

# Complete Genome Sequence of the Industrial Strain *Gluconobacter oxydans* H24

Xin Ge,<sup>a</sup> Yan Zhao,<sup>a,b</sup> Wei Hou,<sup>a,c</sup> Weicai Zhang,<sup>a</sup> Weiwei Chen,<sup>a</sup> Jianhua Wang,<sup>a</sup> Nan Zhao,<sup>a,d</sup> Jian Lin,<sup>a,d</sup> Wenxi Wang,<sup>a,d</sup> Mengxia Chen,<sup>e</sup> Qingge Wang,<sup>e</sup> Yinghui Jiao,<sup>e</sup> Zhigang Yuan,<sup>c</sup> Xianghua Xiong<sup>a</sup>

Laboratory of Microorganism Engineering, Beijing Institute of Biotechnology, Beijing, China<sup>a</sup>; Central Laboratory, First Affiliated Hospital of Xiamen University, Xiamen, China<sup>b</sup>; School of Preclinical Medicine, Guangxi Medical University, Nanning, China<sup>c</sup>; School of Life Science, Biochemical Pharmaceutical, Shenyang Pharmaceutical University, Shenyang, China<sup>d</sup>; Weisheng Pharmaceutical Co. Ltd. and Shijiazhuang Pharma Group, Shijiazhuang, China<sup>e</sup>

X.G., Y.Z., and W.H. contributed equally to this article.

***Gluconobacter oxydans* is characterized by its ability to incompletely oxidize carbohydrates and alcohols. The high yields of its oxidation products and complete secretion into the medium make it important for industrial use. We report the finished genome sequence of *Gluconobacter oxydans* H24, an industrial strain with high L-sorbose productivity.**

Received 4 January 2013 Accepted 7 January 2013 Published 21 February 2013

**Citation** Ge X, Zhao Y, Hou W, Zhang W, Chen W, Wang J, Zhao N, Lin J, Wang W, Chen M, Wang Q, Jiao Y, Yuan Z, Xiong X. 2013. Complete genome sequence of the industrial strain *Gluconobacter oxydans* H24. *Genome Announc.* 1(1):e00003-13. doi:10.1128/genomeA.00003-13.

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

Address correspondence to Xianghua Xiong, [xiongxiaohua@sina.com](mailto:xiongxiaohua@sina.com).

*Gluconobacter oxydans* is a Gram-negative bacterium belonging to the family *Acetobacteraceae*. This bacterium has a number of membrane-bound dehydrogenases involved in many oxidation reactions, which bring about the incomplete oxidation of sugars, alcohols, and acids. Owing to its incomplete oxidation and almost complete secretion, *G. oxydans* is widely used in industrial production (1). *Gluconobacter oxydans* H24, used in the industrial production of vitamin C in China, is responsible for the oxidation of D-sorbitol into L-sorbose. It has been consecutively optimized for several decades by mutation from a wild-type strain to improve its production of L-sorbose and its tolerance to substrate and product. Here, the entire genome of *G. oxydans* H24 was sequenced to elucidate the details of the metabolic pathway of D-sorbitol.

The complete genome sequence of *G. oxydans* H24 was determined at the Beijing Genome Institute (BGI) (Shenzhen, China) using Solexa technology. Draft assemblies were based on 925Mb total reads. All reads provided 242-fold coverage of the genome. The initial assembly of Solexa sequencing data into 24 contigs was provided by BGI, whereas the assembly of contigs into 15 scaffolds was performed using Short Oligonucleotide Alignment Program (SOAP) *denovo* software (<http://soap.genomics.org.cn>) (2). Physical gaps, repeats, and assembly ambiguities were closed and corrected by custom primer walks or by long-distance PCR amplification and Sanger sequencing. Protein-coding genes were predicted using Glimmer 3.0 (3). rRNAs and tRNAs were predicted using rRNAmmer (4) and tRNAscan-SE (5), respectively. Genes were annotated through BLASTp searches against the nonredundant (NR), Swiss-Prot, TrEMBL, Kyoto Encyclopedia of Genes and Genomes (KEGG), Clusters of Orthologous Groups (COG), and Gene Ontology (GO) databases (6–8).

The complete genome of *G. oxydans* H24 consists of a circular chromosome and a plasmid. The chromosome is composed of

3,602,424 bp, with a G+C content of 56.25%. The plasmid contains 213,808 bp, with a G+C content of 56.14%. There are a total of 3,732 putative open reading frames (3,469 in the chromosome and 263 in the plasmid), yielding a coding intensity of 89.86%. A total of 59 tRNA-encoding genes and 5 16S-23S-5S rRNA-encoding operons were identified. The genome sequences of *Gluconobacter oxydans* 621H (GenBank accession no. CP000004 to CP000009) were used as the reference (9).

