Genome Sequence of the Acetogenic Bacterium *Oxobacter pfennigii* DSM 3222^T

Frank R. Bengelsdorf,[a] Anja Poehlein,[b] Bettina Schiel-Bengelsdorf,[a] Rolf Daniel,[b] Peter Dürre[a]

Institut für Mikrobiologie und Biotechnologie, Universität Ulm, Ulm, Germany[a]; Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Georg-August University Göttingen, Göttingen, Germany[b]

F.R.B. and A.P. contributed equally to this work.

Here, we report the draft genome sequence of *Oxobacter pfennigii* DSM 3222^T, an anaerobic, acetogenic, carbon monoxide-oxidizing, and butyrate-producing bacterium. The genome consists of a chromosome with a size of 4.49 Mbp.

Address correspondence to Peter Dürre, peter.duerre@uni-ulm.de.

*O*xobacter pfennigii DSM 3222^T (originally named *Clostridium pfennigii*) is an anaerobic endospore-forming bacterium, which was isolated from rumen fluid of a steer and described by Krumholz and Bryant in 1985 (1). In 1994, Collins et al. (2) compared phenotypic and phylogenetic data of clostridial species and found that *C. pfennigii* clearly merits a new genus. The closest autotrophic acetogenic relatives are, for instance, *Clostridium ljungdahlii* and *Clostridium carboxidivorans* (3). *O. pfennigii* does not use sugars or amino acids as an energy source, but catabolizes methoxylated aromatic compounds (e.g., vanilline, ferulate, caffeate, syringate) to butyrate (1). Moreover, it can reduce carbon monoxide and produce acetate and butyrate (1). As of yet, *O. pfennigii* is still the only member of its genus and not even a closely related bacterium has been described.

Chromosomal DNA of *O. pfennigii* was isolated using the MasterPure complete DNA purification kit (Epicentre, Madison, WI, USA). Illumina shotgun libraries were generated from the extracted DNA according to the protocol of the manufacturer. Sequencing was performed by employing a MiSeq system using MiSeq ReagentKit v3 (600 cycles), as recommended by the manufacturer (Illumina, San Diego, CA, USA), resulting in 4,476,360 paired end reads (300 bp) that were trimmed using Trimmomatic 0.32 (4).

The *de novo* assembly performed with the SPAdes genome assembler software 3.5.0 (5) resulted in 69 contigs (>500 bp) and an average coverage of 192-fold. The genome of *O. pfennigii* probably comprises a circular chromosome (4,510,552 bp) that has an overall G+C content of 39.0%. The software tool Prodigal (6) was used for automatic gene prediction, genes coding for rRNA and tRNA were identified using RNAmmer (7) and tRNAscan (8), respectively. The Integrated Microbial Genomes-Expert Review (IMG-ER) system (9) was used for automatic annotation, which was subsequently manually curated by using the Swiss-Prot, TrEMBL, and InterPro databases (10). The genome contains 9 rRNA genes, 66 tRNA genes, 3,328 protein-coding genes with predicted functions, and 949 genes coding for hypothetical proteins. The genome of *O. pfennigii* harbors genes encoding all proteins of the

glycolysis pathway even though the bacterium does not use sugars as an energy source. Since no PTS (phosphoenolpyruvate [PEP]: carbohydrate phosphotransferase system) for sugars is found in the genome, the respective gene products of the glycolysis pathway probably serve only for gluconeogenesis.

With respect to the autotrophic acetogenic nature of *O. pfennigii*, we found genes encoding proteins of the methyl- and carbonyl-branches of the Wood-Ljungdahl pathway as well as of the Rnf (*Rhodobacter* nitrogen fixation) complex. The genes of the Wood-Ljungdahl pathway are not strictly clustered as shown for members of the genus *Clostridium* (11). Regarding the butyrate metabolism of *O. pfennigii*, genes encoding thiolase, 3-hydroxybutyryl-CoA dehydrogenase, crotonase, butyryl-CoA dehydrogenase, and electron-transfer-flavoproteins (A and B) are located in one cluster as in *Clostridium kluyveri* (12). Additionally, phosphotransbutyrylase- and butyrate kinase-encoding genes are clustered separately. Moreover, genes encoding proteins for nitrogenase, sporulation, and a glycine reductase were found in the genome of *O. pfennigii*.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. LKET00000000. The version described in this paper is version LKET01000000.

REFERENCES

1. **Krumholz LR, Bryant MP.** 1985. *Clostridium pfennigii* sp. nov. uses methoxyl groups of monobenzenoids and produces butyrate. Int J Syst Bacteriol **35:**454–456. http://dx.doi.org/10.1099/00207713-35-4-454.
2. **Collins MD, Lawson PA, Willems A, Cordoba JJ, Fernandez-Garayzabal J, Garcia P, Cai J, Hippe H, Farrow JAE.** 1994. The phylog-

eny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. Int J Syst Bacteriol **44**:812–826. http://dx.doi.org/10.1099/00207713-44-4-812.

3. **Bengelsdorf FR, Straub M, Dürre P**. 2013. Bacterial synthesis gas (syngas) fermentation. Environ Tech **34**:1639–1651. http://dx.doi.org/10.1080/09593330.2013.827747.

4. **Bolger AM, Lohse M, Usadel B**. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30**:2114–2120. http://dx.doi.org/10.1093/bioinformatics/btu170.

5. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA**. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol **19**:455–477. http://dx.doi.org/10.1089/cmb.2012.0021.

6. **Hyatt D, Chen G, 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.

7. **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.

8. **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.

9. **Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Pillay M, Ratner A, Huang J, Woyke T, Huntemann M, Anderson I, Billis K, Varghese N, Mavromatis K, Pati A, Ivanova NN, Kyrpides NC**. 2014. IMG4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res **42**:D560–D567. http://dx.doi.org/10.1093/nar/gkt963.

10. **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.

11. **Poehlein A, Cebulla M, Ilg MM, Bengelsdorf FR, Schiel-Bengelsdorf B, Whited G, Andreesen JR, Gottschalk G, Daniel R, Dürre P**. 2015. The complete genome sequence of *Clostridium aceticum*: a missing link between Rnf- and cytochrome-containing autotrophic acetogens. mBio **6**:e01168-15. http://dx.doi.org/10.1128/mBio.01168-15.

12. **Seedorf H, Fricke WF, Veith B, Brüggemann H, Liesegang H, Strittmatter A, Miethke M, Buckel W, Hinderberger J, Li F, Hagemeier C, Thauer RK, Gottschalk G**. 2008. The genome of *Clostridium kluyveri*, a strict anaerobe with unique metabolic features. Proc Natl Acad Sci USA **105**:2128–2133. http://dx.doi.org/10.1073/pnas.0711093105.