Check for updates

Near-Complete Genome Sequence of *Zygosaccharomyces rouxii* NRRL Y-64007, a Yeast Capable of Growing on Lignocellulosic Hydrolysates

Sujit Sadashiv Jagtap,[a,b] Jing-Jing Liu,[a,b] Hanna E. Walukiewicz,[a,b] Jasmyn Pangilinan,[c] Anna Lipzen,[c] Steven Ahrendt,[c] Maxim Koriabine,[c] Kelly Cobaugh,[c] Asaf Salamov,[c] Yuko Yoshinaga,[c] Vivian Ng,[c] Chris Daum,[c] Igor V. Grigoriev,[c,d] Patricia J. Slininger,[e] Bruce S. Dien,[e] Yong-Su Jin,[b,f] ![ORCID]Christopher V. Rao[a,b]

[a]Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA
[b]DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA
[c]U.S. Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, California, USA
[d]Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, California, USA
[e]Bioenergy Research Unit, National Center for Agricultural Utilization Research, USDA-Agricultural Research Service, Peoria, Illinois, USA
[f]Department of Food Science and Nutrition, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

ABSTRACT The halotolerant and osmotolerant yeast *Zygosaccharomyces rouxii* can produce multiple volatile compounds and has the ability to grow on lignocellulosic hydrolysates. We report the annotated genome sequence of *Z. rouxii* NRRL Y-64007 to support its development as a platform organism for biofuel and bioproduct production.

*Z*ygosaccharomyces rouxii NRRL Y-64007 is a halotolerant, osmotolerant, acidophilic, and fructophilic yeast (1–3). It is used for the production of volatile compounds, organic acids, lipids, and sugar alcohols (4–10). This yeast was initially called *Lipomyces starkeyi* Y-11557 but was renamed *Z. rouxii* NRRL Y-64007 based on the internal transcribed spacer (ITS) sequence (GenBank accession number OM905798) and genome sequencing results (9). In this study, we sequenced the genome and transcriptome of *Z. rouxii* NRRL Y-64007 to facilitate further investigation of its physiology, metabolism, and metabolic engineering to produce biofuels and bioproducts.

YPG medium (10 g/L yeast extract, 20 g/L peptone, and 20 g/L glucose) was used for the growth of *Z. rouxii* NRRL Y-64007. Cell cultures with a total optical density at 600 nm (OD_{600}) of 10 were used for genomic DNA and RNA extraction. The genomic DNA and total RNA of *Z. rouxii* NRRL Y-64007 were extracted using the Dr. GenTLE (from yeast) high recovery kit (TaKaRa Bio Inc., Shiga, Japan) and the RNeasy minikit (Qiagen, Hilden, Germany), respectively (11, 12).

For genome sequencing, libraries of >10 kb were prepared for Pacific Biosciences (PacBio) sequencing using 5 μg of genomic DNA as reported previously and according to the PacBio template preparation and sequencing guide (13, 14). The sheared DNA was treated with exonuclease, followed by end repair and ligation of blunt adapters using the SMRTbell template preparation kit v1.0. The library was purified with AMPure PB beads with a 10-kb cutoff value. The prepared SMRTbell template libraries were then sequenced on a PacBio Sequel II sequencer using v3 sequencing primer, 8M v1 single-molecule real-time (SMRT) cells, and v2.0 sequencing chemistry with 1 × 900-bp sequencing movie run times. BBduk and BBMerge from BBTools v36.63 (https://sourceforge.net/projects/bbmap) were used to remove reads aligning to human, cat, dog, mouse, and common microbial contaminants (15).

