

# Complete Genome Sequence of Enterotoxigenic *Escherichia coli* N4-Like Podophage Pollock

Roosheel S. Patel, Lauren E. Lessor, Adriana C. Hernandez, Gabriel F. Kutty Everett

Center for Phage Technology, Texas A&M University, College Station, Texas, USA

**Enterotoxigenic *Escherichia coli* is a multidrug-resistant bacterium that is well known for its ability to cause diarrhea in humans. Bacteriophages may be used to treat clinical cases involving bacterial dysentery. Here, we present the complete genome sequence of an enterotoxigenic *E. coli* phage, Pollock, an N4-like podophage.**

Received 1 December 2014 Accepted 18 December 2014 Published 29 January 2015

**Citation** Patel RS, Lessor LE, Hernandez AC, Kutty Everett GF. 2015. Complete genome sequence of enterotoxigenic *Escherichia coli* N4-like podophage Pollock. *Genome Announc* 3(1):e01431-14. doi:10.1128/genomeA.01431-14.

**Copyright** © 2015 Patel et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Gabriel F. Kutty Everett, [gabbyeverett@tamu.edu](mailto:gabbyeverett@tamu.edu).

Enterotoxigenic *Escherichia coli* (ETEC) is a pathogenic type of *E. coli* associated with food/water contamination-related traveler's diarrhea (1). With the rise of bacteria that are resistant to antibiotics, the need for alternative treatments is growing (2). One such alternative is the therapeutic use of virulent bacteriophages. Pollock is a newly isolated virulent N4-like podophage that has the potential to be used as a treatment against ETEC.

Bacteriophage Pollock was isolated from a sewage sample collected in College Station, TX. The phage DNA was sequenced using 454 pyrosequencing at the Emory GRA Genome Center (Emory University, GA). The trimmed FLX Titanium reads were assembled to a single contig at 157.7-fold coverage using the Newbler assembler, version 2.5.3 (454 Life Sciences), at the default settings. The contig was confirmed to be complete by PCR using primers that face the upstream and downstream ends of the phage DNA. The products from the PCR amplification of the junctions of concatemeric molecules were sequenced by Sanger sequencing (Eton Bioscience, San Diego, CA). The genes were predicted using GeneMarkS (3) and corrected using the software tools available on the Center for Phage Technology (CPT) Galaxy instance (<https://cpt.tamu.edu/galaxy-public/>). Morphology was determined using transmission electron microscopy performed at the Texas A&M University Microscopy and Imaging Center.

Pollock has a unit genome of 67,755 bp, including 84 predicted coding sequences, a G+C content of 36%, and a coding density of 93%. Four tRNA-coding sequences were also found. An analysis of the raw sequencing reads using PAUSE (<https://cpt.tamu.edu/computer-resources/pause/>) shows that Pollock has a 610-bp terminal repeat. Pollock is syntenic with *E. coli* podophage N4 (accession no. NC\_008720) and has 49.9% identity to N4 via Emboss Stretcher analysis (4, 5).

Pollock encodes N4 homologs involved in DNA replication, transcription, DNA packaging, and morphogenesis. The defining characteristics of N4 are its large virion RNA polymerase (vRNAP), RNA polymerase II (RNAP II), and a single-stranded DNA binding protein (SSB) (6, 7). Pollock encodes the homologs of these proteins, suggesting that it accomplishes DNA injection and early gene expression by a similar mechanism. Unlike N4, how-

ever, the DNA polymerase of Pollock contains an intron-encoded HNH endonuclease (IPR003615) interrupting the coding sequence. The lysis cassette of Pollock is slightly different from that of N4, in that instead of a signal-anchor-release (SAR) endolysin, as encoded by N4 (*N*-acetylmuramidase), Pollock encodes a cytoplasmic endolysin (glycoside hydrolase family 24) (8, 9). Other lysis genes identified in Pollock include a holin/antiholin pair and an embedded inner/outer spanin pair (10).

Curiously, Pollock also encodes a homolog of the bacteriophage P1 protein, TciA (tellurite or colicin resistance or inhibition of cell division) (11). P1 *tciA* is homologous to *terB*, a gene from the *ter* gene cluster on the incompatibility subgroup IncHI2 R478 multiresistance plasmid. The *ter* genes are thought to confer *E. coli* resistance to various stresses, including channel-forming colicins and the propagation of certain phages (12, 13). How this protein plays a role in a presumably lytic N4-like phage is unknown.

