

Draft Genome Sequences of Three *Arcobacter* Strains of Pig and Dairy Cattle Manure Origin

Zaky Adam,^a Kerri Whiteduck-Leveillee,^a Michel Cloutier,^a James T. Tambong,^a Wen Chen,^a Christopher T. Lewis,^a C. André Lévesque,^a Edward Topp,^b David R. Lapen,^a Guylaine Talbot,^c Izhar U. H. Khan^a

Eastern Cereal and Oilseed Research Centre (ECORC), Agriculture and Agri-Food Canada, Ottawa, Ontario, Canada^a; Southern Crop Protection and Food Research Centre, Agriculture and Agri-Food Canada, London, Ontario, Canada^b; Dairy and Swine Research Development Centre, Agriculture and Agri-Food Canada, Sherbrooke, Quebec, Canada^c

The genus *Arcobacter* has been associated with human illness and fecal contamination by humans and animals. Here, we announce the draft genome sequences of three strains of *Arcobacter* species cultured from pig and dairy cattle manure tanks. This information will assist in the characterization of features related to host specificities and identify potential pathogenic health risks to humans and animals.

Received 7 April 2014 Accepted 1 May 2014 Published 15 May 2014

Citation Adam Z, Whiteduck-Leveillee K, Cloutier M, Tambong JT, Chen W, Lewis CT, Lévesque CA, Topp E, Lapen DR, Talbot G, Khan IUH. 2014. Draft genome sequences of three *Arcobacter* strains of pig and dairy cattle manure origin. *Genome Announc.* 2(3):e00377-14. doi:10.1128/genomeA.00377-14.

Copyright © 2014 Adam et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](http://creativecommons.org/licenses/by/3.0/).

Address correspondence to Izhar U. H. Khan, izhar.khan@agr.gc.ca.

Arcobacter species have been implicated in human infections, which has led to these microbes being characterized as potential food-borne and water-borne pathogens (1). Moreover, the presence of *Arcobacter* spp. in livestock indicates that animals may also serve as possible reservoirs for *Arcobacter* species (2). Therefore, the prevalence, relative abundance, sources, and route(s) of human and animal exposures to *Arcobacter* spp. need to be identified and considered health concerns.

In 2010, *Arcobacter* strains AF1430 and AF1440 were recovered from a pig manure tank, and strain AF1581 was isolated from a dairy cattle manure tank from farms located in the southeast region of Ottawa, Ontario, Canada. The genome sequencing project was initiated based on multiple pairwise alignment of 16S rRNA gene, where these *Arcobacter* strains showed high phylogenetic relatedness (91.3 to 97.5% identity) to other known *Arcobacter* species. The whole-genome sequencing of the purified genomic DNA was performed using paired-end sequencing on Illumina HiSeq 2500 with TruSeq version 3 chemistry at the National Research Council Canada (Saskatoon, Saskatchewan, Canada). Following fragmentation, end reparation, and sample tagging, the sequencer produced 101-bp paired-end reads that were obtained from 300-bp inserts, yielding average coverages of 572×, 419×, and 893× for strains AF1430, AF1440, and AF1581, respectively. The program FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) was applied to check the quality of the reads.

De novo assembly was performed using ABySS version 1.3.6 (3). SSPACE version 2.0 (4) was applied to extend and merge the resulting scaffolds based on read-pair information and short overlaps to reduce the number of scaffolds. The gaps between the short scaffolds, which are contained within the large scaffolds, were closed using GapFiller version 1.11 (5) by replacing the unknown nucleotides (Ns) with true nucleotides. The resulting scaffolds of the three *Arcobacter* strains were ordered by alignment to the reference genomes of *Arcobacter butzleri* ED-1 and *Arcobacter cibarius* LMG 21996 using Mauve Contig Mover version 2.3.1 (6–8). The genome G+C contents were relatively low (26.4 to 26.8%) but were close to those of the reference genomes (~27%) (7, 8). The genome information data for each strain are summarized in Table 1.

