

# Genome Sequence of *Klebsiella oxytoca* M5al, a Promising Strain for Nitrogen Fixation and Chemical Production

Guanhui Bao,<sup>a,b</sup> Yanping Zhang,<sup>a</sup> Chenyu Du,<sup>e,f</sup> Zugen Chen,<sup>c,d</sup> Yin Li,<sup>a</sup> Zhu'an Cao,<sup>e</sup> Yanhe Ma<sup>a,d</sup>

Institute of Microbiology, Chinese Academy of Sciences, Beijing, China<sup>a</sup>; University of Chinese Academy of Sciences, Beijing, China<sup>b</sup>; Department of Human Genetics, School of Medicine, University of California Los Angeles, Los Angeles, California, USA<sup>c</sup>; Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, China<sup>d</sup>; Department of Chemical Engineering, Tsinghua University, Beijing, China<sup>e</sup>; School of Biosciences, University of Nottingham, Nottingham, United Kingdom<sup>f</sup>

***Klebsiella oxytoca* is an important microorganism for nitrogen fixation and chemical production. Here, we report an annotated draft genome of *K. oxytoca* strain M5al that contains 5,256 protein-coding genes and 95 structural RNAs, which provides a genetic basis for a better understanding of the physiology of this species.**

Received 4 November 2012 Accepted 19 November 2012 Published 31 January 2013

Citation Bao G, Zhang Y, Du C, Chen Z, Li Y, Cao Z, Ma Y. 2013. Genome sequence of *Klebsiella oxytoca* M5al, a promising strain for nitrogen fixation and chemical production. *Genome Announc.* 1(1):e00074-12. doi:10.1128/genomeA.00074-12.

Copyright © 2013 Bao et al. This is an open-access article distributed under the terms of the Attribution 3.0 Unported Creative Commons License.

Address correspondence to Yanping Zhang, zhangyp@im.ac.cn.

*Klebsiella oxytoca* is a Gram-negative bacterium that has shown promise in utilizing a wide variety of sugars to produce valuable chemicals and biofuels, including ethanol (1), hydrogen (2), 1,3-propanediol, and 2,3-butanediol (3). The species is also of agricultural importance due to its function in nitrogen fixation (4). To date, two complete genomes and seven draft genomes of *K. oxytoca* strains have been stored in the NCBI database, among which only strain KCTC 1686 showed the capability to produce chemicals (2,3-butanediol) (5). Compared with the other *K. oxytoca* strains, strain M5al does not have the regular polysaccharide capsule and is therefore more easily genetically accessed and modified. Moreover, the metabolic engineering of strain M5al for the production of 1,3-propanediol and 2,3-butanediol has been reported (6). Therefore, sequencing M5al will not only enrich the genome sequence database of *K. oxytoca*, but it will help to understand the genetic background of this useful strain. Here, we report the genome sequence of *K. oxytoca* M5al and describe the sequence differences from the genome of strain KCTC 1686.

The genome of *K. oxytoca* M5al was sequenced using the Solexa technology according to Illumina's recommendations. A total of 6,199,416 reads, totaling 223,178,976 bases, were obtained, providing 37× coverage. Assembly was performed using AMOScp (7) and Velvet (8), with comparative genome and *de novo* assemblies, resulting in an assembly of 282 contigs of >500 bp each. The final assembly includes 114 contigs after manual curation. The total size of the assembly was 5.8 Mbp, with an N<sub>50</sub> of 94.98 Kbp and a G+C content of 60%. Compared to the 841-contig draft genome of *K. oxytoca* strain VJSK009, a mutant derived from the wild-type strain M5al (unpublished, [http://genome.wustl.edu/genomes/view/klebsiella\\_oxytoca\\_m5al](http://genome.wustl.edu/genomes/view/klebsiella_oxytoca_m5al)), our draft genome sequence has a significant improvement in assembly quality.

The genome sequence was annotated using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (9). The draft genome sequence of M5al contains 5,351 genes, including 5,256 predicted coding sequences (CDSs), 19 rRNA genes, and 76 tRNA genes. A total of 4,450 orthologous genes between strains

M5al and KCTC 1686 (GenBank accession no. NC\_016612.1) were identified using the Reciprocal Smallest Distance (RSD) algorithm with the default parameters (10). These comparative analyses revealed 806 protein-coding genes that are present in strain M5al but absent in strain KCTC 1686. Among the 806 proteins unique to strain M5al, 231 proteins could be mapped into the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, and most of these genes were related to purine, glyoxylate and dicarboxylate, methane, and nitrogen metabolism; this explains the genetic basis underpinning the function of strain M5al in nitrogen fixation and chemical production.

The genome sequence of *K. oxytoca* M5al serves as a basis for further investigation of the molecular basis of its potential in nitrogen fixation and chemical production. Relatively detailed annotations will reveal physiological difference among the various mutants derived from strain M5al, which showed differing capabilities to produce 1,3-propanediol (6).

**Nucleotide sequence accession numbers.** This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. [AMPJ0000000](https://doi.org/10.1093/seqmap/AMPJ0000000). The version described in this article is the first version, [AMPJ01000000](https://doi.org/10.1093/seqmap/AMPJ01000000).

## ACKNOWLEDGMENTS

This research was supported by the National Natural Science Foundation of China (grant no. 30900012) and the Scientific Equipment R&D Program of the Chinese Academy of Sciences (grant no. YZ200941). Genome sequencing was supported by the National High Technology Program of China (863 Project, grant no. 2012AA022101), the Knowledge Innovation Program of the Chinese Academy of Sciences (no. KSCX2-EW-G-14-1), and the Key Technologies R&D Program of Tianjin (no. 11ZC-ZDSY08400). Y.Z. is supported by the Knowledge Innovation Program of the Chinese Academy of Sciences (no. KSCX2-EW-Q-14).

## REFERENCES

1. Dien BS, Cotta MA, Jeffries TW. 2003. Bacteria engineered for fuel ethanol production: current status. *Appl. Microbiol. Biotechnol.* 63: 258–266.

2. Diz HR, Felder MS, Felder J. February 2007. Method for sustained microbial production of hydrogen gas in a bioreactor using *Klebsiella oxytoca*. WO patent WO/2007/032798.
3. Zhang G, Yang G, Wang X, Guo Q, Li Y, Li J. 2012. Influence of blocking of 2,3-butanediol pathway on glycerol metabolism for 1,3-propanediol production by *Klebsiella oxytoca*. Appl. Biochem. Biotechnol. **168**:116–128.
4. Fujii T, Huang YD, Higashitani A, Nishimura Y, Iyama S, Hirota Y, Yoneyama T, Dixon RA. 1987. Effect of inoculation with *Klebsiella oxytoca* and *Enterobacter cloacae* on dinitrogen fixation by rice bacteria associations. Plant Soil **103**:221–226.
5. 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.
6. Du C, Zhang Y, Li Y, Cao Z. 2007. Novel redox potential-based screening strategy for rapid isolation of *Klebsiella pneumoniae* mutants with enhanced 1,3-propanediol-producing capability. Appl. Environ. Microbiol. **73**:4515–4521.
7. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. Genome Biol. **5**:R12.
8. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.
9. Pruitt KD, Tatusova T, Klimke W, Maglott DR. 2009. NCBI reference sequences: current status, policy and new initiatives. Nucleic Acids Res. **37**:D32–D36.
10. DeLuca TF, Cui J, Jung JY, St. Gabriel KC, Wall DP. 2012. Roundup 2.0: enabling comparative genomics for over 1800 genomes. Bioinformatics **28**:715–716.