



Genome sequencing, annotation of *Citrobacter freundii* strain GTC 09479



Kazuyuki Kimura^{a,1}, Shailesh Kumar^{b,1}, Masahiro Takeo^{a,*}, Shanmugam Mayilraj^{c,**}

^a Department of Materials Science and Chemistry, Graduate School of Engineering, University of Hyogo 2167 Shosha, Himeji, Hyogo 671-2280, Japan

^b Bioinformatics Centre, CSIR-Institute of Microbial Technology, Chandigarh 160036, India

^c Microbial Type Culture Collection and Gene bank (MTCC), CSIR-Institute of Microbial Technology, Chandigarh 160036, India

ARTICLE INFO

Article history:

Received 30 October 2013

Accepted 9 December 2013

Available online 13 March 2014

Keywords:

Citrobacter freundii strain GTC 09479

Illumina-HiSeq

NGS QC toolkit

Rapid annotations using subsystems technology (RAST)

ABSTRACT

We report the 4.9-Mb genome sequence of *Citrobacter freundii* strain GTC 09479, isolated from urine sample collected during the year 1983 at Gifu University Graduate School of Medicine, Japan. This draft genome consist of 4,899,578 bp with 51.62% G + C, 4,574 predicted CDSs, 72 tRNAs and 10 rRNAs.

© 2014 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Specifications	
Organism/cell line/tissue	<i>Citrobacter freundii</i>
Strain(s)	GTC 09479
Sequencer or array type	Sequencer; the Illumina-HiSeq 1000
Data format	Processed
Experimental factors	Microbial strain
Experimental features	Draft genome sequence of <i>Citrobacter freundii</i> strain GTC 09479, assembly and annotation
Consent	n/a

The Genus *Citrobacter* was first proposed by Werkman and Gillen [1]. At present, the genus *Citrobacter* consists of ten recognized species. The organism in this study is *Citrobacter freundii* strain GTC 09479, a Gram-negative, aerobic, short rods bacterium isolated from urine sample. Genomic DNA was extracted from a 48 hour old culture using ZR Fungal/Bacterial DNA MiniPrep™ as per the manufacturer's

instructions. The genome of *C. freundii* strain GTC 09479 was sequenced using the Illumina-HiSeq 1000 paired-end technology produced a total of 17,547,712 paired-end reads (insert size of 350 bp) of 101 bp. We have used NGS QC toolkit v2.3 [2] to filter the data for high quality (Cut off read length for HQ = 70%, Cut off quality score = 20), vector/adaptor free reads for genome assembly. A total of 17,080,322 high quality, vector filtered reads (~345× coverage) were used for assembly with Velvet 1.2.08 (at hash length of 53) [3]. The final assembly contains 31 contigs of total size 4,899,578 bp with an N50 contig length of 370.8 kb; the largest contig assembled measured 773.9 kb. The draft genome (31 contigs) comprising 4,899,578 nt was annotated with the help of RAST (rapid annotation using subsystem technology) server [4]. A total of 4574 predicted coding regions (CDSs), 10 rRNAs and 72 tRNAs were predicted.

