JB Journals.ASM.org

Genome Sequence of the Marine Bacterium *Marinilabilia salmonicolor* JCM 21150^T

Shailesh Kumar, Srikrishna Subramanian, Gajendra Pal Singh Raghava, and Anil Kumar Pinnaka

Council of Scientific and Industrial Research (CSIR)—Institute of Microbial Technology (IMTECH), Chandigarh, India

We report the 4.98-Mb genome sequence of *Marinilabilia salmonicolor* JCM 21150^T, which was isolated from marine mud in the year 1961. The draft genome of strain *Marinilabilia salmonicolor* JCM 21150^T contains 4,982,627 bp with a G+C content of 41.92% and 4,227 protein coding genes, 52 tRNAs, and 3 rRNAs.

Marinilabilia salmonicolor JCM 21150^T was isolated from marine mud in the year 1961, and it is the only type strain of the genus *Marinilabilia* isolated to date. Cells of the strain JCM 21150^T are long slender rods of 0.3 to 0.5 μm in width and 2 to 50 μm in length. The strain is capable of gelatin liquefaction. All the strains of the genus *Marinilabilia* were reported to decompose various biomacromolecules (4).

The genome of *Marinilabilia salmonicolor* JCM 21150^T was sequenced using Illumina-HiSeq 1000 technology. A total of 35,195,778 paired-end reads (insert size, 350 bp) of 101 nucleotides in length were obtained. We used the NGS QC toolkit v2.2.1 (5) to filter the data for high-quality (HQ) (cutoff read length for HQ, 70%; cutoff quality score, 20), vector/adaptor-free reads for genome assembly. A total of 30,926,566 high-quality, vector-filtered reads (~618× coverage) were used for assembly with Velvet 1.2.03 (at a hash length of 65) (6). The final assembly contains 72 contigs with a total size of 4.98 Mb, and an N_{50} contig length of 149.27 kb was submitted to GenBank.

The 72 contigs have a total size of 4,982,627 bp (with a G+C content of 41.92%) and were used for gene prediction and annotation by the RAST (Rapid Annotation using Subsystem Technology) system (1), with tRNAscan-SE-1.23 (3) software and RNAmmer1.2 (2) servers. Out of 4,282 genes predicted, 4,227 were predicted as coding regions (CDSs), 52 as tRNA, and 3 as rRNA genes. A comparison of the genome sequences available in the RAST server shows that *Bacteroides* sp. strain 2_1_7 (score, 530), *Bacteroides* sp. strain 3_1_19 (score, 501), and *Parabacteroides distasonis* ATCC 8503 (score, 491) were the closest neighbors of the strain JCM 21150^T.

Strain JCM 21150^T contains genes for the Entner-Doudoroff pathway, hyaluronidase (not present in *Bacteroides* sp. strain 2_1_7), 2-dehydro-3-deoxygluconate kinase (EC 2.7.1.45), 4-hydroxy-2-oxoglutarate aldolase (EC 4.1.3.16)/2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14), arylsulfatase (EC 3.1.6.1), beta-hexosaminidase (EC 3.2.1.52) (chitin and *N*-acetylglucosamine utilization), alpha-glucosidase (EC 3.2.1.20), beta-galactosidase (EC 3.2.1.23), *N*-acetylmuramoyl-L-alanine amidase (EC 3.5.1.28), ATP-dependent Clp protease proteolytic subunit (EC 3.4.21.92), dipeptidyl carboxypeptidase Dcp (EC 3.4.15.5), purine nucleoside phosphorylase (EC 2.4.2.1), aminoacyl-histidine dipeptidase (peptidase D) (EC 3.4.13.3), membrane-associated zinc metalloprotease, Xaa-Pro aminopeptidase (EC 3.4.11.9), and alpha-*N*-arabinofuranosidase (EC 3.2.1.55), involved in metabolism of biomacromolecules.

Nucleotide sequence accession numbers. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number AJKI00000000. The version described in this paper is the first version, AJKI01000000.

ACKNOWLEDGMENTS

This work was funded by IMTECH-CSIR. S.K. is supported by a research fellowship from the Council of Scientific and Industrial Research (CSIR), India.

We thank the Centre for Cellular and Molecular Platforms (C-CAMP), a Department of Biotechnology (Government of India) initiative, for providing high-quality Illumina-HiSeq 1000 data.

REFERENCES

1. **Aziz RK, et al.** 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics **9**:75.
2. **Lagesen K, et al.** 2007. RNAmmer: consistent annotation of rRNA genes in genomic sequences. Nucleic Acids Res. **35**:3100–3108.
3. **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.
4. **Muller HE, Brenner DJ, Fanning GR, Grimont PA, Kampfer P.** 1996. Emended description of Buttiauxella agrestis with recognition of six new species of Buttiauxella and two new species of Kluyvera: Buttiauxella ferragutiae sp. nov., Buttiauxella gaviniae sp. nov., Buttiauxella brennerae sp. nov., Buttiauxella izardii sp. nov., Buttiauxella noackiae sp. nov., Buttiauxella warmboldiae sp. nov., Kluyvera cochleae sp. nov., and Kluyvera georgiana sp. nov. Int. J. Syst. Bacteriol. **46**:50–63.
5. **Patel RK, Jain M.** 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One **7**:e30619. doi:10.1371/journal.pone.0030619.
6. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.

Received 18 April 2012 **Accepted** 30 April 2012

Address correspondence to Anil Kumar Pinnaka, apinnaka@imtech.res.in.

doi:10.1128/JB.00649-12