AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequence of *Agrobacterium tumefaciens* Myophage Milano

Toni Nittolo,[a] Aravind Ravindran,[a,b] Carlos F. Gonzalez,[a,b] Jolene Ramsey[a]

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

ABSTRACT *Agrobacterium tumefaciens* C58 is a tumor-causing pathogen targeting plants and is ubiquitously found in soil. Here, the complete genome sequence of Milano, a myophage infecting *A. tumefaciens* C58, is presented. Milano encodes 127 proteins, of which 45 can be assigned a predicted function, and it is most similar to the flagellotropic *Agrobacterium* phage 7-7-1.

Agrobacterium tumefaciens C58 is a Gram-negative rod-shaped bacterium ubiquitously found in soil (1). As a plant pathogen, *A. tumefaciens* C58 contains plasmid TiC58 that transfers transfer DNA (T-DNA) to 90 families of dicotyledonous plants, inevitably resulting in crown gall tumors. Bacteriophages may be useful in manipulating this characteristic for engineering *Agrobacterium* strains. Here, we present the genome sequence of myophage Milano.

Milano was isolated from filtered rice stem extracts in Beaumont, TX, on *A. tumefaciens* C58 grown aerobically at 28°C in nutrient broth yeast (NBY) medium without glucose (2) by the soft agar overlay method (3). The high-titer lysate generated via the soft agar overlay method was used for extracting genomic DNA with the phenol-chloroform method, as in reference 4, and then the phage genomic DNA libraries were prepared using a NEBNext Ultra II DNA library prep kit and sequenced on an Illumina MiSeq instrument at the Genome Sequencing and Analysis Facility at the University of Texas at Austin (5). The 820,950 250-bp paired-end sequence reads were quality controlled with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and trimmed with the FASTX-Toolkit 0.0.14 (hannonlab.cshl.edu/fastx_toolkit/). Complete assembly into a single contig using SPAdes v.3.5.0, with default parameters, was confirmed via PCR of the genome ends (forward primer, 5′-GTCCTGAATCCATTTCGTA TGC-3′; reverse primer, 5′-CTCCGTTCTTCAGCTACTATG-3′), coupled with Sanger sequencing (5, 6). Genes were called and annotated using GLIMMER 3.0 and MetaGeneAnnotator 1.0 in the Web Apollo instance hosted by the Center for Phage Technology (https://cpt.tamu.edu/galaxy-pub), and all analyses were performed in their Galaxy instance (7–10). Potential tRNA genes were inspected using ARAGORN 2.36 (11). Gene functions were predicted using domains from InterProScan v.5.22, LipoP, and TMHMM, as well as BLASTp comparisons to the NCBI nonredundant (nr) and UniProtKB Swiss-Prot/TrEMBL databases (12–16). HHpred results were used as confirmatory evidence, in addition to the presence of domains or alignments (17). Rho-independent termination sites were detected using TransTerm (http://transterm.cbcb.umd.edu/). Milano's morphology was determined by negative-stain transmission electron microscopy at the Texas A&M Microscopy and Imaging Center with 2% (wt/vol) uranyl acetate (18).

Milano is a myophage with a 68,451-bp genome, 93.1% coding density, and a G+C content of 52.5%, which is lower than the G+C content of 58% of the host, *A. tumefaciens* C58 (1). Our analysis revealed 127 coding sequences, of which 45 have a

Citation Nittolo T, Ravindran A, Gonzalez CF, Ramsey J. 2019. Complete genome sequence of *Agrobacterium tumefaciens* myophage Milano. Microbiol Resour Announc 8:e00587-19. https://doi.org/10.1128/MRA.00587-19.

Editor Simon Roux, DOE Joint Genome Institute

Address correspondence to Jolene Ramsey, jolenerr@tamu.edu.

Received 20 May 2019
Accepted 29 May 2019
Published 20 June 2019

function predicted by using BLASTp or InterProScan, but no tRNAs. Based on PhageTerm prediction, Milano uses headful DNA packaging (19). While Milano contains mostly hypothetical proteins, there is one closely related phage in the NCBI nr database, namely, *Agrobacterium* myophage 7-7-1 (GenBank accession number JQ312117) (20). By comparison with progressiveMauve, Milano has 61.68% nucleotide identity and shares 94 proteins with *Agrobacterium* phage 7-7-1, a flagellotropic phage (21).

