

# Draft Genome Sequence of *Cronobacter sakazakii* Clonal Complex 45 Strain HPB5174, Isolated from a Powdered Infant Formula Facility in Ireland

Arthur W. Pightling, Franco Pagotto

Listeriosis Reference Service, Research Division, Bureau of Microbial Hazards, Food Directorate, Health Products and Food Branch, Health Canada, Ottawa, Ontario, Canada

***Cronobacter sakazakii* is a food-borne pathogenic bacterium that may cause severe illness in neonates and the elderly. We present the genome sequence of a rare strain (ST40, CC45), commonly found in multiple food processing facilities and in powdered infant formula and only indicted in a single clinical case.**

Received 11 July 2014 Accepted 21 July 2014 Published 7 August 2014

**Citation** Pightling AW, Pagotto F. 2014. Draft genome sequence of *Cronobacter sakazakii* clonal complex 45 strain HPB5174, isolated from a powdered infant formula facility in Ireland. *Genome Announc.* 2(4):e00778-14. doi:10.1128/genomeA.00778-14.

**Copyright** © 2014 Pightling and Pagotto. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Franco Pagotto, [Franco.Pagotto@hc-sc.gc.ca](mailto:Franco.Pagotto@hc-sc.gc.ca).

*Cronobacter sakazakii* (synonym, *Enterobacter sakazakii*) is a food-borne pathogenic bacterium that causes outbreaks and sporadic cases of severe (albeit rare) meningitis, most often in neonates when contaminated powdered infant formula is ingested (1, 2). Multi-locus sequence typing of *C. sakazakii* isolates has revealed that some strains are more virulent than others and at least one group (ST4) is strongly associated with neonatal infections, while other sequence types have not been shown to be implicated in infections (3). Still, little is known about *C. sakazakii* virulence traits (4). Here, we present a high-quality draft genome sequence of a currently unrepresented group of *C. sakazakii* (ST40, CC45). Despite the presence of multiple ST40/CC45 isolates in our collection, obtained from either powdered infant formula or facilities making formula, none in our collection have been connected to any clinical cases and a search reveals that it has been indicted in only one clinical case occurring in China ([http://pubmlst.org/perl/bigsdb/bigsdb.pl?page=info&db=pubmlst\\_cronobacter\\_isolates&id=486](http://pubmlst.org/perl/bigsdb/bigsdb.pl?page=info&db=pubmlst_cronobacter_isolates&id=486)). Thus, this strain may be avirulent or only mildly virulent and may provide a useful outgroup for the identification of virulence factors.

*In silico* analysis of the multi-locus sequence typing (5) targets (*atpD-3*, *fusA-15*, *glnS-28*, *gltB-22*, *gyrB-5*, *infB-38* and *pps-19* (6)) indicates that strain HPB5174 belongs to the sequence-type 40 group and clonal complex 45 (3, 7). This finding is supported by analysis with pulsed-field gel electrophoresis, which yielded extremely rare patterns with both the XbaI and SpeI restriction endonucleases (BOM\_CSXA1.0108 and BOM\_CSSE1.0033). During a search of the *Cronobacter* PubMLST website ([http://pubmlst.org/perl/bigsdb/bigsdb.pl?db=pubmlst\\_cronobacter\\_isolates](http://pubmlst.org/perl/bigsdb/bigsdb.pl?db=pubmlst_cronobacter_isolates)) we found ten other isolates designated ST40/CC45 from India, United Kingdom, Australia, Netherlands, China, and Slovakia.

We generated short-read sequence data by preparing a paired-end library with the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA) and reading the sequences on a MiSeq Benchtop Sequencer (Illumina) for 600 cycles, yielding 1,835,160 reads with an average length of 250.93 bases. Error correction was

performed with BayesHammer (8) and reads were assembled *de novo* into a high-quality draft genome with SPAdes v3.0.0 (9), utilizing the MismatchCorrector tool. The assembly yielded 84 non-overlapping contiguous sequences (contigs) with a total length of 4,444,610 bases, 56.86% G+C content, and 103.61-fold sequencing coverage. The largest contig is 1,457,596 nucleotides in length and the weighted mean contig length ( $N_{50}$ ) is 526,913 nucleotides. Gene predictions and annotations were performed with the National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline (PGAP) (10). A total of 4,136 genes were identified, including 3,993 protein-coding regions, 39 pseudogenes, 2 clustered regularly interspaced short palindromic repeat (CRISPR) arrays, 23 ribosomal RNA, 76 transfer RNA, and 5 non-coding RNA sequences.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number JNBN00000000. The version described in this paper is the first version, JNBN01000000.

## ACKNOWLEDGMENTS

This work was funded by Genomics Research and Development Initiative (GRDI) grant 4500834, awarded to F.P. A.W.P. was supported by a Natural Sciences and Engineering Research Council (NSERC) postdoctoral fellowship award.

## REFERENCES

1. Nazarowec-White M, Farber JM. 1997. *Enterobacter sakazakii*: A review. *Int. J. Food Microbiol.* 34:103–113.
2. Abdesselam K, Pagotto F. 2014. Bacteria: *Cronobacter* (*Enterobacter*) *sakazakii* and other *Cronobacter* spp., p 424–432. *In* Encyclopedia of food Safety, 1st ed. Academic Press, Waltham, MA.
3. Joseph S, Sonbol H, Hariri S, Desai P, McClelland M, Forsythe SJ. 2012. Diversity of the *Cronobacter* genus as revealed by multilocus sequence typing. *J. Clin. Microbiol.* 50:3031–3039. <http://dx.doi.org/10.1128/JCM.00905-12>.
4. Holý O, Forsythe S. 2014. *Cronobacter* spp. as emerging causes of healthcare-associated infection. *J. Hosp. Infect.* 86:169–177. <http://dx.doi.org/10.1016/j.jhin.2013.09.011>.
5. Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang

- Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. U. S. A.* 95:3140–3145.
6. Baldwin A, Loughlin M, Caubilla-Barron J, Kucerova E, Manning G, Dowson C, Forsythe S. 2009. Multilocus sequence typing of *Cronobacter sakazakii* and *Cronobacter malonaticus* reveals stable clonal structures with clinical significance which do not correlate with biotypes. *BMC Microbiol.* 9:223. <http://dx.doi.org/10.1186/1471-2180-9-223>.
  7. Joseph S, Forsythe SJ. 2012. Insights into the emergent bacterial pathogen *Cronobacter* spp., Generated by multilocus sequence typing and analysis. *Front. Microbiol.* 3:397. <http://dx.doi.org/10.3389/fmicb.2012.00397>.
  8. Nikolenko SI, Korobeynikov AI, Alekseyev MA. 2013. BayesHammer: Bayesian clustering for error correction in single-cell sequencing. *BMC Genomics* 14(Suppl 1):S7. <http://dx.doi.org/10.1186/1471-2164-14-S1-S7>.
  9. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19:455–477. <http://dx.doi.org/10.1089/cmb.2012.0021>.
  10. Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of standard operating procedures (SOPs) for (meta)genomic annotation. *Omics* 12:137–141. <http://dx.doi.org/10.1089/omi.2008.0017>.