

Genome Sequence of the Oleaginous Yeast *Rhodotorula glutinis* ATCC 204091

Debarati Paul,^a Zenaida Magbanua,^b Mark Arick II,^b Todd French,^c Susan M. Bridges,^d Shane C. Burgess,^e Mark L. Lawrence^f

Amity Institute of Biotechnology, Amity University, AUUP, Noida, India^a; Institute for Genomics, Biocomputing, and Biotechnology, Mississippi State University, Mississippi State, Mississippi, USA^b; Dave C. Swalm School of Chemical Engineering, Mississippi State University, Mississippi State, Mississippi, USA^c; Department of Computer Science and Engineering, Mississippi State University, Mississippi State, Mississippi, USA^d; College of Agriculture and Life Sciences, University of Arizona, Tucson, Arizona, USA^e; College of Veterinary Medicine, Mississippi State University, Mississippi State, Mississippi, USA^f

***Rhodotorula glutinis* ATCC 204091 is an oleaginous oxidative red yeast that can accumulate lipids to >50% of its biomass when grown with appropriate carbon and nitrogen ratios. It produces a red pigment consisting of useful antioxidants, such as carotenoids, torulene, and torularhodin, when cultivated under carbon-deficient conditions.**

Received 17 January 2014 Accepted 24 January 2014 Published 13 February 2014

Citation Paul D, Magbanua Z, Arick M, II, French T, Bridges SM, Burgess SC, Lawrence ML. 2014. Genome sequence of the oleaginous yeast *Rhodotorula glutinis* ATCC 204091. *Genome Announc.* 2(1):e00046-14. doi:10.1128/genomeA.00046-14.

Copyright © 2014 Paul et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Mark L. Lawrence, lawrence@cvm.msstate.edu.

Rhodotorula glutinis is an oleaginous oxidative red yeast that potentially can be used for biodiesel production as well as other important products, such as carotenoids (torulene, torularhodin, γ -carotene, and β -carotene) when grown on suitable culture medium containing appropriate carbon and nitrogen ratios (1–3). Historically, *R. glutinis* has been used to identify a yeast pathway for L-lysine biosynthesis via L-pipecolic acid (4). *Rhodotorula* secretes α -L-arabinofuranosidase, which is important to the wine industry for aroma of the wine. *R. glutinis* is nonfastidious, making it attractive for the production of phenylalanine ammonia lyase (PAL) (EC 4.3.1.5), which is a nonhydrolytic enzyme produced in high quantities by *R. glutinis* (5). Here, we report the *R. glutinis* genome sequence, which can be used to study the regulation of these important pathways and to explore genes encoding other useful products.

Genomic DNA was isolated using the Wizard genomic DNA purification kit (Promega) with overnight lyticase treatment (20 mg/ml lyticase [catalog no. L2524; Sigma]) to lyse the cells. The purity and the concentration of DNA were measured by NanoDrop (Thermo Scientific). Genome sequencing was conducted using a 454 GS-FLX Titanium platform. One standard run and one paired-end run were completed, resulting in a genome size of 20,476,699 bp that was organized in 384 large contigs (≥ 500 bp) and 29 scaffolds. Approximately 200 scaffolded gaps were amplified by PCR, and subsequent Sanger sequencing closed 176 gaps, resulting in an assembly with 208 contigs. An Illumina sequencing run was also conducted on genomic DNA from *R. glutinis* and assembled with our previous 454 and Sanger data using CAP3 and SSAFE, resulting in an assembly containing 186 contigs in 29 scaffolds.

Expressed sequence tags (ESTs) are not available to assist in annotation. Therefore, GeneMarkS (6), a self-training *ab initio* gene finder, was used for the initial gene prediction. This set of predicted genes was used to train Augustus (7) and Glimmer (8). A comprehensive gene prediction run was performed on the entire genome using GeneMark, Augustus, and Glimmer. Repeat-

Masker was used to screen for repeats and low-complexity sequences. BLAST searches were conducted on predicted genes against the nonredundant NCBI database. Based on these results, annotations were manually curated and evaluated using InterPro (9). To aid in the manual annotation, Apollo was used to view the predicted gene models and repetitive elements.

