Genomics Data 9 (2016) 128–129

Contents lists available at ScienceDirect

Genomics Data

journal homepage: www.elsevier.com/locate/gdata

Data in Brief

The draft genome sequence of *Mangrovibacter* sp. strain MP23, an endophyte isolated from the roots of *Phragmites karka*

Pratiksha Behera [a], Parag Vaishampayan [b], Nitin K. Singh [b], Samir R. Mishra [c], Vishakha Raina [c], Mrutyunjay Suar [c], Ajit K. Pattnaik [a], Gurdeep Rastogi [a,*]

[a] Wetland Research and Training Centre, Chilika Development Authority, Barkul, Balugaon, 752030, Odisha, India
[b] Biotechnology and Planetary Protection Group, Jet Propulsion Laboratory, NASA, California Institute of Technology, Pasadena, CA 91109,USA
[c] School of Biotechnology, KIIT University, Patia, Bhubaneswar, 751024, Odisha, India

ARTICLE INFO

Article history:
Received 29 June 2016
Received in revised form 7 July 2016
Accepted 18 July 2016
Available online 19 July 2016

Keywords:
Mangrovibacter sp.
Phragmites karka
Draft genome
Chilika
Endophyte

ABSTRACT

Till date, only one draft genome has been reported within the genus *Mangrovibacter*. Here, we report the second draft genome shotgun sequence of a *Mangrovibacter* sp. strain MP23 that was isolated from the roots of *Phargmites karka* (*P. karka*), an invasive weed growing in the Chilika Lagoon, Odisha, India. Strain MP23 is a facultative anaerobic, nitrogen-fixing endophytic bacteria that grows optimally at 37 °C, 7.0 pH, and 1% NaCl concentration. The draft genome sequence of strain MP23 contains 4,947,475 bp with an estimated G + C content of 49.9% and total 4392 protein coding genes. The genome sequence has provided information on putative genes that code for proteins involved in oxidative stress, uptake of nutrients, and nitrogen fixation that might offer niche specific ecological fitness and explain the invasive success of *P. karka* in Chilika Lagoon. The draft genome sequence and annotation have been deposited at DDBJ/EMBL/GenBank under the accession number LYRP00000000.

Specifications	
Organism	*Mangrovibacter* strain MP23
Sequencer	Illumina MiSeq sequencing system
Data format	Processed
Experimental factors	Root endophytic bacterium
Experimental features	De nova genome assembly
Consent	Data is publicly available
	Bio Project: PRJNA323358
	Bio Sample: SAMN05177220
Sample source location	Root of *P. karka*, Chilika Lagoon, Odisha, India

1. Direct link to deposited data

http://www.ncbi.nlm.nih.gov/nuccore/LYRP00000000

2. Experimental design, materials, and methods

Members of the genus *Mangrovibacter* are facultative anaerobic and nitrogen fixing bacteria [1]. To date only two species, *Mangrovibacter plantisponsor* DSM 19579 [1] and *Mangrovibacter yixingensis* KCTC 42181 [2] have been described under the genus *Mangrovibacter* (http://www.bacterio.net/mangrovibacter.html). The draft genome of *Mangrovibacter* sp. strain MFB070, isolated from an aquaculture farm in India has been described recently [3] which is the only available genome sequence within genus *Mangrovibacter*. Members of the genus *Mangrovibacter* have been shown to possess plant growth promoting features such as nitrogen fixation [1] and may provide niche-specific advantage to the plant. In context to invasive weeds, this could provide a better ecological fitness in an invaded territory compared to native vegetation. Thus, understanding the microbiota and role they may play during plant invasion could lead to more directed and sustainable management of weeds.

Chilika Lagoon (19°28′-19°54′N; 85°06′-85°35′E) is a brackish water lagoon located in the Odisha State of India [4]. An invasive weed *P. karka* (Retz.) Trin. ex Steud is extensively spreading and threatening the ecological health of lagoon. *P. karka* is a large perennial grass of the family *Poaceae* and occupy most of the northern shoreline of lagoon. In order to understand the microbial basis of the invasive success of this weed, a study was undertaken to investigate the culturable diversity of rhizosphere microbiota associated with *P. karka*. During this study, a gammaproteobacterium, facultative anaerobic, endophytic nitrogen fixing *Mangrovibacter* sp. strain MP23 was isolated from the roots of *P. karka*. Strain MP23 grew at temperatures between 20 °C and 40 °C with an

* Corresponding author.
E-mail address: rastogigurdeep@gmail.com (G. Rastogi).

http://dx.doi.org/10.1016/j.gdata.2016.07.007

Table 1
General features of *Mangrovibacter* sp. strain MP23 draft genome.

