



# Draft Genome Sequence of *Aquicola tertiaricarbonis* MIMtkpLc11, an Aerobic Anoxygenic Phototrophic Bacterial Strain Isolated from Biological Soil Crusts

Kai Tang,<sup>a</sup> Bo Yuan,<sup>a,b</sup> Yonghui Zeng,<sup>c,d</sup> Lijuan Jia,<sup>a</sup> Fuying Feng<sup>a</sup>

<sup>a</sup>Institute for Applied and Environmental Microbiology, Life Sciences College, Inner Mongolia Agriculture University, Huhhot, China

<sup>b</sup>Life Sciences and Technology, College of Inner Mongolia Normal University, Huhhot, China

<sup>c</sup>Aarhus Institute of Advanced Studies, Aarhus University, Aarhus, Denmark

<sup>d</sup>Department of Environmental Science, Aarhus University, Roskilde, Denmark

**ABSTRACT** *Aquicola tertiaricarbonis* strain MIMtkpLc11 was isolated from biological soil crusts in Inner Mongolia, China. The strain contains photosynthesis gene clusters. Here, we report the draft genome sequence of strain MIMtkpLc11, which comprises 98 contigs ( $N_{50}$ , 233,472 bp) and 5,573 protein-coding sequences.

Biological soil crusts (BSCs) are widely distributed in dry lands (1). They show the ability to preserve soil fertility and reduce erosion (2), and they play a significant role in global carbon cycles (3). Aerobic anoxygenic phototrophic bacteria belong to the classes *Alphaproteobacteria*, *Betaproteobacteria*, and *Gammaproteobacteria* and contribute significantly to the recycling of organic matter (4). *Aquicola* is a genus in the class *Betaproteobacteria*, and its members have been reported to degrade tertiary butyl moieties (5) and possess the photosynthesis gene cluster (comprising *bchFNBHL*, *ppaA*, *ppsR*, and *bchG*) (6). Presently, the genus *Aquicola* has only one type species (5), and genomic data on its members are still lacking.

Strain *Aquicola tertiaricarbonis* MIMtkpLc11 was isolated from BSCs in Liangcheng, Inner Mongolia, China. This white-transparent strain was isolated by plating diluted BSCs onto R2A agars at 28°C. The 16S rRNA gene was amplified with the universal bacterial primers 27F and 1492R (7) and assembled using SeqMan version 7 software (8). Similarities between the 16S rRNA gene sequence of *A. tertiaricarbonis* strain MIMtkpLc11 and the closest relative species were calculated with EZBioCloud (9; see also <http://www.ezbiocloud.net>), and the results showed the highest similarity to be 98.36% with the type strain *A. tertiaricarbonis* L10 (GenBank accession number [NR\\_043913](https://doi.org/10.1128/MRA.01085-18)).

Genomic DNA was extracted using a genomic DNA purification kit (Tiangen Biotech Co., Beijing, China) and fragmented to sizes of 300 to 500 bp using the M220 ultrasonicator (Covaris, USA). Paired-end libraries (300-bp length) were prepared using a TruSeq DNA sample prep kit (Illumina, Inc., USA) and then sequenced using the Illumina HiSeq 2000 system. Prior to genome assembly, the raw reads were trimmed using Trimmomatic version 3.02 (10). The clean reads were assembled and corrected with SOAPdenovo version 2.04 (11) and GapCloser version 1.12 (12) under the default settings. The open reading frames (ORFs) of the assembled genome were predicted with Glimmer version 3.02 (13). All predicted ORFs were then searched with BLAST version 2.2.30+ (14) against all proteins from complete microbial genomes in the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) version 3.1 (15).

The raw sequence data comprise a total of 8,901,302 reads and 1,121,564,052 bases, which provide a high sampling coverage of the genome (169-fold coverage). The draft

Received 2 August 2018 Accepted 1 September 2018 Published 27 September 2018

**Citation** Tang K, Yuan B, Zeng Y, Jia L, Feng F. 2018. Draft genome sequence of *Aquicola tertiaricarbonis* MIMtkpLc11, an aerobic anoxygenic phototrophic bacterial strain isolated from biological soil crusts. *Microbiol Resour Announc* 7:e01085-18. <https://doi.org/10.1128/MRA.01085-18>.

