



# Complete Genome and Plasmid Sequences of *Salmonella enterica* subsp. *enterica* Serovar Enteritidis PT1, Obtained from the *Salmonella* Reference Laboratory at Public Health England, Colindale, United Kingdom

 Muhammad Zohaib Anwar,<sup>a</sup> Athanasios Zervas,<sup>a</sup> Lars Hestbjerg Hansen,<sup>b</sup> Gary Barker,<sup>c</sup> Alexandre Magno Anesio,<sup>a</sup> Carsten Suhr Jacobsen<sup>a</sup>

<sup>a</sup>Department of Environmental Science, Aarhus University RISØ Campus, Roskilde, Denmark

<sup>b</sup>Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark

<sup>c</sup>School of Biological Sciences, University of Bristol, Bristol, United Kingdom

Muhammad Zohaib Anwar and Athanasios Zervas are co-first authors with equal contribution. Author order was determined in order of increasing seniority.

**ABSTRACT** Hybrid assembly of Illumina and Oxford Nanopore sequencing was used here to produce the complete circular genome and plasmid sequences of *Salmonella enterica* subsp. *enterica* serovar Enteritidis PT1 (phage type 1). The organism was obtained from the *Salmonella* Reference Laboratory at Public Health England, Colindale, UK.

*Salmonella* members are classified into >2,500 serovars based on the Kauffmann-White scheme (1). *Salmonella enterica* subsp. *enterica* serovar Enteritidis is associated with salmonellosis in humans and domestic animals (1). *Salmonella enterica* subsp. *enterica* serovar Enteritidis has been frequently detected as the causative agent in foodborne outbreaks (FBOs) reported in Europe and North America (2–4). The complete genome sequence of this strain will not only serve as a reference for identifying similar outbreaks with more precision in the future but also aid in the identification of similar organisms in environmental samples. *Salmonella enterica* subsp. *enterica* serovar Enteritidis PT1 was grown at 37°C overnight in LB growth medium. Three samples were taken from one liquid culture and were used for individual DNA extractions. DNA was extracted from 200 µl of each sample using the AllPrep PowerFecal DNA/RNA kit (Qiagen, Hilden, Germany). Short-read sequencing libraries were prepared using the Illumina TrueSeq DNA PCR-free high-throughput library prep kit. Libraries were pooled equimolarly and sequenced (single-end 75 bp) on the NextSeq platform using the NextSeq 500/550 high-throughput kit v2.5 (Illumina, San Diego, CA, USA). A total of 94,938,911 single-end 75-bp reads were subjected to quality trimming and sequence contamination removal (–average\_qual < Q20) using fastp (5). High-molecular-weight DNA for Nanopore sequencing was also extracted from 1-ml cultures using the MasterPure complete DNA and RNA purification kit (catalog number MC85200; Epicentre, Lucigen), according to the manufacturer's instructions. Libraries for Oxford Nanopore sequencing were prepared using the rapid sequencing kit (SQK-RAD004) and loaded on a MinION flow cell (R.9.5.1). Sequencing was stopped manually after 22 hours when 6.44 Gb were obtained, resulting in 359,803 long reads with an  $N_{50}$  value of 30.23 kb. Adapters were trimmed from long reads using Porechop (<https://github.com/rrwick/Porechop>). Trimmed Illumina and Nanopore reads were then used for a hybrid assembly using Unicycler v0.4.4 (6), which uses SPAdes v3.13.1 (7) for short-read assembly and miniasm and minimap (8) for long-read assembly and mapping the bridges with short reads, respectively, and Racon v1.4.0 (9) and Pilon v1.23 (10)

**Citation** Anwar MZ, Zervas A, Hansen LH, Barker G, Anesio AM, Jacobsen CS. 2020. Complete genome and plasmid sequences of *Salmonella enterica* subsp. *enterica* serovar Enteritidis PT1, obtained from the *Salmonella* Reference Laboratory at Public Health England, Colindale, United Kingdom. Microbiol Resour Announc 9:e01064-19. <https://doi.org/10.1128/MRA.01064-19>.

**Editor** Vincent Bruno, University of Maryland School of Medicine

**Copyright** © 2020 Anwar 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 Muhammad Zohaib Anwar, [mzanwar@envs.au.dk](mailto:mzanwar@envs.au.dk).

