Complete Genome Sequence of the Nitrogen-Fixing and Solvent-Producing *Clostridium pasteurianum* DSM 525

Anja Poehlein,[a] Alexander Grosse-Honebrink,[b] Ying Zhang,[b] Nigel P. Minton,[b] Rolf Daniel[a]

Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Georg-August University, Göttingen, Germany[a]; The Clostridia Research Group, BBSRC/EPSRC Synthetic Biology Research Centre, School of Life Sciences, Centre for Biomolecular Sciences, University of Nottingham, Nottingham, United Kingdom[b]

Here, we report on the closed genome sequence of *Clostridium pasteurianum* DSM 525, which is an anaerobic, Gram-positive and endospore-forming organism. *C. pasteurianum* can fix N_2 and produce solvents such as butanol and 1,3-propanediol from carbohydrates. The genome consists of a single 4,350,673-bp replicon.

The Gram-positive anaerobic spore-forming bacterium *Clostridium pasteurianum* DSM 525, is able to produce butanol from carbohydrates (1). In contrast to most other solventogenic clostridia, *C. pasteurianum* is able to grow with glycerol as sole carbon and energy source (1, 2) and couple glycerol breakdown with a highly active butanol-producing pathway The major products during glycerol degradation are 1,3-propanediol, ethanol, and butanol (1, 3).

Strain DSM 525 was derived from the DSMZ (Braunschweig, Germany). Chromosomal DNA of *C. pasteurianum* DSM 525 was isolated using the MasterPure complete DNA purification kit (Epicentre, Madison, WI, USA). Subsequently, 454-shotgun and Illumina paired-end libraries were generated from the isolated DNA as described by the manufacturers. The libraries were sequenced using a 454 GS-FLX system (Titanium GS70 Chemistry, Roche Life Science, Mannheim, Germany) and MiSeq Illumina system (Illumina, San Diego, CA, USA), respectively. Sequencing yielded 201,156 454-shotgun and 1,215,244 paired-end Illumina reads. Assembly of the reads using the Roche Newbler assembly software 2.9 and the MIRA software (4) resulted in 139 contigs. For scaffolding and contig ordering, the move contigs tool of the Mauve genome alignment software (5) was used. The closed genome of *C. pasteurianum* ATCC 6013 (CP009267) served as the reference. Remaining gaps were closed by PCR-based techniques and Sanger sequencing of the products using BigDye 3.0 chemistry and an ABI3730XL capillary sequencer (Applied Biosystems, Life Technologies GmbH, Darmstadt, Germany). For this purpose, the Gap4 (v4.11) software of the Staden package (6) was employed. The complete genome of *C. pasteurianum* DSM 525 consists of a single chromosome of 4,350,673 bp with an overall G+C content of 30%. Automatic gene prediction was performed with the software tool prodigal (Prokaryotic Dynamic Programming Genefinding Algorithm) (7). Identification of rRNA and tRNA genes was done with RNAmmer (8) and tRNAscan (9), respectively. An integrated microbial genomes/expert review (IMG/ER) system (10, 11) was used for automatic annotation, which was subsequently manually curated by using the Swiss-Prot, TREMBL, and InterPro databases (12). We identified 10 rRNA operons, 81 tRNA genes, 3,220 protein-encoding genes with function prediction, and 768 genes coding for hypothetical proteins. Genes coding for key enzymes of butanol fermentation such as butyryl-CoA dehydrogenase (*bcd*), electron transfer flavoprotein (*eftAB*), 3-hydroxybutyryl-CoA dehydrogenase (*hbd*), and 3-hydroxybutyryl-CoA dehydratase (*crt*) form a cluster that is identical to those identified in other solventogenic clostridia, such as *C. acetobutylicum*, *C. saccharoperbutylacetonicum*, or *C. saccharobutylicum* (13–15). In addition, the genome of *C. pasteurianum* DSM 25 harbors a cluster coding for CoA transferase (*ctfAB*), acetoacetate decarboxylase (*adc*), and alcohol/aldehyde dehydrogenase (*adhE*), which showed the identical arrangement as the *sol* operon of *C. acetobutylicum* (16). We also encountered genes encoding acetate kinase (*ackA*), phosphate acetyltransferase (*pta*), butyrate kinase (*buk*), and phosphate butyryltransferase (*ptb*). In addition, the presence of the previously described genes encoding key enzymes for 1,3-propanediol production such as B_{12}-dependent glycerol dehydratase (17) and 1,3-propanediol dehydrogenase (2) was confirmed.

