Draft Genome Sequence of *Actinomyces succiniciruminis* Strain Am4[T], Isolated from Cow Rumen Fluid

Susakul Palakawong Na Ayudthaya,[a] Bastian Hornung,[a,b]
Adithi Ravikumar Varadarajan,[b] Wendy Plugge,[b] Caroline M. Plugge[a]

Laboratory of Microbiology, Wageningen University, Wageningen, The Netherlands[a]; Laboratory of Systems and Synthetic Biology, Wageningen University, Wageningen, The Netherlands[b]

ABSTRACT *Actinomyces succiniciruminis* strain Am4[T], isolated from cow rumen fluid, can metabolize a range of substrates including complex carbohydrates to organic acids. Here, we report a 3.33-Mbp draft genome of *Actinomyces succiniciruminis*.

The genus *Actinomyces* is comprised of many species that can be found in various environments. Here, we report the draft genome sequence of a species belonging to the genus *Actinomyces* that originates from a cow rumen. *Actinomyces succiniciruminis* (Am4[T]) was isolated from an enrichment of amylopectin using rumen microorganisms (1).

The genus *Actinomyces* is one of the largest genera within the order and class *Actinobacteria* (2). This genus currently contains 47 species and two subspecies that are Gram-positive, pleomorphic, nonmotile, anaerobic, and aerotolerant bacteria with high G+C content (2). *Actinomyces succiniciruminis* can convert amylopectin, starch, and starch waste to mainly succinate, lactate, and small amounts of acetate and formate (1).

Genomic DNA of strain Am4[T] was extracted from glucose grown cells (1) using the MasterPure complete DNA and RNA purification kit (Epicenter, Madison, WI). Genome sequencing was performed on an Illumina MiSeq sequencer with a read length of 250 bp and an insert size of 500 at GATC-Biotech, Konstanz, Germany. The genome size was first estimated by using kmerspectrumanalyzer (3) on the complete left side of the paired-end data. The quality of the reads was evaluated by FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The reads were trimmed and then the adapters were removed using Trimmomatic, version 0.32 (4). Afterward the assembly was performed using Velvet v1.2.10 (5), with a k-mer value of 41 and a minimum contig size of 500, followed by gapfilling with Gapfiller v1.11 (6). Annotation was carried out with an in-house pipeline, as described in references 7, 8, followed by additional annotation via PRIAM (9), version March 2013. Mapping of reads was performed with BWA v0.7.7 (10). Identification of CAZymes was performed via dbcan v3.0 (11).

The total draft genome consists of 3.33 Mbp and 91 scaffolds with an N_{50} of 69,311 bp and a G+C content of 69.8%, which includes 99.37% of all filtered reads. The genome contains 50 tRNAs genes, one of each rRNA subunit and encodes 2,897 putative proteins, of which 832 were denoted as hypothetical.

Of the 2,897 proteins, 1,745 could be classified according to the COG database (12). Of these 1,745 proteins, 273 are involved in carbohydrate transport and metabolism, followed by 245 with only a predicted function and 180 for amino acid transport and metabolism.

Metabolic mapping with Pathway Tools v17 (13) showed that the organism is capable of degrading glucose, galactose, mannose, and maltose. Furthermore, 173 CAZymes were identified. They enable the organism to degrade complex carbohydrates like starch or glycogen. Multiple enzyme coding genes were presented with

multiple copies, e.g., eight beta-glucosidases, 21 glycosyltransferases type 4, four alpha-L-rhamnosidases, 16 glycosylhydrolases type 13, including alpha-amylases, oligo-1,6-glucosidases, pullulanases, and neopullulanase, showing the broad capacity of *A. succiniciruminis* (Am4^T) to degrade complex carbohydrates.

Accession number(s). The draft genome sequence of *Actinomyces succiniciruminis* strain Am4^T has been deposited at DDBJ/EMBL/GenBank under the accession no. LK995455 to LK995545. The version described in this paper is the first version.

REFERENCES

1. Palakawong NAS, Pristaš P, Hrehová L, Javorský P, Stams AJM, Plugge CM. 2016. *Actinomyces succiniciruminis* sp. nov. and *Actinomyces glycerinitolerans* sp. nov., two novel organic acid-producing bacteria isolated from rumen. Syst Appl Microbiol 39:445–452. https://doi.org/10.1016/j.syapm.2016.08.001.

2. Schaal KP. 1986. Genus *Actinomyces*, p 1383–1418. *In* Sneath PHA, Mair NS, Sharpe ME, Holt JG (ed), Bergey's manual of systematic bacteriology, vol 2. Lippincott Williams & Wilkins, Baltimore.

3. Williams D, Trimble WL, Shilts M, Meyer F, Ochman H. 2013. Rapid quantification of sequence repeats to resolve the size, structure and contents of bacterial genomes. BMC Genomics 14:537. https://doi.org/10.1186/1471-2164-14-537.

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

5. Zerbino DR, Birney E. 2008. Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.

6. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. Genome Biol 13:R56. https://doi.org/10.1186/gb-2012-13-6-r56.

7. Koehorst JJ, Saccenti E, Schaap PJ, Martins dos Santos VAP, Suarez-Diez M. 2016. Protein domain architectures provide a fast, efficient and scalable alternative to sequence-based methods for comparative functional genomics. F1000Research 5:1987. https://doi.org/10.12688/f1000research.9416.2.

8. Koehorst JJ, van Dam JCJ, van Heck RGA, Saccenti E, Martins dos Santos VAP, Suarez-Diez M, Schaap PJ. 2016. High throughput functional comparison of 432 genome sequences of pseudomonas using a semantic data framework. Sci Rep, in press.

9. Claudel-Renard C, Chevalet C, Faraut T, Kahn D. 2003. Enzyme-specific profiles for genome annotation—PRIAM. Nucleic Acids Res 31:6633–6639. https://doi.org/10.1093/nar/gkg847.

10. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

11. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res 40:W445–W451. https://doi.org/10.1093/nar/gks479.

12. Tatusov RL, Galperin MY, Natale DA, Koonin EV. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28:33–36. https://doi.org/10.1093/nar/28.1.33.

13. Karp PD, Paley S, Romero P. 2002. The Pathway Tools software. Bioinformatics 18:S225–S232. https://doi.org/10.1093/bioinformatics/18.suppl_1.S225.