Complete Genome Sequences of Novel Bovine T4, rv5-Like, and Dhillonviruses Effective against *Escherichia coli* O157

Domonkos Sváb,[a] Linda Falgenhauer,[b,c] Trinad Chakraborty,[c,d] István Tóth[a]

[a]Institute for Veterinary Medical Research, Centre for Agricultural Research, Budapest, Hungary
[b]Institute of Hygiene and Environmental Medicine, Justus Liebig University Giessen, Giessen, Germany
[c]German Centre for Infection Research, Site Giessen-Marburg-Langen, Giessen, Germany
[d]Institute for Medical Microbiology, Justus Liebig University Giessen, Giessen, Germany

ABSTRACT Enterohemorrhagic *Escherichia coli* (EHEC) isolates of serotype O157:H7 are serious foodborne zoonotic pathogens and prime targets for biocontrol using bacteriophages. We report on the complete genome sequences of 11 novel lytic bacteriophages, representing three viral genera, isolated from cattle in Hungary that target *E. coli* O157 strains.

The role of bacteriophages as alternative antimicrobial agents has gained renewed interest in the past years (1). Enterohemorrhagic *Escherichia coli* (EHEC) strains, particularly of the O157:H7 serotype, are serious foodborne zoonotic pathogens, and several phages which show effective lysis of these have previously been characterized (2, 3). Presently, there are authorized products for use against EHEC as biocontrol in foodstuffs (EcoShield [4]). Nevertheless, a strategy for characterizing new phage isolates that target the eradication of these pathogens from various environments remains prudent.

Here, we present the genome sequences of 11 new lytic bacteriophages (Table 1). Nine phages originated from fecal samples of healthy cattle, the main reservoir of EHEC (5), were isolated from dairy and meat-producing farms in Hungary, using the EHEC O157:H7 strain Sakai for propagation as described earlier (3). Two of the phages were isolated from a sample taken from the cattle farm environment using the same protocol. Briefly, after precultivation of the samples in 10 volumes of tryptic soy broth with bile salts overnight at 37°C under aerobic conditions, the bacterium-free filtrate was spotted onto layered Luria-Bertani agar plates with the propagating strain. All phages were active against the reference strains EHEC O157:H7 Sakai (6) and EDL933 (7); the T4-like and rV5-like phages lysed several bovine *E. coli* O157 strains representing EHEC, enteropathogenic *E. coli* (EPEC), and atypical pathotypes isolated earlier in Hungary, as well (8).

Phage DNA from active lytic stocks was isolated using the phenol-chloroform method (9) and concentrated with the DNA Clean and Concentrator kit (Zymo Research, CA, USA) according to the manufacturer's instructions. Genomic DNA sequencing libraries were prepared using the Nextera XT kit (Illumina, Eindhoven, Netherlands). Sequencing was performed using the NextSeq midoutput reagent kit v2.5 (2×150 bp) on an Illumina NextSeq 500 machine. Coverage was between $71\times$ and $3,830\times$, with an average of $1,298\times$ (Table 1). Quality control and clipping of the reads were performed within the framework of the ASA³P pipeline (10). ASA³P quality control/read clipping includes the programs FastQC (https://github.com/s-andrews/FastQC), FastQ Screen (https://www.bioinformatics.babraham.ac.uk/projects/fastq_screen), and Trimmomatic (11). The types of genome termini were determined using PhageTerm with default parameters (12). All phages depicted a linear topology (Table 1). Assembly was performed with SPAdes

TABLE 1 List of phage genomes announced in the current study with their main characteristics