**Editor** Catherine Putonti, Loyola University Chicago

**Copyright** © 2018 Tang 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 Fuying Feng, [foyefeng@hotmail.com](mailto:foyefeng@hotmail.com).

genome comprises 98 contigs with an  $N_{50}$  value of 233,472 bp and an  $N_{90}$  contig length of 54,404 bp. The total length is 6.32 Mb, containing 5,573 protein-coding sequences (CDSs), 3 rRNAs (including 1 16S rRNA, 1 23S rRNA, and 1 5S rRNA), and 43 tRNAs (19 types). The G+C content was 70.18%. The photosynthesis gene cluster contains *pufABCLM*, *bchCFMOXYZ*, *crtBCDFIWXZ*, *pucC*, *puhA*, and *acsF*. The *pufM* protein sequences were compared against the NCBI protein database and showed 85% and 86% similarities to those of *Roseateles terrae* (GenBank accession number [OWQ86667](#)) and *Pelomonas* sp. HMWF004 (GenBank accession number [PTT84989](#)), respectively.

**Data availability.** The genome sequence of *A. tertiarycarbonis* strain MIMtkpLc11 has been deposited in DDBJ/EMBL/GenBank under the accession number [JZUE00000000](#). The version described in this paper is the first version, JZUE01000000. The raw sequences were deposited in the Sequence Read Archive (SRA) under the accession number [SRP159127](#).

## ACKNOWLEDGMENT

This research was supported by the National Natural Science Foundation of China (number 31560030).

## REFERENCES

1. Belnap J. 2003. The world at your feet: desert biological soil crusts. *Front Ecol Environ* 1:181–189. <https://doi.org/10.2307/3868062>.
2. Belnap J. 2003. Biological soil crusts in deserts: a short review of their role in soil fertility, stabilization, and water relations. *Algo Stud* 109: 113–126. <https://doi.org/10.1127/1864-1318/2003/0109-0113>.
3. Porada P, Weber B, Elbert W, Pöschl U, Kleidon A. 2014. Estimating impacts of lichens and bryophytes on global biogeochemical cycles. *Global Biogeochem Cycles* 28:71–85. <https://doi.org/10.1002/2013GB004705>.
4. Koblížek M. 2015. Ecology of aerobic anoxygenic phototrophs in aquatic environments. *FEMS Microbiol Rev* 39:854–870. <https://doi.org/10.1093/femsre/fuv032>.
5. Lechner U, Brodtkorb D, Geyer R, Hause G, Hartig C, Auling G, Fayolle-Guichard F, Piveteau P, Müller RH, Rohwerder T. 2007. *Aquicola tertiarycarbonis* gen. nov., sp. nov., a tertiary butyl moiety-degrading bacterium. *Int J Syst Evol Microbiol* 57:1295–1303. <https://doi.org/10.1099/ijs.0.64663-0>.
6. Rohwerder T, Müller RH, Weichler MT, Schuster J, Hübschmann T, Müller S, Harms H. 2013. Cultivation of *Aquicola tertiarycarbonis* L108 on the fuel oxygenate intermediate tert-butyl alcohol induces aerobic anoxygenic photosynthesis at extremely low feeding rates. *Microbiology* 159: 2180–2190. <https://doi.org/10.1099/mic.0.068957-0>.
7. Klepac-Ceraj V, Lemon KP, Martin TR, Allgaier M, Kembel SW, Knapp AA, Lory S, Brodie EL, Lynch SV, Bohannon BJ, Green JL, Maurer BA, Kolter R. 2010. Relationship between cystic fibrosis respiratory tract bacterial communities and age, genotype, antibiotics and *Pseudomonas aeruginosa*. *Environ Microbiol* 12:1293–1303. <https://doi.org/10.1111/j.1462-2920.2010.02173.x>.
8. Swindell SR, Plasterer TN. 1997. SeqMan: contig assembly, p 70.75–70.89. In Swindell SR (ed), *Sequence data analysis guidebook: methods in molecular medicine*. Springer, Totowa, NJ. <https://doi.org/10.1385/0-89603-358-9:75>.
9. Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, Seo H, Chun J. 2017. Introducing EZBioCloud: a taxonomically united database of 16S rRNA and whole genome assemblies. *Int J Syst Evol Microbiol* 67:1613–1617. <https://doi.org/10.1099/ijsem.0.001755>.
10. 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>.
11. Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res* 20:265–272. <https://doi.org/10.1101/gr.097261.109>.
12. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu S-M, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam T-W, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaSci* 1:18. <https://doi.org/10.1186/2047-217X-1-18>.
13. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* 23: 673–679. <https://doi.org/10.1093/bioinformatics/btm009>.
14. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402. <https://doi.org/10.1093/nar/25.17.3389>.
15. Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity GM, Kodira CD, Kyrpidis N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O. 2008. Toward an online repository of standard operating procedures (SOPs) for (meta) genomic annotation. *OMICS* 12:137–141. <https://doi.org/10.1089/omi.2008.0017>.