



Genome Sequence of *Staphylococcus aureus* PX03, an Acetoin-Producing Strain with a Small-Sized Genome

Ge Zhang,^{a,b,d} Qian Wang,^e Yulong Su,^c Shugui Li,^{a,b} Haobao Liu^{a,b,d}

Key Laboratory of Tobacco Biology and Processing, Tobacco Research Institute, Chinese Academy of Agricultural Sciences, Qingdao, China^a; Hainan Provincial Branch of CNTC, Haikou, China^b; Yuanjiang Branch of Yunnan Tobacco Monopoly Administration Company, Yuxi, China^c; College of Agriculture, Qingdao Agricultural University, Qingdao, China^d; Key Laboratory of Tobacco Gene Resources, Tobacco Institute of Chinese Academy of Agricultural Sciences, Qingdao, China^e

ABSTRACT *Staphylococcus aureus* PX03 can produce acetoin efficiently. Here, we present a 2.38-Mb assembly of its genome sequence, which might provide further insights into the molecular mechanism of its acetoin biosynthesis to further improve its biotechnological applications.

Staphylococcus aureus is a Gram-positive facultative anaerobe which belongs to the *Staphylococcus* family. It is an opportunistic human bacterial pathogen responsible for nosocomial and community-associated infections (1). *Staphylococcus aureus* has been demonstrated to produce acetoin, a flavoring agent widely used in the food, cosmetics, pharmaceutical, and chemical industries (2). Here, we report *Staphylococcus aureus* PX03, a strain which exhibited the potential for use in the industrial production of acetoin. The yield of acetoin reached 41 g/liter with *S. aureus* PX03 in a flask, which was significantly higher than yields for other reported strains. *S. aureus* PX03 was isolated from a soil sample in Hunan Province, China. The size of the genome is about 2.38 Mb, which is smaller than that of all other sequenced *Staphylococcus aureus* genomes (2.65 to 2.95 Mb). The genome data provide useful information for further metabolic reconstruction of the pathways for acetoin production.

The draft genome sequence of *Staphylococcus aureus* strain PX03 was obtained by high-throughput whole-genome shotgun sequencing using Illumina Hiseq 2000 sequencing technology. A 468-bp Illumina paired-end library produced 5,934,226 reads (read length, 90 bp) totaling 502 Mb. The short reads were assembled into genomes using SOAPdenovo, version 2.04 (3), resulting in 17 contigs. The protein-coding open reading frames (ORFs) were predicted using Glimmer version 3.02 (4), tRNA genes were predicted using tRNAscan-SE 1.2 (5), and rRNA genes were predicted using RNAmmer 1.2 (6) and were annotated using the Gene Ontology and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases (7, 8). A comprehensive annotation and comparative genome analysis are under way.

The size of the *S. aureus* PX03 genome is about 2.38 Mb, which is smaller than that of all other sequenced *Staphylococcus* genomes. The GC content is 32.86%. A total of 2,506 coding regions, 57 tRNA genes, and 5 rRNA loci were detected. Over 87% of genes were assigned to specific Clusters of Orthologous Groups (COG) database functional groups.

Accession number(s). This whole-genome shotgun project has been deposited at GenBank under the accession no. [LFOJ00000000](https://www.ncbi.nlm.nih.gov/nuccore/LFOJ00000000). The version described in this paper is the first version, LFOJ01000000.

ACKNOWLEDGMENT

The work was supported by Hainan Provincial Department of Science and Technology (Department of Science and Technology of Hainan Province) (ZDYF2017155).

Received 17 June 2017 Accepted 24 July 2017 Published 14 September 2017

Citation Zhang G, Wang Q, Su Y, Li S, Liu H. 2017. Genome sequence of *Staphylococcus aureus* PX03, an acetoin-producing strain with a small-sized genome. *Genome Announc* 5:e00753-17. <https://doi.org/10.1128/genomeA.00753-17>.

Copyright © 2017 Zhang 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 Haobao Liu, liuhaobao@caas.cn.

REFERENCES

1. Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui L, Oguchi A, Aoki K-i, Nagai Y, Lian J, Ito T, Kanamori M, Matsumaru H, Maruyama A, Murakami H, Hosoyama A, Mizutani-Ui Y, Takahashi NK, Sawano T, Inoue R-i, Kaito C, Sekimizu K, Hirakawa H, Kuhara S, Goto S, Yabuzaki J, Kanehisa M, Yamashita A, Oshima K, Furuya K, Yoshino C, Shiba T, Hattori M, Ogasawara N, Hayashi H, Hiramatsu K. 2001. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *Lancet* 357:1225–1240. [https://doi.org/10.1016/s0140-6736\(00\)04403-2](https://doi.org/10.1016/s0140-6736(00)04403-2).
2. Zhang X, Zhang R, Bao T, Rao Z, Yang T, Xu M, Xu Z, Li H, Yang S. 2014. The rebalanced pathway significantly enhances acetoin production by disruption of acetoin reductase gene and moderate-expression of a new water-forming NADH oxidase in *Bacillus subtilis*. *Metab Eng* 23:34–41. <https://doi.org/10.1016/j.ymben.2014.02.002>.
3. Li RQ, Zhu HM, Ruan J, Qian WB, Fang XD, Shi ZB, Li YR, Li ST, Shan G, Kristiansen K, Li SG, Yang HM, 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>.
4. 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>.
5. 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.
6. 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. <https://doi.org/10.1093/Nar/Gkm160>.
7. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, Katayama T, Araki M, Hirakawa M. 2006. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Res* 34:D354–D357. <https://doi.org/10.1093/Nar/Gkj102>.
8. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4:41. <https://doi.org/10.1186/1471-2105-4-41>.