Phage	GenBank accession no.	Raw read data accession no.	Total no. of reads	Sequence length (nt)[a]	Coverage (×)	Topology	Termini	Predicted genus
vb_EcoM_bov9_1	MT884006	SRR13132432	10,168,954	166,440	494.836	Linear	Fixed	*Tequatrovirus*
vb_EcoM_bov10K1[b]	MT884007	SRR13132431	9,085,150	166,441	78.3038	Linear	Fixed	*Tequatrovirus*
vb_EcoM_bov10K2[b]	MT884008	SRR13132431	9,085,150	135,960	2,261.21	Linear	Circular permutated	*Vequintavirus*
vb_EcoM_bov11CS3	MT884009	SRR13132430	9,449,636	135,960	1,724.25	Linear	Circular permutated	*Vequintavirus*
vb_EcoM_bov22_2[c]	MT884010	SRR13132429	3,442,978	135,961	983.78	Linear	Circular permutated	*Vequintavirus*
vb_EcoM_bov25_3[d]	MT884011	SRR13132428	11,006,422	135,961	3,327.32	Linear	Circular permutated	*Vequintavirus*
vb_EcoS_bov11C2	MT884012	SRR13132430	9,449,636	44,612	3,830.02	Linear	Circular permutated	*Dhillonvirus*
vb_EcoS_bov16_1	MT884013	SRR13132427	15,936,284	44,745	527.902	Linear	Circular permutated	*Dhillonvirus*
vb_EcoS_bov22_1[c]	MT884014	SRR13132429	3,442,978	44,612	71.37	Linear	Circular permutated	*Dhillonvirus*
vb_EcoS_bov25_1D[d]	MT884015	SRR13132428	11,006,422	44,747	374.845	Linear	Fixed	*Dhillonvirus*
vB_EcoM_bov15_1	MT951623	SRR13132426	1,856,660	44,700	607.948	Linear	Circular permutated	*Dhillonvirus*

[a] nt, nucleotides.
[b] These phages were isolated from the same environmental sample.
[c] These phages were isolated from the same bovine fecal sample.
[d] These phages were isolated from the same bovine fecal sample.

3.10.1 (13). Homology searches were conducted against GenBank with the megaBLAST algorithm, available at the NCBI website; annotation was performed with the RAST server using the RASTtk algorithm with default parameters (14). The sequencing showed that some stocks contained multiple phages; therefore, purification of single plaques followed, and PCR for marker genes with primers specifically designed (data not shown) was used to separate single-phage stocks. Phages originating from the same sample are marked as such in Table 1.

Phages vb_EcoM_bov9_1 and vb_EcoM_bov10K1 proved to be T4-like phages, representing the *Tequatrovirus* genus within the *Myoviridae* family, with 95% average nucleotide identity to the type phage T4 assessed by BLAST (GenBank accession number MT984581.1), genome sizes of 166,440 and 166,441 bp, respectively, and a GC content of 35.4%. Phages vb_EcoM_bov10K2, vb_EcoM_bov11CS3, vb_EcoM_bov22_2, and vb_EcoM_bov25_3 are rV5-like phages belonging to the *Vequintavirus* genus of the *Vequintavirinae* subfamily, having a >96% nucleotide identity to the type phage rV5 (DQ832317.1). Their genome sizes were between 135,960 and 135,961 bp, with a 43.7% GC content. Phages vb_EcoS_bov11C2, vb_EcoS_bov15_1, vb_EcoS_bov16_1, vb_EcoS_bov22_1, and vb_EcoS_bov25_1D represent HK578-like phages, officially the *Dhillonvirus* genus from the *Siphoviridae* family, with 86% genome coverage and >90% nucleotide identity to phage HK578 (JQ086375.1). Their genome sizes were between 44,612 and 44,747 bp, and their GC content was 54.5%.

Phages within the same genus were very uniform. The T4-like phages differed in only a 1-bp gap. Of the rV5-like phages, vb_EcoM_bov10K2, vb_EcoM_bov11CS3, and vb_EcoM_bov25_3 were 100% identical except for a 1-bp gap, while vb_EcoM_bov22_2 differed from them in only 2 single nucleotide polymorphisms (SNPs). The dhillonviruses differed at most by 11 SNPs, with a gap of 30 or 89 bp. The genome of vb_EcoS_bov22_1 was assembled as a partial genome sequence, with a different start position compared to the other dhillonviruses.

