Data Article

Genome sequence data of *Mangrovimonas* sp. strain CR14 isolated from mangrove forest at Tanjung Piai National Park, Malaysia

Muhammad Ramziuddin Zakaria[a], Ming Quan Lam[a],
Sye Jinn Chen[a], Mohamad Hamizan Abdul Karim[a], Lili Tokiman[b],
Adibah Yahya[a], Mohd Shahir Shamsir[c], Chun Shiong Chong[a,*]

[a] Department of Biosciences, Faculty of Science, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia
[b] Johor National Parks Corporation, Kota Iskandar, 79575 Iskandar Puteri, Johor, Malaysia
[c] Faculty of Applied Sciences and Technology, Universiti Tun Hussein Onn Malaysia, Pagoh Higher Education Hub, 84600 Muar, Johor, Malaysia

ARTICLE INFO

ABSTRACT

Mangrovimonas sp. strain CR14 is a halophilic bacterium affiliated with family *Flavobacteriaceae* which was successfully isolated from mangrove soil samples obtained from Tanjung Piai National Park, Johor. The whole genome of strain CR14 was sequenced on an Illumina HiSeq 2500 platform (2×150 bp paired end). Herein, we report the genome sequence of *Mangrovimonas* sp. strain CR14 in which its assembled genome consisted 20 contigs with a total size of 3,590,195 bp, 3209 coding sequences, and an average 36.08% G + C content. Genome annotation and gene mining revealed that this bacterium demonstrated proteolytic activity which could be potentially applied in detergent industry. This whole-genome shotgun data of *Mangrovimonas* sp. strain CR14 has been deposited at DDBJ/ENA/GenBank under the accession JAAFZY000000000. The version described in this paper is version JAAFZY010000000.

* Corresponding author.
E-mail address: cschong@utm.my (C.S. Chong).

Specifications table

Subject	Biology
Specific subject area	Microbiology and genomics
Type of data	• Genome sequence data in FASTA format
	• Table
	• Figure
How data were acquired	Whole-genome sequencing using Illumina HiSeq 2500 (2×150 bp paired end) platform
Data format	Raw and assembled genome sequences
Parameters for data collection	Genomic DNA was extracted from a pure culture of *Mangrovimonas* sp. strain CR14. The genome of strain CR14 was sequenced by using Illumina HiSeq 2500 platform (2×150 bp paired end). The genome was *de novo* assembled using SPAdes version 3.11.1 and annotated using PGAP.
Description of data collection	Whole-genome sequencing, assembly and annotation
Data source location	*Mangrovimonas* sp. strain CR14 was isolated from Tanjung Piai National Park, Johor, Malaysia 1°16′01.7″N 103°30′40.2″E
Data accessibility	This whole-genome shotgun data of *Mangrovimonas* sp. strain CR14 has been deposited at DDBJ/ENA/GenBank under the accession JAAFZY000000000 (https://www.ncbi.nlm.nih.gov/nuccore/JAAFZY000000000). The version described in this paper is version JAAFZY010000000. The sequence data have been registered in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database under the accession number SRR11110036 (https://www.ncbi.nlm.nih.gov/sra/SRR11110036).

Value of the data

- The genome sequence of *Mangrovimonas* sp. strain CR14 provides fundamental knowledge about genes related to proteolytic activity.
- The genomic information of this strain CR14 will be useful for comparative genomic analysis with other *Mangrovimonas* species.
- The proteolytic genes encoded in the genome could be further characterized and potentially benefit to detergent industry for effective proteinaceous stain removal.

1. Data description

Mangrovimonas is a genus that belongs to the family *Flavobacteriaceae* [1] of order *Flavobacteriales*. To date, only three species were successfully isolated from the marine environment, namely, *Mangrovimonas spongiae* [2], *Mangrovimonas xylaniphaga* [3] and *Mangrovimonas yunxiaonensis* [1]. The genome sequence and algicidal ability of *M. yunxiaonensis* were reported [4]. Besides that, the genome sequences of *M. xylaniphaga* and another two *Mangrovimonas*-like strains (ST2-L15 and TPB-H4) were also studied with elucidation of their xylan and arabinan utilizing abilities [5]. So far, the proteolytic genes of *Mangrovimonas* which potentially to be useful for food processing and detergent application were not revealed.

