Draft Genome Sequence of *Lactarius deliciosus* Strain CBS 582.63 from the CBS-KNAW Culture Collection

Jia Ean Goh,[a] Ahmad Yamin Rahman,[b] Ranjeev Hari,[c] Mei Peng Lim,[d] Nazalan Najimudin,[b] Wai Sum Yap,[a] Yong Hui Tan,[a] Baskaran Gunasekaran,[a] Hok Chai Yam[a]

[a]Faculty of Applied Sciences, UCSI University, Kuala Lumpur, Malaysia
[b]School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang, Malaysia
[c]School of Data Sciences, Perdana University, Serdang, Malaysia
[d]Robarts Research Institute, University of Western Ontario, London, Canada

Jia Ean Goh and Ahmad Yamin Rahman contributed equally to this work. Author order was determined by drawing straws.

ABSTRACT A type strain of *Lactarius deliciosus* was obtained from the CBS-KNAW culture collection. The mycelium was cultured using potato dextrose agar, and the extracted genomic DNA was subjected to PacBio genome sequencing. Upon assembly and annotation, the genome size was estimated to be 54 Mbp, with 12,753 genes.

Edible *Lactarius* mushrooms, also known as milk cap mushrooms, are commonly found in the Northern Hemisphere. They produce natural rubber (high-molecular-weight *cis*-polyisoprene), as do some plants (1, 2). *Lactarius deliciosus*, commonly known as the saffron milk cap mushroom, belongs to the family *Russulaceae*, of the *Basidiomycota* subdivision of *Fungi*.

Mushroom strain CBS 582.63 was purchased from the CBS-KNAW culture collection (Netherlands). *L. deliciosus* was cultivated in potato dextrose agar (PDA) medium as described by Hall and Wang (3). The mycelium was cultured at room temperature (27°C) under dark conditions until enough mass was obtained.

Genomic DNA (gDNA) was isolated by using Genomic-tips with protocols recommended by the manufacturer (Qiagen, Germany). The isolated gDNA sample was subjected to high-throughput PacBio Sequel sequencing. The sequencing used the v2 chemistry provided by PacBio utilizing 4 single-molecule real-time (SMRT) cells. A 20-kb library with a 5-kb cutoff was prepared using the protocols of the SMRTbell template prep kit v1.0 according to the manufacturer's recommendations (PacBio, Menlo Park, CA). Approximately 3 Gb of sequence data was generated from the sequencing, and a sequence coverage of about 61× was obtained. The SMRT sequencing yielded a mean read length of 7,965 bp and 386,212 total reads.

The raw sequences were assembled by using fast mapping, error correction, and the *de novo* assembly tool MECAT v1.2 (4), followed by quality assessment by QUAST v4.6 (5). Default parameters were used for all software unless otherwise stated. Genome assembly produced a total length of 54 Mbp with 841 contigs, an N_{50} value of 116,430 bp, and a GC content of approximately 51.98%. Fast alignment-free computation of whole-genome average nucleotide identity was performed using FastANI v1.3 (6), and the results indicated that the assembled genome sequence showed the highest identity with the genome sequences of *Lactarius deliciosus* strains deposited in the Joint Genome Institute (JGI).

Annotation of the assembled genomic sequences of *L. deliciosus* was performed with MAKER2 v2.31 (7). This tool runs independent gene callers and creates consensus predictions based on collective evidence. Three separate gene predictors, AUGUSTUS

v2.5.5 (8), SNAP v2013-11-29 (9), and GeneMark-ES v2.3 (10), were also used to predict evidence. *L. deliciosus* expressed sequence tag (EST) data from the Joint Genome Institute (JGI) were used as transcriptomic data, while the protein sequence was used as annotation reference data (11). The annotation resulted in 12,753 genes.

Data availability. The genome of *Lactarius deliciosus* CBS 582.63 has been deposited in DDBJ/ENA/GenBank under the accession number VHKJ00000000. This version of the project can be found under accession number VHKJ01000000. The raw data have been deposited in the SRA under accession number PRJNA550006.

REFERENCES

1. Tanaka Y, Kawahara S, Eng AH, Takei A, Ohya N. 1994. Structure of cis-polyisoprene from Lactarius mushrooms. Acta Biochim Pol 41:303–309. https://doi.org/10.18388/abp.1994_4719.
2. Cornish K, Xie W. 2012. Natural rubber biosynthesis in plants: rubber transferase. Methods Enzymol 515:63–82. https://doi.org/10.1016/B978-0-12-394290-6.00004-5.
3. Hall IR, Wang Y. 1998. Methods for cultivating edible ectomycorrhizal mushrooms, p 99–114. *In* Varma A (ed), Mycorrhiza manual. Springer, Berlin, Germany.
4. Xiao C-L, Chen Y, Xie S-Q, Chen K-N, Wang Y, Han Y, Luo F, Xie Z. 2017. MECAT: fast mapping, error correction, and de novo assembly for single-molecule sequencing reads. Nat Methods 14:1072–1074. https://doi.org/10.1038/nmeth.4432.
5. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
6. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5514. https://doi.org/10.1038/s41467-018-07641-9.
7. Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res 18:188–196. https://doi.org/10.1101/gr.6743907.
8. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res 34:W435–W439. https://doi.org/10.1093/nar/gkl200.
9. Korf I. 2004. Gene finding in novel genomes. BMC Bioinformatics 5:59. https://doi.org/10.1186/1471-2105-5-59.
10. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res 18:1979–1990. https://doi.org/10.1101/gr.081612.108.
11. Martin F, Tang N. 2018. Lactarius deliciosus 48 v1.0. JGI MycoCosm. https://mycocosm.jgi.doe.gov/Lacdel1/Lacdel1.home.html.