



# Draft Genome Sequence of *Escherichia coli* DSM 12242, a Highly Efficient Host Strain for the Isolation of Somatic Coliphages

 Cátia Pacífico,<sup>a,b</sup>  Dmitrij Sofka,<sup>a</sup>  João André Carriço,<sup>c</sup> Friederike Hilbert<sup>a</sup>

<sup>a</sup>Department of Farm Animals and Veterinary Public Health, University of Veterinary Medicine, Vienna, Austria

<sup>b</sup>Karl Landsteiner University of Health Sciences, Krems an der Donau, Austria

<sup>c</sup>Instituto de Medicina Molecular, Faculty of Medicine, University of Lisbon, Lisbon, Portugal

**ABSTRACT** Here, we report the draft genome sequence and characterization of the commercial strain *Escherichia coli* DSM 12242 (=ATCC 13706/60 =NZRM 3262) derived from strain C (ATCC 13706), which is suitable for the isolation of coliphages from environmental and clinical samples.

*Escherichia coli* is a commensal microorganism that colonizes the gut of humans and warm-blooded animals. Strain DSM 12242 (ATCC 13706/60), a nalidixic acid-resistant derivative of strain C (ATCC 13706), allows the supplementation of this antibiotic to the medium, suppressing the background microbes present in the samples and making it a suitable indicator host for the identification of somatic coliphages (1). Derivatives of strain C, such as the better-known nalidixic acid-resistant strain WG5, are commonly used as a bacterial proxy for fecal contamination and the presence of human enteric viruses in water samples, as described in ISO method 10705-2 (2). The strain herein described was previously used as a host for the isolation of coliphages from chicken meat (3), veterinary surgery surfaces (4), and human clinical samples, such as trachea aspirates, urine, and blood (C. Pacífico, M. Hilbert, D. Sofka, N. Dinhopl, I.-J. Pap, C. Aspöck, J. A. Carriço, and F. Hilbert, unpublished data). In fact, ATCC 13706/60 was reported as giving the highest PFU count out of 16 *E. coli* reference culture strains tested on effluents (5), and experiments with sewage samples also showed a 3-fold higher performance for the enumeration of coliphages with ATCC 13706/60 than that with ATCC 13706 (1).

The isolate was acquired from the DMSZ and grown in modified Scholten's agar plates for 24 to 48 h at 37°C. Genomic DNA was isolated from bacterial cells using the DNeasy blood and tissue kit (Qiagen) according to the manufacturer's protocols. Library preparation, DNA sequencing, and sequence trimming were performed at MicrobesNG (Birmingham, United Kingdom). DNA libraries were prepared using the Nextera XT library prep kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions, with slight modifications (2 ng of DNA was used instead of 1 ng, and the PCR elongation time was increased from 30 to 60 s). Libraries were sequenced on the Illumina HiSeq 2500 platform using a 2 × 250-bp paired-end approach. The trimmed reads were analyzed using the INNUca pipeline (6). Briefly, read quality was inspected using FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and improved with Trimmomatic (7) version 0.38, using default parameters. Paired reads were merged with PEAR (8) version 0.9.10, using default parameters, and *de novo* assembled with SPAdes (9) version 3.13 using the "careful" option. The resulting assembly was corrected by Pilon (10) version 1.18. Whole-genome sequencing yielded a total of 1,097,039 reads, with a length between 36 and 251 bp. The genome assembly resulted in 78 contigs with an  $N_{50}$  value of 139,561 bp.

**Citation** Pacífico C, Sofka D, Carriço JA, Hilbert F. 2019. Draft genome sequence of *Escherichia coli* DSM 12242, a highly efficient host strain for the isolation of somatic coliphages. *Microbiol Resour Announc* 8:e00953-19. <https://doi.org/10.1128/MRA.00953-19>.

**Editor** Catherine Putonti, Loyola University Chicago

**Copyright** © 2019 Pacífico 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 Cátia Pacífico, 1545365@students.vetmeduni.ac.at.

All authors contributed equally to this work.

