

Complete Genome Sequence of Bacteriophage EC6, Capable of Lysing *Escherichia coli* O157:H7

Birendra R. Tiwari, Jungmin Kim

Department of Microbiology, School of Medicine, Kyungpook National University, Daegu, Republic of Korea

A bacteriophage, EC6, capable of lysing *Escherichia coli* serotype O157:H7 was isolated and found to be a member of the family *Myoviridae*. The genomic sequence of phage EC6 was composed of 86,231 bp with a G+C content of 38.9% and is very similar to the sequences of *Escherichia* phage wV8 and *Salmonella* phage Felix O1.

Received 1 November 2012 Accepted 19 November 2012 Published 31 January 2013

Citation Tiwari BR, Kim J. 2013. Complete genome sequence of bacteriophage EC6, capable of lysing *Escherichia coli* O157:H7. *Genome Announc.* 1(1):e00085-12. doi:10.1128/genomeA.00085-12.

Copyright © 2013 Tiwari and Kim. This is an open-access article distributed under the terms of the [Attribution 3.0 Unported Creative Commons License](#).

Address correspondence to Jungmin Kim, minkim@knu.ac.kr.

Escherichia coli serotype O157:H7 is a major pathogen that causes diarrhea, colitis, and hemolytic-uremic syndrome (1, 2). We isolated a novel *E. coli* O157:H7 bacteriophage (phage), EC6, from a sewage sample. Several clinical isolates of *E. coli* were tested to determine the phage's host spectrum. Phage EC6 might form lytic plaques against sorbitol-nonfermenting *E. coli* but not against sorbitol-fermenting strains. Transmission electron microscopic images revealed that the EC6 phage belongs to the family *Myoviridae*. The sequencing of the phage EC6 genome was performed by combining the 454 and Illumina sequencing platforms using Genome Sequencer (GS) assembler 2.6 (Roche Diagnostics, Branford, CT) and CLC Genomics workbench 5.5 (CLCbio, Denmark), respectively. The resulting contigs were assembled through manual curation using CodonCode aligner v3.7.1 (CodonCode Corp., Dedham, MA). The final contigs were remapped with raw reads to correct errors, and dubious regions were masked as gaps. Genes, tRNA, and rRNA were identified using Glimmer, tRNAscan-SE, and HMMER programs along with EzTaxon-e rRNA profiles (3–6). The identified genes were annotated with RefSeq, catFam, Clusters of Orthologous Groups (COG), SEED, and GenBank databases (7–11). The genome of phage EC6 was composed of 86,231 bp with a G+C content of 38.9% and 136 open reading frames (ORFs) that encoded 16.9% functionally known proteins and 83.1% hypothetical proteins. In addition to genes for structural and general functions, the genome was composed of genes encoding a full set of enzymes related to replication, recombination, modification, transport, and metabolism. Based on COG database categories, proteins for the following functions can be predicted: structural and general functions (phage capsid and scaffolds, terminase, Mu-like prophage protein, putative phosphatase, muraminidase), replication, recombination, and repair (ATP-dependent DNA ligase, DNA polymerase I, replicative DNA helicase), transcription regulation (predictive transcriptional regulator), posttranslational modification (periplasmic serine protease homologous to putative head maturation protease of phage wV8, glutaredoxin, organic radical activating enzyme, molecular chaperone-HSP90 family), nucleotide transport and metabolism (thymidylate synthase, ribonucle-

otide reductase alpha-subunit, ribonucleotide reductase beta-subunit, oxygen-sensitive ribonucleoside-triphosphate reductase, phosphoribosyl pyrophosphate synthetase), coenzyme transport and metabolism (dihydrofolate reductase, nicotinic acid phosphoribosyl transferase), and an unknown function (putative phage Mu protein). In addition, the hypothetical protein AFU62408, which is adjacent to DNA polymerase, is 99% similar to the DNA polymerase of *Salmonella* phage SPT-1, which is likely the large subunit of DNA polymerase. Based on the functions of the encoded proteins, the phage EC6 genome is closely similar to that of *Escherichia* phage wV8 (GenBank accession no. NC_012749) (6) and that of *Salmonella* phage Felix O1 (GenBank accession no. NC_005282.1) (7). The phage EC6 genome differed from the genomes of both of these phages in its lack of genes encoding tail fiber proteins. However, the phage EC6 capsid and scaffold proteins are 97% similar to the tail fiber protein GP37 of phage wV8, with minor differences at the C-terminal region.

In conclusion, *E. coli* O157:H7-specific phage EC6 is a member of the *Myoviridae* family and is closely related to the *Escherichia* phage wV8 and *Salmonella* phage Felix O1.

Nucleotide sequence accession number. The complete genome sequence of phage EC6 has been stored in GenBank under the accession no. [JX560968](#).

ACKNOWLEDGMENTS

This research was supported by the Kyungpook National University Research Fund, 2012, and a grant from the Korea Healthcare Technology R&D Project of the Ministry for Health, Welfare, and Family Affairs, Republic of Korea (A084442).

REFERENCES

1. Tarr PI. 1995. *Escherichia coli* O157:H7: clinical, diagnostic, and epidemiological aspects of human infection. *Clin. Infect Dis.* 20(1):1–8.
2. Tauxe RV. 1997. Emerging foodborne diseases: an evolving public health challenge. *Emerg. Infect. Dis.* 3(4):4.
3. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* 23(6):673–679.
4. Eddy SR. 1998. Profile hidden Markov models. *Bioinformatics* 14(9):755–763.

5. Kim OS, Cho YJ, Lee K, Yoon SH, Kim M, Na H, Park SC, Jeon YS, Lee JH, Yi H, Won S, Chun J. 2012. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int. J. Syst. Evol. Microbiol.* 62(Pt 3):716–721.
6. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25(5):955–964.
7. Overbeek R, Begley T, Butler RM, Choudhuri JV, Chuang HY, Cohoon M, de Crécy-Lagard V, Diaz N, Disz T, Edwards R, Fonstein M, Frank ED, Gerdes S, Glass EM, Goesmann A, Hanson A, Iwata-Reuyl D, Jensen R, Jamshidi N, Krause L, Kubal M, Larsen N, Linke B, McHardy AC, Meyer F, Neuweger H, Olsen G, Olson R, Osterman A, Portnoy V, Pusch GD, Rodionov DA, Rückert C, Steiner J, Stevens R, Thiele I, Vassieva O, Ye Y, Zagnitko O, Vonstein V. 2005. The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* 33(17):5691–5702.
8. Pruitt KD, Tatusova T, Klimke W, Maglott DR. 2009. NCBI reference sequences: current status, policy and new initiatives. *Nucleic Acids Res.* 37(Database issue):D32–D36.
9. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28(1):33–36.
10. Yu C, Zavaljevski N, Desai V, Reifman J. 2009. Genome-wide enzyme annotation with precision control: catalytic families (CatFam) databases. *Proteins* 74(2):449–460.
11. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2011. GenBank. *Nucleic Acids Res.* 39(Database issue):D32–D37.