated from seawater of the East Sea in Korea. Int J Syst Evol Microbiol 57:898–899. https://doi.org/10.1099/ijs.0.64637-0.
26. Casjens S, Sampson L, Randall S, Eppler K, Wu H, Petri JB, Schmieger H. 1992. Molecular genetic analysis of bacteriophage P22 gene 3 product, a protein involved in the initiation of headful DNA packaging. J Mol Biol 227:1086–1099. https://doi.org/10.1016/0022-2836(92)90523-M.