

# Genome Sequence of Hydrocarbon-Degrading *Cronobacter* sp. Strain DJ34 Isolated from Crude Oil-Containing Sludge from the Duliajan Oil Fields, Assam, India

Siddhartha Pal,<sup>a</sup> Tirtha Das Banerjee,<sup>a</sup> Ajoy Roy,<sup>a</sup> Pinaki Sar,<sup>b</sup> Sufia K. Kazy<sup>a</sup>

Department of Biotechnology, National Institute of Technology Durgapur, Durgapur, India<sup>a</sup>; Department of Biotechnology, Indian Institute of Technology Kharagpur, Kharagpur, India<sup>b</sup>

**We report here the 4,856,096-bp draft genome sequence of hydrocarbon-degrading *Cronobacter* sp. strain DJ34 isolated from crude oil-containing sludge from the Duliajan oil fields, India. DJ34 contains genes that mediate hydrocarbon degradation, metal resistance, and biosurfactant production. This is the first report of the genome sequence of *Cronobacter* sp. inhabiting an oil-contaminated environment.**

Received 24 September 2015 Accepted 28 September 2015 Published 12 November 2015

**Citation** Pal S, Das Banerjee T, Roy A, Sar P, Kazy SK. 2015. Genome sequence of hydrocarbon-degrading *Cronobacter* sp. strain DJ34 isolated from crude oil-containing sludge from the Duliajan oil fields, Assam, India. *Genome Announc* 3(6):e01321-15. doi:10.1128/genomeA.01321-15.

**Copyright** © 2015 Pal et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Sufia K.Kazy, [sufia\\_kazy@yahoo.com](mailto:sufia_kazy@yahoo.com).

Genome sequences of *Cronobacter* spp., commonly described as opportunistic food-borne pathogens causing fatal neonatal diseases, have already been reported (1, 2). To date, no information is available on the genome sequences of members of this genus that inhabit petroleum oil-rich environments. Very few strains of *Cronobacter* spp. have been reported from hydrocarbon- or industrial waste-contaminated habitats (3, 4). Recently, we isolated a Gram-negative, facultative anaerobic, hydrocarbon-degrading strain, *Cronobacter* sp. DJ34, from crude oil-containing sludge in the Duliajan oil fields, Assam, India. The 16S rRNA gene sequence of the isolate has been deposited in NCBI GenBank under the accession no. KM054665, which showed 99% sequence similarity with *Cronobacter pulveris* strain E444 (accession no. EF059835). Strain DJ34 showed multiple heavy metal resistances, growth under a wide range of pH, temperature, and salinity conditions, biosurfactant production, and an ability to utilize various electron acceptors during anaerobic growth. The genome sequence of this bacterium was determined to obtain better insights into the metabolic versatility of the strain. To our knowledge, this is the first report on the genome sequence of a *Cronobacter* sp. inhabiting an oil-contaminated environment.

The genomic DNA of DJ34 was extracted using the Promega Wizard genomic DNA purification kit (Promega, USA), according to the manufacturer's protocol. The draft genome sequence was obtained by using the Illumina HiSeq 2500 platform. The shotgun sequencing generated 5,092,150 high-quality paired-end reads with 100-fold coverage. Filtered reads were assembled by using *de novo* assemblers A5 (version 20140113) (5), ABySS (version 1.3.7) (6), MaSuRCA (version 2.3.2) (7), and SPAdes (version 3.1.1) (8). Integration of contigs was done with CISA (version 1.3). A total of 30 contigs with  $N_{50}$  length of 314,448 bp and average length of 161,869 bp were assembled. Genome annotation was performed by NCBI Prokaryotic Genomes Automatic Annotation Pipeline using Glimmer, tRNAscan-SE, and RNAmmer tools.

Functional annotation was carried out using the RAST server with the SEED database (9).

