Draft Genome Sequence of Type Strain *Clostridium pasteurianum* DSM 525 (ATCC 6013), a Promising Producer of Chemicals and Fuels

Sugima Rappert, Lifu Song, Wael Sabra, Wei Wang, An-Ping Zeng

Institute of Bioprocess and Biosystems Engineering, Hamburg University of Technology, Hamburg, Germany

Clostridium pasteurianum, an anaerobic bacterium able to utilize atmospheric free nitrogen for biosynthesis, has recently been proven to be a promising producer of chemicals and fuels, such as 1,3-propanediol and *n*-butanol. Here, we report the high-quality draft genome sequence of DSM 525, a type strain of *C. pasteurianum*.

Address correspondence to An-Ping Zeng, aze@tu-harburg.de.

*C*lostridium pasteurianum is an anaerobic bacterium that can utilize free nitrogen from the atmosphere for biosynthesis. The nitrogen-fixing mechanism of *C. pasteurianum* has been widely investigated since the 1950s [1–9]. The production of *n*-butanol [10–15] and 1,3-propanediol [11, 13, 16–18] using strains of *C. pasteurianum* has been studied mainly in recent years, showing promising potential applications for *C. pasteurianum* in industrial biotechnology.

Here, we present the first draft genome sequence of the *C. pasteurianum* type strain DSM 525, obtained using the high-throughput Illumina sequencing technology, which was performed by the Beijing Genomics Institute (BGI) in Shenzhen, China, with a paired-end library. The sequence reads were assembled with Short Oligonucleotide Analysis Package 2 (SOAP2) [19] with a k-mer setting of 63, which was determined by the optimal assembly result. The assembled genome sequence was annotated using the NCBI Prokaryotic Genomes Automatic Annotation Pipeline (PGAAP) (http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html). Open reading frames (ORFs) were predicted with Glimmer [20], GeneMark [21], and Prodigal [22]; tRNAs were predicted using tRNAscan-SE [23]; and ribosomal RNAs were predicted by sequence similarity search using BLAST against an RNA sequence database and/or using Infernal and Rfam models. The G+C content was calculated using the genome sequence. Prophage analysis was performed by a BLAST search against the prophage protein database derived from the prediction results of the Prophinder software and A Classification of Mobile Genetic Elements (ACLAME) database (http://aclame.ulb.ac.be/Tools/Prophinder/).

The final genome sequence of the strain DSM 525 comprises 4,285,687 bases, which are assembled into 37 contigs. It has a G+C content of 29.75%. In total, 3,964 genes were predicted using the PGAAP, including 3,887 protein-coding sequences (CDSs), 74 tRNAs, and 3 rRNA genes. According to the annotation, *C. pasteurianum* is predicted to possess a complete glycolysis pathway, while the tricarboxylic acid cycle and the pentose phosphate pathways are incomplete. Of special importance is the ability of this bacterium to produce *n*-butanol and 1,3-propanediol as major

byproducts of glycerol in a minimal medium with low concentration of yeast extract (1 g/liter). Interestingly, unlike the traditional acetone-butanol-ethanol fermentation, acetone is not a byproduct of glycerol fermentation. No prophage elements were detected.

Nucleotide sequence accession numbers. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. ANZB00000000. The version described in this paper is the first version, ANZB01000000.