**Nucleotide sequence accession number.** The genome sequence of phage Pollock was deposited in GenBank under the accession no. [KM236242](https://www.ncbi.nlm.nih.gov/nuccore/KM236242).

## ACKNOWLEDGMENTS

This work was supported primarily by funding from award EF-0949351, "Whole phage genomics: a student-based approach," from the National Science Foundation. Additional support came from the Center for Phage Technology, an Initial University Multidisciplinary Research Initiative supported by Texas A&M University and Texas AgriLife, and from the Department of Biochemistry and Biophysics.

We thank the CPT staff for their advice and support.

This announcement was prepared in partial fulfillment of the requirements for Bich464 Phage Genomics, an undergraduate course at Texas A&M University.

## REFERENCES

1. Qadri F, Svennerholm AM, Faruque AS, Sack RB. 2005. Enterotoxigenic *Escherichia coli* in developing countries: epidemiology, microbiology, clinical features, treatment, and prevention. *Clin Microbiol Rev* 18: 465–483. <http://dx.doi.org/10.1128/CMR.18.3.465-483.2005>.
2. Sulakvelidze A, Alavidze Z, Morris JG, Jr. 2001. Bacteriophage therapy. *Antimicrob Agents Chemother* 45:649–659. <http://dx.doi.org/10.1128/AAC.45.3.649-659.2001>.

3. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* 29:2607–2618. <http://dx.doi.org/10.1093/nar/29.12.2607>.
4. Myers EW, Miller W. 1988. Optimal alignments in linear space. *Comput Appl Biosci* 4:11–17. <http://dx.doi.org/10.1093/bioinformatics/4.1.11>.
5. Grose JH, Casjens SR. 2014. Understanding the enormous diversity of bacteriophages: the tailed phages that infect the bacterial family *Enterobacteriaceae*. *Virology* 468–470C:421–443. <http://dx.doi.org/10.1016/j.virol.2014.08.024>.
6. Zehring WA, Rothman-Denes LB. 1983. Purification and characterization of coliphage N4 RNA polymerase II activity from infected cell extracts. *J Biol Chem* 258:8074–8080.
7. Choi KH, McPartland J, Kaganman I, Bowman VD, Rothman-Denes LB, Rossmann MG. 2008. Insight into DNA and protein transport in double-stranded DNA viruses: the structure of bacteriophage N4. *J Mol Biol* 378:726–736. <http://dx.doi.org/10.1016/j.jmb.2008.02.059>.
8. Stojković EA, Rothman-Denes LB. 2007. Coliphage N4 N-acetylmuramidase defines a new family of murein hydrolases. *J Mol Biol* 366:406–419. <http://dx.doi.org/10.1016/j.jmb.2006.11.028>.
9. Kutyl GF, Xu M, Struck DK, Summer EJ, Young R. 2010. Regulation of a phage endolysin by disulfide caging. *J Bacteriol* 192:5682–5687. <http://dx.doi.org/10.1128/JB.00674-10>.
10. Summer EJ, Berry J, Tran TA, Niu L, Struck DK, Young R. 2007. Rz/Rz1 lysis gene equivalents in phages of Gram-negative hosts. *J Mol Biol* 373:1098–1112. <http://dx.doi.org/10.1016/j.jmb.2007.08.045>.
11. Lobočka MB, Rose DJ, Plunkett G III, Rusin M, Samojedny A, Lehnher H, Yarmolinsky MB, Blattner FR. 2004. Genome of bacteriophage P1. *J Bacteriol* 186:7032–7068. <http://dx.doi.org/10.1128/JB.186.21.7032-7068.2004>.
12. Whelan KF, Sherburne RK, Taylor DE. 1997. Characterization of a region of the IncHI2 plasmid R478 which protects *Escherichia coli* from toxic effects specified by components of the tellurite, phage, and colicin resistance cluster. *J Bacteriol* 179:63–71.
13. Gilmour MW, Thomson NR, Sanders M, Parkhill J, Taylor DE. 2004. The complete nucleotide sequence of the resistance plasmid R478: defining the backbone components of incompatibility group H conjugative plasmids through comparative genomics. *Plasmid* 52:182–202. <http://dx.doi.org/10.1016/j.plasmid.2004.06.006>.