The most significant feature of *G. oxydans* H24 is its high L-sorbose productivity. Two different membrane-bound, and one cytoplasmic, sorbitol dehydrogenases were identified from genome information: pyrroloquinoline quinone-dependent D-sorbitol dehydrogenase (PQQ-SLDH) (*sldhAB*) (10), flavin adenine dinucleotide-dependent D-sorbitol dehydrogenase (FAD-SLDH) (*sldhSLC*) (11), and NADP-dependent D-sorbitol dehydrogenase (NADP-SLDH) (*sldH*) (12). A comparison of three sorbitol dehydrogenase sequences with those of *G. oxydans* 621H indicated that the homology was 79.4%, 63.6%, and 29.2%, respectively. The gene cluster responsible for the synthesis of the cofactor PQQ (*pqqABCDE*, 3,137 bp) was also found (13). In addition, several genes encoding sorbose dehydrogenase (14), sorbose reductase (15), sorbosone dehydrogenase (14), glucose dehydrogenase, and other enzymes were annotated.

In summary, the genome sequence of *G. oxydans* H24 and its curated annotation are important assets for understanding better the physiology and metabolic potential of *G. oxydans*, and they will open up new opportunities to understand the functional genomics of this species.

**Nucleotide sequence accession numbers.** The nucleotide sequence was deposited in the GenBank database under the accession no. [CP003926](https://www.ncbi.nlm.nih.gov/nuccore/CP003926) and [CP003927](https://www.ncbi.nlm.nih.gov/nuccore/CP003927). The version described in this paper is the first version.

## ACKNOWLEDGMENTS

This work was supported by grants from the National Key Technology R&D Program (2007BAI46B01) and the National Nature Science Foundation of China (grant no. 31100024).

We also acknowledge the Beijing Genome Institute for whole-genome sequencing and annotation.

## REFERENCES

1. Deppenmeier U, Hoffmeister M, Prust C. 2002. Biochemistry and biotechnological applications of *Gluconobacter* strains. *Appl. Microbiol. Biotechnol.* **60**:233–242.
2. Li R, Yu C, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J. 2009. SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**:1966–1967.
3. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. *Bioinformatics* **23**:673–679.
4. 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.
5. Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res.* **33**:W686–W689.
6. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, Gasteiger E, Martin MJ, Michoud K, O'Donovan C, Phan I, Pilboud S, Schneider M. 2003. The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* **31**:365–370.
7. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. 2004. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**:D277–D280.
8. Tatusov RL, Natale DA, Garkavtsev IV, Tatusova TA, Shankavaram UT, Rao BS, Kiryutin B, Galperin MY, Fedorova ND, Koonin EV. 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res.* **29**:22–28.
9. Prust C, Hoffmeister M, Liesegang H, Wiezer A, Fricke WF, Ehrenreich A, Gottschalk G, Deppenmeier U. 2005. Complete genome sequence of the acetic acid bacterium *Gluconobacter oxydans*. *Nat. Biotechnol.* **23**:195–200.
10. Miyazaki T, Tomiyama N, Shinjoh M, Hoshino T. 2002. Molecular cloning and functional expression of D-sorbitol dehydrogenase from *Gluconobacter suboxydans* IF03255, which requires pyrroloquinoline quinone and hydrophobic protein SldB for activity development in *E. coli*. *Biosci. Biotechnol. Biochem.* **66**:262–270.
11. Toyama H, Soemphol W, Moonmangmee D, Adachi O, Matsushita K. 2005. Molecular properties of membrane-bound FAD-containing D-sorbitol dehydrogenase from thermotolerant *Gluconobacter frateurii* isolated from Thailand. *Biosci. Biotechnol. Biochem.* **69**:1120–1129.
12. Sugisawa T, Hoshino T. 2002. Purification and properties of membrane-bound D-sorbitol dehydrogenase from *Gluconobacter suboxydans* IFO 3255. *Biosci. Biotechnol. Biochem.* **66**:57–64.
13. Felder M, Gupta A, Verma V, Kumar A, Qazi GN, Cullum J. 2000. The pyrroloquinoline quinone synthesis genes of *Gluconobacter oxydans*. *FEMS Microbiol. Lett.* **193**:231–236.
14. Saito Y, Ishii Y, Hayashi H, Imao Y, Akashi T, Yoshikawa K, Noguchi Y, Soeda S, Yoshida M, Niwa M, Hosoda J, Shimomura K. 1997. Cloning of genes coding for L-sorbose and L-sorbosone dehydrogenases from *Gluconobacter oxydans* and microbial production of 2-keto-L-gulonate, a precursor of L-ascorbic acid, in a recombinant *G. oxydans* strain. *Appl. Environ. Microbiol.* **63**:454–460.
15. Shinjoh M, Tazoe M, Hoshino T. 2002. NADPH-dependent L-sorbose reductase is responsible for L-sorbose assimilation in *Gluconobacter suboxydans* IFO. *J. Bacteriol.* **184**:861–863.