



Genome Sequence of Hardyhis2, a Gammaleolipovirus Infecting *Haloarcula hispanica*

 Mike Dyll-Smith,^{a,b}  Friedhelm Pfeiffer,^a Pei-Wen Chiang,^c Sen-Lin Tang^c

^aComputational Biology Group, Max Planck Institute of Biochemistry, Martinsried, Germany

^bVeterinary Biosciences, Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville, Australia

^cBiodiversity Research Center, Academia Sinica, Taipei, Taiwan

ABSTRACT Hardyhis2 virus infects the halophilic archaeon *Haloarcula hispanica* DSM 4426^T and is closely related to His2 (*Pleolipoviridae* family). The viral genome is 16,133 bp long, with terminal inverted repeats of 599 bp. The predicted spike protein is only weakly similar (32% amino acid identity) to that of His2.

Hardyhis2 is a novel halovirus that is related to the gammaleolipovirus His2 (1) and was recovered in 1998 from a water sample from the hypersaline Lake Hardy, Victoria, Australia (35°3'58.27"S, 141°44'12.50"E). Originally labeled HH2, it produced turbid plaques of 0.5 to 1 mm (2 to 3 days at 37°C) in soft-layer overlay lawns of the isolating host *Haloarcula hispanica* DSM 4426^T. The isolation and plaque purification followed the methods described for halovirus His2 (1). A stock of this virus that had been stored at –80°C since 1999 was thawed, and DNA was extracted using the cetyltrimethylammonium bromide (CTAB)/phenol-chloroform method (2). After the quality and quantity were checked by agarose gel electrophoresis, total DNA was amplified using the Qiagen REPLI-g minikit; the products were checked for suitable size by agarose gel electrophoresis and then purified using QIAamp DNA minikits (Qiagen, Inc., Valencia, CA, USA). One microgram of DNA was sent to the Yourgene Health Co. (Taipei, Taiwan) for sequencing using the Illumina HiSeq 2500 platform (2 × 250-bp paired-end reads). For library construction, DNA was fragmented by sonication, end repaired, and A tailed (Klenow fragment), indexed adapters were added (DNA ligase), and PCR amplification (10 cycles) was performed to enrich for adapter-modified DNA. Sequencing returned 102,128 paired-end Illumina reads (confidence mean, Q36.8; reads with a score of at least Q30, 95.6%) that were quality trimmed (quality trimming criteria, minimum length of 35 bp and error probability of <0.05), resulting in 102,049 reads (average length, 248.6 bp). These reads were assembled using the *de novo* assembler tool within Geneious version 10.2.6 (3). This and other software tools were run with default parameters. A 16,133-bp linear, double-stranded DNA (dsDNA) contig was assembled (521-fold coverage) with an inverted terminal repeat (ITR) of 599 bp. The ITR was very similar to the 525-bp ITR of halovirus His2, with the final 130 bp being 91% identical and the terminal 23 bp being completely identical. The average G+C content was 39.8%, considerably less than that of the host (62.7%) (4). The Hardyhis2 genome is similar in size to that of His2 (16,067 bp) (1) and shares 82.3% average nucleotide identity (5). The His2 genome carries terminal proteins, but this was not determined for Hardyhis2. The two viruses probably represent distinct species within the *Gammaleolipovirus* genus (6, 7). Annotation used gene prediction by GeneMarkS-2 (8) and BLAST searches (9) of the NCBI nucleotide/nonredundant databases and predicted 33 coding sequences (CDSs), with most being similar in sequence and gene arrangement to those of His2. The conserved block of pleolipovirus genes specifying structural proteins and a putative NTPase (6) (HdyHp2_120 to HdyHp2_155) includes the gene for the major spike protein (HdyHp2_135), a virion surface protein

Citation Dyll-Smith M, Pfeiffer F, Chiang P-W, Tang S-L. 2021. Genome sequence of Hardyhis2, a gammaleolipovirus infecting *Haloarcula hispanica*. *Microbiol Resour Announc* 10:e00226-21. <https://doi.org/10.1128/MRA.00226-21>.

