



Genome Sequence of *Torulaspota microellipsoides* CLIB 830^T

Virginie Galeote,^a Frédéric Bigey,^a Hugo Devillers,^b Raúl A. Ortiz-Merino,^c Sylvie Dequin,^a Kenneth H. Wolfe,^c Cécile Neuvéglise^b

^aSPO, INRA, SupAgro, Université Montpellier, Montpellier, France

^bMicalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France

^cUCD Conway Institute and School of Medicine, University College Dublin, Dublin, Ireland

ABSTRACT We report here the genome sequence of the ascomycetous yeast *Torulaspota microellipsoides* CLIB 830^T. A reference genome for this species, which has been found as a donor of genetic material in wine strains of *Saccharomyces cerevisiae*, will undoubtedly give clues to our understanding of horizontal transfer mechanisms between species in the wine environment.

The genus *Torulaspota* belongs to the Saccharomycotina subphylum and is now composed of eight species (*T. delbrueckii*, *T. globosa*, *T. franciscaea*, *T. pretoriensis*, *T. microellipsoides*, *T. maleeae*, *T. quercuum*, and *T. incana*); *T. maleeae*, *T. quercuum*, and *T. incana* have been described in the past decade (1–4). *Torulaspota delbrueckii* is probably the most widely distributed species, found in nature and in anthropic environments requiring fermentation, such as wine or bread, whereas other species, such as *T. microellipsoides*, appear more sporadically. However, their presence in fermentation vats may play an important role. Indeed, *T. microellipsoides* has been shown to be the donor of a DNA region of at least 158 kb to *Saccharomyces cerevisiae* wine strains, which confers to them an adaptive advantage during wine fermentation (5, 6). Yet, only the *T. delbrueckii* genome sequence is available (7). Sequencing the genome of *T. microellipsoides* strain CLIB 830^T (=CBS 427) may therefore contribute to an increase in the genomics knowledge of this clade.

Total genomic DNA was used to construct a paired-end (PE) 500-bp insert library and a mate pair (MP) 6-kb insert library. Both libraries were sequenced using the Illumina HiSeq 2000 platform, resulting in raw sequencing coverage depths of 197× (PE) and 177× (MP). Sequencing reads were cleaned using Trimmomatic version 0.32 (8), resulting in a sequencing depth of 319×. A preliminary assembly of all reads was obtained using SOAPdenovo2 version 2.04 (9). MP reads were mapped with BWA version 0.6.2 (10), and inward read pairs were removed using BamTools version 2.2.3 (11). A second assembly of PE and cleaned MP reads was performed using SOAPdenovo2, with a k-mer of 61. Gap closing was performed using GapCloser version 1.12 (9). The final assembly was made of 46 scaffolds (N_{50} , 1.2 Mb) larger than 1 kb; 8 of them corresponded to mitochondrial DNA. The remaining 38 scaffolds were suitable for automatic annotation.

The structural annotation of protein-coding genes was performed using the Amadea Annotation transfer tool (Isoft, France), with the *Lachancea kluyveri* CBS3082^T genome as a reference (revised version available at <http://gryc.inra.fr> [12]). Missing genes were investigated through a BLASTX search against the NCBI RefSeq database, with a comparison to an annotation from the YGAP pipeline (13), and manual curation. In total, 5,239 protein-coding genes were predicted, including 145 pseudogenes. tRNA genes were identified using tRNAscan-SE version 1.3.1 (14). Transposable elements were identified by BLAST with known yeast elements from different families, such as *Ty1-copia*, *Ty3-gypsy*, and *hAT*. A family of 38 *Rover* elements was identified, including intact and degenerate copies and miniature inverted-repeat transposable elements

Received 30 May 2018 Accepted 30 May 2018 Published 28 June 2018

Citation Galeote V, Bigey F, Devillers H, Ortiz-Merino RA, Dequin S, Wolfe KH, Neuvéglise C. 2018. Genome sequence of *Torulaspota microellipsoides* CLIB 830^T. Genome Announc 6:e00615-18. <https://doi.org/10.1128/genomeA.00615-18>.

Copyright © 2018 Galeote et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Virginie Galeote, virginie.galeote@inra.fr.

(MITEs) (15). Interestingly, 3 tandem *Rover* elements and the surrounding region encoding 4 proteins in scaffold TOMI504 were found to be almost identical (99.8% identity over 15,389 nucleotides [nt]) to contig 1098 of *S. cerevisiae* strain AWRI1631 (16). This clearly indicates that this sequence has been acquired by horizontal transfer into AWRI1631 from *T. microellipsoides* and not from *Lachancea* species, as previously thought (15).

A reference genome for this species will undoubtedly give clues to aid in our understanding of the transfer of genetic material between species in the wine environment and the mechanisms involved.