Plate-based RNA sample preparation was performed with the PerkinElmer Sciclone next-generation sequencing (NGS) robotic liquid handling system using the Illumina

TruSeq stranded mRNA high-throughput (HT) sample preparation kit with poly(A) selection of mRNA, following the protocol outlined by Illumina in the user guide (16); 1 μg RNA per sample and eight cycles of PCR were used for library amplification. The prepared libraries were quantified using the KAPA Biosystems NGS library quantitative PCR (qPCR) kit and run on a Roche LightCycler 480 real-time PCR instrument. Sequencing was performed using NovaSeq XP v1 reagent kits and an S4 flow cell. An Illumina library was constructed and sequenced (2 $\times$ 151 bp) using the Illumina NovaSeq S4 system. BBDuk v38.79 (http://bbtools.jgi.doe.gov) was used to remove contaminants, to trim reads that contained adapter sequences, and to quality trim reads with quality scores of <6. Reads mapped with BBMap to common contaminants and rRNA reads were removed. Filtered fastq files were used as input for *de novo* assembly of RNA contigs. Reads filtered and trimmed for quality and contamination were assembled into consensus sequences using Trinity v2.8.5 (17).

The 9.95-Mb genome assembly contained 8 contigs (N_{50}, 1.53 Mb), with sequencing read coverage depth of 181.58$\times$ and a GC content of 39.12%. The genome was annotated using the JGI annotation pipeline (18, 19) to predict 5,001 protein-coding genes. A noncanonical telomere consensus sequence was checked to assess the near-complete genome; it was present at both the start and the end of the first 4 scaffolds and only at the end of scaffolds 5, 6, 7, and 8.

Data availability. The whole-genome assembly and annotation are available from MycoCosm (https://mycocosm.jgi.doe.gov/Zygrou1) (17). This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession number JAKETS000000000. The version described in this paper is JAKETS010000000. The accession numbers for the BioProject and reads are PRJNA784295 and SRR17438072, respectively.

REFERENCES

1. Leandro MJ, Sychrová H, Prista C, Loureiro-Dias MC. 2011. The osmotolerant fructophilic yeast *Zygosaccharomyces rouxii* employs two plasma-membrane fructose uptake systems belonging to a new family of yeast sugar transporters. Microbiology (Reading) 157:601–608. https://doi.org/10.1099/mic.0.044446-0.

2. Dakal TC, Solieri L, Giudici P. 2014. Adaptive response and tolerance to sugar and salt stress in the food yeast *Zygosaccharomyces rouxii*. Int J Food Microbiol 185:140–157. https://doi.org/10.1016/j.ijfoodmicro.2014.05.015.

3. Martorell P, Stratford M, Steels H, Fernández-Espinar MT, Querol A. 2007. Physiological characterization of spoilage strains of *Zygosaccharomyces bailii* and *Zygosaccharomyces rouxii* isolated from high sugar environments. Int J Food Microbiol 114:234–242. https://doi.org/10.1016/j.ijfoodmicro.2006.09.014.

4. De Francesco G, Turchetti B, Sileoni V, Marconi O, Perretti G. 2015. Screening of new strains of *Saccharomycodes ludwigii* and *Zygosaccharomyces rouxii* to produce low-alcohol beer. J Inst Brew 121:113–121. https://doi.org/10.1002/jib.185.

5. Groleau D, Chevalier P, Yuen T. 1995. Production of polyols and ethanol by the osmophilic yeast *Zygosaccharomyces rouxii*. Biotechnol Lett 17:315–320. https://doi.org/10.1007/BF01190645.

6. Jansen M, Veurink JH, Euverink G-JW, Dijkhuizen L. 2003. Growth of the salt-tolerant yeast *Zygosaccharomyces rouxii* in microtiter plates: effects of NaCl, pH and temperature on growth and fusel alcohol production from branched-chain amino acids. FEMS Yeast Res 3:313–318. https://doi.org/10.1111/j.1567-1364.2003.tb00172.x.

7. Kashyap P, Sabu A, Pandey A, Szakacs G, Soccol CR. 2002. Extra-cellular L-glutaminase production by *Zygosaccharomyces rouxii* under solid-state fermentation. Process Biochem 38:307–312. https://doi.org/10.1016/S0032-9592(02)00060-2.

8. Saha BC, Sakakibara Y, Cotta MA. 2007. Production of D-arabitol by a newly isolated *Zygosaccharomyces rouxii*. J Ind Microbiol Biotechnol 34:519–523. https://doi.org/10.1007/s10295-007-0211-y.

