

## Genome Sequence of the Oleaginous Red Yeast *Rhodospiridium toruloides* MTCC 457

Shailesh Kumar, Hariom Kushwaha, Anand Kumar Bachhawat, Gajendra Pal Singh Raghava and Kaliannan Ganesan

*Eukaryotic Cell* 2012, 11(8):1083. DOI: 10.1128/EC.00156-12.

---

Updated information and services can be found at:  
<http://ec.asm.org/content/11/8/1083>

---

### REFERENCES

*These include:*

This article cites 15 articles, 7 of which can be accessed free at:  
<http://ec.asm.org/content/11/8/1083#ref-list-1>

### CONTENT ALERTS

Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), [more»](#)

---

---

Information about commercial reprint orders: <http://journals.asm.org/site/misc/reprints.xhtml>  
To subscribe to to another ASM Journal go to: <http://journals.asm.org/site/subscriptions/>

---

# Genome Sequence of the Oleaginous Red Yeast *Rhodospiridium toruloides* MTCC 457

Shailesh Kumar,<sup>a</sup> Hariom Kushwaha,<sup>a</sup> Anand Kumar Bachhawat,<sup>a,b</sup> Gajendra Pal Singh Raghava,<sup>a</sup> and Kaliannan Ganesan<sup>a</sup>

CSIR-Institute of Microbial Technology, Chandigarh, India,<sup>a</sup> and Indian Institute of Science Education and Research (IISER), Knowledge City, SAS Nagar, Manauli, Punjab, India<sup>b</sup>

**We report the *de novo* assembled 20.05-Mb draft genome of the red yeast *Rhodospiridium toruloides* MTCC 457, predicted to encode 5,993 proteins, 4 rRNAs, and 125 tRNAs. Proteins known to be unique to oleaginous fungi are present among the predicted proteins. The genome sequence will be valuable for molecular genetic analysis and manipulation of lipid accumulation in this yeast and for developing it as a potential host for biofuel production.**

*Rhodospiridium toruloides* (synonym, *Rhodotorula gracilis*; anamorph, *Rhodotorula rubescens*) is an oleaginous yeast with characteristic red color due to the presence of carotenoids. It can accumulate lipids to a higher level (~75% of dry weight under certain conditions) than most other oleaginous yeasts and fungi (2). *R. toruloides* strains are haploids and exist in two mating types, a and A, making them potentially more amenable to genetic and molecular analysis (1, 5, 13). Thus, *R. toruloides* offers many opportunities for being developed as an additional yeast model and synthetic biology platform to *Saccharomyces cerevisiae*, which, despite being extremely well studied, lacks several biochemical features of biotechnological importance that are present in the red yeasts. As a first step toward this end, and to better understand the oleaginous nature of this yeast, we report here its draft genome sequence.

The genome of the yeast *Rhodospiridium toruloides* MTCC 457 was sequenced using Illumina GAIIX at Genotypic Technology, Bangalore, India. The SeqQC (<http://genotypic.co.in/SeqQC.html?mnu=1>) and Fastx toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)) were used to filter the data for high quality (cutoff read length = 70%; cutoff quality score = 20) and to remove adaptor sequences from raw reads. A total of 14,041,098 paired-end reads with a length of 66 nucleotides (nt) and 20,837,494 reads with a length of 72 nt were used for *de novo* genome assembly with SOAPdenovo v. 1.04 (7) to yield 644 scaffolds 20,090,673 bp in size with an  $N_{50}$  contig length of 177 kb (at a hash length of 39). These scaffolds were split into 689 contigs (at gaps of >10 unknown nt) with an  $N_{50}$  contig length of 74 kb. The draft genome with 62% G+C content is predicted to code for 125 tRNAs by tRNAscan-SE v. 1.23 (8) and 4 rRNAs (5S, 5.8S, 18S, and 28S) by nucleotide BLAST (15) and RNAmmer 1.2 (6). We have also sequenced pooled RNA from cells grown under a variety of conditions with Illumina GAIIX technology and assembled the sequences into 8,412 transcripts with the help of genome scaffolds by TopHat (11) and Cufflinks (12).

