



Complete Genome Sequence of a Multidrug-Resistant Strain, *Escherichia coli* ATCC BAA-196, as a Model for Studying Induced Antibiotic Resistance Reversion

Ilya S. Korotetskiy,^a Monique Joubert,^b Setshaba Taukobong,^b Ardak B. Jumagazyieva,^a Sergey V. Shilov,^a Sergey V. Shvidko,^a Natalya A. Sulдина,^a Sabina T. Kenesheva,^{a,c} Anna Yssel,^b Oleg N. Reva,^b Aleksandr I. Ilin^a

^aScientific Center for Anti-Infectious Drugs, Almaty, Kazakhstan

^bCentre for Bioinformatics and Computational Biology, Department of Biochemistry, Genetics, and Microbiology, University of Pretoria, Pretoria, South Africa

^cFaculty of Biology and Biotechnology, al-Farabi Kazakh National University, Almaty, Kazakhstan

ABSTRACT Here, we report the complete genome sequence of the multidrug-resistant *Escherichia coli* strain ATCC BAA-196, a model organism used for studying possible antibiotic resistance reversion induced by FS-1, an iodine-containing complex. Two genomes, representing FS-1-treated and negative-control variants and composed of a chromosome and several plasmids, were assembled.

Escherichia coli ATCC BAA-196 was isolated in 1988 at a chronic-care facility in Massachusetts and at first was misidentified as *Klebsiella pneumoniae* (1, 2). It produces extended-spectrum beta-lactamases (3, 4). This strain was used as a model of nosocomial drug-resistant infections for studying the effect of the iodine-containing complex FS-1 in reverting drug resistance (5, 6).

The strain was cultivated for 10 daily passages in Mueller-Hinton broth (HiMedia, India) with FS-1 (500 $\mu\text{g/ml}$) (FS genome) or without FS-1 (negative-control [NC] genome), in three repeats. DNA was extracted using the PureLink genomic DNA kit (Thermo Fisher). Samples were prepared according to the SMRTbell preparation guide for the PacBio RS II system. Sequencing was performed at Macrogen (South Korea) with SMRT Cell 8Pac v3 cells using the DNA polymerase binding kit P6, following the SMRTbell 20-kb library preparation protocol. For the NC and FS genomes, 334,150 and 429,631 reads, respectively, were generated (N_{50} , 9,500 kb). After read-quality trimming using the UGENE v1.32.0 raw DNA-seq processing pipeline with default settings (7), the genomes were assembled with SMRT Link v5.0.1 with default parameters (8). The FS and NC genome assemblies are 4,682,561 and 4,682,572 bp, respectively (GC content, 51%; coverage, 250-fold); the genomes also include large plasmids of 266,396 and 279,992 bp, respectively (GC content, 47%), showing 90 to 99% sequence similarity to *Klebsiella pneumoniae* plasmid pKP64477b. The FS genome plasmid has an insertion of a prophage flanked by two copies of *insH* transposases. Moreover, the FS-1-treated strain contains two smaller plasmids (44,240 and 11,153 bp), which are excision products of the large plasmid. A plasmid-destabilizing effect of FS-1 was hypothesized.

Genomes were annotated with the RAST server (9) and manually curated. Phylogenetic inference based on concatenated alignments of 3,179 orthologous genes identified by OrthoFinder (10) (default settings) that were shared by *E. coli* reference genomes (Fig. 1) showed clustering of BAA-196 with the *E. coli* K12-related strains K12 (GenBank accession number NC_000913), K12 substrain W3110 (GenBank accession number NC_007779), and K12 substrain DH10B (GenBank accession number NC_010473).

Horizontally transferred genomic islands (Fig. 1) were identified by SeqWord Sniffer (11). Genomic islands and plasmids contain genetic determinants associated with antibiotic resistance, including beta-lactamases of the A and D classes, the tellurium

Citation Korotetskiy IS, Joubert M, Taukobong S, Jumagazyieva AB, Shilov SV, Shvidko SV, Sulдина NA, Kenesheva ST, Yssel A, Reva ON, Ilin AI. 2019. Complete genome sequence of a multidrug-resistant strain, *Escherichia coli* ATCC BAA-196, as a model for studying induced antibiotic resistance reversion. *Microbiol Resour Announc* 8:e01118-19. <https://doi.org/10.1128/MRA.01118-19>.

Editor Vincent Bruno, University of Maryland School of Medicine

Copyright © 2019 Korotetskiy 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 Ilya S. Korotetskiy, laeda81@gmail.com.