9. Slininger PJ, Dien BS, Kurtzman CP, Moser BR, Bakota EL, Thompson SR, O'Bryan PJ, Cotta MA, Balan V, Jin M, Sousa LDC, Dale BE. 2016. Comparative lipid production by oleaginous yeasts in hydrolyzates of lignocellulosic biomass and process strategy for high titers. Biotechnol Bioeng 113:1676–1690. https://doi.org/10.1002/bit.25928.

10. Taing O, Taing K. 2006. Production of malic and succinic acids by sugar-tolerant yeast *Zygosaccharomyces rouxii*. Eur Food Res Technol 224:343–347. https://doi.org/10.1007/s00217-006-0323-z.

11. Jagtap SS, Bedekar AA, Singh V, Jin Y-S, Rao CV. 2021. Metabolic engineering of the oleaginous yeast *Yarrowia lipolytica* PO1f for production of erythritol from glycerol. Biotechnol Biofuels 14:188. https://doi.org/10.1186/s13068-021-02039-0.

12. Jagtap SS, Deewan A, Liu J-J, Walukiewicz HE, Yun EJ, Jin Y-S, Rao CV. 2021. Integrating transcriptomic and metabolomic analysis of the oleaginous yeast *Rhodosporidium toruloides* IFO0880 during growth under different

carbon sources. Appl Microbiol Biotechnol 105:7411–7425. https://doi.org/10.1007/s00253-021-11549-8.

13. Kim KE, Peluso P, Babayan P, Yeadon PJ, Yu C, Fisher WW, Chin C-S, Rapicavoli NA, Rank DR, Li J, Catcheside DEA, Celniker SE, Phillippy AM, Bergman CM, Landolin JM. 2014. Long-read, whole-genome shotgun sequence data for five model organisms. Sci Data 1:140045. https://doi.org/10.1038/sdata.2014.45.

14. Travers KJ, Chin C-S, Rank DR, Eid JS, Turner SW. 2010. A flexible and efficient template format for circular consensus sequencing and SNP detection. Nucleic Acids Res 38:e159. https://doi.org/10.1093/nar/gkq543.

15. Bushnell B, Rood J, Singer E. 2017. BBMerge: accurate paired shotgun read merging via overlap. PLoS One 12:e0185056. https://doi.org/10.1371/journal.pone.0185056.

16. Palomares M-A, Dalmasso C, Bonnet E, Derbois C, Brohard-Julien S, Ambroise C, Battail C, Deleuze J-F, Olaso R. 2019. Systematic analysis of TruSeq, SMARTer and SMARTer Ultra-Low RNA-seq kits for standard, low and ultra-low quantity samples. Sci Rep 9:7550. https://doi.org/10.1038/s41598-019-43983-0.

17. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I. 2014. MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res 42:D699–D704. https://doi.org/10.1093/nar/gkt1183.

18. Génolevures Consortium, Souciet J-L, Dujon B, Gaillardin C, Johnston M, Baret PV, Cliften P, Sherman DJ, Weissenbach J, Westhof E, Wincker P, Jubin C, Poulain J, Barbe V, Ségurens B, Artiguenave F, Anthouard V, Vacherie B, Val M-E, Fulton RS, Minx P, Wilson R, Durrens P, Jean G, Marck C, Martin T, Nikolski M, Rolland T, Seret M-L, Casarégola S, Despons L, Fairhead C, Fischer G, Lafontaine I, Leh V, Lemaire M, de Montigny J, Neuvéglise C, Thierry A, Blanc-Lenfle I, Bleykasten C, Diffels J, Fritsch E, Frangeul L, Goëffon A, Jauniaux N, Kachouri-Lafond R, Payen C, Potier S, Pribylova L, Ozanne C, Richard GF, Sacerdot C, Straub ML, Talla E. 2009. Comparative genomics of protoploid *Saccharomycetaceae*. Genome Res 19:1696–1709. https://doi.org/10.1101/gr.091546.109.

19. Pryszcz LP, Gabaldón T. 2016. Redundans: an assembly pipeline for highly heterozygous genomes. Nucleic Acids Res 44:e113. https://doi.org/10.1093/nar/gkw294.