

Draft genome sequence of *Rhodotorula* sp. BUB8, an oleaginous yeast isolated from Mt. Makiling Forest Reserve, Laguna, Philippines

Irene G. Pajares,¹ Kristine Rose M. Ramos,¹ Andrew D. Montecillo,² Francisco B. Elegado,¹ Asuncion K. Raymundo²

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT This paper reports the draft genome sequence of *Rhodotorula* sp. BUB8, an oleaginous yeast isolated from the forest canopy of Mt. Makiling Forest Reserve in the Philippines. The draft genome is 20,394,133 bp with 64.3% guanine-cytosine (GC) content, 6,604 coding genes, 109 tRNA coding, and 6 snRNA-coding genes.

KEYWORDS oleaginous yeast, biofuel, biodiesel

Oleaginous yeasts are being explored for their potential to provide unique platforms for the sustainable production of fatty acid-derived oleochemicals due to their capability to synthesize and accumulate lipids mainly as triacylglycerides (1). *Rhodotorula* sp. BUB8 can produce significant amounts of microbial lipids from glucose and glycerol and share an almost similar fatty acid profile to vegetable oils (1).

Rhodotorula sp. BUB8 was isolated from the upper bark of a Bagtikan tree (*Parashorea malaanonan*) found in the Mt. Makiling Forest Reserve (2). The strain was cultured on yeast extract-peptone-dextrose (YEPD) media for 24 h at 30°C and was deposited to the Philippine Culture Collection of Microorganisms at the National Institute of Molecular Biology and Biotechnology, University of the Philippines Los Baños, with accession number 2345. Using the Basic Local Alignment Search Tool of the National Center for Biotechnology Information database, the ITS-5.8s rRNA gene (GenBank accession number [OP808033.1](#)) of BUB8 displayed 99% identity with *Rhodotorula paludigena* (2).

The genomic DNA of BUB8 was extracted from a 10-mL YEPD overnight culture using Quick DNA Fungal/Bacterial DNA Extraction Kit according to the manufacturer's instructions (Zymo Research, Irvine, CA). Sequencing was conducted by Novogene Co., Ltd. (China), with the DNA library prepared using a Novogene NGS DNA library prep set (catalog no. Pt004), in which the DNA was randomly sheared into short fragments, end-repaired, A-tailed, then ligated with an Illumina adapter. The fragments with adapters were PCR amplified, size selected, and purified with solid-phase reversible immobilization beads (SPRI) beads. The library was checked with Qubit and quantified by real-time PCR. Fragment size distribution was determined using a bioanalyzer. Quantified libraries were pooled and sequenced on an Illumina MiSeq (Novogene Co., Ltd.). Reads were assessed using FastQC, and reads with low-quality nucleotides (Q value ≤38) exceeding 40 bp, containing N nucleotides exceeding 10 bp, and whose overlap with the adapter exceeded 15 bp by default were eliminated. The assembly was performed using SOAPdenovo (v.2.04) (3–7). Genome size was estimated by K-mer analysis using [K-mer Analysis Toolkit](#) (8), and the assembly quality was assessed using QUAST (Galaxy v.5.0.2 + Galaxy 5) (9). Contigs with less than 500 bp were removed. Default parameters were used with all tools unless otherwise specified.

A total of 13,788,111 and 12,998,269 of raw and filtered paired-end reads were generated, respectively. The reads assembled into 54 scaffolds with a total length

Editor Jason E. Stajich, University of California
Riverside, Riverside, California, USA

Address correspondence to Irene G. Pajares,
igpajares@up.edu.ph.

The authors declare no conflict of interest.

