# First Insights into the Genome of the Gram-Negative, Endospore-Forming Organism *Sporomusa ovata* Strain H1 DSM 2662

Anja Poehlein, Gerhard Gottschalk, Rolf Daniel

Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University Göttingen, Göttingen, Germany

The genome of *Sporomusa ovata* strain H1 DSM 2662, an anaerobic, Gram-negative endospore-forming bacterium, was sequenced. *S. ovata* uses *N*-methyl compounds, primary alcohols, fatty acids, and $H_2$ and $CO_2$ as energy and carbon sources to produce acetate. The genome harbors one chromosome, which encodes proteins typical for sporulation.

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

The Gram-negative endospore-forming bacterium *Sporomusa ovata* belongs to the class *Negativicutes* within the *Firmicutes*. This class comprises only a few genera, which are Gram negative and form endospores. *S. ovata* was one of the first described species with this feature (1). Based on genomic comparisons of Gram-negative members of the *Firmicutes*, the assignment of *Sporomusa* to the new family *Sporomusaceae* was recommended (2). *S. ovata* ferments *N*-methyl compounds, such as betaine, *N,N*-dimethylglycine, and sarcosine, but also primary alcohols, hydroxy fatty acids, and 2,3-butanediol. The main product is acetate, which is also produced from $H_2$ and $CO_2$.

Genomic DNA of *S. ovata* strain H1 DSM 2662 was isolated with the MasterPure complete DNA purification kit (Epicenter, Madison, WI). The extracted DNA was used to generate 454-shotgun, paired-end, and Illumina-shotgun libraries according to the manufacturer's protocols. The libraries were sequenced using a 454 GS-FLX system (Titanium GS70 chemistry; Roche Life Sciences, Mannheim, Germany) and Genome Analyzer II (Illumina, San Diego, CA). Sequencing resulted in coverages of 17.99 and 101.75, respectively, with the two sequencing systems. Assembly of the reads using Roche Newbler assembly software 2.6 for scaffolding and MIRA software (3) resulted in 37 scaffolds with 60 contigs. The remaining gaps were closed with PCR-based techniques and Sanger sequencing of the products (4) employing the Gap4 (v.4.11) software of the Staden package (5). The draft genome of *S. ovata* H1 DSM 2662 comprised one circular chromosome of 5.38 Mb with an overall G+C content of 42.25 mol%. Functional annotation of the 5,110 predicted protein-encoding genes was initially carried out with the IMG/ER (Intergrated Microbial Genomes/Expert Review) system (6, 7). Subsequently, annotations were manually curated by using the Swiss-Prot, TREMBL, and InterPro databases (8). The genome harbored at least 13 rRNA operons and 127 tRNA genes, which were identified with RNAmmer and tRNAscan, respectively (9, 10).

Analysis of the genome sequence revealed the presence of various sensory histidine kinase (KinACDE) transcription and sigma factors such as Spo0A, $\sigma^H$, $\sigma^F$, $\sigma^E$, $\sigma^G$, and $\sigma^K$, which are essential for initiation of sporulation (11, 12). At least 83 genes coding for proteins involved in the various stages of sporulation were identified, and all such proteins were orthologous to known proteins involved in sporulation of *Clostridia* and *Bacilli* (13).

Genes coding for outer membrane proteins, chaperones, and outer membrane efflux proteins were detected, as well as genes for lipid A biosynthesis acetyl transferases and lipid A disaccharide synthetases. In addition, a putative *pylTScBCDSn* gene cluster encoding proteins necessary for incorporation of pyrrolysine into proteins was present (14). Upstream of this cluster, putative genes encoding corrinoid-dependent and pyrrolysine-containing methylamine methyltransferases (15) were located. Besides those in the *Methanosarcinaceae*, in which the *pyl* genes were discovered, we identified these genes by genome comparisons in only a few genera belonging to the *Peptococcaceae*, *Halobacteroidaceae*, and *Thermoanaerobacteriaceae*, which are, as is *S. ovata*, members of the *Firmicutes*.

**Nucleotide sequence accession numbers.** The draft genome sequence of *Sporomusa ovata* H1 DSM 2662 has been deposited at DDBJ/EMBL/GenBank under the accession number ASXP00000000. The version described is version ASXP01000000.

## REFERENCES

1. **Möller B, Oßmer R, Howard BH, Gottschalk G, Hippe H.** 1984. *Sporomusa*, a new genus of gram-negative anaerobic bacteria including *Sporomusa sphaeroides* spec. nov. and Sporomusa ovata spec. nov. Arch. Microbiol. **139:**388–396.
2. **Yutin N, Galperin MY.** 2013. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced *Clostridia*. Environ. Microbiol. [Epub ahead of print.] doi:10.1111/1462-2920.12173.
3. **Chevreux B, Wetter T, Suhai S.** 1999. Genome sequence assembly using trace signals and additional sequence information, p 45–56. *In* Computer science and biology: proceedings of the German Conference on Bioinformatics. GCB, Hannover, Germany.
4. **Sanger F, Nicklen S, Coulson AR.** 1992. DNA sequencing with chain-terminating inhibitors. 1977. Biotechnology **24:**104–108.

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

6. **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.

7. **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. doi:10.1093/nar/gkr1044.

8. **Zdobnov EM, Apweiler R.** 2001. InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17:** 847–848.

9. **Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108.

10. **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.

11. **Piggot PJ, Hilbert DW.** 2004. Sporulation of *Bacillus subtilis*. Curr. Opin. Microbiol. **6:**579–586.

12. **Steil L, Serrano M, Henriques AO, Völker U.** 2005. Genome-wide analysis of temporally regulated and compartment-specific gene expression in sporulating cells of *Bacillus subtilis*. Microbiology **151:**399–420.

13. **Paredes CJ, Alsaker KV, Papoutsakis ET.** 2005. A comparative genomic view of clostridial sporulation and physiology. Nat. Rev. Microbiol. **12:**969–978.

14. **Zhang Y, Baranov PV, Atkins JF, Gladyshev VN.** 2005. Pyrrolysine and selenocysteine use dissimilar decoding strategies. J. Biol. Chem. **280:** 20740–20751.

15. **Krzycki JA.** 2004. Function of genetically encoded pyrrolysine in corrinoid-dependent methylamine methyltransferases. Curr. Opin. Chem. Biol. **8:**484–491.