

Draft Genome Sequence of *Clostridium pasteurianum* NRRL B-598, a Potential Butanol or Hydrogen Producer

Jan Kolek,^a Karel Sedlář,^b Ivo Provazník,^b Petra Patáková^a

Department of Biotechnology, Institute of Chemical Technology, Prague, Czech Republic^a; Department of Biomedical Engineering, Brno University of Technology, Brno, Czech Republic^b

We present a draft genome sequence of *Clostridium pasteurianum* NRRL B-598. This strain ferments saccharides by two-stage acetone-butanol (AB) fermentation, is oxygen tolerant, and has high hydrogen yields.

Received 21 February 2014 Accepted 27 February 2014 Published 20 March 2014

Citation Kolek J, Sedlář K, Provazník I, Patáková P. 2014. Draft genome sequence of *Clostridium pasteurianum* NRRL B-598, a potential butanol or hydrogen producer. *Genome Announc.* 2(2):e00192-14. doi:10.1128/genomeA.00192-14.

Copyright © 2014 Kolek 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 Petra Patáková, petra.patakova@vscht.cz.

The strain *Clostridium pasteurianum* NRRL B-598 is a spore-forming, anaerobic, mesophilic, heterofermentative, rod-shaped (young cells are motile) bacterium that differs from the recently sequenced *C. pasteurianum* DSM 525 (1), especially in its inability to utilize glycerol as a substrate and its negligible formation of ethanol and production of acetone instead of 1,3-propanediol. This strain has been used in only a few studies (2–7); however, it might be a useful platform for further genetic modification because it is not sensitive to oxygen, has versatile sugar-fermenting and proteolytic abilities, seems to be genetically stable in comparison with other clostridia, and tolerates minor changes in fermentation conditions.

Based on DNA isolation, no plasmids were present and only chromosomal DNA was obtained. For *C. pasteurianum* NRRL B-598, a single-end library was sequenced with the GS Junior System (Roche). Two sequencing runs were performed. The sequence reads from both runs were assembled with a GS De Novo Assembler 2.9 (Roche), which provided the most acceptable assembly. Annotation was added by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (http://www.ncbi.nlm.nih.gov/genome/annotation_prok/). ProSplign (<http://www.ncbi.nlm.nih.gov/sutils/static/prosplign/prosplign.html>) and GeneMarkS+ (8) were used for open reading frame (ORF) detection; tRNAscan-SE (9) was used for tRNA prediction, and rRNAs were predicted by a 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 draft genome sequence. The resulting draft genome sequence of *C. pasteurianum* NRRL B-598 comprises 6,041,878 bases that are split into 138 contigs. The G+C content is 29.6%. In total, 5,547 genes were predicted by PGAP, including 5,367 protein-coding sequences (CDSs). The genome of *C. pasteurianum* NRRL B-598 is larger than that of type strain *Clostridium pasteurianum* DSM 525 (4.29 Mb) (1) as well as those of other solvent producers, e.g., *Clostridium acetobutylicum* ATCC 824 (4.13 Mb) (10) and *Clostridium acetobutylicum* DSM 1731 (11), but smaller than that of *Clostridium saccharoperbutylacetonicum* N1-4 (6.67 Mb) (12). In

total, 29 rRNA and 76 tRNA genes were identified in the genome sequence.

The genome will be subjected to thorough gene mining in the near future; however, some interesting genes have already been identified, e.g., the *spo0A* gene coding for protein sporulation initiator or catalase and superoxide dismutase genes corresponding with oxygen tolerance. Also, genes involved in solvent production (*ald*, *ctfA*, *ctfB*, and *adc*) have been identified. Genes are probably clustered in operons, and all of them are highly similar to equivalent genes which were found in the genome of *Clostridium beijerinckii* NCIMB 8052.

Nucleotide sequence accession numbers. Data from this whole-genome shotgun project have been deposited at DDBJ/EMBL/GenBank under the accession no. [AYXR000000000](https://www.ncbi.nlm.nih.gov/nuccore/AYXR000000000). Version AYXR01000000 is described in this paper.

