Whole-Genome Sequence of *Escherichia coli* Serotype O157:H7 Strain B6914-ARS

Gaylen A. Uhlich, Erin R. Reichenberger, Bryan J. Cottrell, Pina Fratamico, Elisa Andreozzi
Molecular Characterization of Foodborne Pathogens Research Unit, USDA, Agriculture Research Service, Eastern Regional Research Center, Wyndmoor, Pennsylvania, USA

**ABSTRACT** *Escherichia coli* serotype O157:H7 strain B6914-MS1 is an isolate from the Centers for Disease Control and Prevention that is missing both Shiga toxin genes and has been used extensively in applied research studies. Here we report the genome sequence of strain B6914-ARS, a B6914-MS1 clone that has unique biofilm properties.

Shiga toxin-producing *Escherichia coli* (STEC) organisms cause intestinal disease characterized by hemorrhagic colitis (HC) that may progress to hemolytic uremic syndrome (1). STEC O157:H7 is the serotype most frequently associated with both sporadic cases and large outbreaks of HC in the United States (2). The Shiga toxins (Stx1 and Stx2) are important virulence factors expressed by STEC strains, but Stx2 exhibits the highest correlation to severe disease (3). The stx genes are generally carried on prophage integrated in the host bacterium chromosome.

*E. coli* O157:H7 strain B6914-MS1 (B6914) is a human fecal isolate that was identified by the Centers for Disease Control and Prevention (CDC), deposited into the American Type Culture Collection (ATCC), and made publicly available as strain ATCC 43888. Strain B6914 does not produce or carry the genes for Stx1 or Stx2 toxins and has been classified as a biosafety level one (BSL-1) organism according to the *Biosafety in Microbiological and Biomedical Laboratories* (BMBL) guidelines (4). Strain B6914 has been used extensively in applied research studies due its low potential for pathogenicity.

Clinical strains of serotype O157:H7 generally have attenuated biofilm properties as a result of prophage insertions in the mlrA gene, a transcription factor required for full expression of the central biofilm regulator, CsgD (5). A B6914 isolate from our collection (designated B6914-ARS) carries the mlrA prophage, but a recent study showed that biofilm formation was restored in that strain by prophage excisions at a higher than expected rate by an undetermined mechanism (6). Therefore, we have sequenced the entire genome of strain B6914-ARS to investigate the genetic mechanisms involved with biofilm production and/or other relevant phenotypes described in studies of B6914-MS1 derivatives.

Large DNA fragments for single-molecule real-time (SMRT) sequencing were extracted from a frozen cell pellet (from 5 mL of overnight culture) of B6914-ARS using the Qiagen Genomic-tip 100/G kit (Qiagen, Valencia, CA). SMRT sequencing was done at the University of Delaware Sequencing and Genotyping Center using the PacBio RS II SMRT DNA sequencing system (Pacific Biosciences, Menlo Park, CA). The *de novo* genome assembly was done with Hierarchical Genome Assembly Process 3.0 (HGAP 3.0) software, and the genome annotation was performed via the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (7).

The unclosed B6914-ARS draft assembly consists of a predicted 5,537,960 paired-nucleotide chromosome in three contigs with a combined G+C content of 50.3% and a single large plasmid of 115,027 paired nucleotides with a G+C content of 48.2%. The
PHASTER prophage locator and annotation tool identified 11 intact, 3 incomplete, and 3 questionable prophage regions (8, 9). A locus of enterocyte effacement (LEE) pathogenicity island with high similarity in genetic content and arrangement to that of O157:H7 outbreak strain Sakai (10) (GenBank accession number NC_002695) was also identified. The genes encoding the A and B subunits of both Stx1 and Stx2 are missing from the B6914-ARS genome. However, a prophage similar to SP15 (which carries Stx1 in Sakai) was identified in the proximal portion of the mlfA gene.

**Accession number(s).** This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession numbers NOKN01000001 to NOKN01000003 (chromosome) and NOKN01000004 (plasmid). The versions described in this paper are the first versions.

**ACKNOWLEDGMENTS**

Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. The USDA is an equal opportunity provider and employer.

**REFERENCES**

1. Tarr PI, Gordon CA, Chandler WL. 2005. Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. Lancet 365:1073–1086. https://doi.org/10.1016/S0140-6736(05)71144-2.

2. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM. 2011. Foodborne illness acquired in the United States—major pathogens. Emerg Infect Dis 17:7–15. https://wwwnc.cdc.gov/eid/article/17/1/P1-1101_article.

3. Boerlin P, McEwen SA, Boerlin-Petzold F, Wilson JB, Johnson RP, Gyles CL. 1999. Associations between virulence factors of Shiga toxin-producing *Escherichia coli* and disease in humans. J Clin Microbiol 37:497–503.

4. U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institutes of Health. 2009. Biosafety in microbiological and biomedical laboratories, 5th ed. Centers for Disease Control and Prevention, Atlanta, GA. [https://www.cdc.gov/biosafety/publications/bmbl5/BMBL.pdf](https://www.cdc.gov/biosafety/publications/bmbl5/BMBL.pdf).

5. Uhlich GA, Chen CY, Cottrell BJ, Hofmann CS, Dudley EG, Strobaugh TP, Jr., Nguyen LH. 2013. Phage insertion in mlfA and variations in rpoS limit curli expression and biofilm formation in *Escherichia coli* serotype O157: H7. Microbiology 159:1586–1596. https://doi.org/10.1099/mic.0.066118-0.

6. Chen CY, Nguyen LH, Cottrell BJ, Irwin PL, Uhlich GA. 2016. Multiple mechanisms responsible for strong Congo-red-binding variants of *Escherichia coli* O157:H7 strains. Pathog Dis 74:ftv123. https://doi.org/10.1093/femspd/ftv123.

7. Tatusova T, DiCuccio M, Badretdin A, Chevtverin 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.

8. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. Nucleic Acids Res 39:W347–W352. https://doi.org/10.1093/nar/gkr485.

9. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W347–W352. https://doi.org/10.1093/nar/gkw387.

10. Hayashi T, Makino K, Ohnishi M, Kurokawa K, Ishii K, Yokoyama K, Han CG, Ohitsubo E, Nakayama K, Murata T, Tanaka M, Tobe T, Iida T, Takami H, Honda T, Sasakawa C, Ogasawara N, Yasunaga T, Kuhara S, Shibata T, Hattori M, Shinagawa H. 2001. Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. DNA Res 8:11–22. [https://doi.org/10.1093/dnares/8.1.11](https://doi.org/10.1093/dnares/8.1.11).