genomeAnnouncements™

Check for updates

Draft Genome Sequence of *Marinomonas fungiae* Strain AN44ᵀ (JCM 18476ᵀ), Isolated from the Coral *Fungia echinata* from the Andaman Sea

Prabla Kumari,ᵃ* Jhasketan Badhai,ᵃ Subrata K. Dasᵃ

ᵃDepartment of Biotechnology, Institute of Life Sciences, Bhubaneswar, Odisha, India

ABSTRACT *Marinomonas fungiae* strain AN44ᵀ was isolated from mucus of the coral *Fungia echinata*. Optimum growth occurs at 3 to 5% NaCl. The draft genome is 4.2 Mb, with 3,776 protein-coding genes. It harbors genes for the degradation of aromatic compounds, such as quinate, ferulate, *p*-coumarate, protocatechuate, and *p*-hydroxyphenylacetate.

Marinomonas fungiae strain AN44ᵀ (JCM 18476ᵀ) is an aerobic, Gram-negative, motile, and rod-shaped bacterium belonging to the class *Gammaproteobacteria* and was isolated from mucus of the coral *Fungia echinata* from the Andaman Sea (1). The draft genome sequence will help in understanding the plasticity and diversity of metabolic pathways for the degradation of aromatic compounds/pollutants in marine bacteria. Further, the strain harbors two unique mobile genetic elements belonging to the SXT integrating conjugative element (ICE) family (2).

The draft genome sequence of *M. fungiae* strain AN44ᵀ was generated at the DOE Joint Genome Institute (JGI) using the Illumina HiSeq 2000 platform (3). An Illumina standard shotgun library was constructed and sequenced using the Illumina HiSeq 2000 platform, which generated 10,961,766 reads totaling 1,655.2 Mb. The filtered Illumina reads were assembled using the Velvet (4), wgsim (https://github.com/lh3/wgsim), and AllPaths-LG (5) tools. The final draft assembly contains 42 contigs in 36 scaffolds, totaling 4.2 Mb, with an input read coverage of 268.9-fold. The largest and N_{50} contigs are 554.8 kb and 268.5 kb, respectively, with a G+C content of 46.2%.

The genome was annotated using the JGI Microbial Genome Annotation Pipeline (6). Genes were identified using the Prodigal (7) and GenePRIMP (8) programs. The predicted coding sequences (CDSs) were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant, UniProt, TIGRFAM, Pfam, KEGG, COG, and InterPro databases. The tRNAs, rRNAs, and other noncoding RNA genes were identified by searching the genome using the tRNAscan-SE tool (9), rRNA gene models built from SILVA (10), and Infernal (http://infernal.janelia.org), respectively. The draft genome contains a total of 3,776 CDSs, 78 pseudogenes, 53 tRNAs, 16 rRNAs (5 5S rRNAs, 9 16S rRNAs, and 2 23S rRNAs), 5 noncoding RNAs (ncRNAs), and 1 clustered regularly interspaced short palindromic repeat (CRISPR). Based on COG functional categories, the CDSs of the *M. fungiae* genome were distributed into categories of amino acid transport and metabolism (9.84%), carbohydrate transport and metabolism (5.92%), cell cycle control, cell division, and chromosome partitioning (1.08%), cell motility (3.51%), cell wall/membrane/envelope biogenesis (5.06%), chromatin structure and dynamics (0.09%), coenzyme transport and metabolism (5.47%), defense mechanisms (2.18%), energy production and conversion (6.61%), extracellular structures (0.44%), function unknown (5.41%), general function prediction only (7.31%), inorganic ion transport and metabolism (5.19%), intracellular trafficking, secretion, and vesicular transport (1.55%), lipid transport and metabolism (3.35%), mobilome prophages and

Address correspondence to Subrata K. Das, subratkdas@hotmail.com.

* Present address: Prabla Kumari, Division of Biochemical Sciences, CSIR-National Chemical Laboratory, Pashan, Pune, Maharashtra, India.

transposons (1.08%), nucleotide transport and metabolism (2.47%), posttranslational modification, protein turnover, and chaperones (4.15%), RNA processing and modification (0.03%), replication, recombination, and repair (3.83%), secondary metabolite biosynthesis, transport, and catabolism (2.12%), signal transduction mechanisms (8.13%), transcription (7.97%), translation, ribosomal structure, and biogenesis (7.18%), and unassigned (28.88%).

Further, putative genes or gene clusters for peripheral aromatic compound degradation pathways of quinate, ferulate, and *p*-coumarate, along with a single operonic gene cluster for central aromatic compound degradation via the protocatechuate *meta*-cleavage pathway, were found in the genome. Additionally, a single operonic gene cluster for the degradation of *p*-hydroxyphenylacetate was found.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number LIQF00000000. The version described in this paper is the first version, LIQF01000000, and consists of sequences LIQF01000001 to LIQF01000036.

REFERENCES

1. Kumari P, Poddar A, Das SK. 2014. *Marinomonas fungiae* sp. nov., isolated from a coral *Fungia echinata* from Andaman sea. Int J Syst Evol Microbiol 64:487–494. https://doi.org/10.1099/ijs.0.054809-0.

2. Badhai J, Das SK. 2016. Characterization of three novel SXT/R391 integrating conjugative elements ICE*Mfu*Ind1a and ICE*Mfu*Ind1b, and ICE*M-pr*Chn1 identified in the genomes of *Marinomonas fungiae* JCM 18476ᵀ and *Marinomonas profundimaris* strain D104. Front Microbiol 7:1896. https://doi.org/10.3389/fmicb.2016.01896.

3. Bennett S. 2004. Solexa Ltd. Pharmacogenomics 5:433–438. https://doi.org/10.1517/14622416.5.4.433.

4. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.

5. Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci U S A 108:1513–1518. https://doi.org/10.1073/pnas.1017351108.

6. Mavromatis K, Ivanova NN, Chen IM, Szeto E, Markowitz VM, Kyrpides NC. 2009. The DOE-JGI standard operating procedure for the annotations of microbial genomes. Stand Genomic Sci 1:63–67. https://doi.org/10.4056/sigs.632.

7. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119. https://doi.org/10.1186/1471-2105-11-119.

8. Pati A, Ivanova NN, Mikhailova N, Ovchinnikova G, Hooper SD, Lykidis A, Kyrpides NC. 2010. GenePRIMP: a gene prediction improvement pipeline for prokaryotic genomes. Nat Methods 7:455–457. https://doi.org/10.1038/nmeth.1457.

9. 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.

10. Pruesse E, Quast C, Knittel K, Fuchs B, Ludwig W, Peplies J, Glckner FO. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35:7188–7196. https://doi.org/10.1093/nar/gkm864.