

UC Irvine

UC Irvine Previously Published Works

Title

Complete Genome Sequence of Salmonella enterica subsp. enterica Serovar Kottbus Strain Kharkiv, Isolated from a Commercial Pork Production Facility in Ukraine.

Permalink

<https://escholarship.org/uc/item/6617m1z8>

Journal

Microbiology Resource Announcements, 9(49)

Authors

Arefiev, Vasiliy

Kovalenko, Ganna

Frant, Maciej

et al.

Publication Date

2020-12-03


DOI

10.1128/MRA.01171-20

Peer reviewed



Complete Genome Sequence of *Salmonella enterica* subsp. *enterica* Serovar Kottbus Strain Kharkiv, Isolated from a Commercial Pork Production Facility in Ukraine

Vasiliy Arefiev,^a Ganna Kovalenko,^b Maciej Frant,^c Tetiana Chumachenko,^d Yuliia Polyvianna,^d Svitlana Pivnenko,^d Vitaliy Bolotin,^a Olga Mayboroda,^a Oleksii Solodiankin,^a Oleksandr Tarasov,^b Maksym Bezymenni,^b Cora Lyon,^e Matthew Redlinger,^e Maryna Sapachova,^f Andrii A. Mezhenyskyi,^f Anne-Lise Ducluzeau,^g Eric Bortz,^{b,e} Anton Gerilovych,^a  Devin M. Drown^{g,h}

^aNational Scientific Center, Institute of Experimental and Clinical Veterinary Medicine, Kharkiv, Ukraine

^bNational Academy of Agrarian Sciences of Ukraine, Institute for Veterinary Medicine, Kyiv, Ukraine

^cNational Veterinary Research Institute, Pulawy, Poland

^dKharkiv National Medical University, State Institution Kharkiv Oblast Laboratory Center of the Ministry of Health of Ukraine, Kharkiv, Ukraine

^eDepartment of Biological Sciences, University of Alaska Anchorage, Anchorage, Alaska, USA

^fState Scientific Research Institute of Laboratory Diagnostics and Veterinary and Sanitary Expertise, Kyiv, Ukraine

^gDepartment of Biology and Wildlife, University of Alaska Fairbanks, Fairbanks, Alaska, USA

^hInstitute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, Alaska, USA

ABSTRACT The complete genome of *Salmonella enterica* subsp. *enterica* serovar Kottbus strain Kharkiv (serogroup C2-C3), which was isolated from a commercial pork production facility in Kharkiv, Ukraine, was assembled using long-read Nanopore sequences. A single circular contig (4,799,045 bp) comprised a complete chromosome encoding antibiotic resistance, highlighting the risk of cross-species livestock and human infection.

As part of a scientific-capacity-building initiative in Ukraine to better understand African swine fever outbreaks, the genomes of concurrent or coinfecting pathogens of commercial and backyard swine and wild boar were sequenced in regional veterinary laboratories. For this study, the original sample source was a commercial pork production facility (chopping boards) identified by the State Institution Kharkiv Oblast Laboratory Center of the Ministry of Health of Ukraine. A *Salmonella* sp. strain was isolated on 6 June 2018 from the Sakhnovshchanskyi Rayon district of Kharkiv Oblast (province). The primary enrichment of the microbial culture was carried out using selenite broth. After 1 day of incubation, parallel inoculation was carried out on Endo agar and bismuth sulfite agar (37°C). On 5 July 2018, the strain was transferred to the National Scientific Center Institute of Experimental and Clinical Veterinary Medicine (IECVM) bacteriological laboratory. A tryptic soy broth (TSB) culture grown from a single colony incubated overnight at 22°C was cryopreserved at –70°C in cryovials at IECVM prior to DNA isolation. To obtain DNA for sequencing, we inoculated liquid cultures of TSB, incubated them overnight at 22°C, and used 1.8 ml as input for the DNeasy UltraClean microbial kit (Qiagen).

We used 1 µg of DNA as input for a rapid sequencing library (SQK-RAD004; Oxford Nanopore Technologies [ONT]) and sequenced it on an R9.4.1 flow cell (FLO-MIN106, flow cell identification no. FAK94576) for 26 h using a MinION Mk1B device (ONT). We base called raw data with Guppy v3.6.1 (ONT) using the high-accuracy model (c, dna_r9.4.1_450bps_hac.cfg) and default parameters. This run generated a total of 9,547,083,284 bp in 3,536,512 reads with a read length N_{50} value of 5,727 bp. We used FilTlong v0.2.0 (<https://github.com/rrwick/FilTlong>) to filter by a length of $\geq 1,000$ bp (min_length, 1000) and quality (Q) score of ≥ 10 (min_mean_q, 90). After filtering, we

Citation Arefiev V, Kovalenko G, Frant M, Chumachenko T, Polyvianna Y, Pivnenko S, Bolotin V, Mayboroda O, Solodiankin O, Tarasov O, Bezymenni M, Lyon C, Redlinger M, Sapachova M, Mezhenyskyi AA, Ducluzeau A-L, Bortz E, Gerilovych A, Drown DM. 2020. Complete genome sequence of *Salmonella enterica* subsp. *enterica* serovar Kottbus strain Kharkiv, isolated from a commercial pork production facility in Ukraine. *Microbiol Resour Announc* 9:e01171-20. <https://doi.org/10.1128/MRA.01171-20>.

Editor Vincent Bruno, University of Maryland School of Medicine

Copyright © 2020 Arefiev 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 Devin M. Drown, dmdrown@alaska.edu.

