Complete genome sequences of two T4-like Escherichia coli bacteriophages

Costa, Ana Rita; Brouns, Stan J.J.; Nobrega, Franklin L.

DOI
10.1128/genomeA.00586-18

Publication date
2018

Document Version
Final published version

Published in
Genome Announcements

AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

VIRUSES

Complete Genome Sequences of Two T4-Like *Escherichia coli* Bacteriophages

Ana Rita Costa,[a] Stan J. J. Brouns,[a,b] Franklin L. Nobrega[a]

[a]Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Delft, The Netherlands

[b]Department of Microbiology, Wageningen University and Research Centre, Wageningen, The Netherlands

ABSTRACT Bacteriophages and their proteins have potential applications in biotechnology for the detection and control of bacterial diseases. Here, we describe the sequencing and genome annotations of two strictly virulent *Escherichia coli* bacteriophages that may be explored for biocontrol strategies and to expand the understanding of phage-host interactions.

Escherichia coli is a commensal Gram-negative microorganism inhabiting the gastrointestinal tract of humans and animals, but it is also known for causing fatal infections (1–3). With the emergence of antibiotic-resistant bacteria, alternatives to antibiotics are lacking; bacteriophages and their derived proteins are a promising solution (4, 5).

Here, we isolated phages from ditch samples (Wageningen, The Netherlands) using a multihost approach of 32 *E. coli* strains (including O157:H7) and propagation in *E. coli* BL21. Both phages belong to the *Myoviridae* family and have broad lytic activity (in 15/32 strains for vB_CEB_NBG1 and 12/32 strains for vB_CEB_NBG2). Phage DNA was extracted as previously described (6). Sequencing was performed by Nucleomics Core using the next-generation sequencing (NGS) Illumina MiSeq platform and NEBNext Ultra DNA library prep kit. Sequencing reads (≈300 bp) with more than 100-fold coverage were *de novo* assembled using CLC Genomics Workbench version 7.0 (Aarhus, Denmark) and manually inspected. Annotation was performed using the RAST server (7), followed by manual screening of all predicted proteins against the NCBI protein database using BLASTp (8) and a Pfam domain search (9). tRNAs, promoters, and terminators were predicted with tRNAscan-SE version 2.0 (10), Geneious version 9.1.3 using motif TTGACAN(15,18)TATAAT with a maximum of one mismatch, and ARNold (11), respectively. The genome packaging strategy was predicted by phylogenetic analysis of the large terminase subunit (12).

Phages vB_EcoM_NBG1 and vB_EcoM_NBG2 have linear double-stranded DNA, with genome sizes of 168,869 bp and 166,083 bp (33.8% homologous), 2 (Arg and Met) and 10 (Gln, Leu, Gly, Pro, Ser, Thr, Met, Tyr, Asn, and Arg) tRNAs, 8 promoters each, and 27 and 20 Rho-independent terminators, respectively. Their GC contents of 37.7% and 35.4% are lower than that of *E. coli* (≈50%) (12).

Phages vB_EcoM_NBG1 and vB_EcoM_NBG2 have 269 and 261 predicted open reading frames (ORFs), of which 125 (46.5%) and 134 (51.3%) could be assigned a function, respectively. None of the predicted proteins exhibit homology toward virulence factors, integration-related proteins, or antibiotic resistance determinants; no genomic markers were found indicating a temperate lifestyle. These genetic features make both phages suitable candidates for phage therapy. Also, proteins were identified with exploratory interest, such as tail fibers (genes 165, 173, 216, 246, and 248 to 250 of vB_EcoM_NBG1 and genes 148, 200, 201, and 234 to 238 of vB_EcoM_NBG2), endolysins (gene 124 of vB_EcoM_NBG1), holins (genes 251 of vB_EcoM_NBG1 and 239

Received 27 May 2018 **Accepted** 1 June 2018 **Published** 28 June 2018

Citation Costa AR, Brouns SJJ, Nobrega FL. 2018. Complete genome sequences of two T4-like *Escherichia coli* bacteriophages. Genome Announc 6:e00586-18. https://doi.org/10.1128/genomeA.00586-18.