Gene prediction and annotation of the draft genome were done by the MAKER pipeline (4) using transcripts as evidence. Among the 5,993 predicted proteins (with a minimum length of 100 amino acids), 4,222 showed homologs in the Swiss-Prot database (at an  $E$  value of  $10^{-6}$ ). Malic enzyme and ATP-citrate lyase, reported to be critical for lipid accumulation (3, 9, 10), are also present. Recent comparative genomic analysis has revealed that oleaginous fungi have additional pathways for synthesis of

acetyl coenzyme A (acetyl-CoA), a key intermediate in fatty acid synthesis (14). Orthologs of representative enzymes in these pathways, such as acyl-CoA dehydrogenase (fatty acid  $\beta$ -oxidation), isovaleryl-CoA dehydrogenase (leucine degradation), glutaryl-CoA dehydrogenase, and enoyl-CoA hydratase (lysine degradation), among others, are present in *R. toruloides*, suggesting that this yeast also has these pathways unique to oleaginous fungi.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number [AJM000000000](http://ajm000000000). The version described in this paper is the first version, [AJM010000000](http://ajm010000000).

## ACKNOWLEDGMENTS

Financial support for this work is from CSIR, India. S.K. is a recipient of the Senior Research Fellowship from CSIR, India.

Genome assembly and annotation data of this project can be downloaded at our genomics web portal (<http://crdd.osdd.net/raghava/genomesrs/>).

## REFERENCES

1. Abe K, Kusaka I, Fukui S. 1975. Morphological change in the early stages of the mating process of *Rhodospiridium toruloides*. *J. Bacteriol.* 122:710–718.
2. Ageitos JM, Vallejo JA, Veiga-Crespo P, Villa TG. 2011. Oily yeasts as oleaginous cell factories. *Appl. Microbiol. Biotechnol.* 90:1219–1227.
3. Beopoulos A, Nicaud JM, Gaillardin C. 2011. An overview of lipid metabolism in yeasts and its impact on biotechnological processes. *Appl. Microbiol. Biotechnol.* 90:1193–1206.
4. Cantarel BL, et al. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 18:188–196.
5. Coelho MA, Rosa A, Rodrigues N, Fonseca A, Goncalves P. 2008. Identification of mating type genes in the bipolar basidiomycetous yeast *Rhodospiridium toruloides*: first insight into the MAT locus structure of the *Sporidiobolales*. *Eukaryot. Cell* 7:1053–1061.
6. Lagesen K, et al. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35:3100–3108.
7. Li R, et al. 2010. *De novo* assembly of human genomes with massively parallel short read sequencing. *Genome Res.* 20:265–272.

Received 5 June 2012 Accepted 5 June 2012

Address correspondence to Kaliannan Ganesan, [ganesan@imtech.res.in](mailto:ganesan@imtech.res.in), or Anand Kumar Bachhawat, [anand@iisermohali.ac.in](mailto:anand@iisermohali.ac.in).

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/EC.00156-12

8. 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.
9. Ratledge C, Wynn JP. 2002. The biochemistry and molecular biology of lipid accumulation in oleaginous microorganisms. *Adv. Appl. Microbiol.* 51:1–51.
10. Shashi K, Bachhawat AK, Joseph R. 1990. ATP:citrate lyase of *Rhodotula gracilis*: purification and properties. *Biochim. Biophys. Acta* 1033: 23–30.
11. Trapnell C, Pachter L, Salzberg SL. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25:1105–1111.
12. Trapnell C, et al. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28:511–515.
13. Tully M, Gilbert HJ. 1985. Transformation of *Rhodospiridium toruloides*. *Gene* 36:235–240.
14. Vorapreeda T, Thammarongtham C, Cheevadhanarak S, Laoteng K. 2012. Alternative routes of acetyl-CoA synthesis identified by comparative genomic analysis: involvement in the lipid production of oleaginous yeast and fungi. *Microbiology* 158:217–228.
15. Zhang Z, Schwartz S, Wagner L, Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* 7:203–214.