Received 7 October 2020

Accepted 6 November 2020

Published 3 December 2020

TABLE 1 Antimicrobial resistance genes

Antimicrobial resistance mechanism	Gene(s) or gene product(s)
Antibiotic activation enzyme	KatG
Antibiotic inactivation enzyme	AAC(6')-Ic, AAC(6')-If, AAC(6')-Ig, AAC(6')-Ih, AAC(6')-Ij, AAC(6')-Ik, AAC(6')-Il, AAC(6')-Ir, AAC(6')-Is, AAC(6')-It, AAC(6')-Iu, AAC(6')-Iv, AAC(6')-Iw, AAC(6')-Ix, AAC(6')-Iy, AAC(6')-Iz
Antibiotic resistance gene cluster, cassette, or operon	MarA, MarB, MarR
Antibiotic target in susceptible species	Alr, Ddl, <i>dxr</i> , EF-G, EF-Tu, <i>folA</i> , Dfr, <i>folP</i> , <i>gyrA</i> , <i>gyrB</i> , <i>inhA</i> , <i>fabI</i> , Iso-tRNA, <i>kasA</i> , <i>MurA</i> , <i>rho</i> , <i>rpoB</i> , <i>rpoC</i> , S10p, S12p
Antibiotic target protection protein	BcrC
Efflux pump conferring antibiotic resistance	AcrAB-TolC, AcrAD-TolC, AcrEF-TolC, AcrZ, EmrAB-TolC, MacA, MacB, MdfA/Cmr, MdtABC-TolC, MdtL, MdtM, MexPQ-OpmE, OprM/OprM family, SugE, TolC/OpmH
Gene conferring resistance via absence	<i>gidB</i>
Protein altering cell wall charge, conferring antibiotic resistance	GdpD, PgsA
Regulator modulating expression of antibiotic resistance genes	AcrAB-TolC, EmrAB-TolC, H-NS, OxyR

had 7,261,568,727 bp in 1,612,614 reads with a read length N_{50} value of 6,283 bp and a median Q score of 13.4.

We assembled quality-controlled read data *de novo* using Flye v2.7 (1) with default parameters, specifying the estimated genome size (-genomesize=5m), Nanopore reads (-nanopore-raw), and subsampling for initial disjointing assembly (--asm-coverage 100). We corrected the single contig, marked as circular by Flye, using two rounds of Racon v1.3.1 (2) polishing with the following parameters: score for matching bases: match, 8; score for mismatching bases: score for matching bases, --match 8; score for mismatching bases, --mismatch -6; threshold for average base quality of windows, --quality-threshold -1; default gap penalty, --gap -8; and default window --window-length 500. We ran a final polish with Medaka v1.0.3 (<https://github.com/nanoporetech/medaka>) specifying the base-caller model (-m r941_min_high_g360) and using default parameters. For polishing, we used the entire 7.2-Gb quality-controlled data set (coverage, 970 \times). Our 4,799,045-bp polished assembly (N_{50} , 4,799,045 bp) consists of one circular contig (GC content, 52.21%).

To annotate the genome, we used PATRIC v3.6.5 (3, 4) to assign 84 tRNAs, 22 rRNAs, 104 repeat regions, and 4,870 coding sequences. PATRIC reported a 100% completeness score with minor contamination (1.6%) (5). PATRIC identified a total of 57 antibiotic resistance genes (Table 1) and 232 virulence factors. PATRIC assigned the isolate as a member of the *Salmonella enterica* group with multilocus sequence typing (MLST) sequence type 212. Using the *Salmonella in Silico* Typing Resource (SISTR) command line tool (6), the genome was assigned to serogroup C2-C3 and serovar Kottbus. SISTR provided O-antigen (types 6 and 8) and H-antigen (types e and h and types 1 and 5) types consistent with determined serology. We assign this isolate as *Salmonella enterica* subsp. *enterica* serovar Kottbus strain Kharkiv. The genome deposited in GenBank was annotated with PGAP (7) during submission.

Data availability. This whole-genome shotgun project has been deposited in GenBank under the accession no. CP062220. The version described in this paper is the first version, CP062220.1. Raw data for this project can be found in the GenBank SRA under PRJNA666173.

ACKNOWLEDGMENTS

Many thanks go to Ralf Dagdag and Elaina Milton for technical assistance and to Nataliya Mykhalovska and Olga Fedorenko for project support.

This work was funded by the U.S. Defense Threat Reduction Agency (DTRA) through the Biological Threat Reduction Program in Ukraine (Ukraine Project 9). Research reported in this publication was in part supported by an NIH NIGMS Institutional Development Award (grant P20GM103395) (Alaska INBRE). We also acknowledge the generous support of the Institute of Arctic Biology at the University of Alaska Fairbanks and the University of Alaska Anchorage College of Arts and Sciences.

The contents of this publication are the responsibility of the authors and do not necessarily reflect the views of the Department of Defense, the NIH, or the United States Government.

REFERENCES

1. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
2. Vaser R, Sovic I, Nagarajan N, Sikic 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>.
3. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia F, Yoo H, Stevens RL. 2017. Improvements to PATRIC, the all-bacterial Bioinformatics Database and Analysis Resource Center. *Nucleic Acids Res* 45:D535–D542. <https://doi.org/10.1093/nar/gkw1017>.
4. 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>.
5. 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>.
6. Yoshida CE, Kruczkiewicz P, Laing CR, Lingohr EJ, Gannon VP, Nash JH, Taboada EN. 2016. The *Salmonella* In Silico Typing Resource (SISTR): an open web-accessible tool for rapidly typing and subtyping draft *Salmonella* genome assemblies. *PLoS One* 11:e0147101. <https://doi.org/10.1371/journal.pone.0147101>.
7. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.