Genome Sequence of the Wine Yeast *Saccharomycodes ludwigii* UTAD17

Maria J. Tavares,a,b Ulrich Güldener,c Marcos Esteves,d,e Arlete Mendes-Faia,d,e Ana Mendes-Ferreira,d,e Nuno P. Miraa,b

a Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
b iBB–Institute for Bioengineering and Biosciences, Biological Sciences Research Group, Instituto Superior Técnico, Lisbon, Portugal
c Department of Bioinformatics, Wissenschaftszentrum Weihenstephan, Technische Universität München, Freising, Germany
d Escola de Ciências da Vida e Ambiente, Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal
e BioISI-Biosystems and Integrative Sciences Institute, Lisbon, Portugal

**ABSTRACT** This work describes, for the first time, the genome sequence of a *Saccharomycodes ludwigii* strain. Although usually seen as a wine spoilage yeast, *S. ludwigii* has been of interest for the production of fermented beverages because it harbors several interesting properties, including the production of beneficial aroma compounds.

*Saccharomycodes ludwigii* is one of the many non-*Saccharomyces* (NSY) species present in the wine must mycobiome (1–3). Phylogenetically, the *Saccharomycodes* genus is placed in the *Saccharomycoidea* family, in the phylum Ascomycota and subphylum *Saccharomycotina*, being considered a sister genus to *Hanseniaspora* (4). Unlike the genus *Hanseniaspora*, which is seen as having a positive effect in vinification (5), the presence of *S. ludwigii* is of concern since it is considered a spoilage agent, reducing the organoleptic properties of wine and interfering with the clarification process (4, 6). Despite this negative effect of *S. ludwigii* in wine production, this species has been explored for the production of other fermented beverages (7–10), and it has been reported to be an interesting flavoring agent that is able to produce several relevant aroma compounds (9).

Little is known concerning the genetics and physiology of the *S. ludwigii* species. In particular, no genomic sequence has yet been reported for a strain of this species, which impedes our better understanding and exploration of it. As such, in this work, we have obtained the genomic sequence of an *S. ludwigii* isolate (UTAD17) recovered from wine must in the Douro region of Portugal by using a selective medium for NSY species. To obtain the genome sequence of *S. ludwigii* UTAD17, we cultivated cells in rich medium and extracted the DNA as described previously (11). The DNA libraries were prepared using the ThruPLEX DNA-seq kit, and paired-end sequencing of the generated DNA fragments was performed on a MiSeq platform. After two sequencing rounds, 20,333,547 reads of 250 bp on average were obtained and de novo assembled into 1,360 contigs (N50 length of 17,540 bp; filtered to have a coverage above 300× and a size above 1,000 nucleotides). The sum of the assembled contigs totaled 10,785,241 bp, which is in line with the genome size predicted for *Hanseniaspora osmophila* (4, 12), another species in the *Saccharomycodaceae* family. Automatic annotation of the *S. ludwigii* UTAD17 genomic sequence was undertaken using (i) Fgenesh trained on *Aspergillus nidulans*, *Neurospora crassa*, and a mixed matrix based on different species (13); (ii) GeneMark-ES (14); and (iii) Augustus (15). The different gene models proposed by the algorithms were displayed in the Generic Genome Browser (GBrowse) (16), allowing individual manual validation. Gene models showing the highest similarity with homologues described in other yeast species were selected. If needed, gene structures are available on the Generic Genome Browser at https://moby.etsi.ulisboa.pt/ for the *S. ludwigii* genome assembly (SACLUU00000000) under the project *Saccharomycodes ludwigii* UTAD17.

Received 6 September 2018 Accepted 9 October 2018 Published 8 November 2018


**Editor** Christina A. Cuomo, Broad Institute of MIT and Harvard

**Copyright** © 2018 Tavares et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Ana Mendes-Ferreira, anamf@utad.pt, or Nuno P. Mira, nuno.mira@tecnico.ulisboa.pt.
were adjusted by splitting or fusing the gene models or redefining exon-intron boundaries. The predicted complete set of open reading frames (ORFeome) of S. ludwigii UTAD17 is estimated to be 4,015 protein-coding genes. BLASTP analysis (using the nonredundant protein sequence databases as a reference) (17) revealed that S. ludwigii proteins share a high degree of homology with H. osmophila and Lachancea fermentati. This is an interesting observation since both species are indigenous to wine musts and influence the wine fermentation process (12, 18, 19). It is expected that the S. ludwigii UTAD17 genome sequence reported here can foster research focused on this species, contributing particularly to overcoming its activity as a wine spoilage agent and improving its utilization in the production of fermented beverages.

**Data availability.** The reads and the assembled genome sequences of Saccharomyces ludwigii UTAD17 have been deposited in ENA under accession number UFAJ01000000 (contigs UFAJ01000001 through UFAJ01001360; study accession number PRJEB27462; read accession number SAMEA4945973).

**ACKNOWLEDGMENTS**

This work was financed by FEDER through POCl-COMPETE 2020 and by Fundação Ciência e Tecnologia (contracts PD/BD/142945/2018 and BI/PDTC/AGR-TEC/3315/2014, 2016, project SMARTWINE—“Smarter wine fermentations: integrating omic-tools for development of novel mixed-starter cultures for tailoring wine production”). The Biosystems and Integrative Sciences Institute (BioSL; FCT/UID/Multi/04046/2013) and the iBB-Institute for Bioengineering and Biosciences (through contract UID/BIO/04565/2013) received funding from FCT-Portuguese Foundation for Science and Technology. Programa Operacional Regional de Lisboa 2020 is also acknowledged for its financial support to iBB (project number 007317).

The authors thank Rogério Tenreiro and Filipa Antunes Silva for support with the preliminary pulsed-field gel electrophoresis (PFGE) analysis of the S. ludwigii strain UTAD17 genome.

**REFERENCES**