



Complete Genome Sequence of EtG, the First Phage Sequenced from *Erwinia tracheiphila*

Andrés Andrade-Domínguez,^a Roberto Kolter,^a Lori R. Shapiro^{a,b}

^aDepartment of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA

^bDepartment of Applied Ecology, North Carolina State University, Raleigh, North Carolina, USA

ABSTRACT *Erwinia tracheiphila* is the causal agent of bacterial wilt of cucurbits. Here, we report the genome sequence of the temperate phage EtG, which was isolated from an *E. tracheiphila*-infected cucumber plant. Phage EtG has a linear 30,413-bp double-stranded DNA genome with cohesive ends and 45 predicted open reading frames.

Erwinia tracheiphila (family *Enterobacteriaceae*) is an economically important phytopathogen that infects cultivated *Cucumis melo* (muskmelon), *Cucumis sativus* (cucumber), and *Cucurbita* spp. (squash, pumpkin, and gourds) in temperate northeastern and midwestern North America (1). The reference genome (2) shows structural changes associated with its recent emergence into a novel ecological niche (3). This includes the absence of clustered regularly interspaced short palindromic repeat (CRISPR) loci associated with *cas* genes, as well as an extraordinarily high number of predicted prophage regions (4). The high number of predicted prophages suggests that polylysogeny may be important for *E. tracheiphila* pathogenicity and/or ecological interactions.

Here, we sought to isolate phages induced *in vivo* from an *E. tracheiphila*-infected cucumber plant. In this infected cucumber plant, we found a phage that we have named EtG. EtG is able to infect 8 of 30 *E. tracheiphila* environmental isolates. A strain of *E. tracheiphila* carrying EtG was isolated from the same plant from which we recovered the phage EtG (unpublished data), suggesting that phage EtG is a temperate phage. The genome of phage EtG was sequenced with the Illumina MiSeq platform to ~100× coverage. The 100-bp reads were *de novo* assembled into a single contig using Velvet (5), and phage genome ends were determined through closure PCR and Sanger sequencing.

Annotation of the open reading frames was performed with the Rapid Annotations using Subsystems Technology (RAST) (6) and PHAge Search Tool (PHAST) (7) servers, which predicted 45 coding sequences (CDSs). Of these, 34 CDSs were assigned a predicted function and 9 have no assigned function. Sequence similarity searches were performed with the translation of each predicted CDS against the NCBI protein database using BLASTp (8) in order to assign putative protein functions. tRNAscan-SE (9) was used to search for tRNAs, but none were found.

Phage EtG has a linear double-stranded DNA genome of 30,413 bp, with 19 base-cohesive ends and a G+C content of 54.1%. The phage EtG genome showed complete collinearity and 66.9% overall nucleotide identity to *Escherichia virus 186* (GenBank accession number NC_001317). The two subunits of the EtG terminase (EtG_02 and EtG_03) share 98.8% amino acid identity with the *Escherichia virus 186* terminase subunits, suggesting a common packaging process.

The genome of phage EtG contains a protein for DNA replication (EtG_43) and an integrase gene (EtG_33). For DNA packaging, phage EtG has a packaging protein (EtG_01) and two subunits of terminase (EtG_02 and EtG_03). There were 17 genes identified for the head, baseplate, and tail morphogenesis. Other notable genes are the

Received 31 January 2018 Accepted 1 February 2018 Published 22 February 2018

Citation Andrade-Domínguez A, Kolter R, Shapiro LR. 2018. Complete genome sequence of EtG, the first phage sequenced from *Erwinia tracheiphila*. *Genome Announc* 6:e00127-18. <https://doi.org/10.1128/genomeA.00127-18>.

Copyright © 2018 Andrade-Domínguez 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 Lori R. Shapiro, lori.r.shapiro@gmail.com.

repressor protein CI for the maintenance of immunity (EtG_34), the regulatory protein CII, required for the establishment of lysogeny (EtG_36), and the *Apl* repressor (EtG_35) that down-regulates lytic transcription. The genes for host-cell lysis are spanins, holins, and a lysozyme (EtG_10). This is the first of dozens of *Erwinia tracheiphila* phage genomes, many of which are likely to be novel and likely to impact the host's ecology.

Accession number(s). The complete genome sequence of *E. tracheiphila* phage EtG was deposited in GenBank under the accession number [MF276773](https://doi.org/10.1093/nar/gkz366).

ACKNOWLEDGMENTS

This study was supported by postdoctoral research fellowships to A.A.-D. from the Consejo Nacional de Ciencia y Tecnología (Mexico) and the Fundación México en Harvard, A.C.; from NIH grant GM58213 to R.K.; and from an NSF postdoctoral fellowship (DBI-1202736) to L.R.S.

REFERENCES

1. Rojas ES, Batzer JC, Beattie GA, Fleischer SJ, Shapiro LR, Williams MA, Bessin R, Bruton BD, Boucher TJ, Jesse LCH, Gleason ML. 2015. Bacterial wilt of cucurbits: resurrecting a classic pathosystem. *Plant Dis* 99:564–574. <https://doi.org/10.1094/PDIS-10-14-1068-FE>.
2. Shapiro LR, Scully ED, Roberts D, Straub TJ, Geib SM, Park J, Stephenson AG, Salaau Rojas E, Liu Q, Beattie G, Gleason M, De Moraes CM, Mescher MC, Fleischer SG, Kolter R, Pierce N, Zhaxybayeva O. 2015. Draft genome sequence of *Erwinia tracheiphila*, an economically important bacterial pathogen of cucurbits. *Genome Announc* 3(3):e00482-15. <https://doi.org/10.1128/genomeA.00482-15>.
3. Moran NA, Plague GR. 2004. Genomic changes following host restriction in bacteria. *Curr Opin Genet Dev* 14:627–633. <https://doi.org/10.1016/j.gde.2004.09.003>.
4. Shapiro RL, Scully ED, Straub TJ, Park J, Stephenson AG, Beattie GA, Gleason ML, Kolter R, Coelho MC, De Moraes CM, Mescher MC, Zhaxybayeva O. 2016. Horizontal gene acquisitions, mobile element proliferation, and genome decay in the host-restricted plant pathogen *Erwinia tracheiphila*. *Genome Biol Evol* 8:649–664. <https://doi.org/10.1093/gbe/evw016>.
5. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* 18:821–829. <https://doi.org/10.1101/gr.074492.107>.
6. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. *BMC Genomics* 9:75. <https://doi.org/10.1186/1471-2164-9-75>.
7. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Res* 39:W347–W352. <https://doi.org/10.1093/nar/gkr485>.
8. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2).
9. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS Web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* 33:W686–W689. <https://doi.org/10.1093/nar/gki366>.