JB Journals.ASM.org

GENOME ANNOUNCEMENT

Complete Genome Sequence of the Ethanol-Producing *Zymomonas mobilis* subsp. *mobilis* Centrotype ATCC 29191

Andreas Desiniotis,[a] Vassili N. Kouvelis,[a] Karen Davenport,[b] David Bruce,[c] Chris Detter,[b] Roxanne Tapia,[b] Cliff Han,[b] Lynne A. Goodwin,[b] Tanja Woyke,[b] Nikos C. Kyrpides,[c] Milton A. Typas,[a] and Katherine M. Pappas[a]

Department of Genetics & Biotechnology, Faculty of Biology, University of Athens, Panepistimiopolis, Athens, Greece[a]; DOE Joint Genome Institute, Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA[b]; and DOE Joint Genome Institute, Walnut Creek, California, USA[c]

***Zymomonas mobilis* is an ethanologenic bacterium that has been studied for use in biofuel production. Of the sequenced *Zymomonas* strains, ATCC 29191 has been described as the phenotypic centrotype of *Zymomonas mobilis* subsp. *mobilis*, the taxon that harbors the highest ethanol-producing *Z. mobilis* strains. ATCC 29191 was isolated in Kinshasa, Congo, from palm wine fermentations. This strain is reported to be a robust levan producer, while in recent years it has been employed in studies addressing *Z. mobilis* respiration. Here we announce the finishing and annotation of the ATCC 29191 genome, which comprises one chromosome and three plasmids.**

Zymomonas mobilis is an alphaproteobacterium that ferments sugars to ethanol and carbon dioxide to almost-perfect yields (14, 15). A comparative analysis of several strains of the species is under way at the U.S. DOE Joint Genome Institute, in collaboration with the University of Athens (http://www.jgi.gov). Strain ATCC 29191 (Z6; NCIMB 11199) was included in the analysis, as it is regarded the most representative strain of *Z. mobilis* subsp. *mobilis* (4). ATCC 29191 was isolated from Zairian *Elaeis* sp. sap fermentations (18), and it is superior to other *Z. mobilis* strains in levan (polyfructan) production. It is also highly comparable to the biotechnologically important fast-growing *Zymomonas* sp. ATCC 31821 derivatives in overall growth and yields, particularly on sucrose substrates (16). Notably, energy metabolism in *Z. mobilis*, i.e., electron transport and oxidative phosphorylation, has been intensely studied with this strain (9, 10, 17).

Total and plasmid DNA from ATCC 29191 were prepared as described previously (12) and used for whole-genome shotgun sequencing at the DOE JGI, using a combination of Illumina (2) and 454 (11) technologies. For this, we constructed an Illumina GAii library (generating 6,353,828 sequence reads, totaling 228.7 Mb) and two 454 libraries, a Titanium standard and a paired-end library with a 22-kb average insert size (515,697 and 168,806 reads, respectively, totaling 121.6 Mb) (http://www.jgi.doe.gov). The 454 and Illumina data were assembled with Newbler version 2.3 and Velvet version 0.7.63, correspondingly (20). Final data integration made use of parallel phrap, version SPS 4.24 (High Performance Software, LLC). Illumina data were used to correct base errors and increase consensus quality using Polisher (A. Lapidus, unpublished data). Misassemblies were corrected by using gapResolution or Dupfinisher (8; C. Han, unpublished data), or sequencing bridging PCR fragments. Gaps between contigs were closed by editing in Consed (5, 6, 7), by PCR and by Bubble PCR primer walks (J.-F. Cheng, unpublished data). A total of 217 additional reactions and 2 Shatter libraries closed gaps and raised the quality to 0.00 errors per 10 kb. The final assembly was based on 38.3× and 115.8× genome coverages for the 454 and Illumina data, respectively. Coding gene prediction, functional gene assignment, and tRNA/rRNA gene identification were conducted as described before (13). Genome structure comparisons relied on ACT (3), BLASTN (1), and MegaBLAST (21).

ATCC 29191 contains a circular chromosome of 1,961,307 bp and three plasmids, p29191_1 to p29191_3, of 18,350 bp, 14,947 bp, and 13,742 bp, respectively (GC contents of 46.21% and of 41.02%, 42.19%, and 44.21%, correspondingly). The entire genome has 1,765 protein-coding genes, 51 tRNA genes, and 3 rRNA gene clusters.