Data availability. The nucleotide sequences of the phages are deposited in GenBank under the accession numbers MT884006 through MT884015 and MT951623.

REFERENCES

1. Danis-Wlodarczyk K, Dąbrowska K, Abedon ST. 2020. Phage therapy: the pharmacology of antibacterial viruses. Curr Issues Mol Biol 40:81–164. https://doi.org/10.21775/cimb.040.081.
2. Sabouri S, Sepehrizadeh Z, Amirpour-Rostami S, Skurnik M. 2017. A mini-review on the in vitro and in vivo experiments with anti-Escherichia coli O157:H7 phages as potential biocontrol and phage therapy agents. Int J Food Microbiol 243:52–57. https://doi.org/10.1016/j.ijfoodmicro.2016.12.004.
3. Sváb D, Falgenhauer L, Rohde M, Chakraborty T, Tóth I. 2018. Identification and characterization of new broad host-range rV5-like coliphages C203 and P206 directed against enterobacteria. Infect Genet Evol 64:254–261. https://doi.org/10.1016/j.meegid.2018.07.004.
4. Ferguson S, Roberts C, Handy E, Sharma M. 2013. Lytic bacteriophages reduce Escherichia coli O157:H7 on fresh cut lettuce introduced through cross-contamination. Bacteriophage 3:e24323. https://doi.org/10.4161/bact.24323.
5. Gyles CL. 2007. Shiga toxin-producing Escherichia coli: an overview. J Anim Sci 85:E45–E62. https://doi.org/10.2527/jas.2006-508.
6. Hayashi T, Makino K, Ohnishi M, Kurokawa K, Ishii K, Yokoyama K, Han CG, Ohtsubo E, Nakayama K, Murata T, Tanaka M, Tobe T, Iida T, Takami H, Honda T, Sasakawa C, Ogasawara N, Yasunaga T, Kuhara S, Shiba T, Hattori M, Shinagawa H. 2001. Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12. DNA Res 8:11–22. https://doi.org/10.1093/dnares/8.1.11.
7. Perna NT, Plunkett G, III, Burland V, Mau B, Glasner JD, Rose DJ, Mayhew GF, Evans PS, Gregor J, Kirkpatrick HA, Pósfai G, Hackett J, Klink S, Boutin A, Shao Y, Miller L, Grotbeck EJ, Davis NW, Lim A, Dimalanta ET, Potamousis KD, Apodaca J, Anantharaman TS, Lin J, Yen G, Schwartz DC, Welch RA, Blattner FR. 2001. Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. Nature 409:529–533. https://doi.org/10.1038/35054089.
8. Tóth I, Schmidt H, Kardos G, Lancz Z, Creuzburg K, Damjanova I, Pászti J, Beutin L, Nagy B. 2009. Virulence genes and molecular typing of different groups of Escherichia coli O157 strains in cattle. Appl Environ Microbiol 75:6282–6291. https://doi.org/10.1128/AEM.00873-09.
9. Sambrook JF, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
10. Schwengers O, Hoek A, Fritzenwanker M, Falgenhauer L, Hain T, Chakraborty T, Goesmann A. 2020. ASA³P: an automatic and scalable pipeline for the assembly, annotation and higher level analysis of closely related bacterial isolates. PLoS Comput Biol 16:e1007134. https://doi.org/10.1371/journal.pcbi.1007134.
11. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for

Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

12. Garneau JR, Depardieu F, Fortier L-C, Bikard D, Monot M. 2017. Phage-Term: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. Sci Rep 7:8292. https://doi.org/10.1038/s41598-017-07910-5.

13. 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. https://doi.org/10.1089/cmb.2012.0021.

14. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA, III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.