REFERENCES

1. **Zinoni F, Robson RM, Robson RL.** 1993. Organization of potential alternative nitrogenase genes from *Clostridium pasteurianum*. Biochim. Biophys. Acta **1174**:83–86.
2. **Kasap M, Chen JS.** 2005. *Clostridium pasteurianum* W5 synthesizes two NifH-related polypeptides under nitrogen-fixing conditions. Microbiology **151**:2353–2362.
3. **Wang SZ, Chen JS, Johnson JL.** 1990. A nitrogen-fixation gene (nifC) in *Clostridium pasteurianum* with sequence similarity to chlJ of *Escherichia coli*. Biochem. Biophys. Res. Commun. **169**:1122–1128.
4. **Westlake DW, Wilson PW.** 1959. Molecular hydrogen and nitrogen fixation by *Clostridium pasteurianum*. Can. J. Microbiol. **5**:617–620.
5. **Rosenblum ED, Wilson PW.** 1951. The utilization of nitrogen in various compounds by *Clostridium pasteurianum*. J. Bacteriol. **61**:475–480.
6. **Zelitch I.** 1951. Simultaneous use of molecular nitrogen and ammonia by *Clostridium pasteurianum*. Proc. Natl. Acad. Sci. U. S. A. **37**:559–565.
7. **Knörzer P, Silakov A, Foster CE, Armstrong FA, Lubitz W, Happe T.** 2012. Importance of the protein framework for catalytic activity of [FeFe]-hydrogenases. J. Biol. Chem. **287**:1489–1499.
8. **Park IY, Youn B, Harley JL, Eidsness MK, Smith E, Ichiye T, Kang C.** 2004. The unique hydrogen bonded water in the reduced form of *Clostridium pasteurianum* rubredoxin and its possible role in electron transfer. J. Biol. Inorg. Chem. **9**:423–428.
9. **Meyer J, Gagnon J.** 1991. Primary structure of hydrogenase I from *Clostridium pasteurianum*. Biochemistry **30**:9697–9704.
10. **Kao WC, Lin DS, Cheng CL, Chen BY, Lin CY, Chang JS.** Enhancing butanol production with *Clostridium pasteurianum* CH4 using sequential

glucose-glycerol addition and simultaneous dual-substrate cultivation strategies. Bioresour. Technol., in press. doi:10.1016/j.biortech.2012.09.108.

11. **Moon C, Lee CH, Sang BI, Um Y.** 2011. Optimization of medium compositions favoring butanol and 1,3-propanediol production from glycerol by *Clostridium pasteurianum*. Bioresour. Technol. **102:** 10561–10568.

12. **Malaviya A, Jang YS, Lee SY.** 2012. Continuous butanol production with reduced byproducts formation from glycerol by a hyper producing mutant of *Clostridium pasteurianum*. Appl. Microbiol. Biotechnol. **93:** 1485–1494.

13. **Jensen TO, Kvist T, Mikkelsen MJ, Westermann P.** 2012. Production of 1,3-PDO and butanol by a mutant strain of *Clostridium pasteurianum* with increased tolerance towards crude glycerol. AMB Express **2:**44.

14. **Jensen TO, Kvist T, Mikkelsen MJ, Christensen PV, Westermann P.** 2012. Fermentation of crude glycerol from biodiesel production by *Clostridium pasteurianum*. J. Ind. Microbiol. Biotechnol. **39:**709–717.

15. **Venkataramanan KP, Boatman JJ, Kurniawan Y, Taconi KA, Bothun GD, Scholz C.** 2012. Impact of impurities in biodiesel-derived crude glycerol on the fermentation by *Clostridium pasteurianum* ATCC 6013. Appl. Microbiol. Biotechnol. **93:**1325–1335.

16. **Luers F, Seyfried M, Daniel R, Gottschalk G.** 1997. Glycerol conversion to 1,3-propanediol by *Clostridium pasteurianum*: cloning and expression of the gene encoding 1,3-propanediol dehydrogenase. FEMS Microbiol. Lett. **154:**337–345.

17. **Zeng AP, Biebl H.** 2002. Bulk chemicals from biotechnology: the case of 1,3-propanediol production and the new trends. Adv. Biochem. Eng. Biotechnol. **74:**239–259.

18. **Biebl H, Sproer C.** 2002. Taxonomy of the glycerol fermenting clostridia and description of *Clostridium diolis* sp. nov. Syst. Appl. Microbiol. **25:**491–7.

19. **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.

20. **Delcher AL, Bratke KA, Powers EC, Salzberg SL.** 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. Bioinformatics **23:**673–679.

21. **Besemer J, Borodovsky M.** 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res. **33:** W451–W454.

22. **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.

23. **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.