The draft genomes were further annotated for gene prediction using the RAST annotation server (9). We identified similar, but not identical, numbers of protein-coding sequences in all three strains; strain AF1430 contains 2,223 predicted protein-coding sequences, including 1,636 functional, 241 proposed functional, and 346 hypothetical proteins with 42 predicted noncoding RNAs (39 tRNAs, 1 pseudo-tRNAs, and 2 rRNAs consisting of 1 copy each of 16S rRNA and 23S rRNA genes), compared to strain AF1440, which contains 2,245 predicted protein-coding sequences, including 1,655 functional, 241 proposed functional, and 349 hypothetical proteins with 64 predicted noncoding RNAs

TABLE 1 Summary of data for newly sequenced genomes of three strains of *Arcobacter* species

<i>Arcobacter</i> sp. strain	Source	Accession no.	Genome size (bp)	N_{50} (bp)	No. of scaffolds (>300 bp)	G+C content (%)
AF1430	Pig manure	JATO00000000	2,236,207	338,260	27	26.41
AF1440	Pig manure	JARU00000000	2,287,768	466,048	29	26.66
AF1581	Dairy cattle manure	JARV00000000	2,260,089	414,791	22	26.84

(51 tRNAs, 1 pseudo-tRNAs, and 12 rRNAs consisting of 6 copies each of 16S rRNA and 23S rRNA genes). On the other hand, strain AF1581 contains 2,224 predicted protein-coding sequences, including 1,625 functional, 223 proposed functional, and 376 hypothetical proteins with 75 predicted noncoding RNAs (55 tRNAs, 1 pseudo-tRNAs, and 19 rRNAs consisting of 10 copies of 16S rRNA and 9 copies of 23S rRNA genes).

Nucleotide sequence accession numbers. The draft genome sequences of the *Arcobacter* strains AF1430, AF1440, and AF1581 in this study have been deposited as whole-genome shotgun projects at DDBJ/EMBL/GenBank under the accession no. [JATO00000000](https://www.ncbi.nlm.nih.gov/nuclink/JATO00000000), [JARU00000000](https://www.ncbi.nlm.nih.gov/nuclink/JARU00000000), and [JARV00000000](https://www.ncbi.nlm.nih.gov/nuclink/JARV00000000), respectively. The version of each strain described in this paper is the first version, with accession no. JATO01000000, JARU01000000, and JARV01000000.

ACKNOWLEDGMENTS

Funding for this study was provided by the Canadian Safety and Security Program (CSSP) (project CRTI 09-462RD) and Agriculture and Agri-Food Canada (project no. 1800).

We thank lab technicians and co-op students for their dedicated help.

REFERENCES

- Samie A, Obi CL, Barrett LJ, Powell SM, Guerrant RL. 2007. Prevalence of *Campylobacter* species, *Helicobacter pylori* and *Arcobacter* species in stool samples from the Venda region, Limpopo, South Africa: studies using molecular diagnostic methods. *J. Infect.* 54:558–566. <http://dx.doi.org/10.1016/j.jinf.2006.10.047>.
- Kabaya H, Maruyama S, Morita Y, Kubo M, Yamamoto K, Arai S, Izumi T, Kobayashi Y, Katsube Y, Mikami T. 2003. Distribution of *Arcobacter* species among livestock in Japan. *Vet. Microbiol.* 93:153–158. [http://dx.doi.org/10.1016/S0378-1135\(02\)00312-7](http://dx.doi.org/10.1016/S0378-1135(02)00312-7).
- Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. *Genome Res.* 19:1117–1123. <http://dx.doi.org/10.1101/gr.089532.108>.
- Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27:578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
- Boetzer M, Pirovano W. 2012. Toward almost closed genomes with Gap-Filler. *Genome Biol.* 13:R56. <http://dx.doi.org/10.1186/gb-2012-13-6-r56>.
- Rissman AI, Mau B, Biehl BS, Darling AE, Glasner JD, Perna NT. 2009. Reordering contigs of draft genomes using the Mauve Aligner. *Bioinformatics* 25:2071–2073. <http://dx.doi.org/10.1093/bioinformatics/btp356>.
- 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. <http://dx.doi.org/10.1128/JB.06084-11>.
- Adam Z, Whiteduck-Leveillee K, Cloutier M, Chen W, Lewis CT, Lévesque CA, Topp E, Lapen DR, Tambong JT, Talbot G, Khan IUH. 2014. Draft genome sequence of *Arcobacter cibarius* strain LMG21996^T, isolated from broiler carcasses. *Genome Announc.* 2(1):e00034-14.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formisano 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. <http://dx.doi.org/10.1186/1471-2164-9-75>.