

# Complete Genome Sequence of *Citrobacter freundii* Myophage Merlin

Kayla C. LeSage, Emily C. Hargrove, Jesse L. Cahill, Eric S. Rasche, Gabriel F. Kutyl Everett

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

**Bacteriophage Merlin is a T4-like myophage which infects *Citrobacter freundii*, a member of the *Enterobacteriaceae* family. *C. freundii* is an opportunistic pathogen that is a common cause of nosocomial infections. This report announces the complete genome of myophage Merlin and describes its features.**

Received 17 August 2015 Accepted 19 August 2015 Published 1 October 2015

**Citation** LeSage KC, Hargrove EC, Cahill JL, Rasche ES, Kutyl Everett GF. 2015. Complete genome sequence of *Citrobacter freundii* myophage Merlin. *Genome Announc* 3(5): e01133-15. doi:10.1128/genomeA.01133-15.

**Copyright** © 2015 LeSage 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/4.0/).

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

*Citrobacter freundii* is an opportunistic pathogen that is commonly found in the environment, food, and the intestinal tracts of animals, including humans (1). *C. freundii* infections have been reported to be acquired nosocomially. The bacteria can cross the blood-brain barrier, which most commonly causes brain abscesses, neonatal meningitis, neonatal sepsis, and damage to the central nervous system, all of which have a high fatality rate (2). The persistence of *C. freundii* is primarily due to its antimicrobial resistance and imperfect hospital practices (1). To that end bacteriophages, particularly myophage Merlin, may be useful for the treatment of *C. freundii*.

Bacteriophage Merlin was isolated from a water sample collected in College Station, TX. Phage DNA was sequenced in an Illumina MiSeq 250-bp paired-end run with a 550-bp insert library at the Genomic Sequencing and Analysis Facility at the University of Texas (Austin, TX, USA). Quality controlled, trimmed reads were assembled into a single contig of circular assembly at 25.9-fold coverage using SPAdes version 3.5.0 (3). The contig was confirmed to be complete by PCR using primers that face the upstream and downstream ends of the contig. 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/>). Morphology was determined using transmission electron microscopy performed at the Texas A&M University Microscopy and Imaging Center.

Merlin is T4-like phage with a 172,733-bp genome that is circularly permuted. For annotation purposes Merlin was opened to the *rIIa* gene (5). Merlin and T4 have 55.9% nucleotide sequence identity across the genome, as determined by Emboss Stretcher (6). Most of the differences between the two phages are found in hypothetical proteins of unknown function. Merlin has 303 predicted coding sequences, of which 119 (39.3%) had homologs predicted by BLASTp and InterPro Scan analysis (7, 8), 183 were conserved hypothetical proteins (60.4%), and 1 was a hypothetical novel protein. The G+C content is 38.8% with a coding density of 94.4%. Eleven tRNAs were identified in phage Merlin compared to the 8 tRNAs present in T4 (5).

T4-like core genes involved in replication, recombination, DNA packaging, lysis, and virion morphogenesis were identified with few deviations. Merlin encodes no homing endonucleases, in contrast to the 14 homing endonucleases present in T4. Merlin contains no homologs to the *ipI*, *ipII*, or *ipIII* internal head proteins of T4 and no RepEA/B replication initiation protein homologs (5). A DksA/TraR domain protein that contains a Cys4 zinc finger motif was identified. This protein is prevalent in many T4-like phages but not present in T4 itself. The primase proteins of bacteriophages T4 and T7 have a Cys4 motif that mediates protein-DNA interactions; however, the function of the Cys 4 motif in the DksA/TraR protein described here is unknown (9).

**Nucleotide sequence accession number.** The genome sequence of phage Merlin was contributed to GenBank with the accession number [KT001915](https://www.ncbi.nlm.nih.gov/nuclseq/KT001915).

## ACKNOWLEDGMENTS

This work was supported primarily by funding from award number 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

- Wang JT, Chang SC, Chen YC, Luh KT. 2000. Comparison of antimicrobial susceptibility of *Citrobacter freundii* isolates in two different time periods. *J Microbiol Immunol Infect* 33:258–262.
- Badger JL, Stins MF, Kim KS. 1999. *Citrobacter freundii* invades and replicates in human brain microvascular endothelial cells. *Infect Immun* 67:4208–4215.
- 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. <http://dx.doi.org/10.1089/cmb.2012.0021>.
- 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. Miller ES, Kutter E, Mosig G, Arisaka F, Kunisawa T, Rüger W. 2003. Bacteriophage T4 genome. *Microbiol Mol Biol Rev* 67:86–156, table of contents. <http://dx.doi.org/10.1128/MMBR.67.1.86-156.2003>.
6. 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>.
7. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <http://dx.doi.org/10.1186/1471-2105-10-421>.
8. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laugraud A, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Sigrist CJ, Thimma M, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C. 2009. InterPro: the integrative protein signature database. *Nucleic Acids Res* 37: D211–D215. <http://dx.doi.org/10.1093/nar/gkn785>.
9. Kusakabe T, Hine AV, Hyberts SG, Richardson CC. 1999. The Cys4 zinc finger of bacteriophage T7 primase in sequence-specific single-stranded DNA recognition. *Proc Natl Acad Sci USA* 96:4295–4300. <http://dx.doi.org/10.1073/pnas.96.8.4295>.