Accession number(s). The sequence of the *T. microellipsoides* genome has been deposited at the European Nucleotide Archive (ENA) under BioProject no. PRJEB7632. The accession numbers of the 46 scaffolds are [FYBL01000001](#) to [FYBL01000046](#). The version described in this paper is the first version.

ACKNOWLEDGMENTS

The development of the annotation transfer tool based on Amadea Biopack was financially supported by the ANR project GB-3G (2010 BLAN 1606). We thank the FP7 Marie Curie Initial Training Networks (ITN) (FP7-PEOPLE-2013-ITN, YEASTCELL).

We thank Jean-Luc Legras for his helpful advice and discussions.

REFERENCES

- Kurtzman CP. 2011. *Torulaspota* Lindner (1904), p 867–874. In Kurtzman CP, Fell JW, Boekhout T (ed), *The yeasts, a taxonomic study*, vol 2, 5th ed. Elsevier, New York, NY.
- Saluja P, Yelchuri RK, Sohal SK, Bhagat G, Paramjit, Prasad GS. 2012. *Torulaspota indica* a novel yeast species isolated from coal mine soils. *Antonie Van Leeuwenhoek* 101:733–742. <https://doi.org/10.1007/s10482-011-9687-6>.
- Limtong S, Imanishi Y, Jindamorakot S, Ninomiya S, Yongmanitchai W, Nakase T. 2008. *Torulaspota maleeae* sp. nov., a novel ascomycetous yeast species from Japan and Thailand. *FEMS Yeast Res* 8:337–343. <https://doi.org/10.1111/j.1567-1364.2007.00324.x>.
- Kurtzman CP, Robnett CJ. 2003. Phylogenetic relationships among yeasts of the “*Saccharomyces complex*” determined from multigene sequence analyses. *FEMS Yeast Res* 3:417–432. [https://doi.org/10.1016/S1567-1356\(03\)00012-6](https://doi.org/10.1016/S1567-1356(03)00012-6).
- Marsit S, Mena A, Bigey F, Sauvage F-X, Couloux A, Guy J, Legras J-L, Barrio E, Dequin S, Galeote V. 2015. Evolutionary advantage conferred by an eukaryote-to-eukaryote gene transfer event in wine yeasts. *Mol Biol Evol* 32:1695–1707. <https://doi.org/10.1093/molbev/msv057>.
- Marsit S, Sanchez I, Galeote V, Dequin S. 2016. Horizontally acquired oligopeptide transporters favour adaptation of *Saccharomyces cerevisiae* wine yeast to oenological environment. *Environ Microbiol* 18: 1148–1161. <https://doi.org/10.1111/1462-2920.13117>.
- Gordon JL, Armisén D, Proux-Wéra E, ÓhÉigeartaigh SS, Byrne KP, Wolfe KH. 2011. Evolutionary erosion of yeast sex chromosomes by mating-type switching accidents. *Proc Natl Acad Sci U S A* 108:20024–20029. <https://doi.org/10.1073/pnas.1112808108>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu S-M, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam T-W, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience* 1:1–6. <https://doi.org/10.1186/2047-217X-1-18>.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>.
- Barnett DW, Garrison EK, Quinlan AR, Strömberg MP, Marth GT. 2011. BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* 27:1691–1692. <https://doi.org/10.1093/bioinformatics/btr174>.
- Vakirlis N, Sarilar V, Drillon G, Fleiss A, Agier N, Meyniel J-P, Blanpain L, Carbone A, Devillers H, Dubois K, Gillet-Markowska A, Graziani S, Huu-Vang N, Poirel M, Reisser C, Schott J, Schacherer J, Lafontaine I, Llorente B, Neuvéglise C, Fischer G. 2016. Reconstruction of ancestral chromosome architecture and gene repertoire reveals principles of genome evolution in a model yeast genus. *Genome Res* 26:918–932. <https://doi.org/10.1101/gr.204420.116>.
- Proux-Wéra E, Armisén D, Byrne KP, Wolfe KH. 2012. A pipeline for automated annotation of yeast genome sequences by a conserved synteny approach. *BMC Bioinfo* 13:237–248. <https://doi.org/10.1186/1471-2105-13-237>.
- 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.
- Sarilar V, Bleykasten-Grosshans C, Neuvéglise C. 2014. Evolutionary dynamics of *hAT* DNA transposon families in *Saccharomycetaceae*. *Genome Biol Evol* 7:172–190. <https://doi.org/10.1093/gbe/evu273>.
- Borneman AR, Forgan AH, Pretorius IS, Chambers PJ. 2008. Comparative genome analysis of a *Saccharomyces cerevisiae* wine strain. *FEMS Yeast Res* 8:1185–1195. <https://doi.org/10.1111/j.1567-1364.2008.00434.x>.