Complete Genome Sequence of Carbendazim-Degrading *Mycobacterium* sp. Strain djl-10

Ji Zhang, Qiaoyun Yuan, Wei Yang, Xinfeng Wang

Jiangsu Collaborative Innovation Center of Regional Modern Agriculture and Environmental Protection, Jiangsu Key Laboratory for Eco-Agricultural Biotechnology around Hongze Lake, School of Life Science, Huaiyin Normal University, Hua’ian, People’s Republic of China

**ABSTRACT** *Mycobacterium* sp. strain djl-10, an efficient degrader of carbendazim, was isolated from a carbendazim manufacturing wastewater treatment system. Here, we report the complete genome sequence of djl-10, which consists of a chromosome and three plasmids.

Carbendazim is a fungicide widely used around the world to control a broad range of fungal diseases in agricultural crops (1, 2). *Mycobacterium* sp. strain djl-10, a carbendazim-degrading strain, was isolated from a carbendazim manufacturing wastewater treatment system. Our clarification of the complete genome sequence of this strain is expected to provide insights into the mechanisms of carbendazim degradation by the strain and provide help for its application in corresponding carbendazim-contaminated environment remediation in the future.

The complete genome of *Mycobacterium* sp. djl-10 was generated using the Illumina HiSeq and PacBio RS single-molecule sequencing platforms. For this purpose, two libraries (500-bp paired-end library for Illumina HiSeq and 8- to ~10-kb SMRT Bell library for PacBio RS) were constructed. A total of 3,308,392,500 bp of Illumina raw data, 3,132,093,644 bp of Illumina high-quality data, and 134,281,312 bp of PacBio raw data were generated. The sequencing data of Illumina high-quality reads were assembled by SOAPdenovo (version 2.04) and used to align with PacBio sequencing data by blasR to reduce the single-base and insert-missing errors of long single-molecule sequence. The corrected single-molecule sequencing data of PacBio were connected using the overlap relationship between sequence scaffolds, and then Celera assembler 8.0 was used for subsequent assembly. After completing all the scaffold connection, the Illumina data were used again to verify and close gaps of the assembled scaffolds by GapCloser version 1.12 (SOAPdenovo-related software) (3).

The genome of djl-10 consists of a circular chromosome of 6,395,946 bp, with 67.93% G+C content, and three plasmids (plasmid1, 134,689 bp; plasmid2, 21,705 bp; plasmid3, 26,982 bp). A total of 46 tRNA genes encompassing all 20 amino acids and two copies of rRNA operons (5S-23S-16S) were identified by using tRNAscan-SE version 1.3.1 (4) and Barrnap 0.4.2 (RNAmmer-1.2) (5), respectively. Genes were predicted by Glimmer 3.02 ([http://ccb.jhu.edu/software/glimmer/index.shtml](http://ccb.jhu.edu/software/glimmer/index.shtml)). There were 6,392 genes in total, with a length of 6,120,036 bp with 68.2% G+C content, which accounts for 92.4% of the entire genome. To annotate the genes, all the corresponding protein sequences were aligned in the databases of Nr, genes, string, and GO, by BLASTp (BLAST2.2.28+). Among all the predicted genes, 6,020 genes hit top target sequences and achieved annotation information.

A putative methyl-1H-benzimidazol-2-ylcarbamate (MBC)-hydrolyzing esterase (*mhe*) gene involved in carbendazim utilization located on plasmid1 was found with the CDS of BAC37_RS31035, and it showed 100% and 99% sequence identity with *mheI*.
genes cloned from *Rhodococcus erythropolis* djl-11 (6) and *Nocardioides* sp. SG-4G (7), respectively. The genes responsible for the subsequent degradation of carbendazim in djl-10 are still unknown and will be clarified in our future work. The availability of the *Mycobacterium* sp. djl-10 genome sequence will act as an invaluable supplement to the ongoing research efforts toward understanding several unanswered questions associated with the degradation of carbendazim and would thus aid in the development of *in situ* bioremediation in the future.

**Accession number(s).** The genome and three plasmid sequences have been deposited in GenBank under accession numbers CP016640, CP016641, CP016642, and CP016643, respectively.

**ACKNOWLEDGMENT**

This work was supported by grants from the Chinese National Natural Science Foundation (31300099).

**REFERENCES**

1. Chen Y, Zhou MG. 2009. Characterization of *Fusarium graminearum* isolates resistant to both carbendazim and a new fungicide JS399-19. *Phytopathology* 99:441–446. https://doi.org/10.1094/PHYTO-99-4.0441.

2. Papadopoulou-Mourkidou E. 1991. Postharvest-applied agrochemicals and their residues in fresh fruits and vegetables. *J Assoc Off Anal Chem* 74:745–765.

3. Koren S, Schatz MC, Walenz BP, Martin J, Howard JT, Ganapathy G, Wang Z, Rasko DA, McCombie WR, Jarvis ED, Adam M, Phillippy AM. 2012. Hybrid error correction and *de novo* assembly of single-molecule sequencing reads. *Nat Biotechnol* 30:693–700. https://doi.org/10.1038/nbt.2280.

4. Lowe TM, Eddy SR. 1997. TRNAScan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25:955–964. https://doi.org/10.1093/nar/25.5.0955.

5. Lagesen K, Hallin P, Redlund EA, Stærfeldt H-H, Rognes T, Ussery DW. 2007. RNAmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 35:3100–3108. https://doi.org/10.1093/nar/gkm160.

6. Zhang X, Huang Y, Harvey PR, Li H, Ren Y, Li J, Wang J, Yang H. 2013. Isolation and characterization of carbendazim degrading *Rhodococcus erythropolis* djl-11. *PLoS One* 8:e74810. https://doi.org/10.1371/journal.pone.0074810.

7. Pandey G, Dorrian SJ, Russell RJ, Brearley C, Kotsonis S, Oakeshott JG. 2010. Cloning and biochemical characterization of a novel carbendazim (methyl-1-H-nenzimidazol-2-ylcarbamate)-hydrolyzing esterase from the newly isolated *Nocardioides* sp. strain SG-4G and its potential for use in enzymatic bioremediation. *Appl Environ Microbiol* 76:2940–2945. https://doi.org/10.1128/AEM.02990-09.