Strain CR14, a bacterium that affiliated to genus *Mangrovimonas* (99% 16S rRNA gene similarity to *Mangrovimonas* sp. strain ST2-L15), was isolated from mangrove soil samples obtained from Tanjung Piai National Park, Johor. Colony of strain CR14 was orange-pigmented, with 1 - 2 mm in diameter, round shape, smooth surface, entire margin, convex elevation and small size after 48 h of incubation on Marine agar plate. Its genome was sequenced and the proteolytic genes were mined. The genome features of strain CR14 were summarized in Table 1. The assembled genome of *Mangrovimonas* sp. strain CR14 consisted 20 contigs with a total size of 3590195 bp, while the N_{75} value and G+C content was 367,190 bp and 36.08% respectively. Based on the genome annotation, a total of 3209 genes were predicted in which, 3152 of them were responsible for coding specific proteins while 46 and 11 of them were coded for RNA genes (39 tRNAs, 4 ncRNAs and 1 for each 16S-23S-5S rRNA operon) and pseudo genes, respectively.

Table 1
General genome statistics of *Mangrovimonas* sp. strain CR14.

Category	Strain CR14	
	Number	Total percentage (%)
Number of contigs	20	–
Genome size (bp)	3590,195	100.00
G+C content	1295,342	36.08%
Total genes predicted	3209	100.00
Protein coding genes	3152	98.22
Non-coding RNA genes	46	1.43
rRNA genes		
5S rRNA	1	0.03
16S rRNA	1	0.03
23S rRNA	1	0.03
tRNA	39	1.22
ncRNA	4	0.12
Pseudogenes	11	0.34

Fig. 1. Mangrovimonas sp. strain CR14 positive hydrolysis on skim milk containing agar showing ability of this bacterium to produce extracellular proteolytic enzymes.

Furthermore, a total of 100 genes in the genome of strain CR14 were predicted to be involved in proteolytic activity. In details, 1, 42, 6, 15 and 36 gene(s) were encoded for aspartic, metallo-, zinc metallo-, serine and ATP-dependent proteases respectively. After being subjected to SignalP version 5.0 server [6], 10 proteolytic genes were predicted to be secreted extracellularly. These proteases were matrixin family metalloprotease (locus tag: GZ212_06430), ATP-dependent zinc metalloprotease (locus tag: GZ212_13440), M1 family metalloproteases (locus tag: GZ212_02095, GZ212_11920, GZ212_15430 and GZ212_13780), S8 family serine proteases (locus tag: GZ212_02100 and GZ212_10035), M23 family metalloprotease (locus tag: GZ212_14610) and P1 family proteases (locus tag: GZ212_15520). In addition, *Mangrovimonas* sp. strain CR14 was also tested positive on skim milk containing agar (clear zone was shown) further proved the ability of this bacterium to produce extracellular proteolytic enzymes (Fig. 1).

2. Experimental design, materials, and methods

The inoculum of *Mangrovimonas* sp. strain CR14 was prepared. A 0.5% (v/v) of strain CR14 glycerol stock was aseptically cultured in Marine broth 2216 at 30 °C for 18 h. Inoculum with optical density (OD_{600}) of 0.7 was then streaked onto Marine agar 2216 and cultured at 30 °C for 48 h. Genomic DNA of strain CR14 was extracted and purified by using the Quick-DNA Miniprep Plus kit (Zymo Research) and DNA Clean and Concentrator™-25 (Zymo Research) respectively according to manufacturer's instructions. The quantity and quality of the purified genomic DNA of strain CR14 were then checked using both Nanodrop™ spectrophotometer and Qubit® fluorometer respectively. The library was constructed by using Nextera sample preparation kit [7]. The whole genome of strain CR14 was sequenced on an Illumina HiSeq 2500 platform (2×150 bp paired end) with a genome coverage of $155 \times$. Sequence adaptors and reads with low quality scores were filtered using BBDuk of the BBTools Packages [8]. The filtered reads were *de novo* assembled using SPAdes version 3.11.1 [9]. The genome of strain CR14 was then annotated using National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline (PGAP) version 4.11 [10].