**Received** 6 August 2019

**Accepted** 27 August 2019

**Published** 19 September 2019

Strain DSM 12242 has a draft genome of 4,542,748 bp and a GC content of 50.9%. The genome coverage was calculated as 82.2×. Annotation with Prokka 1.13.3 (11) revealed the presence of 4,476 genes, 4,174 coding sequences, 11 rRNAs, 81 tRNAs, and 2 CRISPR repeats. Based on the *in silico* data, the multilocus sequence type (ST) and predicted serotype were ST 1721 (12, 13) and NT:H19 (14), respectively. ResFinder (15) identified the presence of *mdfa* (99.92% identity), a gene that encodes a multidrug resistance efflux pump. No plasmid sequences were detected using PlasmidFinder (16). The PHASTER prophage finder (17) identified two incomplete prophage regions of 8.2 kb and 16.7 kb. OrthoANI (18) revealed that this bacterium shares an average nucleotide identity of 98.99% with *E. coli* C (GenBank accession no. CP029371) (19) and 99.97% with strain WG5 (GenBank accession no. CP024090) (20). Nalidixic acid resistance is conferred to strain DSM 12242 by a single point mutation (Asp87→Gly) in the *gyrA* gene, while in strain WG5, it is associated with two point mutations (Asp87→Gly and Gln-106→His) (15).

**Data availability.** The genome sequence has been deposited in GenBank under the accession no. VNIO00000000, and the version described in this paper is the first version (VNIO01000000). The Illumina raw reads have been deposited in the SRA under the accession no. SRR9717097.

## ACKNOWLEDGMENT

This work was in part supported by the Niederösterreichische Forschungs- und Bildungsgesellschaft (NFB) grant no. LS14-006.

## REFERENCES

- Sinton LW, Finlay RK, Reid AJ. 1996. A simple membrane filtration-elution method for the enumeration of F-RNA, F-DNA and somatic coliphages in 100-ml water samples. *J Microbiol Methods* 25:257–269. [https://doi.org/10.1016/0167-7012\(95\)00100-X](https://doi.org/10.1016/0167-7012(95)00100-X).
- International Organization for Standardization. 2000. Water quality—detection and enumeration of bacteriophages. Part 2. Enumeration of somatic coliphages. ISO 10705-2:2000. International Organization for Standardization, Geneva, Switzerland. <https://www.iso.org/standard/20127.html>.
- Shousha A, Awaianont N, Sofka D, Smulders FJM, Paulsen P, Szostak MP, Humphrey T, Hilbert F. 2015. Bacteriophages isolated from chicken meat and the horizontal transfer of antimicrobial resistance genes. *Appl Environ Microbiol* 81:4600–4606. <https://doi.org/10.1128/AEM.00872-15>.
- Hilbert M, Csadek I, Auer U, Hilbert F. 2017. Antimicrobial resistance-transducing bacteriophages isolated from surfaces of equine surgery clinics—a pilot study. *Eur J Microbiol Immunol* 7:296–302. <https://doi.org/10.1556/1886.2017.00032>.
- Sinton LW, Ching SB. 1987. An evaluation of two bacteriophages as sewage tracers. *Water Air Soil Pollut* 35:347–356. <https://doi.org/10.1007/BF00290941>.
- Machado MP, Halkilahti J, Jaakkonen A, Silva DN, Mendes I, Nalbantoglu Y, Borges V, Ramirez M, Rossi M, Carrico JC. 2019. INNUca. <https://github.com/B-UMMI/INNUca>.
- 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>.
- Zhang J, Kobert K, Flouri T, Stamatakis A. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30:614–620. <https://doi.org/10.1093/bioinformatics/btt593>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. <https://doi.org/10.1371/journal.pone.0112963>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Seemann T. 2019. mlst. <https://github.com/tseemann/mlst>.
- Jolley KA, Maiden MCJ. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 10:595. <https://doi.org/10.1186/1471-2105-11-595>.
- Machado MP, Halkilahti J, Mendes I, Pinto M, Lizarazo E, Gomes JP, Ramirez M, Rossi M, Carrico JC. 2019. seq\_typing. [https://github.com/B-UMMI/seq\\_typing](https://github.com/B-UMMI/seq_typing).
- Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 67:2640–2644. <https://doi.org/10.1093/jac/dks261>.
- Carattoli A, Zankari E, García-Fernández A, Larsen MV, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 58:3895–3903. <https://doi.org/10.1128/AAC.02412-14>.
- Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21. <https://doi.org/10.1093/nar/gkw387>.
- Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>.
- Pekar JE, Phaneuf P, Szubin R, Palsson B, Feist A, Monk M. 2018. Gapless, unambiguous genome sequence for *Escherichia coli* C, a workhorse of industrial biology. *Microbiol Resour Annot* 7:e00890-18. <https://doi.org/10.1128/MRA.00890-18>.
- Imamovic L, Misiakou M, Van Der Helm E, Panagiotou G, Muniesa M, Sommer M. 2018. Complete genome sequence of *Escherichia coli* strain WG5. *Genome Annot* 6:e01403-17. <https://doi.org/10.1128/genomeA.01403-17>.