Received 15 September 2019

Accepted 12 November 2019

Published 12 December 2019

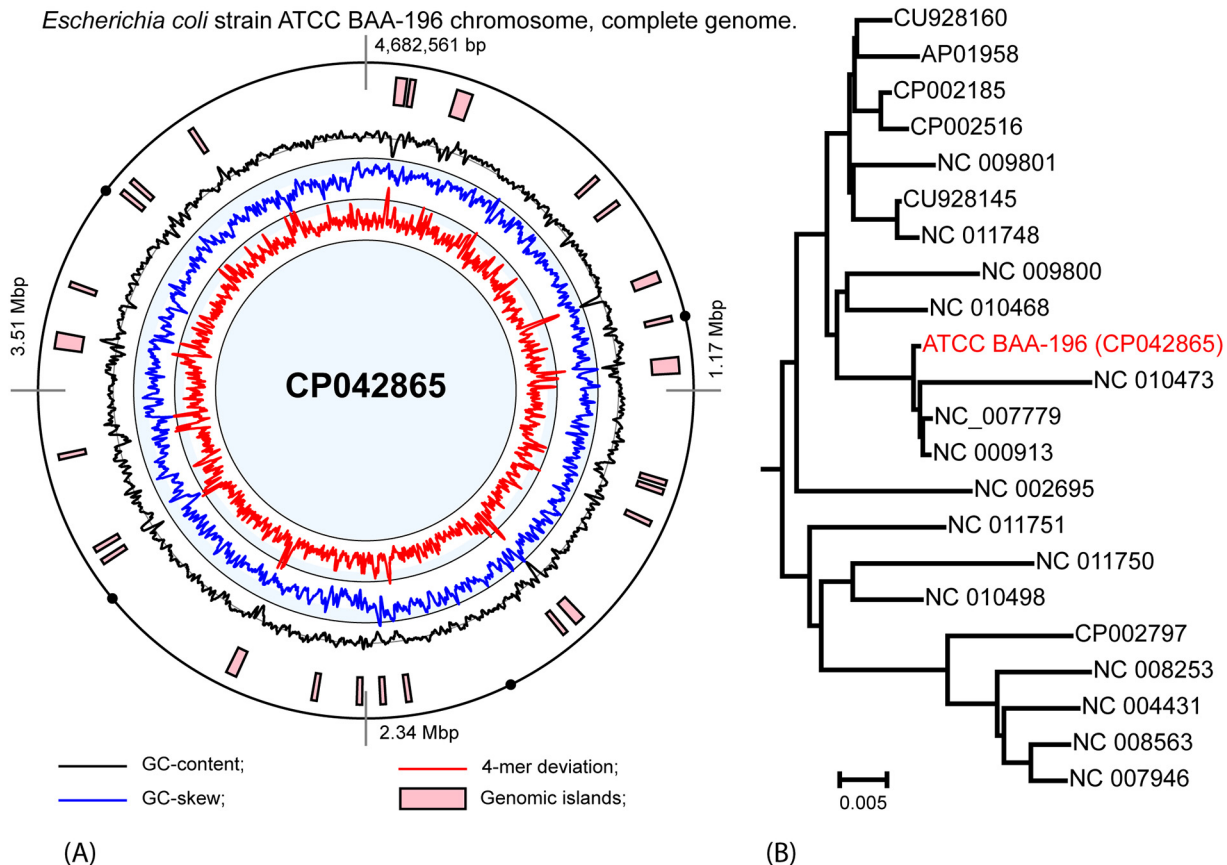


FIG 1 (A) Circular map of the chromosome, with predicted genomic islands and histograms of GC content, GC skew, and tetranucleotide pattern deviations calculated in an 8-kb sliding window stepping 2 kb. (B) Neighbor-joining phylogenetic tree of *E. coli* genomes, based on concatenated alignments of orthologous proteins.

resistance operon *terABCDW*, the arsenic resistance gene *arsR*, chloramphenicol and aminoglycoside acetyltransferases, and several other genes for antibiotic-modifying enzymes, drug resistance regulators, and multidrug efflux proteins. Virulence-associated fimbrial adhesin genes *ecpD* and *fimH**BGFE* were acquired horizontally.

The SMRT Link DNA modification pipeline (12) was used to profile epigenetic modifications in bacterial genomes. The most abundant DNA modification was *N*⁶-adenosine methylation in both strands, at GATC and GCAC(*N*⁶)GTT restriction sites (methylated nucleotides are underlined, and thymidine nucleotides opposing methylated ones on the complement strand are in italic type), corresponding to typical findings for *E. coli* DAM methyltransferases associated with EcoRV and EcoKI restriction-modification complexes (13, 14). Methylated GATC sequences often occurred in tandem with cytosine methylation in CRGKGATC motifs. Two other cytosine methylation motifs, CCAGGRAH and WCCCTGGYR, controlled by EcoRII family restriction-modification genes (15) showed alternative distribution patterns in the NC and FS genomes.

