

UC Berkeley

UC Berkeley Previously Published Works

Title

Draft Genome Sequence of Mycobacterium sp. Strain JC1 DSM 3803.

Permalink

<https://escholarship.org/uc/item/3g36b7mr>

Journal

Microbiology resource announcements, 10(19)

ISSN

2576-098X

Authors

Thompson, Mitchell G
Eiben, Christopher B
Pearson, Allison N
et al.

Publication Date

2021-05-01

DOI

10.1128/mra.00150-21

Peer reviewed



Draft Genome Sequence of *Mycobacterium* sp. Strain JC1 DSM 3803

Mitchell G. Thompson,^{a,b,c} Christopher B. Eiben,^d Allison N. Pearson,^{a,b}  Patrick M. Shih^{a,b,c,e,f}

^aFeedstocks Division, Joint BioEnergy Institute, Emeryville, California, USA

^bBiological Systems & Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

^cDepartment of Plant Biology, University of California, Davis, Davis, California, USA

^dPerlumi Chemicals, Inc., Berkeley, California, USA

^eGenome Center, University of California, Davis, Davis, California, USA

^fEnvironmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA

ABSTRACT *Mycobacterium* sp. strain JC1 DSM 3803 is one of the few known bacteria predicted to possess the xylulose monophosphate (XuMP) pathway of C₁ assimilation. The draft genome is 7,921,603 bp with a GC content of 66.88% and will allow more in-depth investigation of this bacterium's unique metabolism.

Mycobacterium sp. strain JC1 DSM 3803 has piqued the interest of researchers due to the diversity of C₁ assimilation pathways it employs. Isolated from soil in Seoul, South Korea, the bacterium was originally misclassified as an *Acinetobacter* sp. before 16S rRNA sequencing showed it to be a *Mycobacterium* sp. (1). Subsequent work has revealed that the bacterium does not appear to employ the canonical prokaryotic ribulose monophosphate (RuMP) cycle (2), but it possesses functional ribulose biphosphate (RuBP) (2, 3) and xylulose monophosphate (XuMP) cycles (4). Although some short fragments of the gene clusters responsible for these carbon assimilation pathways have been sequenced, the lack of a full-genome sequence prevents a complete understanding of this organism's metabolic potential.

We obtained *Mycobacterium* sp. JC1 DSM 3803 from the DSMZ-German Collection of Microorganisms and Cell Cultures (DSM 3803) and passaged pure cultures once before isolating the genomic DNA. To prepare the genomic DNA, 10 ml of bacterial culture was first grown overnight in brain heart infusion medium (BD Biosciences, San Jose, CA, USA) at 30°C, from which 1 ml was pelleted and stored at -80°C. Genomic DNA was isolated via phenol-chloroform extraction followed by ethanol precipitation, as described previously (3). Illumina library preparation and sequencing were performed by the Vincent J. Coates Genomics Sequencing Laboratory. Genomic DNA was fragmented using a Covaris M220 sonicator and size selected using AMPure XP beads to isolate fragments of 300 bp. Libraries were prepared using library preparation kits from Kapa Biosystems (Wilmington, MA, USA) and sequenced with a 150-bp paired-end NovaSeq S4 flow cell (Illumina, Inc., San Diego, CA, USA). Paired-end reads were then checked for quality with FastQC 0.11.9 (4) and trimmed using Trimmomatic 0.36 with the settings LEADING:30 TRAILING:30 MINLEN:120, resulting in 54,496,719 surviving read pairs (5). The genome was assembled *de novo* using SPades version 3.10.1 (6), and the assembly quality was assessed with QUAST 5.0.2 (7). The assembly resulted in 72 contigs over 2,000 bp (N_{50} , 224,424 bp; L_{50} , 10), comprising a genome with a total size of 7,921,603 bp, a GC content of 66.88%, and an average read coverage of 1,365 \times . Contigs were annotated via the Prokaryotic Genome Annotation Pipeline version 5.1 (8). Unless otherwise stated, all software was run using default settings.

The PGAP annotation pipeline predicts that *Mycobacterium* sp. JC1 DSM 3803 contains 7,493 coding sequences, of which 1,127 are annotated as hypothetical proteins (8, 9).

Citation Thompson MG, Eiben CB, Pearson AN, Shih PM. 2021. Draft genome sequence of *Mycobacterium* sp. strain JC1 DSM 3803. *Microbiol Resour Annot* 10:e00150-21. <https://doi.org/10.1128/MRA.00150-21>.

Editor David Rasko, University of Maryland School of Medicine

Copyright © 2021 Thompson et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Patrick M. Shih, pmsih@ucdavis.edu.

Received 26 February 2021

Accepted 14 April 2021

Published 13 May 2021

Interestingly, while previous work could not detect any RuMP activity (2), there are homologs of both 3-hexulose-6-phosphate synthase (JNN96_36860) and 6-phospho-3-hexuloisomerase (JNN96_36865) in the genome, suggesting that the bacterium may be able to utilize the RuMP pathway under certain conditions. Further work dissecting the complex metabolism of this bacterium may reveal the reasons for such a diversity of methylotrophic pathways in a single organism.

Data availability. This whole-genome sequencing project has been deposited in NCBI GenBank under the accession no. [PRJNA694986](https://www.ncbi.nlm.nih.gov/nuccore/PRJNA694986), and the Illumina short-read data have been deposited in the SRA under the accession no. [SRR13553746](https://www.ncbi.nlm.nih.gov/sra/SRR13553746).

ACKNOWLEDGMENTS

We declare the following competing financial interest: C.B.E. has a financial and managerial interest in Perlumi Chemicals, Inc.

This work was part of the DOE Joint BioEnergy Institute (<https://www.jbei.org>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.

REFERENCES

1. Song T, Lee H, Park YH, Kim E, Ro YT, Kim SW, Kim YM. 2002. Reclassification of a carboxydobacterium, *Acinetobacter* sp. strain JC1 DSM 3803, as *Mycobacterium* sp. strain JC1 DSM 3803. *J Microbiol* 40:237–240.
2. Park SW, Hwang EH, Park H, Kim JA, Heo J, Lee KH, Song T, Kim E, Ro YT, Kim SW, Kim YM. 2003. Growth of mycobacteria on carbon monoxide and methanol. *J Bacteriol* 185:142–147. <https://doi.org/10.1128/jb.185.1.142-147.2003>.
3. Sambrook J, Russell DW. 2006. Purification of nucleic acids by extraction with phenol:chloroform. *Cold Spring Harb Protoc* 2006:pdb.prot4455. <https://doi.org/10.1101/pdb.prot4455>.
4. Andrews S. 2015. FastQC. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
6. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
7. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
8. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
9. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res* 42:D206–D214. <https://doi.org/10.1093/nar/gkt1226>.