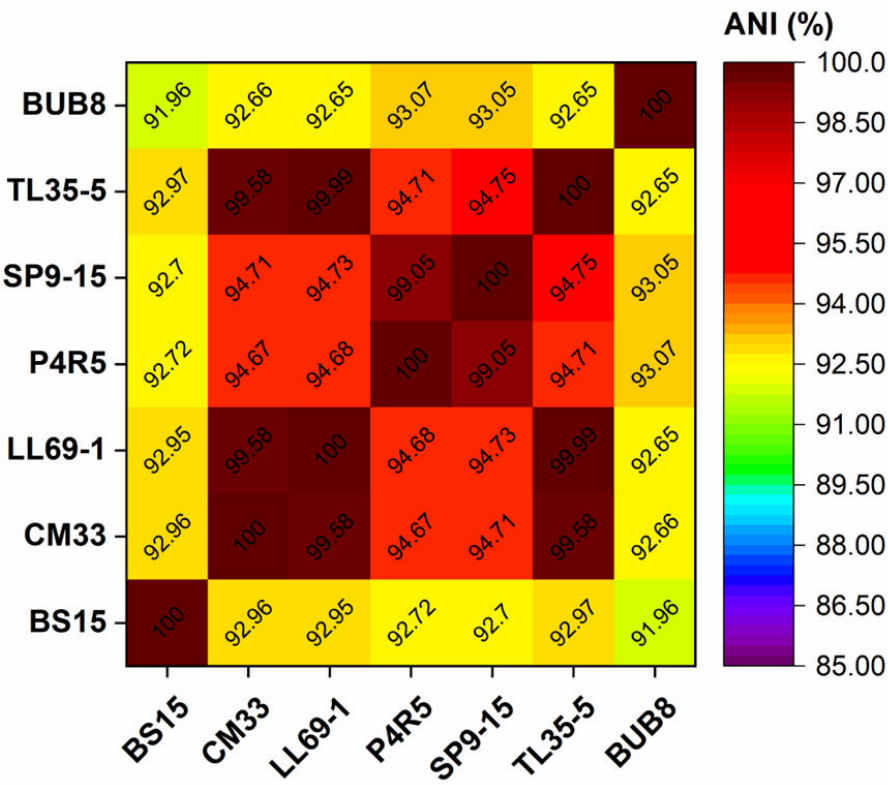
Received 11 April 2024

Accepted 12 August 2024

Published 30 August 2024

Copyright © 2024 Pajares et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

A



B

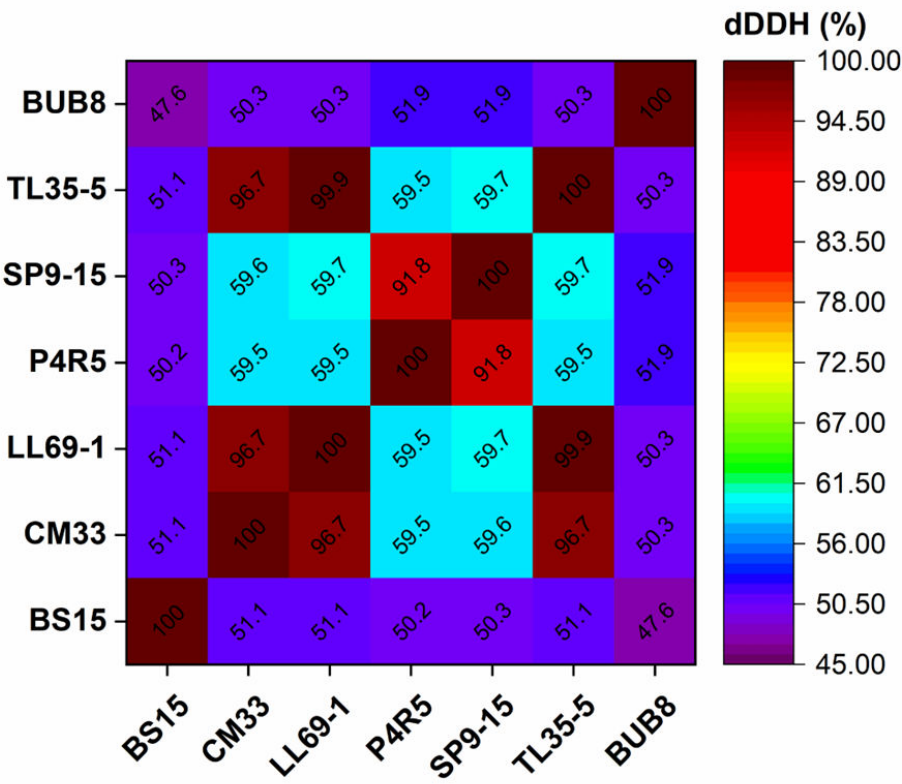


FIG 1 Comparative genomics average nucleotide identity (ANI) (A) and digital DNA-DNA hybridization (dDDH) (B) analyses of BUB8 with related *R. paludigena* strains. Genome accession numbers for *R. paludigena* strains: BS15 ([GCA_036244875.1](#)), CM33 ([GCA_005281665.1](#)), LL69-1 ([GCA_033458335.1](#)), P4R5 ([GCA_019671115.1](#)), SP9-15 ([GCA_029873615.1](#)), and TL35-5 ([GCA_028828615.1](#)).

of 20,394,133 bp, N50 of 536,656 bp, GC content of 64.3%, and an estimated mean sequencing depth of 162×. The nearest neighbor identity of BUB8 genome based on genome-relatedness indices (i.e., average nucleotide identity and digital DNA-DNA hybridization) was identified using the Enveomics Average Nucleotide Identity Calculator (<http://enve-omics.ce.gatech.edu/ani/>) and ANI-independent Genome-to-Genome Distance Calculator (v.3.0) online tools (<https://ggdc.dsmz.de/ggdc.php#>) (10, 11). BUB8 was most closely related to P4R5 (GCA_019671115.1) with 93.07% and 51.9% identities, respectively, among *R. paludigena* strains (12–15). These values are below the generally accepted cut-off for species boundary of 95% and 70%, suggesting BUB8 is a novel species of *Rhodotorula* (Fig. 1).