Attributes	Values
Assembly size (bp)	4,947,475
Contigs	50
G + C content (%)	49.9
Total genes	4592
tRNA genes	77
rRNA genes	28
Protein coding genes	4392
Fold coverage	310×

optimum at 37 °C in presence of 1% NaCl. The 16S rRNA gene phylogenetic analysis showed that strain MP23 was most closely (99.71% similarity) related to *M. yixingensis* KCTC 42181 and *M. plantisponsor* DSM 19579 indicating that it belong to genus *Mangrovibacter*. Here, we described the draft whole genome shotgun sequence of strain *Mangrovibacter* sp. strain MP23 (DSM 100250^T = KCTC 42580^T), which will provide genetic insights into the nitrogen fixation, stress tolerance, plant niche adaptation, and comparative evolution of this species.

The genome of *Mangrovibacter* sp. strain MP23 was sequenced by a shotgun sequencing method using the Illumina MiSeq sequencing system with a paired-end module. The NGS QC Toolkit v 2.3 [5] was used to filter the data for high-quality (HQ) vector- and adapter-free reads for genome assembly. A total of 6,895,374 paired end reads were generated out of which 6,702,332 high quality vector-filtered reads were considered, representing approximately 310 fold coverage of the genome. These reads were assembled using MaSuRCA v. 3.1.3 [6] and resulted into 50 contigs with a total size of 4,947,475 bp and an N_{50}contig length of 428,946 bp. The largest assembled contig measured 1,277,690 bp. Annotations of protein-coding genes, as well as other functional genome units were carried out through NCBI Prokaryotic Genome Annotation Pipeline (PGAP) [7]. The genome contained a total of 4592 genes and 4392 protein-coding genes with predicted function.

The complete genome of strain MP23 was 4,947,475 bp in length with an estimated G + C content of 49.9% and consists of 77 tRNA genes and 28 rRNA (8 = 5S, 5 = 16S, 15 = 23S) genes (Table 1).

3. Nucleotide sequence accession number

This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number LYRP00000000. The version described in this paper is version LYRP01000000.

Conflict of interest

The authors declare that there is no conflict of interests with respect to the work published in this paper.

Acknowledgements

The authors would like to acknowledge the funding support (CREDIT NO.-4765-IN) received from the World Bank through the Integrated Coastal Zone Management Project (ICZMP) of Odisha.

References

[1] N. Rameshkumar, E. Lang, S. Nair, *Mangrovibacter plantisponsor* gen. nov., sp. nov., a nitrogen-fixing bacterium isolated from a mangrove-associated wild rice (*Porteresia coarctata* Tateoka). Int. J. Syst. Evol. Microbiol. 60 (2010) 179–186.
[2] H. Zhang, S. Guo, B. Sun, J. Zhang, M. Cheng, Q. Li, Q. Hong, X. Huang, *Mangrovibacter yixingensis* sp. nov., isolated from farmland soil. Int. J. Syst. Evol. Microbiol. 65 (2015) 2447–2452.
[3] T.C. Joseph, A.M. Varghese, A. Baby, D. Reghunathan, V. Murugadas, K.V. Lalitha, First draft genome sequence of a member of the genus mangrovibacter, isolated from an aquaculture farm in India. Genome Announc. 6 (2014) (pii: e01209-14).
[4] S. Srichandan, J.Y. Kim, A. Kumar, D.R. Mishra, P. Bhadury, P.R. Muduli, A.K. Pattnaik, G. Rastogi, Interannual and cyclone-driven variability in phytoplankton communities of a tropical coastal lagoon. Mar. Pollut. Bull. 101 (2015) 39–52.
[5] R.K. Patel, M. Jain, NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One 2 (2012) e30619.
[6] A.V. Zimin, G. Marcais, D. Puiu, M. Roberts, S.L. Salzberg, J.A. Yorke, The MaSuRCA genome assembler. Bioinformatics 29 (2013) 2669–2677.
[7] S.V. Angiuoli, A. Gussman, W. Klimke, G. Cochrane, D. Field, G.M. Garrity, et al., Toward an online repository of standard operating procedures (SOPs) for (meta) genomic annotation. OMICS: J. Integr. Biol. 12 (2008) 137–141.