Address correspondence to Franklin L. Nobrega, F.LuziadeNobrega@tudelft.nl.

of vB_EcoM_NBG2), spanin complex (genes 228 and 229 of vB_EcoM_NBG1 and genes 212 and 213 of vB_EcoM_NBG2), and tail-associated lysozymes (gene 156 of vB_EcoM_NBG1 and genes 118 and 139 of vB_EcoM_NBG2).

Phylogenetic analysis of the large terminase subunit suggests both phages use a T4-like mechanism of headful packaging, with no preferred packaging signal (12). As this implies random phage termini, phages were zeroed using phage T4 as a reference.

Comparative genomics revealed that phage vB_EcoM_NBG1 is closely related to *Escherichia* phage APCEc01 (accession number NC_029091), sharing 98% identity over 98% of its sequence, whereas phage vB_CEB_NBG2 is closely related to *Escherichia* phage PEC04 (accession number KR233165), with 97% identity over 95% of its sequence. The genome comparisons revealed regions with pronounced differences located mainly in the tail fiber proteins, which likely confer the phages a distinct host lytic spectrum.

Accession number(s). The genome sequences have been deposited in GenBank under the accession numbers MH243438 (vB_EcoM_NBG1) and MH243439 (vB_EcoM_NBG2).

ACKNOWLEDGMENTS

F.L.N. is supported by the Netherlands Organization for Scientific Research (NWO) Veni grant 016.Veni.181.092. S.J.J.B. is supported by NWO Vidi grant 864.11.005, European Research Council (ERC) Stg grant 639707, and a TU Delft startup grant.

The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Janny S, Bert F, Dondero F, Nicolas Chanoine MH, Belghiti J, Mantz J, Paugam-Burtz C. 2013. Fatal *Escherichia coli* skin and soft tissue infections in liver transplant recipients: report of three cases. Transpl Infect Dis 15:E49–E53. https://doi.org/10.1111/tid.12046.
2. Vigil KJ, Johnson JR, Johnston BD, Kontoyiannis DP, Mulanovich VE, Raad II, DuPont HL, Adachi JA. 2010. *Escherichia coli* pyomyositis: an emerging infectious disease among patients with hematologic malignancies. Clin Infect Dis 50:374–380. https://doi.org/10.1086/649866.
3. Wagner S, Gally DL, Argyle SA. 2014. Multidrug-resistant *Escherichia coli* from canine urinary tract infections tend to have commensal phylotypes, lower prevalence of virulence determinants and *ampC*-replicons. Vet Microbiol 169:171–178. https://doi.org/10.1016/j.vetmic.2014.01.003.
4. Drulis-Kawa Z, Majkowska-Skrobek G, Maciejewska B. 2015. Bacteriophages and phage-derived proteins—application approaches. Curr Med Chem 22:1757–1773. https://doi.org/10.2174/0929867322666150209152851.
5. Roach DR, Donovan DM. 2015. Antimicrobial bacteriophage-derived proteins and therapeutic applications. Bacteriophage 5:e1062590. https://doi.org/10.1080/21597081.2015.1062590.
6. Sambrook J, Russell D. 2001. Molecular cloning: a laboratory manual. Cold Spring Harbor Press, Cold Spring Harbor, NY.
7. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.
8. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.
9. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285. https://doi.org/10.1093/nar/gkv1344.
10. Lowe TM, Chan PP. 2016. tRNAscan-SE on-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44:W54–W57. https://doi.org/10.1093/nar/gkw413.
11. Naville M, Ghuillot-Gaudeffroy A, Marchais A, Gautheret D. 2011. ARNold: a Web tool for the prediction of Rho-independent transcription terminators. RNA Biol 8:11–13. https://doi.org/10.4161/rna.8.1.13346.
12. Iguchi A, Thomson NR, Ogura Y, Saunders D, Ooka T, Henderson IR, Harris D, Asadulghani M, Kurokawa K, Dean P, Kenny B, Quail MA, Thurston S, Dougan G, Hayashi T, Parkhill J, Frankel G. 2009. Complete genome sequence and comparative genome analysis of enteropathogenic *Escherichia coli* O127:H6 strain E2348/69. J Bacteriol 191:347–354. https://doi.org/10.1128/JB.01238-08.