

Genome Sequence of *Klebsiella oxytoca* SA2, an Endophytic Nitrogen-Fixing Bacterium Isolated from the Pioneer Grass *Psammochloa villosa*

Mingyue Chen,^a Li Lin,^b Yanming Zhang,^a Li Sun,^a Qianli An^a

State Key Laboratory of Rice Biology, Institute of Biotechnology, Zhejiang University, Hangzhou, China^a; Guangxi Key Laboratory of Sugarcane Genetic Improvement, Sugarcane Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, China^b

***Klebsiella oxytoca* strain SA2 is an endophytic nitrogen-fixing bacterium isolated from the pioneer grass *Psammochloa villosa*, which grows in the moving sand dunes of Ordos Plateau, China. The SA2 genome sequence provides the genetic background for understanding its endophytic lifestyle and survival in association with grass in nitrogen-poor environments.**

Received 5 July 2013 Accepted 17 July 2013 Published 15 August 2013

Citation Chen M, Lin L, Zhang Y, Sun L, An Q. 2013. Genome sequence of *Klebsiella oxytoca* SA2, an endophytic nitrogen-fixing bacterium isolated from the pioneer grass *Psammochloa villosa*. *Genome Announc.* 1(4):e00601-13. doi:10.1128/genomeA.00601-13.

Copyright © 2013 Chen et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Qianli An, an@zju.edu.cn.

Klebsiella oxytoca is a known opportunistic pathogen that causes septicemia, pneumonia, and urinary tract infections in humans (1–3). Some *K. oxytoca* strains are known due to their roles in biofuel production (e.g., *K. oxytoca* strain KCTC 1686 [4]) and biological nitrogen fixation (e.g., *K. oxytoca* strain M5a1 [5]).

K. oxytoca strain SA2 was isolated from roots of the pioneer grass *Psammochloa villosa*, which grows in the moving sand dunes of Ordos Plateau, China (6). It is able to fix N₂ in free-living states and in association with plants (7), as well as to colonize tissues of rice, maize, and sugarcane plants (7–9).

The genomic DNA of strain SA2 was extracted and constructed into a 500-bp-insert library, and it was sequenced by using the Illumina HiSeq 2000 sequencing system. Whole-genome sequencing resulted in 45,849,616 high-quality filtered paired-end reads, with an average length of 100 bp and about 800-fold coverage. The filtered reads were assembled *in silico* with the Velvet program (10) and the CLC Genomic Workbench 5.5.1 (CLC bio, Aarhus, Denmark). Gaps between scaffolds and within scaffolds were closed with SSPACE basic 2.0 (11) and GapFiller 1.0 (12). A draft genome containing 72 contigs was obtained based on the strain KCTC 1686 genome (accession no. CP003218) (4).

The draft genome sequence of strain SA2 comprises 5,768,574 bases representing >99.9% of the estimated genome size, and it has a G+C content of 55.9%. A total of 5,449 coding sequences (CDSs) were predicted by using Prodigal 2.60 (13) with the default parameters. Putative CDS functions were identified by using the GenDB annotation engine (14). The chromosome has 3 rRNA operons and 75 tRNAs as predicted with RNAmmer (15) and tRNAscan (16), respectively.

The SA2 genome contains genes for nitrogen fixation, siderophore production, and indoleacetic acid synthesis, which likely function in the survival of the associated bacterium and grass in the nutrition-poor sand dunes. The SA2 genome encodes multiple cell wall-degrading enzymes and multiple enzymes that mediate the elimination of reactive oxygen species generated by the plant hosts, which likely facilitate the bacterial endophytic lifestyle. The SA2 genome encodes components of multiple protein secretion

systems but not the type III and type VI secretion systems, which likely cut down the mechanisms for secreting effector proteins and virulence factors into plant and animal hosts. Together, the genome sequence of strain SA2 provides the genetic background for understanding the association between the endophytic bacterium and the pioneer grass in nitrogen-poor environments and for future application of the beneficial bacterium to enhance crop nutrition and production.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. ATNG00000000. The version described in this paper is the first version, ATNG01000000.

ACKNOWLEDGMENTS

This work was supported by the Qianjiang Talents Project grant no. 2007R10037 from the Science and Technology Department of Zhejiang Province, grants no. 31171504 and 31240056 from the National Natural Science Foundation of China, and grant no. 12-K-03-01 from the Guangxi Key Laboratory of Sugarcane Genetic Improvement.

REFERENCES

1. Hazen TH, Robinson GL, Harris AD, Rasko DA, Johnson JK. 2012. Genome sequence of *Klebsiella oxytoca* 11492-1, a nosocomial isolate possessing a fox-5 AmpC β -lactamase. *J. Bacteriol.* 194:3028–3029.
2. Liao TL, Lin AC, Chen E, Huang TW, Liu YM, Chang YH, Lai JF, Lauderdale TL, Wang JT, Chang SC, Tsai SF, Chen YT. 2012. Complete genome sequence of *Klebsiella oxytoca* E718, a New Delhi metallo- β -lactamase-1-producing nosocomial strain. *J. Bacteriol.* 194:5454. doi:10.1128/JB.01216-12.
3. Podschun R, Ullmann U. 1998. *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin. Microbiol. Rev.* 11:589–603.
4. Shin SH, Kim S, Kim JY, Lee S, Um Y, Oh MK, Kim YR, Lee J, Yang KS. 2012. Complete genome sequence of *Klebsiella oxytoca* KCTC 1686, used in production of 2,3-butanediol. *J. Bacteriol.* 194:2371–2372.
5. Bao G, Zhang Y, Du C, Chen Z, Li Y, Cao Z, Ma Y. 2013. Genome sequence of *Klebsiella oxytoca* M5a1, a promising strain for nitrogen fixation and chemical production. *Genome Announc.* 1(1):e00074-12. doi:10.1128/genomeA.00074-12.
6. Li J-D, Sun J, Kuang B-J, Chen Z-X. 1997. Isolation and characterization of endophytic diazotrophs in *Psammochloa mongolica*. *Chin. Bull. Bot.* 14:52.

7. An Q-L, Kuang B-J, Mu X-J, Li J-D. 1999. Colonization and nitrogenase production by *Klebsiella oxytoca* SA2 in rice roots. *Prog. Nat. Sci.* 9:1262–1268.
8. An Q-L, Yang X-J, Dong Y-M, Feng L-J, Kuang B-J, Li J-D. 2001. Using confocal laser scanning microscope to visualize the infection of rice roots by GFP-labelled *Klebsiella oxytoca* SA2, an endophytic diazotroph. *Acta Bot. Sin.* 43:558–564.
9. An Q-L, Yang X-J, Feng L, Mao B-Z, Kuang B-J, Li J-D. 2010. Using orthogonal sectioning with a laser scanning confocal microscopic system to determine endophytic colonization of plant roots by GFP-tagged bacteria. *J. Chin. Electr. Microsc. Soc.* 29:354–360.
10. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res.* 18:821–829.
11. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27: 578–579.
12. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. *Genome Biol.* 13:R56. doi:10.1186/gb-2012-13-6-r56.
13. 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. doi:10.1186/1471-2105-11-119.
14. Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, Kalinowski J, Linke B, Rupp O, Giegerich R, Pühler A. 2003. GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res.* 31:2187–2195.
15. Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.
16. 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.