




Complete Genome Sequence of *Escherichia coli* BE104, an MC4100 Derivative Lacking the Methionine Reductive Pathway

Brian P. Anton,^a Richard D. Morgan,^a Benjamin Ezraty,^b Bruno Manta,^a Frédéric Barras,^{b,c}  Mehmet Berkmen^a

^aNew England BioLabs, Ipswich, Massachusetts, USA

^bAix-Marseille Université, CNRS, Laboratoire de Chimie Bactérienne, Marseille, France

^cSAMe Unit, Department of Microbiology, Institut Pasteur, Paris, France

ABSTRACT In this announcement, we present the complete annotated genome sequence of an *Escherichia coli* MC4100 mutant strain, BE104. This strain has several methionine sulfoxide reductase gene deletions, making it ideal for studying enzymes that alter the redox state of methionine.

Escherichia coli is the predominant prokaryotic model organism with two popular ancestral strains, K-12 and B (1). In general, *E. coli* K-12 is used for DNA manipulation and cloning, while *E. coli* B is used mostly for protein expression. There are many *E. coli* K-12 derivatives currently used in research and production laboratories worldwide, including MC4100, engineered by Malcolm Casadaban (2). Of the 801 completely sequenced *E. coli* genomes, only 33 are from K-12-derived laboratory strains, and of those, only 1 is derived from MC4100. MC4100 has been used extensively since its inception (3), resulting in thousands of publications. Its genome has been sequenced (4), and its relationship to other *E. coli* K-12 strains has been analyzed (5).

The strain BE104, a derivative of an MC4100 methionine auxotroph mutant lacking five methionine sulfoxide reductase (Msr)-encoding genes, was previously constructed in order to characterize an enzymatic system (MsrPQ) responsible for repairing proteins containing methionine sulfoxide in the bacterial periplasm (6). Briefly, BE104 was derived from MC4100, a methionine auxotroph mutant, by (i) a series of P1vir crosses to delete all cytoplasmic Msr-encoding genes (*msrA*, *msrB*, *msrC*, and *bisC*) by replacement with corresponding alleles from Keio knockout (KO) strains (7), (ii) selection for suppressor strains that could reduce methionine sulfoxide, and (iii) deletion of *msrP*, a consequently discovered periplasmic MsrP enzyme (5).

As BE104 is being used in our research and will be further engineered, we sequenced its genome using the Pacific Biosciences (PacBio) RS II sequencing platform, as described previously (8). BE104 was grown in standard rich medium (10 g/liter tryptone, 5 g/liter yeast extract, 5 g/liter NaCl, NaOH [pH 7.2]) at 30°C, and its genomic DNA (gDNA) was isolated using the Monarch gDNA kit (New England BioLabs). A SMRTbell library was constructed from 5 µg gDNA and sheared to ~10 kb using a g-TUBE (Covaris). The library was sequenced on two single-molecule real-time (SMRT) cells using P6-C4 chemistry. The first cell yielded a 258-Mb sequence from 22,107 (15%) P1 reads, with a mean polymerase read length of 11,672 bases and a mean read insert length of 6,096 bases (180-minute data collection time). A second cell was sequenced to increase coverage, yielding 905 Mb of sequence from 63,262 (42%) P1 reads with a mean polymerase read length of 14,318 bases and mean read insert of 4,594 bases (240-minute data collection time). Sequencing reads were processed and assembled with the Pacific Biosciences SMRT Analysis v2.3.0 software using the HGAP3 protocol (expected genome size, 5 Mb; filters set to minimum subread length, 1,000 bp; minimum polymerase read length, 2,000 bp; minimum read quality, 0.80) and polished

Citation Anton BP, Morgan RD, Ezraty B, Manta B, Barras F, Berkmen M. 2019. Complete genome sequence of *Escherichia coli* BE104, an MC4100 derivative lacking the methionine reductive pathway. Microbiol Resour Announc 8:e00721-19. <https://doi.org/10.1128/MRA.00721-19>.

Editor Kenneth M. Stedman, Portland State University

Copyright © 2019 Anton 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 Mehmet Berkmen, berkmen@neb.com.

Received 17 June 2019

Accepted 19 June 2019

Published 11 July 2019

using Quiver (8). The 1.1 Gb of sequence assembled into a single closed circular genome of 4,775,122 bp with a mean coverage of 200-fold and a GC content of 50.8%. The assembled sequence was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.8.

The following expected deletions were observed: *metB* (*O*-succinyl homoserine lyase) is disrupted by a 2-base deletion/frameshift; *msrA* is disrupted by insertion with a spectinomycin cassette; and the *msr* loci *msrB* (formerly *yeaA*), *msrC* (formerly *yebR*), and *msrP* (formerly *yedY*) are all deleted, as is *bisC* (biotin sulfoxide reductase) (9). As expected, the *yedV::IS2* insertion was not observed, as the *msrP::kan* deletion was transduced using the kanamycin marker from the Keio collection (JW1954), cotransducing the wild-type (wt) *yedV* allele.

This strain should be of general use to the research community for studying protein redox states and would be an important addition to the repertoire of sequenced *E. coli* genomes.

Data availability. The complete *E. coli* BE104 genome sequence has been deposited in GenBank with the accession number [CP040643](https://doi.org/10.1093/genbank/CP040643). The raw data are available in the NCBI Sequence Read Archive (SRA) with the accession number [PRJNA544505](https://doi.org/10.1093/bioinformatics/PRJNA544505).

REFERENCES

- Studier FW, Daegelen P, Lenski RE, Maslov S, Kim JF. 2009. Understanding the differences between genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3) and comparison of the *E. coli* B and K-12 genomes. *J Mol Biol* 394:653–680. <https://doi.org/10.1016/j.jmb.2009.09.021>.
- Journal of Bacteriology. 2010. Remembering Malcolm J. Casadaban. *J Bacteriol* 192:4261–4263. <https://doi.org/10.1128/JB.00484-10>.
- Casadaban MJ. 1976. Transposition and fusion of the lac genes to selected promoters in *Escherichia coli* using bacteriophage lambda and Mu. *J Mol Biol* 104:541–555. [https://doi.org/10.1016/0022-2836\(76\)90119-4](https://doi.org/10.1016/0022-2836(76)90119-4).
- Ferenci T, Zhou Z, Betteridge T, Ren Y, Liu Y, Feng L, Reeves PR, Wang L. 2009. Genomic sequencing reveals regulatory mutations and recombinational events in the widely used MC4100 lineage of *Escherichia coli* K-12. *J Bacteriol* 191:4025–4029. <https://doi.org/10.1128/JB.00118-09>.
- Peters JE, Thate TE, Craig NL. 2003. Definition of the *Escherichia coli* MC4100 genome by use of a DNA array. *J Bacteriol* 185:2017–2021. <https://doi.org/10.1128/JB.185.6.2017-2021.2003>.
- Gennaris A, Ezraty B, Henry C, Agrebi R, Vergnes A, Oheix E, Bos J, Leverrier P, Espinosa L, Szewczyk J, Vertommen D, Iranzo O, Collet JF, Barras F. 2015. Repairing oxidized proteins in the bacterial envelope using respiratory chain electrons. *Nature* 528:409–412. <https://doi.org/10.1038/nature15764>.
- Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2:2006.0008. <https://doi.org/10.1038/msb4100050>.
- 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>.
- Ezraty B, Bos J, Barras F, Aussel L. 2005. Methionine sulfoxide reduction and assimilation in *Escherichia coli*: new role for the biotin sulfoxide reductase BisC. *J Bacteriol* 187:231–237. <https://doi.org/10.1128/JB.187.1.231-237.2005>.