

# Genome Sequence of the Enterohemorrhagic *Escherichia coli* Bacteriophage UFV-AREG1

Maryoris E. Soto Lopez,<sup>a</sup> Laís Silva Batalha,<sup>a</sup> Pedro Marcus Pereira Vidigal,<sup>b</sup> Luiz Augusto A. Albino,<sup>a</sup> Delaine Meireles Gouveia Boggione,<sup>a</sup> Marco Tulio Pardini Gontijo,<sup>a</sup> Denise M. Soares Bazzolli,<sup>c</sup> Regina C. Santos Mendonça<sup>a</sup>

Departamento de Tecnologia de Alimentos (DTA), Universidade Federal de Viçosa, Minas Gerais, Brazil<sup>a</sup>; Núcleo de Análise de Biomoléculas (NuBioMol), Centro de Ciências Biológicas (CCB), Universidade Federal de Viçosa, Minas Gerais, Brazil<sup>b</sup>; Departamento de Microbiologia (DMB), Universidade Federal de Viçosa, Minas Gerais, Brazil<sup>c</sup>

Here, we present the genome sequence of the *Escherichia coli* bacteriophage UFV-AREG1. This phage was isolated from cowshed wastewater and showed specificity for enterohemorrhagic *E. coli* O157:H7 (ATCC 43895), *E. coli* 0111 (CDC O11ab) and *E. coli* (ATCC 23229).

Received 1 April 2016 Accepted 26 August 2016 Published 13 October 2016

Citation Lopez MES, Batalha LS, Vidigal PMP, Albino LAA, Boggione DMG, Gontijo MTP, Bazzolli DMS, Mendonça RCS. 2016. Genome sequence of the enterohemorrhagic *Escherichia coli* bacteriophage UFV-AREG1. *Genome Announc* 4(5):e00412-16. doi:10.1128/genomeA.00412-16.

Copyright © 2016 Lopez 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 Maryoris E. Soto Lopez, [dramicrocta@hotmail.com](mailto:dramicrocta@hotmail.com).

In general, bacteriophages show great diversity, morphology, biochemistry, and genetic content (1). The study of the T4-like bacteriophage genomic sequences has provided countless contributions for these fields, and specially to functional genomics and proteomics (2, 3).

The bacteriophage UFV-AREG1 was isolated from cowshed wastewater showing specificity to *Escherichia coli* O157:H7 (CDC EDL-933), *E. coli* 0111 (CDC O11ab), and *E. coli* (ATCC 23229). Bacteriophage UFV-AREG1 can reach concentrations of over 10<sup>12</sup> PFU/ml in laboratory conditions when the host is *E. coli* O157:H7. Transmission electron microscopy showed a phage with an icosahedral head of 114 by 86 nm and a contractile tail of 117 by 23 nm. These parameters allow the classification of UFV-AREG1 as a member of the *Myoviridae* family (4, 5). UFV-AREG1 can remain viable after simulation of sanitization processes with sodium dichloroisocyanurate (219 ± 13 mg/l) and hydrogen peroxide (58.1 ± 5 mg/l). On the other hand, this phage lost viability after contact with peracetic acid (91 ± 17 mg/l).

The genome of UFV-AREG1 was sequenced using Illumina HiSeq by Macrogen (Seoul, Republic of Korea), yielding a total of 10,319,880 paired-end reads 100 bp in length. The sequenced reads were trimmed for quality (minimum Q30 score), and the genome sequence was assembled with a coverage of 5,600× using Geneious version 9.1.2 (Biomatters) with the genome of *Escherichia* phage HY01 (KF925357) as the reference. Genes were predicted from the genome using Prodigal version 2.50 (6), which revealed 274 open reading frames (ORFs). The proteins encoded by these ORFs were functionally annotated using BLAST searches (<http://blast.ncbi.nlm.nih.gov>) against the GenBank and UniProt databases.

The UFV-AREG1 genome comprises 170,788 bp of a linear dsDNA with a density of 1.60 gene/Kb and a G+C content of 35.3%, which is within the range observed in T4-like phages (35.0% to 43.0%). UFV-AREG1 is more similar to *Escherichia* phages HY01 and PEC4 (KR233165), sharing 91.8% and 91.5%

nucleotide sequence identities, respectively. The 274 predicted proteins show high similarity to previously described enteric phages and other T4-like phages. As a T4-like phage, the genome of UFV-AREG1 was also opened to the *rIIa* gene. The BLASTx searches showed that the genome of UFV-AREG1 differs from the *Enterobacteria* T4 (NC\_000866) phage in 30 ORFs. Most of the proteins encoded by these ORFs do not have putative conserved domains. Among these proteins, four annotated ones in UFV-AREG1 related to the T4 phage were found: one receptor recognition protein (ORF 244), one transferase protein (ORF 42), and two structural proteins (ORFs 29 and 121).

The best understanding of the gene and proteins encoded by the UFV-AREG1 bacteriophage can provide new information about *E. coli* phages and their genetic diversity.

**Accession number(s).** The genome sequence of UFV-AREG1 is available in GenBank under the accession number [KX009778](https://www.ncbi.nlm.nih.gov/nuclot/KX009778).

## ACKNOWLEDGMENTS

We thank the Núcleo de Análise de Biomoléculas (NuBioMol; <http://www.nubiomol.ufv.br>) of the Universidade Federal de Viçosa for supporting the data analysis.

The first author thanks CAPES for scholarship PEC-PG. This work was funded by the Departamento de Tecnologia de Alimentos (DTA), Universidade Federal de Viçosa, Minas Gerais—Brazil.

## FUNDING INFORMATION

This work, including the efforts of Maryoris Elisa Soto Lopez, was funded by CAPES (scholarship).

## REFERENCES

- Comeau AM, Bertrand C, Letarov A, Tétart F, Krisch HM. 2007. Modular architecture of the T4 phage superfamily: a conserved core genome and a plastic periphery. *Virology* 362:384–396. <http://dx.doi.org/10.1016/j.virol.2006.12.031>.
- Shahrbabak SS, Khodabandehlou Z, Shahverdi AR, Skurnik M, Ackermann HW, Varjosalo M, Yazdi MT, Sepehrizadeh Z. 2013. Isolation, characterization and complete genome sequence of Phax1: a phage of *Esch-*

- erichia coli* O157:H7. Microbiology 159:1629–1638 <http://dx.doi.org/10.1099/mic.0.063776-0>.
3. Miller ES, Kutter E, Mosig G, Arisaka F, Kunisawa T, R ger W. 2003. Bacteriophage T4 genome. Microbiol Mol Biol Rev 67:86–156. <http://dx.doi.org/10.1128/MMBR.67.1.86-156.2003>.
  4. Yaung SJ, Esvelt KM, Church GM. 2015. Complete genome sequences of T4-like bacteriophages RB3, RB5, RB6, RB7, RB9, RB10, RB27, RB33, RB55, RB59, and RB68. Genome Announc 3(1):e01122-14. <http://dx.doi.org/10.1128/genomeA.01122-14>.
  5. Coffey B, Ross RP, O’Flynn G, O’Sullivan O, Casey A, Callanan M, Coffey A, McAuliffe O. 2014. Complete genome sequence of vB\_E-coM\_112, a T-even-type bacteriophage specific for *Escherichia coli* O157:H7. Genome Announc 2(6):e00393-14. <http://dx.doi.org/10.1128/genomeA.00393-14>.
  6. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. <http://dx.doi.org/10.1186/1471-2105-11-119>.