AMERICAN SOCIETY FOR MICROBIOLOGY

Microbiology®
Resource Announcements

Check for updates

Complete Genome Sequence of *Bradyrhizobium japonicum* Podophage Paso

Dylan McBee,[a] Aravind Ravindran,[b] Heather Newkirk,[a] Ben Burrowes,[c] Ry Young,[a,c] Carlos Gonzalez[b,c]

[a]Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas, USA
[b]Department of Plant Pathology and Microbiology, Texas A&M University, College Station, Texas, USA
[c]Center for Phage Technology, Texas A&M University, College Station, Texas, USA

ABSTRACT *Bradyrhizobium japonicum* is a nitrogen-fixing, Gram-negative bacterium that forms a symbiotic relationship with leguminous plants. This announcement describes the isolation and genome annotation of *B. japonicum* T7-like podophage Paso. Genomic analysis reveals genes that are associated with both the T5 and T7 modes of genomic DNA entry into the host.

*B*radyrhizobium japonicum* is found in leguminous root tips, where it ultimately stimulates soybean growth (1). The nitrogen-fixing properties of *B. japonicum* have made it a popular inoculant in crop production, with recent efforts to identify remedial applications (2, 3). Characterizing Paso may accelerate efforts to understand this bacterium and the larger ecosystem that it inhabits.

Phage Paso was isolated by plaque purification (4) from unidentified weed samples from Uvalde, TX, in August 2017 and propagated aerobically on *B. japonicum* D409 (ATCC 10324) at 28°C in L-arabinose medium. Genomic DNA was purified using a Wizard DNA kit (5). Libraries were prepared with 550-bp inserts using a TruSeq Nano kit. The paired-end 500-bp reads were sequenced on an Illumina MiSeq instrument with v2 500-cycle chemistry. The 1,997,626 reads were quality controlled and trimmed with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc) and then FastX v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/download.html) before assembly into a single contig at 1,031.0-fold coverage with SPAdes v3.5.0 (6), which was closed based on the circular assembly generated by the assembler. GLIMMER v3 and MetaGeneAnnotator v1.0 structural annotation outputs were manually verified to predict protein-coding genes, while tRNA genes were predicted with ARAGORN v2.36 (7–9). Protein functions were predicted using outputs from BLAST v2.9.0 and InterProScan v5.33 (10, 11). Putative transmembrane domains were detected using TMHMM v2.0 (12). BLAST queries were compared against the NCBI nonredundant, UniProtKB Swiss-Prot, and TrEMBL databases with a 0.001 maximum expectation value cutoff (13). The genome-wide DNA sequence similarity between Paso and other phages was calculated by progressiveMauve v2.4 (14). All annotation tools were accessed via the CPT Galaxy interface hosted at https://cpt.tamu.edu/galaxy-pub (15–17). Unless otherwise stated, all tools were executed using default parameters.

The 47,808-bp DNA genome sequence of Paso has a predicted 95% coding density and a GC content of 55% compared to 64% in the host (1). PhageTerm (18) predicts the genome sequence to start with a T7-like 371-bp direct terminal repeat. Overall, 55 protein-encoding genes were identified, but no tRNA genes were located; all the coding DNA sequences (CDS) are on one strand. The first 1.5 kb of the genome sequence appears to contain no genes.

Agrobacterium phage Atu_ph03 is the closest characterized phage to Paso, with 29 of the 55 genes sharing amino acid sequence similarity, although the overall nucleotide

sequence similarity is only 14.7%. Genomic analysis and amino acid and nucleotide similarities suggest that Paso is a podophage, which was confirmed visually by transmission electron microscopy (TEM). Interestingly, Paso contains a close homolog of the T5 A1 protein and the T7 RNA polymerase. Analysis by HHpred identified an unusual Clp protease speculated to play a role in capsid degradation (19, 20). Most of the BLASTp hits for Paso genes are from bacterial genomes, suggesting a temperate life cycle for Paso or its close relatives.

Data availability. The Paso genome is deposited as GenBank accession number MT708546.1. The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR11558344, and SAMN14609645, respectively.

REFERENCES

1. Sánchez C, Minamisawa K. 2019. Nitrogen cycling in soybean rhizosphere: sources and sinks of nitrous oxide (N$_2$O). Front Microbiol 10:1943. https://doi.org/10.3389/fmicb.2019.01943.
2. Naamala J, Jaiswal SK, Dakora FD. 2016. Antibiotics resistance in *Rhizobium*: type, process, mechanism and benefit for agriculture. Curr Microbiol 72:804–816. https://doi.org/10.1007/s00284-016-1005-0.
3. Figueredo MS, Tonelli ML, Ibáñez F, Morla F, Cerioni G, Del Carmen Tordable M, Fabra A. 2017. Induced systemic resistance and symbiotic performance of peanut plants challenged with fungal pathogens and co-inoculated with the biocontrol agent Bacillus sp. CHEP5 and Bradyrhizobium sp. SEMIA6144. Microbiol Res 197:65–73. https://doi.org/10.1016/j.micres.2017.01.002.
4. van Charante F, Holtappels D, Blasdel B, Burrowes B. 2019. Isolation of bacteriophages, p 1–32. *In* Harper D, Abedon S, Burrowes B, McConville M (ed), Bacteriophages. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-319-40598-8_14-1.
5. Summer EJ. 2009. Preparation of a phage DNA fragment library for whole genome shotgun sequencing. Methods Mol Biol 502:27–46. https://doi.org/10.1007/978-1-60327-565-1_4.
6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, 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.
7. 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.
8. 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.
9. 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.
10. 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.
11. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240. https://doi.org/10.1093/bioinformatics/btu031.
12. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305:567–580. https://doi.org/10.1006/jmbi.2000.4315.
13. The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.
14. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147. https://doi.org/10.1371/journal.pone.0011147.
15. Jalili V, Afgan E, Gu Q, Clements D, Blankenberg D, Goecks J, Taylor J, Nekrutenko A. 2020. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update. Nucleic Acids Res 48:W395–W402. https://doi.org/10.1093/nar/gkaa434.
16. Ramsey J, Rasche H, Maughmer C, Criscione A, Mijalis E, Liu M, Hu JC, Young R, Gill JJ. 2020. Galaxy and Apollo as a biologist-friendly interface for high-quality cooperative phage genome annotation. PLoS Comput Biol 16:e1008214. https://doi.org/10.1371/journal.pcbi.1008214.
17. Dunn NA, Unni DR, Diesh C, Munoz-Torres M, Harris NL, Yao E, Rasche H, Holmes IH, Elsik CG, Lewis SE. 2019. Apollo: democratizing genome annotation. PLoS Comput Biol 15:e1006790. https://doi.org/10.1371/journal.pcbi.1006790.
18. Garneau JR, Depardieu F, Fortier L-C, 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.
19. Allison GE, Angeles DC, Thi Huan P, Verma NK. 2003. Morphology of temperate bacteriophage SfV and characterisation of the DNA packaging and capsid genes: the structural genes evolved from two different phage families. Virology 308:114–127. https://doi.org/10.1016/S0042-6822(03)00198-3.
20. Zimmermann 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.