Data availability. NCBI accession numbers are [CP042865](#) and [CP042866](#) for the chromosome and plasmid, respectively, of the variant NC genome (PacBio reads [SRR10112463](#), [SRR10112464](#), and [SRR10112472](#)) and [CP042867](#) to [CP042870](#) for the chromosome and three plasmids of the variant FS genome (PacBio reads [SRR10112466](#) to [SRR10112468](#)).

ACKNOWLEDGMENTS

Sequencing was funded by grant O.0776 from the Industrial Development and Industrial Safety Committee of the Ministry of Industry and Infrastructural Development of the Republic of Kazakhstan. Genome assembly, annotation, and bioinformatic anal-

ysis, as well as student support of M.J. and S.T., were funded by the South African National Research Foundation (grant 105996). A.Y. acknowledges the University of Pretoria for her postdoctoral fellowship.

REFERENCES

- Rice LB, Willey SH, Papanicolaou GA, Medeiros AA, Eliopoulos GM, Moellering RC, Jr, Jacoby GA. 1990. Outbreak of ceftazidime resistance caused by extended-spectrum β -lactamases at a Massachusetts chronic-care facility. *Antimicrob Agents Chemother* 34:2193–2199. <https://doi.org/10.1128/aac.34.11.2193>.
- ATCC. 2019. *Escherichia coli* (Migula) Castellani and Chalmers (ATCC BAA-196): history. https://www.lgcstandards-atcc.org/products/all/BAA-196.aspx?geo_country=de#history.
- Rice LB, Marshall SH, Carias LL, Sutton L, Jacoby GA. 1993. Sequences of MGH-1, YOU-1, and YOU-2 extended-spectrum β -lactamase genes. *Antimicrob Agents Chemother* 37:2760–2761. <https://doi.org/10.1128/aac.37.12.2760>.
- Hsu BB, Ouyang J, Wong SY, Hammond PT, Klivanov AM. 2011. On structural damage incurred by bacteria upon exposure to hydrophobic polycationic coatings. *Biotechnol Lett* 33:411–416. <https://doi.org/10.1007/s10529-010-0419-1>.
- Ilin AI, Kulmanov ME, Korotetskiy IS, Islamov RA, Akhmetova GK, Lankina MV, Reva ON. 2017. Genomic insight into mechanisms of reversion of antibiotic resistance in multidrug resistant *Mycobacterium tuberculosis* induced by a nanomolecular iodine-containing complex FS-1. *Front Cell Infect Microbiol* 7:151. <https://doi.org/10.3389/fcimb.2017.00151>.
- Volodina GV, Davtyan TK, Kulmanov ME, Dzhumagazieva AB, Tursunova SK, Abekova AO, Bishimova IE, Abramova ZS, Kenzhebekova RT, Murzageldinova SG, Korotetskiy IS, Ilin AI. 2017. The effect of antibiotic-resistant and sensitive *Escherichia coli* on the production of pro-inflammatory cytokine response by human peripheral blood mononuclear cells. *J Clin Cell Immunol* 8:522. <https://doi.org/10.4172/2155-9899.1000522>.
- Okonechnikov K, Golosova O, Fursov M. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics* 28:1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>.
- Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Non-hybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* 10:563–569. <https://doi.org/10.1038/nmeth.2474>.
- 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>.
- Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 16:157. <https://doi.org/10.1186/s13059-015-0721-2>.
- Bezuidt O, Lima-Mendez G, Reva ON. 2009. SeqWord Gene Island Sniffer: a program to study the lateral genetic exchange among bacteria. *World Acad Sci Eng Technol* 3:2399–2404. <https://doi.org/10.5281/zenodo.1071430>.
- Joubert M, Reva ON, Korotetskiy IS, Shvidko SV, Shilov SV, Jumagaziyeva AB, Kenesheva ST, Suldina NA, Ilin AI. 2019. Assembly of complete genome sequences of negative-control and experimental strain variants of *Staphylococcus aureus* ATCC BAA-39 selected under the effect of the drug FS-1, which induces antibiotic resistance reversion. *Microbiol Resour Announc* 8:e00579-19. <https://doi.org/10.1128/MRA.00579-19>.
- May MS, Hattman S. 1975. Analysis of bacteriophage deoxyribonucleic acid sequences methylated by host- and R-factor-controlled enzymes. *J Bacteriol* 123:768–770.
- Geier GE, Modrich P. 1979. Recognition sequence of the *dam* methylase of *Escherichia coli* K12 and mode of cleavage of *Dpn* I endonuclease. *J Biol Chem* 254:1408–1413.
- Golovenko D, Manakova E, Tamulaitiene G, Grazulis S, Siksnyus V. 2009. Structural mechanisms for the 5'-CCWGG sequence recognition by the N- and C-terminal domains of EcoRII. *Nucleic Acids Res* 37:6613–6624. <https://doi.org/10.1093/nar/gkp699>.