Editor Kenneth M. Stedman, Portland State University

Copyright © 2021 Dyll-Smith 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 Mike Dyll-Smith, mike.dyllsmith@gmail.com.

Received 4 March 2021

Accepted 15 April 2021

Published 13 May 2021

probably involved in host cell attachment and membrane fusion (10). The spike proteins of Hardyhis2 and His2 (His2V_gp29) share only 32% amino acid identity, whereas the proteins specified by nearby genes share $\geq 85\%$ amino acid similarity, suggesting that the two viruses could use distinct binding sites on the cell surface of *H. hispanica*.

Data availability. This genome sequence has been deposited in GenBank with the accession number [MW557853](https://doi.org/10.1093/mbe/mz001). The associated BioProject, SRA, and BioSample accession numbers are [PRJNA698715](https://doi.org/10.1093/bioinformatics/bts199), [SRR13683896](https://doi.org/10.1093/bioinformatics/bts199), and [SAMN17734765](https://doi.org/10.1093/bioinformatics/bts199), respectively.

ACKNOWLEDGMENT

This study was supported by Academia Sinica, Taiwan.

REFERENCES

1. Bath C, Cukalac T, Porter K, Dyall-Smith ML. 2006. His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, *Salterprovirus*. *Virology* 350:228–239. <https://doi.org/10.1016/j.virol.2006.02.005>.
2. Wilson K. 2001. Preparation of genomic DNA from bacteria. *Curr Protoc Mol Biol* 56:2.4.1–2.4.5. <https://doi.org/10.1002/0471142727.mb0204s56>.
3. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28:1647–1649. <https://doi.org/10.1093/bioinformatics/bts199>.
4. Juez G, Rodriguez-Valera F, Ventosa A, Kushner DJ. 1986. *Haloarcula hispanica* spec. nov. and *Haloferax gibbonsii* spec. nov., two new species of extremely halophilic archaeobacteria. *Syst Appl Microbiol* 8:75–79. [https://doi.org/10.1016/S0723-2020\(86\)80152-7](https://doi.org/10.1016/S0723-2020(86)80152-7).
5. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* 30:772–780. <https://doi.org/10.1093/molbev/mst010>.
6. Demina TA, Oksanen HM. 2020. Pleomorphic archaeal viruses: the family *Pleolipoviridae* is expanding by seven new species. *Arch Virol* 165:2723–2731. <https://doi.org/10.1007/s00705-020-04689-1>.
7. Bamford DH, Pietila MK, Roine E, Atanasova NS, Dienstbier A, Oksanen HM, ICTV Report Consortium. 2017. ICTV virus taxonomy profile: *Pleolipoviridae*. *J Gen Virol* 98:2916–2917. <https://doi.org/10.1099/jgv.0.000972>.
8. Lomsadze A, Gemayel K, Tang S, Borodovsky M. 2018. Modeling leaderless transcription and atypical genes results in more accurate gene prediction in prokaryotes. *Genome Res* 28:1079–1089. <https://doi.org/10.1101/gr.230615.117>.
9. Johnson M, Zaretskaya I, Raytselis Y, Merezuk Y, McGinnis S, Madden TL. 2008. NCBI BLAST: a better Web interface. *Nucleic Acids Res* 36:W5–W9. <https://doi.org/10.1093/nar/gkn201>.
10. El Omari K, Li S, Kotecha A, Walter TS, Bignon EA, Harlos K, Somerharju P, De Haas F, Clare DK, Molin M, Hurtado F, Li M, Grimes JM, Bamford DH, Tischler ND, Huiskonen JT, Stuart DI, Roine E. 2019. The structure of a prokaryotic viral envelope protein expands the landscape of membrane fusion proteins. *Nat Commun* 10:846. <https://doi.org/10.1038/s41467-019-08728-7>.