The draft genome is composed of 4,856,096 bp, with a G + C content of 56.5%. A total of 4,382 coding sequences (CDSs), 60 pseudogenes, and 14 noncoding RNA (ncRNA), 76 tRNA, and 23 rRNA genes were identified. The CDSs included many genes previously described for other *Cronobacter* spp. such as genes for secretion systems, superoxide dismutase, hemolysin, O antigen, enterobactin, biofilm formation, metalloprotease, adhesins, antibiotic resistance, heavy metal resistance, iron acquisition, universal stress proteins (*uspA* and *uspB*), starvation-sensing protein (*rspA*), flagella, fimbriae, phages, etc. (10). The genome of *Cronobacter* sp. DJ34 also contained different hydrocarbon degradation-related genes, including alkane-1-monoxygenase, alkane sulfonate monoxygenase, benzoylformate decarboxylase, 4-carboxymuconolactone decarboxylase, 3-carboxy-*cis,cis*-muconate cycloisomerase, 3-oxoadipate-enol-lactonase, 3-oxoadipate coenzyme A (CoA)-transferase, *bph*, protocatechuate 3,4 dioxygenase, salicylaldehyde dehydrogenase, etc. In addition to these, dissimilatory nitrate, nitrite and sulfite reductase, and biosurfactant synthesis genes have been noted. A detailed genome sequence analysis may provide deep insights into the genetic and metabolic potential of this organism for hydrocarbon degradation and survival in an oil-contaminated habitat.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project of *Cronobacter* sp. DJ34 has been deposited at DDBJ/EMBL/GenBank under the accession no. [LFEJ00000000](https://www.ncbi.nlm.nih.gov/nuclink/LFEJ00000000). The version described here is the first version, [LFEJ01000000](https://www.ncbi.nlm.nih.gov/nuclink/LFEJ01000000).

## ACKNOWLEDGMENTS

This work was supported by NER TWINNING Project (grant no. BT/226/NE/TBP/2011) from the Department of Biotechnology, Government of India.

We acknowledge the generous support of Oil India Limited, Duliajan, Assam, India, in sample collection.

## REFERENCES

- Stephan R, Lehner A, Tischler P, Rattei T. 2011. Complete genome sequence of *Cronobacter turicensis* LMG 23827, a foodborne pathogen causing deaths in neonates. *J Bacteriol* 193:309–310. <http://dx.doi.org/10.1128/JB.01162-10>.
- Masood N, Moore K, Farbos A, Hariri S, Paszkiewicz K, Dickins B, McNally A, Forsythe S. 2013. Draft genome sequences of three newly identified species in the genus *Cronobacter*, *C. helveticus* LMG23732<sup>T</sup>, *C. pulveris* LMG24059, and *C. zurichensis* LMG23730<sup>T</sup>. *Genome Announc* 1(5):e00783-13. <http://dx.doi.org/10.1128/genomeA.00783-13>.
- Jaradat ZW, Ababneh QO, Saadoun IM, Samara NA, Rashdan AM. 2009. Isolation of *Cronobacter* spp. (formerly *Enterobacter sakazakii*) from infant food, herbs and environmental samples and the subsequent identification and confirmation of the isolates using biochemical, chromogenic assays, PCR and 16S rRNA sequencing. *BMC Microbiol* 9:225. <http://dx.doi.org/10.1186/1471-2180-9-225>.
- Jain RM, Mody K, Mishra A, Jha B. 2012. Isolation and structural characterization of biosurfactant produced by an alkaliphilic bacterium *Cronobacter sakazakii* isolated from oil contaminated wastewater. *Carbohydr Polym* 87:2320–2326. <http://dx.doi.org/10.1016/j.carbpol.2011.10.065>.
- Tritt A, Eisen JA, Facciotti MT, Darling AE. 2012. An integrated pipeline for *de novo* assembly of microbial genomes. *PLoS One* 7:e42304. <http://dx.doi.org/10.1371/journal.pone.0042304>.
- Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, 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>.
- Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. *Bioinformatics* 29:2669–2677. <http://dx.doi.org/10.1093/bioinformatics/btt476>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshtkin 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>.
- Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma 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>.
- Joseph S, Desai P, Ji Y, Cummings CA, Shih R, Degoricija L, Rico A, Brzoska P, Hamby SE, Masood N, Hariri S, Sonbol H, Chuzhanova N, McClelland M, Furtado MR, Forsythe SJ. 2012. Comparative analysis of genome sequences covering the seven *Cronobacter* species. *PLoS One* 7:e49455. <http://dx.doi.org/10.1371/journal.pone.0049455>.