genomeA
Journals.ASM.org

Complete Genome Sequence of the *Arcobacter butzleri* Cattle Isolate 7h1h

J. Yvette Merga,[a] Craig Winstanley,[a] Nicola J. Williams,[a] Emma Yee,[b] William G. Miller[b]

Institute of Infection and Global Health, University of Liverpool, Liverpool, United Kingdom[a]; Produce Safety and Microbiology Research Unit, U.S. Department of Agriculture, Albany, California, USA[b]

***Arcobacter butzleri* strain 7h1h was isolated in the United Kingdom from the feces of a clinically healthy dairy cow. The genome of this isolate was sequenced to completion. Here, we present the annotation and analysis of the completed 7h1h genome, along with a comparison of this genome to the existing *A. butzleri* genomes.**

Received 23 July 2013 **Accepted** 24 July 2013 **Published** 22 August 2013

Citation Merga JY, Winstanley C, Williams NJ, Yee E, Miller WG. 2013. Complete genome sequence of the *Arcobacter butzleri* cattle isolate 7h1h. Genome Announc. 1(4): e00655-13. doi:10.1128/genomeA.00655-13.

Address correspondence to J. Yvette Merga, ysloane@liv.ac.uk, or William G. Miller, William.Miller@ars.usda.gov.

Arcobacter butzleri is a member of the *Epsilonproteobacteria*, a taxonomic division that also contains the established pathogens *Campylobacter jejuni* and *Helicobacter pylori*. *A. butzleri* has been isolated from water, food animals, and multiple food sources (1). *A. butzleri* has also been associated with human gastroenteritis (1). *A. butzleri* strain 7h1h originated from a clinically healthy dairy cow in Cheshire, United Kingdom. It was assigned to a species using a previously used PCR (2) and was selected for sequencing based on the high quality of the amplicon that was produced. Multilocus sequence typing of the isolate (3) revealed it was a novel sequence type (ST-347), which was closely related to sequence type 303 (ST-303), isolated in the same study (with 6 allele matches). No other isolates shared >3 alleles with 7h1h. The complete genome sequences of two *A. butzleri* strains, the human clinical isolate RM4018 (4) and ED-1, isolated from a microbial fuel cell (5), were determined previously, and an incomplete 7h1h genome sequence was compared with that of RM4018 (6); here, we present the complete genome sequence of strain 7h1h.

Genome sequencing was performed using general and paired-end (8 to 12 kb) libraries and was generated on a Roche 454 FLX+ genome sequencer with Titanium chemistry. Newbler assembler (v2.6) was used to assemble 186,913 shotgun and 108,665 paired-end reads into a single scaffold of 27 contigs, providing 52× coverage. Scaffold gaps were filled using the 454 repeat contigs and the Perlscript Contig_extender3. Contig junctions were validated using amplification and Sanger sequencing. All 454 base calls were validated using 1,164,896 Illumina MiSeq reads, providing an additional 78× coverage.

The *A. butzleri* 7h1h genome size is 2,253,233 bp, with a G+C content of 27.06%. Protein-coding, rRNA-coding, and tRNA-coding genes were identified using GeneMark (v2.8; http://exon.gatech.edu/GeneMark/gmhmm2_prok.cgi), RNAmmer v1.2 (7), and tRNAscan-SE (8), respectively. The gene start points were curated using Artemis (9). Final annotation was performed by BLASTp comparison to the proteomes of RM4018 and/or ED-1 or to proteins in the NCBI nonredundant database, and by identification of Pfam domains (v.26.0 [10]). The 7h1h genome is predicted to carry 2,199 genes, 5 ribosomal RNA operons, and 54 tRNAs.

