genomeA
Journals.ASM.org

Complete Genome Sequence of Enterotoxigenic *Escherichia coli* Siphophage Seurat

Dung P. Doan, Lauren E. Lessor, Adriana C. Hernandez, Gabriel F. Kuty Everett

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

Enterotoxigenic *Escherichia coli* (ETEC) is one of the leading causes of diarrhea in developing countries. Bacteriophage therapy has the potential to aid in the prevention and treatment of ETEC-related illness. To that end, we present here the complete genome of ETEC siphophage Seurat and describe its major features.

Received 14 January 2015 **Accepted** 16 January 2015 **Published** 26 February 2015

Citation Doan DP, Lessor LE, Hernandez AC, Kuty Everett GF. 2015. Complete genome sequence of enterotoxigenic *Escherichia coli* siphophage Seurat. Genome Announc 3(1): e00044-15. doi:10.1128/genomeA.00044-15.

Address correspondence to Gabriel F. Kuty Everett, gabbyeverett@tamu.edu.

Enterotoxigenic *Escherichia coli* (ETEC) is one of the leading causes for infectious gastroenteritis that results in approximately 120,800 annual deaths worldwide (1). The rise in antibiotic resistance has prompted the need for alternative treatments of pathogenic bacterial infections, such as bacteriophage therapy (2, 3). Here, we describe the complete genome of bacteriophage Seurat, a siphophage infecting ETEC. Genome analysis suggests Seurat is a lytic phage and therefore could play a role in phage therapy to reduce the number of deaths due to ETEC infection per year.

Seurat was isolated from a sewage sample collected in College Station, TX. Phage DNA was sequenced using 454 pyrosequencing at the Emory GRA Genome Center (Emory University, GA). Trimmed FLX Titanium reads were assembled to a single contig at 107.3-fold coverage using Newbler assembler, version 2.5.3 (454 Life Sciences) at default settings. The contig was confirmed to be complete by PCR using primers that face the upstream and downstream ends of the phage DNA. Products from the PCR amplification of the junctions of concatemeric molecules were sequenced by Sanger sequencing (Eton Bioscience, San Diego, CA). Genes were predicted using GeneMarkS (4) and corrected using software tools available on the Center for Phage Technology (CPT) Galaxy instance (https://cpt.tamu.edu/galaxy-public/). Transmission electron microscopy was performed at the Texas A&M University Microscopy and Imaging Center.

Seurat has a genome of 56,776 bp with 88 coding sequences, a 44.6% G+C content, and a 93.2% coding density. It shows 10.5% average nucleotide identity to *E. coli* phage 9g (GenBank accession no. NC_024146), which is not enough to justify clustering as described by Grose and Casjens (5). Seurat can therefore be designated a novel singleton cluster containing a single phage.

DNA synthesis machinery includes genes coding for DNA polymerase, sliding clamp subunit, helicase, and primase. Genes for morphogenesis proteins were identified (capsid, tape measure, tape measure chaperone, tail fibers). DNA packaging genes include *terS* and *terL*. Comparison of the Seurat TerL protein to TerL proteins of known packaging strategies suggests that Seurat uses a pac-headful DNA packaging mechanism. For annotation purposes, the genome has been opened to the small terminase gene (6). Seurat also encodes a dUTPase to ensure low levels of cellular dUTP through the hydrolysis of dUTP into dUMP and PP_i (7).

Seurat encodes queuosine biosynthesis proteins, QueC, QueD, QueE, and FolE. Queuosine, a hyper-modified guanine nucleoside, plays an important role in stabilizing the anticodon-codon interaction at the wobble position of certain tRNAs to improve translation accuracy (8). QueCDE and FolE are involved in the biosynthesis of queuosine precursors $preQ_0$ and $preQ_1$ (9, 10). The most complete queuosine biosynthesis pathway reported in a bacteriophage thus far has been in the *Streptococcus pneumoniae* phage Dp-1 (NC_015274) (11). Comparison of the queuosine pathways between the two phages shows that Seurat is missing QueT, a transporter, and QueF, a precursor enzyme.

Nucleotide sequence accession number. The genome sequence of phage Seurat was contributed as accession no. KM236243 to GenBank.

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 are grateful for the advice and support of the CPT staff.

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

REFERENCES

1. **Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Barker-Collo S, Bartels DH, Bell ML, Benjamin EJ, Bennett D, Bhalla K, Bikbov B, Bin Abdulhak A, Birbeck G, Blyth F, Bolliger I, Boufous S, Bucello C, Burch M, Burney P, Carapetis J, Chen H, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC,**

Criqui MH, Cross M, et al. 2012. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet **380:**2095–2128. http://dx.doi.org/10.1016/S0140-6736(12)61728-0.

2. **Tadesse DA, Zhao S, Tong E, Ayers S, Singh A, Bartholomew MJ, McDermott PF.** 2012. Antimicrobial drug resistance in *Escherichia coli* from humans and food animals, United States, 1950–2002. Emerg Infect Dis **18:**741–749. http://dx.doi.org/10.3201/eid1805.111153.
3. **Abedon ST, Kuhl SJ, Blasdel BG, Kutter EM.** 2011. Phage treatment of human infections. Bacteriophage **1:**66–85. http://dx.doi.org/10.4161/bact.1.2.15845.
4. **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.
5. **Grose JH, Casjens SR.** 2014. Understanding the enormous diversity of bacteriophages: the tailed phages that infect the bacterial family *Enterobacteriaceae*. Virology **468–470:**421–443. http://dx.doi.org/10.1016/ j.virol.2014.08.024.
6. **Casjens SR, Gilcrease EB.** 2009. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. Methods Mol Biol **502:**91–111. http://dx.doi.org/10.1007/978-1-60327-565-1_7.
7. **Castillo-Acosta VM, Estévez AM, Vidal AE, Ruiz-Perez LM, González-Pacanowska D.** 2008. Depletion of dimeric all-alpha dUTPase induces DNA strand breaks and impairs cell cycle progression in *Trypanosoma brucei*. Int J Biochem Cell Biol **40:**2901–2913. http://dx.doi.org/10.1016/j.biocel.2008.06.009.
8. **Eichhorn CD, Kang M, Feigon J.** 2014. Structure and function of preQ$_1$ riboswitches. Biochim Biophys Acta **1839:**939–950. http://dx.doi.org/10.1016/j.bbagrm.2014.04.019.
9. **Phillips G, El Yacoubi B, Lyons B, Alvarez S, Iwata-Reuyl D, de Crécy-Lagard V.** 2008. Biosynthesis of 7-deazaguanosine-modified tRNA nucleosides: a new role for GTP cyclohydrolase I. J Bacteriol **190:**7876–7884. http://dx.doi.org/10.1128/JB.00874-08.
10. **Reader JS, Metzgar D, Schimmel P, de Crécy-Lagard V.** 2004. Identification of four genes necessary for biosynthesis of the modified nucleoside queuosine. J Biol Chem **279:**6280–6285. http://dx.doi.org/10.1074/jbc.M310858200.
11. **Sabri M, Häuser R, Ouellette M, Liu J, Dehbi M, Moeck G, García E, Titz B, Uetz P, Moineau S.** 2011. Genome annotation and intraviral interactome for the *Streptococcus pneumoniae* virulent phage Dp-1. J Bacteriol **193:**551–562. http://dx.doi.org/10.1128/JB.01117-10.