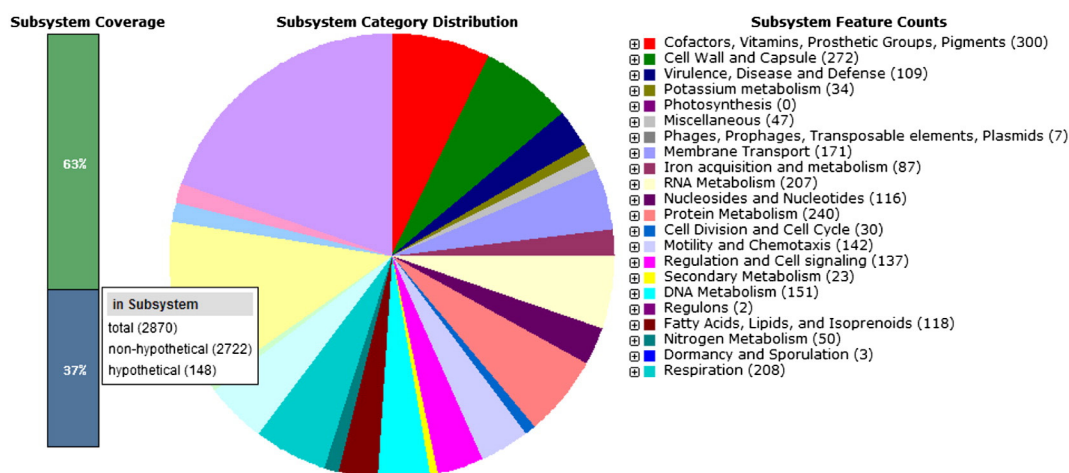
Whole genome annotation available at the RAST server shows that strain GTC 09479 contains genes of glycolysis and gluconeogenesis, TCA cycle, pentose phosphate pathway and lactose and galactose uptake and utilization. Some important genes like alkaline phosphatase (EC 3.1.3.1), galactosidase (alpha and beta both) [EC 3.2.1.22], alpha-xylosidase (EC 3.2.1.—), xylanase, xylose kinase (EC 2.7.1.17), xyloside transporter XynT, catalase (EC 1.11.1.6) and ferroxidase (EC 1.16.3.1). Genes involved in decomposition of urea i.e. urea ABC transporter (ATPase protein *UrtB*, *UrtC*, *UrtD* and *UrtE*), urease accessory protein (*UreD*, *UreE*, *UreF* and *UreG*) and urease (alpha, beta and gamma subunit) [EC 3.5.1.5] are also present in the genome annotation. Polysaccharide export lipoprotein *Wza* gene is also found in the annotation.

* Corresponding author. Tel./fax: +81 79 267 4893.

** Correspondence to: S. Mayilraj, Institute of Microbial Technology (IMTECH), Sector 39-A, Chandigarh, 160036, India. Tel.: +91 172 6665166; fax: +91 172 2695215.

E-mail addresses: takeo@eng.u-hyogo.ac.jp (M. Takeo), mayil@imtech.res.in (S. Mayilraj).

¹ Both are first authors, equally contributed.



Sub-system distribution of *C. freundii* strain GTC 09479 (based on RAST annotation server).

Nucleotide sequence accession number

The *C. freundii* strain GTC 09479 whole genome shot gun (WGS) project has been deposited at DDBJ/EMBL/GenBank under the project accession AOMS00000000 of the project (01) has the accession number AOMS01000000 and consists of sequences AOMS01000001-AOMS01000031.

Direct link: <http://www.ncbi.nlm.nih.gov/nuccore/AOMS01000000>.

Conflict of interest

The authors declare that there is no conflict of interest on any work published in this paper.

Acknowledgments

This work was partly supported by The Regional Innovation Strategy Support Programme (2012-Ministry of Education, Cultures, Sports,

Science and Technology, Japan and the JSPS invitation fellowship (Fellow's ID No: L-12546) for research in Japan granted to Dr. Shanmugam Mayilraj). KK is supported by Hyogo Analysis Center Co., Ltd., Himeji, Japan and SK is supported by a research fellowship from Council of Scientific and Industrial Research (CSIR-INDIA). We thank C-CAMP (<http://www.ccamp.res.in/>) next generation genomics facility for help in obtaining the genome sequence. Genome assembly and annotation data can be downloaded from <http://crdd.osdd.net/raghava/genomesrs/>. This is IMTECH communication number 55/2013.

References

- [1] C.H. Werkman, G.F. Gillen, Bacteria producing trimethylene glycol. *J. Bacteriol.* 23 (1932) 167.
- [2] R.K. Patel, M. Jain, NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7 (2012) e30619.
- [3] D.R. Zerbino, E. Birney, Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18 (2008) 821–829.
- [4] R.K. Aziz, D. Bartels, A.A. Best, M. DeJongh, T. Disz, R.A. Edwards, K. Formsma, S. Gerdes, E.M. Glass, M. Kubal, F. Meyer, G.J. Olsen, R. Olson, A.L. Osterman, R.A. Overbeek, L.K. McNeil, D. Paarmann, T. Paczian, B. Parrello, G.D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein, A. Wilke, O. Zagnitko, The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9 (2008) 75.