AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Complete Genome Sequence of *Escherichia coli* ME8067, an Azide-Resistant Laboratory Strain Used for Conjugation Experiments

Yasufumi Matsumura,[a] Masaki Yamamoto,[a] Satoshi Nakano,[a] Miki Nagao[a]

[a]Department of Clinical Laboratory Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

ABSTRACT We report here the complete genome sequence of *Escherichia coli* ME8067, an azide-resistant laboratory strain used for conjugation experiments. The ME8067 genome was closely related to *E. coli* strain K-12 substrain W3110. This genome sequence will support further genetic analysis of conjugative elements.

Escherichia coli strain ME8067 (strain WD8014) is a laboratory *E. coli* strain that cannot ferment galactose due to its lack of UDP glucose pyrophosphorylase activity (1). This strain has been used for conjugation experiments as a recipient strain (2–4) because it is negative for fertility factors and is resistant to sodium azide. Recent advances in sequencing technology have enabled the sequencing of transconjugants, which contribute to a comprehensive understanding of transferable elements. Here, we present a complete genome sequence of *E. coli* ME8067, which is important for accurate nucleotide sequence analysis.

Whole-genome DNA was sequenced using both Illumina NextSeq 500 (150-bp paired-end) and Oxford Nanopore Technologies MinION (R9.4 flow cell) systems. A circular chromosome was obtained using a *de novo* hybrid assembly pipeline of Unicycler version 0.4.4 (5). The assembly resulted in average coverages of 186× with the NextSeq 500 system and 190× with the MinION system. The genome was annotated using the NCBI Prokaryotic Genome Annotation Pipeline (6).

The complete genome contains a 4,614,635-bp chromosome with a GC content of 50.8%, 4,413 coding sequences, 81 tRNA-coding genes, and 22 rRNA-coding operons. The Achtman multilocus sequence typing scheme (http://mlst.warwick.ac.uk/mlst/dbs/Ecoli) classified ME8067 into sequence type 10, which is a lineage of *E. coli* K-12 laboratory strains. Using the 41 *E. coli* K-12 genomes available on the NCBI assembly resource (https://www.ncbi.nlm.nih.gov/assembly) as of April 28, 2018, we constructed a phylogenetic tree based on core single-nucleotide polymorphisms (SNPs) with kSNP3 (7). ME8067 belonged to a cluster made up of the W3110 substrain and its derivatives (HMS174 and ZK126). This cluster was closely related to a cluster that included 15 MG1655 substrains.

Genomic comparison with the W3110 genome (GenBank accession no. AP009048) (8) using progressiveMauve (9) revealed that ME8067 had deletions totaling 53,580 bp, including *lacZYA*, IS*5* (6 copies), IS*2* (3 copies), and tRNAs. Deficiency of the *lac* operon can explain the lactose-nonfermenting characteristic of ME8067. Insertions totaling 34,648 bp included IS*5* (3 copies), tRNAs, and RtT small RNAs (sRNAs). Among the 10 prophages present in MG1655 (GenBank accession no. U00096), CPZ-55 was absent in W3110 but present in ME8067. All of the other 9 prophages were present in W3110, though ME8067 had a partial deletion of CP4-6, an insertion of IS*3* to DLP12, and an inversion of an internal segment within e14. Single-nucleotide polymorphisms (SNPs) and small nucleotide insertions/deletions (indels) were sought using SNIT (10) and annotated using snpEff (11). Between the W3110 and ME8067 genomes, 693 SNPs and

Received 5 May 2018 **Accepted** 8 May 2018 **Published** 21 June 2018

Citation Matsumura Y, Yamamoto M, Nakano S, Nagao M. 2018. Complete genome sequence of *Escherichia coli* ME8067, an azide-resistant laboratory strain used for conjugation experiments. Genome Announc 6:e00515-18. https://doi.org/10.1128/genomeA.00515-18.

Address correspondence to Yasufumi Matsumura, yazblood@kuhp.kyoto-u.ac.jp.

30 indels were present, resulting in 406 nonsynonymous mutations. An SNP in the *secA* gene resulting in the amino acid alteration A112T was found and may be associated with azide resistance (12). Another SNP in *galU* (amino acid alteration, P14S) may be associated with the lack of UDP glucose pyrophosphorylase activity.

The complete genome sequence of strain ME8067 revealed that it belongs to a lineage of *E. coli* strain K-12 substrain W3110 and will aid in comprehensive genetic analysis for conjugation experiments.

Accession number(s). The complete genome sequence of the chromosome of *E. coli* ME8067 has been deposited at GenBank under accession no. CP028703.

ACKNOWLEDGMENTS

We thank the National BioResource Project (National Institute of Genetics, Japan) for providing the *E. coli* ME8067 strain.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Hossain SA, Tanizawa K, Kazuta Y, Fukui T. 1994. Overproduction and characterization of recombinant UDP-glucose pyrophosphorylase from *Escherichia coli* K-12. J Biochem 115:965–972. https://doi.org/10.1093/oxfordjournals.jbchem.a124446.
2. Yamamoto M, Nagao M, Hotta G, Matsumura Y, Matsushima A, Ito Y, Takakura S, Ichiyama S. 2012. Molecular characterization of IMP-type metallo-β-lactamases among multidrug-resistant *Achromobacter xylosoxidans*. J Antimicrob Chemother 67:2110–2113. https://doi.org/10.1093/jac/dks179.
3. Yamamoto M, Matsumura Y, Gomi R, Matsuda T, Tanaka M, Nagao M, Takakura S, Uemoto S, Ichiyama S. 2016. Interspecies dissemination of a mobilizable plasmid harboring bla_{IMP-19} and the possibility of horizontal gene transfer in a single patient. Antimicrob Agents Chemother 60:5412–5419. https://doi.org/10.1128/AAC.00933-16.
4. Matsumura Y, Tanaka M, Yamamoto M, Nagao M, Machida K, Ito Y, Takakura S, Ogawa K, Yoshizawa A, Fujimoto Y, Okamoto S, Uemoto S, Ichiyama S. 2015. High prevalence of carbapenem resistance among plasmid-mediated AmpC β-lactamase-producing *Klebsiella pneumoniae* during outbreaks in liver transplantation units. Int J Antimicrob Agents 45:33–40. https://doi.org/10.1016/j.ijantimicag.2014.08.015.
5. 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.
6. 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.
7. Gardner SN, Slezak T, Hall BG. 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics 31:2877–2878. https://doi.org/10.1093/bioinformatics/btv271.
8. Hayashi K, Morooka N, Yamamoto Y, Fujita K, Isono K, Choi S, Ohtsubo E, Baba T, Wanner BL, Mori H, Horiuchi T. 2006. Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. Mol Syst Biol 2:2006.0007. https://doi.org/10.1038/msb4100049.
9. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147. https://doi.org/10.1371/journal.pone.0011147.
10. Vijaya Satya R, Zavaljevski N, Reifman J. 2011. SNIT: SNP identification for strain typing. Source Code Biol Med 6:14. https://doi.org/10.1186/1751-0473-6-14.
11. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly 6:80–92. https://doi.org/10.4161/fly.19695.
12. Oliver DB, Cabelli RJ, Dolan KM, Jarosik GP. 1990. Azide-resistant mutants of *Escherichia coli* alter the SecA protein, an azide-sensitive component of the protein export machinery. Proc Natl Acad Sci U S A 87:8227–8231. https://doi.org/10.1073/pnas.87.21.8227.