The Milano genome is organized in a modular fashion, with predicted structural, replication, and lysis proteins grouped together. The predicted tape measure protein (GenBank accession number QBQ72047) is preceded by the likely tail assembly chaperones (QBQ72045) with a putative frameshifted protein product (QBQ72046), analogous to the well-studied lambda G/GT chaperone system (22). The predicted endolysin (QBQ72055), i-spanin (QBQ72056), and embedded o-spanin (QCQ78506) are encoded consecutively; however, the holin was not identified. Additionally, a putative nucleoid occlusion-like protein (QBQ72073) with a ParB domain (InterProScan IPR003115) and many BLASTp hits to Noc proteins and two potential ribosome modulation factor domain superfamily proteins (QBQ72082 and QBQ72098) were found.

Data availability. The genome sequence and associated data for phage Milano were deposited under GenBank accession number MK637516, BioProject accession number PRJNA222858, SRA accession number SRR8869236, and BioSample accession number SAMN11360273.

ACKNOWLEDGMENTS

This work was supported primarily with funding from the National Science Foundation (awards EF-0949351 and DBI-1565146) and the Texas A&M AgriLife Research Zebra Chip Disease Program (grant number ZC 06-407697) to C.F.G. The Center for Phage Technology received an Initial University Multidisciplinary Research Initiative from Texas A&M University and Texas AgriLife.

We are grateful to the Texas A&M Department of Biochemistry and Biophysics and the CPT staff.

This announcement was prepared in partial fulfillment of the requirements for BICH464 Bacteriophage Genomics, an undergraduate course at Texas A&M University.

REFERENCES

1. Wood DW, Setubal JC, Kaul R, Monks DE, Kitajima JP, Okura VK, Zhou Y, Chen L, Wood GE, Almeida NF, Woo L, Chen Y, Paulsen IT, Eisen JA, Karp PD, Bovee D, Chapman P, Clendenning J, Deatherage G, Gillet W, Grant C, Kutyavin T, Levy R, Li MJ, McClelland E, Palmieri A, Raymond C, Rouse G, Saenphimmachak C, Wu Z, Romero P, Gordon D, Zhang S, Yoo H, Tao Y, Biddle P, Jung M, Krespan W, Perry M, Gordon-Kamm B, Liao L, Kim S, Hendrick C, Zhao ZY, Dolan M, Chumley F, Tingey SV, Tomb JF, Gordon MP, Olson MV, Nester EW. 2001. The genome of the natural genetic engineer Agrobacterium tumefaciens C58. Science 294:2317–2323. https://doi.org/10.1126/science.1066804.
2. Vidaver AK. 1967. Synthetic and complex media for the rapid detection of fluorescence of phytopathogenic pseudomonads: effect of the carbon source. Appl Microbiol 15:1523–1524.
3. Adams MH. 1956. Bacteriophages. Interscience Publishers, Inc., New York, NY.
4. Sambrook J, Russell DR. 2006. Extraction of bacteriophage λ DNA from large-scale preparations using proteinase K and SDS, p 98–99. *In* The condensed protocols from molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
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. 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.
9. Lee E, Helt GA, Reese JT, Munoz-Torres MC, Childers CP, Buels RM, Stein L, Holmes IH, Elsik CG, Lewis SE. 2013. Web Apollo: a Web-based genomic annotation editing platform. Genome Biol 14:R93. https://doi.org/10.1186/gb-2013-14-8-r93.
10. Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Cech M, Chilton J, Clements D, Coraor N, Grüning BA, Guerler A, Hillman-Jackson J, Hiltemann S, Jalili V, Rasche H, Soranzo N, Goecks J, Taylor J, Nekrutenko A, Blankenberg D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res 46:W537–W544. https://doi.org/10.1093/nar/gky379.
11. 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.
12. 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.
13. Juncker AS, Willenbrock H, Heijne Von 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. Krogh A, Larsson B, Heijne von G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: appli-

cation to complete genomes. J Mol Biol 305:567–580. https://doi.org/10.1006/jmbi.2000.4315.
15. 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.
16. The UniProt Consortium. 2018. UniProt: the universal protein knowledgebase. Nucleic Acids Res 46:2699. https://doi.org/10.1093/nar/gky092.
17. 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.
18. Valentine RC, Shapiro BM, Stadtman ER. 1968. Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from Escherichia coli. Biochemistry 7:2143–2152. https://doi.org/10.1021/bi00846a017.
19. 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.
20. Kropinski AM, Van den Bossche A, Lavigne R, Noben J-P, Babinger P, Schmitt R. 2012. Genome and proteome analysis of 7-7-1, a flagellotropic phage infecting Agrobacterium sp. H13-3. Virol J 9:102. https://doi.org/10.1186/1743-422X-9-102.
21. 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.
22. Xu J, Hendrix RW, Duda RL. 2014. Chaperone-protein interactions that mediate assembly of the bacteriophage lambda tail to the correct length. J Mol Biol 426:1004–1018. https://doi.org/10.1016/j.jmb.2013.06.040.