Check for updates

Draft Genome Sequence of KCTC 12630, the Type Strain of the Novel Species *Sphingomonas ginsengisoli*

Yan-Qiong Li,[a] Hui Zhang,[a] Manik Prabhu Narsing Rao,[b] Zhou-Yan Dong,[b] Dalal Hussien M. Alkhalifah,[c] Wael N. Hozzein,[d] Min Xiao,[b] Wen-Jun Li[b]

[a]Kunming Medical University Haiyuan College, Kunming, People's Republic of China
[b]State Key Laboratory of Biocontrol and Guangdong Provincial Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-Sen University, Guangzhou, People's Republic of China
[c]Biology Department, Faculty of Science, Princess Nourah bint Abdulrahman University, Riyadh, Kingdom of Saudi Arabia
[d]Zoology Department, College of Science, King Saud University, Riyadh, Kingdom of Saudi Arabia

ABSTRACT We report here the draft genome sequence of *Sphingomonas ginsengisoli* KCTC 12630^T. The draft genome has a size of 3,045,889 bp and a G+C content of 67.1%. The availability of the genome sequence will provide a better understanding of strain KCTC 12630^T and the genus *Sphingomonas*.

The genus *Sphingomonas* was first described by Yabuuchi et al. (1), and at the time of writing, the genus comprised 126 recognized species (http://www.bacterio.net/sphingomonas.html) (2). The genus *Sphingomonas* is very important for biotechnological applications due to the ability of the members to produce several bioactive molecules (3–5).

Sphingomonas ginsengisoli was reported by An et al. (6) and was isolated from soil collected in a ginseng field in Pocheon Province, South Korea. Here, we report the draft genome sequence of KCTC 12630, the type strain of the novel species *Sphingomonas ginsengisoli*, to provide a better understanding of this strain and the genus *Sphingomonas.* Strain KCTC 12630^T was cultured on Reasoner's 2A agar incubated at 30°C. The biomass obtained after 4 days of incubation was collected. The genomic DNA from the biomass for whole-genome sequencing (WGS) was extracted employing the sodium dodecyl sulfate-cetyltrimethylammonium bromide (SDS-CTAB) method (7). Sequencing libraries were generated using the NEBNext Ultra DNA library prep kit for Illumina (New England Biolabs, USA) following manufacturer recommendations. The library quality was assessed on the Qubit version 3.0 fluorometer (Bio-Rad) and Agilent 4200 system. Whole-genome sequencing of strain KCTC 12630^T was performed by sequencing paired-end reads on a HiSeq X platform (Illumina, San Diego, CA) at Magigene Company (Guangzhou, People's Republic of China). The quality of the sequence reads (19,885,796 raw reads and 2,982,869,400 raw bases) was analyzed using the FastQC tool version 0.11.5 (8).

The raw data were filtered (1,923,2872 clean reads and 2,817,875,153 clean bases) using the Fastx-toolkit (9). All good-quality paired reads were assembled using SPAdes version 3.6.0 (10). The draft genome size of strain KCTC 12630^T is 3,045,889 bp with an N_{50} scaffold length of 1,581,088 bp. The G+C content is 67.1%. Gene prediction was carried out using Glimmer version 3.02 (11). The identification of tRNAs and rRNAs was carried out using tRNAscan-SE version 1.23 (12) and RNAmmer version 1.2 (13). A total of 3,067 genes were predicted with 3 rRNA and 45 tRNA genes. The functional annotation of the genome was performed with Rapid Annotations using Subsystems Technology (RAST) (14), which revealed 182 genes for carbohydrate metabolism and 235 genes involved in protein metabolism.

Address correspondence to Wen-Jun Li, liwenjun3@mail.sysu.edu.cn.

ACKNOWLEDGMENT

REFERENCES

1. Yabuuchi E, Yano I, Oyaizu H, Hashimoto Y, Ezaki T, Yamamoto H. 1990. Proposals of *Sphingomonas paucimobilis* gen. nov. and comb. nov., *Sphingomonas parapaucimobilis* sp. nov., *Sphingomonas yanoikuyae* sp. nov., *Sphingomonas adhaesiva* sp. nov., *Sphingomonas capsulata* comb. nov., and two genospecies of the genus *Sphingomonas*. Microbiol Immunol 34:99–119. https://doi.org/10.1111/j.1348-0421.1990.tb00996.x.

2. Parte AC. 2014. LPSN—list of prokaryotic names with standing in nomenclature. Nucleic Acids Res 42:D613–D616. https://doi.org/10.1093/nar/gkt1111.

3. Stolz A, Schmidt-Maag C, Denner E, Busse HJ, Egli T, Kampfer P. 2000. Description of *Sphingomonas xenophaga* sp. nov. for strains BN6ᵀ and N,N which degrade xenobiotic aromatic compounds. Int J Syst Evol Microbiol 50:35–41. https://doi.org/10.1099/00207713-50-1-35.

4. Kim JH, Kim SH, Kim KH, Lee PC. 2015. *Sphingomonas lacus* sp. nov., an astaxanthin-dideoxyglycoside-producing species isolated from soil near a pond. Int J Syst Evol Microbiol 65:2824–2830. https://doi.org/10.1099/ijs.0.000337.

5. Xie CH, Yokota A. 2006. *Sphingomonas azotifigens* sp. nov., a nitrogen-fixing bacterium isolated from the roots of *Oryzasativa*. Int J Syst Evol Microbiol 56:889–893. https://doi.org/10.1099/ijs.0.64056-0.

6. An DS, Liu QM, Lee HG, Jung MS, Kim SC, Lee ST, Im WT. 2013. *Sphingomonas ginsengisoli* sp. nov. and *Sphingomonas sediminicola* sp. nov. Int J Syst Evol Microbiol 63:496–501. https://doi.org/10.1099/ijs.0.039578-0.

7. Moller EM, Bahnweg G, Sandermann H, Geiger HH. 1992. A simple and efficient protocol for isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant tissues. Nucleic Acids Res 20:6115–6116. https://doi.org/10.1093/nar/20.22.6115.

8. Babraham Bioinformatics. 2010. FastQC: a quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

9. Gordon A, Hannon G. 2010. Fastx-Toolkit. FASTQ: a short-reads pre-processing tools. http://hannonlab.cshl.edu/fastx_toolkit/.

10. 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. https://doi.org/10.1089/cmb.2012.0021.

11. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics 23:673–679. https://doi.org/10.1093/bioinformatics/btm009.

12. 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. https://doi.org/10.1093/nar/25.5.955.

13. 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. https://doi.org/10.1093/nar/gkm160.

14. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.