The 7h1h genome is highly syntenic to both the RM4018 and ED-1 genomes; no large-scale rearrangements were observed with respect to the other two genomes. Of the 2,199 genes carried by strain 7h1h, 1,946 (88%) were also identified in either RM4018 or ED-1. Of the remainder, 111 genes are either contained within an integrated element or are predicted to encode surface structure-associated proteins. Three integrated elements, bounded by 13- to 45-bp direct repeats and adjacent to tRNA-coding genes, were identified in the 7h1h genome. Unique to strain 7h1h are four toxin-antitoxin family gene pairs and two families of insertion sequences (11 insertion sequence [IS] elements in total) that are unrelated to the mobile element identified in strain ED-1. Also present in strain 7h1h are genes encoding ATP-independent (urease) and ATP-dependent (urea carboxylase/allophanate hydrolase [11]) urea degradation pathways.

Nucleotide sequence accession number. The genome sequence of *A. butzleri* strain 7h1h has been deposited in GenBank under the accession no. CP006615.

ACKNOWLEDGMENTS

Illumina MiSeq reads were provided by SeqMatic (Union City, CA).

C.W. and N.J.W. acknowledge funding from the Higher Education Funding Council for England (HEFCE) and the United Kingdom Department for Environment, Food, and Rural Affairs, as part of the Veterinary Training and Research Initiative.

REFERENCES

1. **Wesley IV, Miller WG.** 2010. *Arcobacter*: an opportunistic human foodborne pathogen?, p 185–211. *In* Scheld WM, Grayson ML, Hughes JM (ed), Emerging infections, vol 9. American Society for Microbiology, Washington, DC.
2. **Houf K, Tutenel A, De Zutter L, Van Hoof J, Vandamme P.** 2000. Development of a multiplex PCR assay for the simultaneous detection and identification of *Arcobacter butzleri, Arcobacter cryaerophilus* and *Arcobacter skirrowii*. FEMS Microbiol. Lett. **193**:89–94.
3. **Miller WG, Wesley IV, On S, Houf K, Megraud F, Wang G, Yee E,**

Srijan A, Mason C. 2009. First multi-locus sequence typing scheme for *Arcobacter* spp. BMC Microbiol. **9:**196. doi:10.1186/1471-2180-9-196.
4. **Miller WG, Parker CT, Rubenfield M, Mendz GL, Wösten MM, Ussery DW, Stolz JF, Binnewies TT, Hallin PF, Wang G, Malek JA, Rogosin A, Stanker LH, Mandrell RE.** 2007. The complete genome sequence and analysis of the *Epsilonproteobacterium Arcobacter butzleri*. PLoS One **2:**e1358. doi:10.1371/journal.pone.0001358.
5. **Toh H, Sharma VK, Oshima K, Kondo S, Hattori M, Ward FB, Free A, Taylor TD.** 2011. Complete genome sequences of *Arcobacter butzleri* ED-1 and *Arcobacter* sp. strain L, both isolated from a microbial fuel cell. J. Bacteriol. **193:**6411–6412.
6. **Merga JY, Williams NJ, Miller WG, Leatherbarrow AJ, Bennett M, Hall N, Ashelford KE, Winstanley C.** 2013. Exploring the diversity of *Arcobacter butzleri* from cattle in the UK using MLST and whole genome sequencing. PLoS One **8:**e55240. doi:10.1371/journal.pone.0055240.
7. **Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108.
8. **Schattner P, Brooks AN, Lowe TM.** 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. Nucleic Acids Res. **33:**W686–W689. doi:10.1093/nar/gki366.
9. **Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B.** 2000. Artemis: sequence visualization and annotation. Bioinformatics **16:**944–945.
10. **Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J, Heger A, Holm L, Sonnhammer EL, Eddy SR, Bateman A, Finn RD.** 2012. The Pfam protein families database. Nucleic Acids Res. **40:**D290–D301. doi:10.1093/nar/gkr1065.
11. **Kanamori T, Kanou N, Kusakabe S, Atomi H, Imanaka T.** 2005. Allophanate hydrolase of *Oleomonas sagaranensis* involved in an ATP-dependent degradation pathway specific to urea. FEMS Microbiol. Lett. **245:**61–65.