The *R. glutinis* genome is 20,478,880 bp, with a G+C content of 61.9%. The length of putative genic regions is 17,229,291 bp, and there are 3,359 putative genes coding for 2,817 proteins. By comparison, the genome of the systematically similar yeast *Ustilago maydis* 521 is 19,742,445 nucleotides (nt) in length and has a 54% G+C content with 6,631 genes. The *R. glutinis* genome sequence will enable pathway analysis and functional genomics investigations, which will allow for the identification of regulatory mechanisms controlling lipid accumulation. By improving the efficiency of carotenoid/lipid accumulation in *R. glutinis*, this yeast has the potential to produce useful antioxidants and biodiesel.

Nucleotide sequence accession numbers. The *R. glutinis* whole-genome shotgun (WGS) project (PID-59971) has been deposited at DBJ/EMBL/GenBank under the accession no. [AEVR000000000](https://www.ncbi.nlm.nih.gov/nuccore/AEVR000000000) and consists of (WGS) sequences [AEVR02000001](https://www.ncbi.nlm.nih.gov/nuccore/AEVR02000001) to [AEVR02000029](https://www.ncbi.nlm.nih.gov/nuccore/AEVR02000029). The version described in this paper is [AEVR020000000](https://www.ncbi.nlm.nih.gov/nuccore/AEVR020000000).

ACKNOWLEDGMENTS

This research was supported by the U.S. Department of Energy under award no. DE-FG3606GO86025.

We thank Amanda Cooksey, Philippe Chouvarine, and Michelle Banes for technical assistance.

REFERENCES

- Bhosale P, Gadre RV. 2001. Production of β -carotene by a *Rhodotorula glutinis* mutant in sea water medium. *Bioresour. Technol.* 76:53–55. [http://dx.doi.org/10.1016/S0960-8524\(00\)00075-4](https://doi.org/10.1016/S0960-8524(00)00075-4).
- Chen X, Li Z, Zhang X, Hu F, Ryu DD, Bao J. 2009. Screening of oleaginous yeast strains tolerant to lignocellulose degradation compounds. *Appl. Biochem. Biotechnol.* 159:591–604. [http://dx.doi.org/10.1007/s12010-008-8491-x](https://doi.org/10.1007/s12010-008-8491-x).

3. Xue F, Miao J, Zhang X, Tan T. 2010. A new strategy for lipid production by mix cultivation of *Spirulina platensis* and *Rhodotorula glutinis*. *Appl. Biochem. Biotechnol.* 160:498–503. <http://dx.doi.org/10.1007/s12010-008-8376-z>.
4. Kurtz M, Bhattacharjee JK. 1975. Biosynthesis of lysine in *Rhodotorula glutinis*: role of pipercolic acid. *J. Gen. Microbiol.* 86:103–110. <http://dx.doi.org/10.1099/00221287-86-1-103>.
5. D’Cunha GB, Satyanarayan V, Nair M P. 1996. Purification of phenylalanine ammonia lyase from *Rhodotorula glutinis*. *Phytochemistry* 42:17–20. [http://dx.doi.org/10.1016/0031-9422\(95\)00914-0](http://dx.doi.org/10.1016/0031-9422(95)00914-0).
6. Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res.* 33:W451–W454. <http://dx.doi.org/10.1093/nar/gki487>.
7. Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. Augustus: a Web server for gene finding in eukaryotes. *Nucleic Acids Res.* 32:W309–W312. <http://dx.doi.org/10.1093/nar/gkh379>.
8. Aggarwal G, Ramaswamy R. 2002. *Ab initio* gene identification: prokaryote genome annotation with GeneScan and Glimmer. *J. Biosci.* 27: 7–14. <http://dx.doi.org/10.1007/BF02703679>.
9. Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, Corpet F, Croning MDR, Durbin R, Falquet L, Fleischmann W, Gouzy J, Hermjakob H, Hulo N, Jonassen I, Kahn D, Kanapin A, Karavidopoulou Y, Lopez R, Marx B, Mulder NJ, Oinn TM, Pagni M, Servant F, Sigrist CJ, Zdobnov EM. 2001. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res.* 29:37–40. <http://dx.doi.org/10.1093/nar/29.1.37>.