Acknowledgments

Conflict of Interest

The authors declare that there is no conflict of interest.

Author contribution statement

Muhammad Ramziuddin Zakaria carried out the experiment, analysed the data and drafted the manuscript. Ming Quan Lam, Sye Jinn Chen and Mohamad Hamizan Abdul Karim designed

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2020.105658.

References

[1] Y. Li, S. Bai, C. Yang, Q. Lai, H. Zhang, Z. Chen, J. Wei, W. Zheng, Y. Tian, T. Zheng, *Mangrovimonas yunxiaonensis* gen. nov., sp. nov., isolated from mangrove sediment, Int. J. Syst. Evol. Microbiol 63 (2013) 2043–2048, doi:10.1099/ijs.0.046193-0.

[2] L. Zhuang, B. Lin, L. Luo, *Mangrovimonas spongiae* sp. nov., a novel member of the genus *Mangrovimonas* isolated from marine sponge, Int. J. Syst. Evol. Microbiol. (2020), doi:10.1099/ijsem.0.004006.

[3] B. Dinesh, G. Furusawa, A.A. Amirul, *Mangrovimonas xylaniphaga* sp. nov. isolated from estuarine mangrove sediment of Matang Mangrove Forest, Malaysia, Arch. Microbiol. 199 (2017) 63–67, doi:10.1007/s00203-016-1275-8.

[4] Y. Li, H. Zhu, C. Li, H. Zhang, Z. Chen, W. Zheng, H. Xu, T. Zheng, Draft genome sequence of the algicidal bacterium *Mangrovimonas yunxiaonensis* strain LY01, Genome Announc 2 (2014) 1–2, doi:10.1128/genomeA.01234-14.

[5] B. Dinesh, N.S. Lau, G. Furusawa, S.W. Kim, T.D. Taylor, S.Y. Foong, A.C. Shu-Chien, Comparative genome analyses of novel *Mangrovimonas*-like strains isolated from estuarine mangrove sediments reveal xylan and arabinan utilization genes, Mar. Genomics. 25 (2016) 115–121, doi:10.1016/j.margen.2015.12.006.

[6] J.J. Almagro Armenteros, K.D. Tsirigos, C.K. Sønderby, T.N. Petersen, O. Winther, S. Brunak, G. von Heijne, H. Nielsen, SignalP 5.0 improves signal peptide predictions using deep neural networks, Nat. Biotechnol. 37 (2019) 420–423, doi:10.1038/s41587-019-0036-z.

[7] J.D. Ring, K. Sturk-Andreaggi, M.A. Peck, C. Marshall, A performance evaluation of Nextera XT and KAPA HyperPlus for rapid Illumina library preparation of long-range mitogenome amplicons, Forensic Sci. Int. Genet. 29 (2017) 174–180, doi:10.1016/j.fsigen.2017.04.003.

[8] A. Mannion, N. Fabian, M. Stair, J. Dzink-Fox, S.E. Carrasco, E. Buckley-Jordan, D. Annamalai, J.G. Fox, Draft genome sequences of *Klebsiella pneumoniae* strains isolated from immunocompromised NOD-scid gamma research mice, Microbiol. Resour. Announc (2019) 8, doi:10.1128/mra.00869-19.

[9] A. Bankevich, S. Nurk, D. Antipov, A.A. Gurevich, M. Dvorkin, A.S. Kulikov, V.M. Lesin, S.I. Nikolenko, S. Pham, A.D. Prjibelski, A.V. Pyshkin, A.V. Sirotkin, N. Vyahhi, G. Tesler, M.A. Alekseyev, P.A. Pevzner, SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing, J. Comput. Biol. 19 (2012) 455–477, doi:10.1089/cmb.2012.0021.

[10] T. Tatusova, M. Dicuccio, A. Badretdin, V. Chetvernin, E.P. Nawrocki, L. Zaslavsky, A. Lomsadze, K.D. Pruitt, M. Borodovsky, J. Ostell, NCBI prokaryotic genome annotation pipeline, Nucleic Acids Res. 44 (2016) 6614–6624, doi:10.1093/nar/gkw569.