**Received** 9 September 2019

**Accepted** 6 December 2019

**Published** 9 January 2020

for polishing the assembly. The complete genome sequence (100% coverage) consists of one circular chromosome (4,685,587 bp) with an overall G+C content of 52.17% and one circular plasmid sequence (59,372 bp) with a G+C content of 51.95%. The genome and plasmid were annotated using rapid prokaryotic genome annotation (Prokka) v1.13.7 (11) with genome and plasmid parameters, respectively. A total of 4,463 predicted genes were identified on the genome, including 22 rRNA, 84 tRNA, 1 transfer-messenger RNA (tmRNA), and 4,354 protein-coding genes, while 75 coding DNA sequences (CDSs) were identified on the plasmid. The genome was then compared with 6 other *Salmonella enterica* genomes and the *Sphingobium herbicidovorans* MH (12) genome based on marker gene lineage analysis using CheckM v1.0.7 (13), which indicated 100% completeness of the genome and 0.00 strain heterogeneity with *Salmonella enterica* subsp. *enterica* serovar Enteritidis. All bioinformatics software programs were used with default parameters, unless otherwise stated.

**Data availability.** Illumina short reads were deposited at the SRA and are available under BioSample accession numbers [SAMN12634776](#), [SAMN12634778](#), and [SAMN12634782](#) under BioProject accession number [PRJNA562128](#). Oxford Nanopore long reads were also deposited at the SRA and are available under BioProject accession number [PRJNA562126](#). The complete genome and plasmid sequences are available at GenBank under accession numbers [CP043433](#) and [CP043434](#), respectively, and under BioProject accession number [PRJNA562154](#).

## ACKNOWLEDGMENTS

This work was supported by a grant from the European Commission's Marie Skłodowska Curie Actions program under project number 675546 (MicroArctic). A.Z. and A.M.A. were supported by a starting grant from the Aarhus University Research Foundation (AUFF).

We thank Tanja Begovic for capable laboratory assistance. We also thank Tim Dallman and Elizabeth De Pinna at the *Salmonella* Reference Laboratory, Gastrointestinal Bacteria Reference Unit at the National Infection Service, Public Health England, United Kingdom (<https://www.gov.uk/guidance/gbru-reference-and-diagnostic-services>), for providing the strain.

## REFERENCES

- Park SH, Kim HJ, Cho WH, Kim JH, Oh MH, Kim SH, Lee BK, Ricke SC, Kim HY. 2009. Identification of *Salmonella enterica* subspecies I, *Salmonella enterica* serovars Typhimurium, Enteritidis and Typhi using multiplex PCR. *FEMS Microbiol Lett* 301:137–146. <https://doi.org/10.1111/j.1574-6968.2009.01809.x>.
- Vencia W, Gariano GR, Bianchi DM, Zuccon F, Sommariva M, Nguon B, Malabaila A, Gallina S, Decastelli L. 2015. A *Salmonella enterica* subsp. *enterica* serovar Enteritidis foodborne outbreak after consumption of homemade lasagne. *Ital J Food Saf* 4:5127. <https://doi.org/10.4081/ijfs.2015.5127>.
- Much P, Pichler J, Kasper S, Lassnig H, Kornschöber C, Buchner A, König C, Allerberger F. 2009. A foodborne outbreak of *Salmonella* Enteritidis phage type 6 in Austria, 2008. *Wien Klin Wochenschr* 121:132–136. <https://doi.org/10.1007/s00508-008-1134-y>.
- Middleton D, Savage R, Tighe MK, Vrbova L, Walton R, Whitfield Y, Varga C, Lee B, Rosella L, Dhar B, Johnson C, Ahmed R, Allen VG, Crowcroft NS. 2014. Risk factors for sporadic domestically acquired *Salmonella* serovar Enteritidis infections: a case-control study in Ontario, Canada, 2011. *Epidemiol Infect* 142:1411–1421. <https://doi.org/10.1017/S0950268813001945>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- 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>.
- Li H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 32:2103–2110. <https://doi.org/10.1093/bioinformatics/btw152>.
- Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res* 27:737–746. <https://doi.org/10.1101/gr.214270.116>.
- 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>.
- Nielsen TK, Rasmussen M, Demanèche S, Cecillon S, Vogel TM, Hansen LH. 2017. Evolution of sphingomonad gene clusters related to pesticide catabolism revealed by genome sequence and mobilomics of *Sphingobium herbicidovorans* MH. *Genome Biol Evol* 9:2477–2490. <https://doi.org/10.1093/gbe/evx185>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.