ACKNOWLEDGMENTS

Special thanks belong to Jaroslav Vohanka from Roche CZ for assistance with the sequencing and Heinrich Klefenz from the RTM company (Germany) for inspiration and financial support. This work was also supported by the TACR BIORAF (TE01020080) and GACR NanoBioTECell (P102/11/1068) projects.

REFERENCES

1. Rappert S, Song L, Sabra W, Wang W, Zeng AP. 2013. Draft genome sequence of type strain *Clostridium pasteurianum* DSM 525 (ATCC 6013), a promising producer of chemicals and fuels. *Genome Announc.* 1(1):e00232-12. <http://dx.doi.org/10.1128/genomeA.00232-12>.
2. Patáková P, Lipovsky J, Cizkova H, Fortova J, Rychtera M, Melzoch K. 2009. Exploitation of food feedstock and waste for production of biobutanol. *Czech J. Food Sci.* 27:276–283.
3. Linhová M, Patáková P, Lipovský J, Friber P, Paulová L, Rychtera M, Melzoch K. 2010. Development of flow cytometry technique for detection of thinning of peptidoglycan layer as a result of solvent production by *Clostridium pasteurianum*. *Folia Microbiol.* 55:340–344. <http://dx.doi.org/10.1007/s12223-010-0054-1>.
4. Patáková P, Maxa D, Rychtera M, Linhova L, Friber P, Muzikova Z, Lipovsky J, Paulova L, Pospisil M, Sebor G, Melzoch K. 2011. Perspectives of biobutanol production and use, p 243–266. In Mads B (ed), *Bio-fuel's engineering process technology*. InTech, Rijeka, Croatia.
5. Patáková P, Lipovsky J, Paulova L, Linhova M, Friber P, Rychtera M,

- Melzoch K. 2011. Continuous production of butanol by bacteria of the genus *Clostridium*. *J. Chem. Chem. Eng.* 5:124–131.
6. Linhová M, Branská B, Patáková P, Lipovský J, Friber P, Rychtera M, Melzoch K. 2012. Rapid flow cytometric method for viability determination of solventogenic clostridia. *Folia Microbiol.* 57:307–311. <http://dx.doi.org/10.1007/s12223-012-0131-8>.
 7. Patakova P, Linhova M, Rychtera M, Paulova L, Melzoch K. 2013. Novel and neglected issues of acetone-butanol-ethanol (ABE) fermentation by clostridia: *Clostridium* metabolic diversity, tools for process mapping and continuous fermentation systems. *Biotechnol. Adv.* 31:58–67. <http://dx.doi.org/10.1016/j.biotechadv.2012.01.010>.
 8. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29:2607–2618. <http://dx.doi.org/10.1093/nar/29.12.2607>.
 9. 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. <http://dx.doi.org/10.1093/nar/25.5.0955>.
 10. Nölling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J, Wolf YI, Tatusov RL, Sabathe F, Doucette-Stamm L, Soucaille P, Daly MJ, Bennett GN, Koonin EV, Smith DR. 2001. Genome sequence and comparative analysis of the solvent-producing bacterium *Clostridium acetobutylicum*. *J. Bacteriol.* 183:4823–4838. <http://dx.doi.org/10.1128/JB.183.16.4823-4838.2001>.
 11. Bao G, Wang R, Zhu Y, Dong H, Mao S, Zhang Y, Chen Z, Li Y, Ma Y. 2011. Complete genome sequence of *Clostridium acetobutylicum* DSM 1731, a solvent-producing strain with multireplicon genome architecture. *J. Bacteriol.* 193:5007–5008. <http://dx.doi.org/10.1128/JB.05596-11>.
 12. del Cerro C, Felpeto-Santero C, Rojas A, Tortajada M, Ramón D, García JL. 2013. Genome sequence of the butanol hyperproducer *Clostridium saccharoperbutylacetonicum* N1-4. *Genome Announc.* 1(2):e00070-13. <http://dx.doi.org/10.1128/genomeA.00070-13>.