ACKNOWLEDGMENTS

The authors thank the Forest Canopy Observation Positioning and Investigation Program.

This project was funded by the University of the Philippines Los Baños under the Office of the Vice-Chancellor for Research and Extension (project ID number 16214).

AUTHOR AFFILIATIONS

¹National Institute of Molecular Biology and Biotechnology, University of the Philippines Los Baños, Los Baños, Laguna, Philippines

²Institute of Biological Sciences, University of the Philippines Los Baños, Los Baños, Laguna, Philippines

AUTHOR ORCIDS

Irene G. Pajares  <http://orcid.org/0000-0002-9119-8969>

Kristine Rose M. Ramos  <http://orcid.org/0000-0002-5978-2604>

Andrew D. Montecillo  <http://orcid.org/0000-0003-0627-1920>

DATA AVAILABILITY

The draft genome sequence of *Rhodotorula* sp. BUB8 has been deposited in GenBank under accession number JAZHGP000000000, BioProject number PRJNA1051797, and Sequence Read Archive accession number SRX24721207.

REFERENCES

1. Ratledge C. 2004. Fatty acid biosynthesis in microorganisms being used for single cell oil production. *Biochimie* 86:807–815. <https://doi.org/10.1016/j.biochi.2004.09.017>
2. Pajares IG, Requiso PJ, Fabro Jr LM, Ramos KRM, Raymundo AK. 2024. Biodiesel production of oleaginous yeast isolated from the Mount Makiling Forest Reserve. *J App Biol Biotech* 12:67–75. <https://doi.org/10.7324/JABB.2024.153703>
3. Wingett SW, Andrews S. 2018. FastQ screen: a tool for multi-genome mapping and quality control. *F1000Res* 7:1338. <https://doi.org/10.12688/f1000research.15931.2>
4. Li R, Li Y, Kristiansen K, Wang J. 2008. SOAP: short oligonucleotide alignment program. *J Bioinform* 24:713–714. <https://doi.org/10.1093/bioinformatics/btn025>
5. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>
6. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. *Genome Res* 19:1117–1123. <https://doi.org/10.1101/gr.089532.108>
7. Lin S-H, Liao Y-C. 2013. CISA: contig integrator for sequence assembly of bacterial genomes. *PLoS One* 8:e60843. <https://doi.org/10.1371/journal.pone.0060843>
8. Mapleson D, Garcia Accinelli G, Kettleborough G, Wright J, Clavijo BJ. 2017. KAT: a K-mer analysis toolkit to quality control NGS datasets and genome assemblies. *Bioinformatics* 33:574–576. <https://doi.org/10.1093/bioinformatics/btt086>
9. 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>
10. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. 2013. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. <https://doi.org/10.1186/1471-2105-14-60>
11. Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M. 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 50:D801–D807. <https://doi.org/10.1093/nar/gkab902>
12. Wang M, Mao W, Wang X, Li F, Wang J, Chi Z, Chi Z, Liu G. 2019. Efficient simultaneous production of extracellular polyol esters of fatty acids and intracellular lipids from inulin by a deep-sea yeast *Rhodotorula paludigena* P4R5. *Microb Cell Fact* 18:149. <https://doi.org/10.1186/s12934-019-1200-3>

13. Mano J, Sushida H, Tanaka T, Naito K, Ono H, Ike M, Tokuyasu K, Kitaoka M. 2023. Extracellular oil production by *Rhodotorula paludigena* BS15 for biorefinery without complex downstream processes. *Appl Microbiol Biotechnol* 107:6799–6809. <https://doi.org/10.1007/s00253-023-12762-3>
14. Gosalawit C, Imsoonthornruksa S, Udomsil N, Ketudat-Cairns M. 2020. Genome sequence of the oleaginous yeast *Rhodotorula paludigena* strain CM33, a potential strain for biofuel production. *Microbiol Resour Announc* 9:e00286–20. <https://doi.org/10.1128/MRA.00286-20>
15. Phuengjayaem S, Kingkaew E, Hoondet P, Rojsitthisak P, Sritularak B, Thitikornpong W, Thompho S, Pornputtapong N, Tanasupawat S. 2023. Diversity, astaxanthin production, and genomic analysis of *Rhodotorula paludigena* SP9-15. *Heliyon* 9:e18280. <https://doi.org/10.1016/j.heliyon.2023.e18280>