The ATCC 29191 genome is 95,057 bp smaller than that of reference strain ATCC 31821 (ZM4) (19) and shares an average 97% identity with it. Synteny is retained for the largest part, with the exception of local shuffling, plausibly due to transposase presence (27 annotated transposases in total). Forty-one genes—assigned, pseudogenes, or hypothetical—are unique to ATCC 29191 compared to ZM4, whereas for the latter, 115 genes are unique. Many of the ATCC 29191-unique genes are located on three stretches bearing the lowest resemblance to ZM4 (coordinates 110991 to 115129, 1298096 to 1304030, and 1861489 to 1864804), including helicase, transporter, and tellurium resistance genes. The plasmids harbor replicon maintenance, metabolism, regulation, transposition, and DNA restriction/modification genes.

Nucleotide sequence accession numbers. The ATCC 29191 genome was assigned GenBank accession numbers CP003704 for the chromosome and CP003705 to CP003707 for the plasmids.

ACKNOWLEDGMENTS

We sincerely acknowledge all DOE-JGI collaborators who contributed to the sequencing, assembly, and automated annotation.

Work at the DOE JGI is financed by the U.S. DOE Office of Science, contract number DE-AC02-05CH11231. K.M.P. thanks the NKUA Research Committee for award 70/4/7809. A.D. thanks the Vienna Medical Society/Sanofi for the Wilhelm-Auerswald-Preis 2012 award.

REFERENCES

1. **Altschul SF, et al.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389–3402.
2. **Bennett S.** 2004. Solexa Ltd. Pharmacogenomics **5**:433–438.
3. **Carver TJ, et al.** 2005. ACT: the Artemis comparison tool. Bioinformatics **21**:3422–3423.

Received 7 August 2012 **Accepted** 17 August 2012

Address correspondence to Katherine M. Pappas, kmpappas@biol.uoa.gr.

doi:10.1128/JB.01398-12

4. **De Ley J, Swings J.** 1976. Phenotypic description, numerical analysis and a proposal for improved taxonomy and nomenclature of the genus *Zymomonas* Kluyver and van Niel 1936. Int. J. Syst. Bacterol. **26**:146–147.
5. **Ewing B, Green P.** 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. **8**:186–194.
6. **Ewing B, Hillier L, Wendl MC, Green P.** 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. **8**:175–185.
7. **Gordon D, Abajian C, Green P.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8**:195–202.
8. **Han CS, Chain P.** 2006. Finishing repeat regions automatically with Dupfinisher, p 141–146. *In* Arabnia HR, Valafar H (ed), Proceedings of the 2006 International Conference on Bioinformatics and Computational Biology. CSREA Press, Las Vegas, NV.
9. **Kalnenieks U.** 2006. Physiology of *Zymomonas mobilis*: some unanswered questions. Adv. Microb. Physiol. **51**:73–117.
10. **Kalnenieks U, et al.** 2008. NADH dehydrogenase deficiency results in low respiration rate and improved aerobic growth of *Zymomonas mobilis*. Microbiology **154**:989–994.
11. **Margulies M, et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437**:376–380.
12. **Pappas KM, Galani I, Typas MA.** 1997. Transposon mutagenesis and strain construction in *Zymomonas mobilis*. J. Appl. Microbiol. **82**:379–388.
13. **Pappas KM, et al.** 2011. Genome sequence of the ethanol-producing *Zymomonas mobilis* subsp. *mobilis* lectotype ATCC 10988. J. Bacteriol. **193**:5051–5052.
14. **Rogers PL, Jeon YJ, Lee KJ, Lawford HG.** 2007. *Zymomonas mobilis* for fuel ethanol and higher value products. Adv. Biochem. Engin. Biotechnol. **108**:263–288.
15. **Sahm H, Bringer-Meyer S, Sprenger G.** 2006. The genus *Zymomonas*, p 201–221. *In* Dworkin M, Falkow S, Rosenberg E, Schleifer K-H (ed), The prokaryotes: a hand book on the biology of bacteria, 3rd ed. Springer, New York, NY.
16. **Skotnicki ML, Lee KJ, Tribe DE, Rogers PL.** 1981. Comparison of ethanol production by different *Zymomonas* strains. Appl. Environ. Microbiol. **41**:889–893.
17. **Strazdina I, et al.** 2012. Electron transport and oxidative stress in Zymomonas mobilis respiratory mutants. Arch. Microbiol. **194**:461–471.
18. **Swings J, De Ley J.** 1977. The biology of *Zymomonas*. Bacteriol. Rev. **41**:1–46.
19. **Yang S, et al.** 2009. Improved genome annotation for *Zymomonas mobilis*. Nat. Biotechnol. **27**:893–894.
20. **Zerbino D, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.
21. **Zhang Z, Schwartz S, Wagner L, Miller W.** 2000. A greedy algorithm for aligning DNA sequences. J. Comput. Biol. **7**:203–214.