Nucleotide sequence accession number. The complete genome sequence has been deposited in GenBank under the accession no. CP009268.

REFERENCES

1. **Dabrock B, Bahl H, Gottschalk G.** 1992. Parameters affecting solvent production by *Clostridium pasteurianum*. Appl Environ Microbiol **58**:1233–1239.
2. **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. http://dx.doi.org/10.1111/j.1574-6968.1997.tb12665.x.

3. **Biebl H.** 2001. Fermentation of glycerol by *Clostridium pasteurianum*—batch and continuous culture studies. J Ind Microbiol Biotechnol **27:**18–26. http://dx.doi.org/10.1038/sj/jim/7000155.

4. **Chevreux B, Wetter T, Suhai S.** 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. *In* Wingender E (ed), Computer science and biology: proceedings of the German conference on bioinformatics (GCB) 1999. GBF-Braunschweig, Hannover, Germany. Department of Bioinformatics, Braunschweig, Germany.

5. **Darling AE, Mau B, Perna NT.** 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One **5:**e11147. http://dx.doi.org/10.1371/journal.pone.0011147.

6. **Staden R, Beal KF, Bonfield JK.** 2000. The Staden package, 1998. Methods Mol Biol **132:**115–130.

7. **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. http://dx.doi.org/10.1186/1471-2105-11-119.

8. **Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res **35:**3100–3108. http://dx.doi.org/10.1093/nar/gkm160.

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. **Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC.** 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271–2278. http://dx.doi.org/10.1093/bioinformatics/btp393.

11. **Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC.** 2012. IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Res **40:**D115–D122. http://dx.doi.org/10.1093/nar/gkr1044.

12. **Zdobnov EM, Apweiler R.** 2001. InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17:**847–848. http://dx.doi.org/10.1093/bioinformatics/17.9.847.

13. **Nölling J, Breton G, Omelchenko MV, Makarova KS, Zeng Q, Gibson R, Lee HM, Dubois J, Qiu D, Hitti J, GTC Sequencing Center Production Finishing, and Bioinformatics Teams, 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.

14. **Poehlein A, Krabben P, Dürre P, Daniel R.** 2014. Complete genome sequence of the solvent producer *Clostridium saccharoperbutylacetonicum* strain DSM 14923. Genome Announc **2**(5):e01056-14. http://dx.doi.org/10.1128/genomeA.01056-14.

15. **Poehlein A, Hartwich K, Krabben P, Ehrenreich A, Liebl W, Dürre P, Gottschalk G, Daniel R.** 2013. Complete genome sequence of the solvent producer *Clostridium saccharobutylicum* NCP262 (DSM 13864). Genome Announc **1**(6):e00997-13. http://dx.doi.org/10.1128/genomeA.00997-13.

16. **Fischer RJ, Helms J, Dürre P.** 1993. Cloning, sequencing, and molecular analysis of the *sol* operon of *Clostridium acetobutylicum*, a chromosomal locus involved in solventogenesis. J Bacteriol **175:**6959–6969.

17. **Macis L, Daniel R, Gottschalk G.** 1998. Properties and sequence of the coenzyme B_{12}-dependent glycerol dehydratase of *Clostridium pasteurianum.* FEMS Microbiol Lett **164:**21–28. http://dx.doi.org/10.1111/j.1574-6968.1998.tb13062.x.