Complete Genome Sequence of *Rhizobium japonicum* Podophage Pasto

Nikhil Manuel,[a] Leika Rushing,[b] Aravind Ravindran,[b] Heather Newkirk,[a] ⓘ Ben Burrowes,[c] ⓘ Ry Young,[a,c] ⓘ Carlos Gonzalez[b,c]

aDepartment of Biochemistry and Biophysics, Texas A&M University, College Station, Texas, USA
bDepartment of Plant Pathology and Microbiology, Texas A&M University, College Station, Texas, USA
cCenter for Phage Technology, Texas A&M University, College Station, Texas, USA

ABSTRACT *Rhizobium japonicum* is a Gram-negative bacterium of interest for research into nitrogen fixation in legumes. This article describes the isolation, sequencing, and annotation of *R. japonicum* podophage Pasto. While it shows no significant similarity to identified phages, genomic analysis indicates that Pasto may be temperate and is a novel T7-like podophage.

*R*hizobium japonicum is a Gram-negative, motile bacterium with the ability to create a symbiotic relationship with legumes (1), in which the bacterium inhabits root nodules to perform nitrogen fixation (2). *R. japonicum* phages can be utilized to better understand the bacterial host and possibly to facilitate understanding of the transfer of genetic material in the rhizosphere through transduction (2).

R. japonicum phage Pasto was isolated from potato root samples obtained in Olton, Texas, in August 2017. Plaques were observed on *R. japonicum* strain D409 (ATCC 10324) cultured on L-arabinose agar at 28°C (3). The phage DNA was isolated with a Wizard DNA cleanup kit as described before (4), and libraries were prepared with 550-bp inserts using a TruSeq Nano kit and sequenced on an Illumina MiSeq v2 system (500 bp). The 1,066,802 total reads were quality controlled and manually trimmed with FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and FastX v0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/download.html) before assembly into a single contig at 41.8× coverage with SPAdes v3.5.0 (5). The genome was closed by PCR performed off the end of the contig (forward primer, 5′-GGCAGACACACGAGAGATAAA-3′; reverse primer, 5′-TTGTTCCGTTCGTCTTGTGTG-3′) and Sanger sequencing of the resulting product. Structural annotation was performed using GLIMMER v3 (6) and MetaGeneAnnotator v1.0 (7), and tRNAs were predicted with ARAGORN v2.36 (8). Gene functions were predicted by searching through conserved domains with InterProScan v5.33 (9) and determining sequence similarity with BLAST v2.9.0 (10) by comparison with the NCBI nonredundant, Swiss-Prot, and TrEMBL databases (11). TMHMM v2.0 was also utilized (12), and the LipoP v1.0 tool evaluated lipoylation signals (13). progressiveMauve v2.4 was used to calculate genome-wide DNA sequence similarity (14). All annotation tools were run with default parameters and accessed through the Center for Phage Technology (CPT) Galaxy/Apollo Web platform (https://cpt.tamu.edu/galaxy-pub) (15–17).

Pasto is a 42,407-bp podophage with a G+C content of 58.6%, which is lower than the average host G+C content of 64.1% (3). Genomic analysis predicted 1 tRNA and 56 protein-coding genes, yielding an overall coding density of 92.8% despite the first 1.0 kb being devoid of coding sequences (CDSs). PhageTerm (18) was used to predict the phage termini, which were identified as 271-bp direct terminal repeats. Thirty-one of the 56 genes were assigned putative functions, of which 12 appear to be similar (BLASTp or BLASTn E values of <0.001) to those encoded by *Escherichia coli* phage T7.

This finding suggested that Pasto is a podophage, which was confirmed visually by transmission electron microscopy. Surprisingly, Pasto has many hypothetical proteins that are similar to bacterial proteins, suggesting that it is related to temperate phages despite the prevalence of genes similar to those of T7. A single tRNA gene was found embedded with an antisense orientation within a CDS encoding a hypothetical protein. This tRNA sequence may be a prophage integration site, as previously observed for a mobile genetic element in *Rhizobium* (19, 20).

Data availability. The Pasto genome was deposited in GenBank with accession number MT708545.1. The associated BioProject, SRA, and BioSample accession numbers are PRJNA222858, SRR11558345, and SAMN14609644, respectively.

REFERENCES

1. Parker J. 2001. Rhizobium, p 1715–1716. *In* Brenner S, Miller JH (ed), Encyclopedia of genetics. Academic Press, New York, NY.
2. Halmillawewa AP, Restrepo-Córdoba M, Yost CK, Hynes MF. 2015. Genomic and phenotypic characterization of *Rhizobium gallicum* phage vB_RglS_P106B. Microbiology (Reading) 161:611–620. https://doi.org/10.1099/mic.0.000022.
3. Kaneko T, Nakamura Y, Sato S, Minamisawa K, Uchiumi T, Sasamoto S, Watanabe A, Idesawa K, Iriguchi M, Kawashima K, Kohara M, Matsumoto M, Shimpo S, Tsuruoka H, Wada T, Yamada M, Tabata S. 2002. Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA110. DNA Res 9:189–197. https://doi.org/10.1093/dnares/9.6.189.
4. 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.
5. 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.
6. 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.
7. 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.
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. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong SY, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. Bioinformatics 30:1236–1240. https://doi.org/10.1093/bioinformatics/btu031.
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. The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.
12. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 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. Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. 2003. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Sci 12:1652–1662. https://doi.org/10.1110/ps.0303703.
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 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.
19. Sullivan JT, Ronson CW. 1998. Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a Phe-tRNA gene. Proc Natl Acad Sci U S A 95:5145–5149. https://doi.org/10.1073/pnas.95.9.5145.
20. Abebe HM, Sadowsky MJ, Kinkle BK, Schmidt EL. 1992. Lysogeny in *Bradyrhizobium japonicum* and its effect on soybean nodulation. Appl Environ Microbiol 58:3360–3366. https://doi.org/10.1128/AEM.58.10.3360-3366.1992.