genomeA
Journals.ASM.org

Complete Genome Sequence of the Solvent Producer *Clostridium saccharoperbutylacetonicum* Strain DSM 14923

Anja Poehlein,[a] Preben Krabben,[b] Peter Dürre,[c] Rolf Daniel[a]

Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Georg-August University, Göttingen, Germany[a]; Green Biologics Ltd., Abingdon, Oxfordshire, United Kingdom[b]; Institut für Mikrobiologie und Biotechnologie, Universität Ulm, Ulm, Germany[c]

Clostridium saccharoperbutylacetonicum **strain DSM 14923 is known as a butanol-producing bacterium. Various organic compounds such as glucose, fructose, sucrose, mannose, and cellobiose are fermented. The genome consists of one chromosome and one circular megaplasmid. *C. saccharoperbutylacetonicum* was used in industrial fermentation processes to produce the solvents acetone, butanol, and ethanol.**

Received 8 September 2014 **Accepted** 10 September 2014 **Published** 16 October 2014

Citation 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. doi:10.1128/genomeA.01056-14.

Address correspondence to Rolf Daniel, rdaniel@gwdg.de.

Clostridium saccharoperbutylacetonicum is an anaerobic, spore-forming, Gram-positive bacterium that produces solvents and hydrogen. The strain was first cultured from soil by Hongo et al. in 1959 as strain 97 (1, 2) and was subsequently used by the Sanraku Distillers Company for butanol production in the early 1960s, until production was stopped due to phage issues (3). Here, we report the closed and fully annotated genome sequence of *C. saccharoperbutylacetonicum* strain DSM 14923. This strain is also designated N1-4 (HMT). Another genome sequence of this organism is publicly available, but it is a draft sequence with 210 contigs (4).

The MasterPure complete DNA purification kit (Epicentre, Madison, WI, USA) was used to isolate chromosomal DNA of *C. saccharoperbutylacetonicum* DSM 14923. For whole-genome sequencing the Genome Analyzer II (Illumina, San Diego, CA, USA) and the 454 GS-FLX Titanium XL pyrosequencing system (Titanium Chemistry, Roche Life Science, Mannheim, Germany) were used. Preparation of shotgun libraries was performed according to the manufacturers' protocols and resulted in 3,087,596 paired-end Illumina reads and 402,533 pyrosequencing reads. Initial hybrid *de novo* assembly using the MIRA software (5) resulted in 327 contigs and an average coverage of 58.29×. The remaining gaps were closed by PCR-based techniques and Sanger sequencing using BigDye 3.0 chemistry and an ABI3730XL capillary sequencer (Applied Biosystems, Life Technologies GmbH, Darmstadt, Germany). For this purpose, the Gap4 (version 4.11) software of the Staden package was employed (6). The complete genome of *C. saccharoperbutylacetonicum* DSM14923 comprises two replicons, a chromosome (6.53 Mb), and a circular megaplasmid (136 kb) with an overall GC content of 29.54%. This represents the highest genome size of all completely sequenced clostridial genomes. Automatic gene prediction was performed using YACOP and GLIMMER software (7). Identification of rRNA and tRNA genes was done with RNAmmer (8) and tRNAscan (9), respectively. The IMG/ER (Integrated Microbial Genomes/Expert Review) 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 11 rRNA operons, 70 tRNA genes, 4,287 protein-encoding genes with function prediction, 1,534 genes coding for hypothetical proteins, and 11 pseudogenes. In contrast to *C. acetobutylicum* ATCC 824 (13), the *sol* operon of *C. saccharoperbutylacetonicum* is located on the chromosome and not on the megaplasmid. It consists of aldehyde dehydrogenase (*ald*), CoA transferase (*ctfAB*) and acetoacetate decarboxylase (*adc*). The *sol* operon of *C. saccharoperbutylacetonicum* shows the same arrangement as that of *C. beijerinckii* NCIMB8052 (14, 15) and *C. saccharobutylicum* (16) but differs from that of *C. acetobutylicum* ATCC 824, in which the *ald* gene is replaced by an alcohol/aldehyde dehydrogenase-encoding gene (*adhE*) and *adc* forms a separate operon. Based on a substrate preference for butyraldehyde, the aldehyde dehydrogenase of *C. saccharoperbutylacetonicum* was described as butyraldehyde dehydrogenase (17). Genes encoding acetyl-CoA acetyltransferase, crotonase, butyryl-CoA dehydrogenase, phosphate butyltransferase, butyrate kinase, phosphate acetyltransferase, acetate kinase, and several alcohol dehydrogenases, key enzymes for solvent production, were identified.

Nucleotide sequence accession numbers. The genome sequence was deposited in GenBank under accession numbers CP004121 (chromosome) and CP004122 (megaplasmid).

ACKNOWLEDGMENTS

We thank the Bundesministerium für Bildung und Forschung (BMBF), the Bundesministerium für Ernährung, Landwirtschaft und Verbraucherschutz (BMEL), and the Innovate UK/Technology Strategy Board for support. We thank Frauke-Dorothee Meyer and Kathleen Gollnow for technical support.

REFERENCES

1. **Hongo M, Murata A, Kono K, Kato F.** 1968. Lysogeny and bacteriocinogeny in strains of *Clostridium* species. Agric. Biol. Chem. **32:**27–33. http://dx.doi.org/10.1271/bbb1961.32.27.
2. **Hongo M.** July 1960. Process for producing butanol by fermentation. US patent 2,945,786.

3. **Jones D.** 2013. The strategic importance of biobutanol for Japan during WWII: a case study of the butanol fermentation process in Taiwan and Japan, p 220–272. *In* Dürre P (ed), Systems biology of clostridium. Imperial College Press, London.
4. **del Cerro C, Felpeto-Santero C, Rojas A, Tortajada M, Ramón D, García JL.** 2013. Genome sequence of the butanolhyperproducer *Clostridium saccharoperbutylacetonicum* N1-4. Genome Announc. **1**(2)**:**e00070-13. http://dx.doi.org/10.1128/genomeA.00070-13.
5. **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) '99 Hannover, Germany. GBF-Braunschweig, Department of Bioinformatics, Braunschweig, Germany.
6. **Staden R, Beal KF, Bonfield JK.** 2000. The Staden package, 1998. Methods Mol. Biol. **132:**115–130. http://dx.doi.org/10.1385/1-59259-192-2:115.
7. **Tech M, Merkl R.** 2003. YACOP: Enhanced gene prediction obtained by a combination of existing methods. In Silico Biol. **3:**441–451.
8. **Lagesen K, Hallin P, Rødland EA, Staerfeldt 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. **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.
14. **Chen CK, Blaschek HP.** 1999. Examination of physiological and molecular factors involved in enhanced solvent production by *Clostridium beijerinckii* BA101. Appl. Environ. Microbiol. **65:**2269–2271.
15. **Wang Y, Li X, Mao Y, Blaschek HP.** 2011. Single-nucleotide resolution analysis of the transcriptome structure of *Clostridium beijerinckii* NCIMB 8052 using RNA-Seq. BMC Genomics **12:**479. http://dx.doi.org/10.1186/1471-2164-12-479.
16. **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.
17. **Kosaka T, Nakayama S, Nakaya K, Yoshino S, Furukawa K.** 2007. Characterization of the *sol* operon in butanol-hyperproducing *Clostridium saccharoperbutylacetonicum* strain N1-4 and its degeneration mechanism. Biosci. Biotechnol. Biochem. **71:**58–68. http://dx.doi.org/10.1271/bbb.60370.