

Draft genome sequence of the yeast *Schwanniomyces capriottii* strain UCD805, isolated from soil in Ireland

Padraic G. Heneghan,¹ Adam P. Ryan,² Eoin Ó Cinnéide,¹ Jordan Davies,² Cathal Bracken,² Victoria Ogundipe,² Maria Doheny,¹ Liam Lenihan,¹ Anastasia Passalaris,¹ Rosalind Walker,¹ Kenneth H. Wolfe,¹ Geraldine Butler,² Kevin P. Byrne^{1,2}

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT *Schwanniomyces capriottii* is a member of the Debaryomycetaceae family in the order Saccharomycetales. Here, we present the genome sequence of *S. capriottii* UCD805, which was isolated from soil in Dublin, Ireland. This genome is 12.2 Mb and was assembled into 14 scaffolds plus a mitochondrial genome scaffold.

KEYWORDS yeasts, genome analysis

Schwanniomyces capriottii is a species in the *Debaryomyces* clade of the budding yeast family Debaryomycetaceae (1). It was previously known as *Debaryomyces castellii* but renamed in 2010 (2). It was first isolated from soil in Sweden (3). It has no known food or biotechnological applications, and as it does not grow at 37°C, it is unlikely to be a human pathogen.

S. capriottii UCD805 was isolated from soil collected on the University College Dublin campus in Dublin, Ireland (GPS coordinates 53.310884, -6.223159) as part of an undergraduate research module (4). The sample was collected from under a willow tree beside a forest walk.

Soil material was passaged twice at room temperature in 9 mL liquid yeast extract-peptone-dextrose (YPD) containing chloramphenicol (30 µg/mL) and ampicillin (100 µg/mL) and cultured on YPD agar plates. The species was identified by PCR and Sanger sequencing of the ribosomal DNA internal transcribed spacer (ITS) and D1/D2 regions (accession numbers [OR578565](#) and [OR578569](#)). Sequence identity was 100% (591/591 bp) in the ITS and 99% (568/571 bp) in the D1/D2 region, to the type strain of *S. capriottii* (accession numbers [KY105380](#) and [KY109604](#)).

DNA for genome sequencing was isolated from liquid YPD cultures grown at 20°C. For short-read sequencing, DNA was isolated by phenol/chloroform extraction. Illumina library construction (300-cycle v1.5 kit) and sequencing were done by Novogene (UK) Company Ltd. using a NovaSeq 6000 instrument with S4 flowcell and yielded 7.4 million read pairs (2 × 150 bp). For long-read sequencing, DNA was extracted using a Bioscience Technology Masterpure Yeast DNA Purification Kit (MPY80010). Oxford Nanopore sequencing was done on a MinION MK1C instrument with flowcell FLO-MIN112 (R10.4) and barcoding kit SQK-NBD112-24. NanoFilt (v2.3.0) (5) retained 44,648 reads with quality ≥ 7 and length ≥ 1,000 bp (reads N50 = 17,687 bp), which were then assembled using Canu (v2.0.0) (6), followed by five rounds of error correction using NextPolish (v.1.4.1) with the Illumina reads (7).

The UCD805 assembly consisted of 14 nuclear scaffolds (total 12.15 Mb; N50 = 888,568 bp) and the mitochondrial genome (31,651 bp; accession number [JAV-SEF010000001](#)). Three scaffolds terminate with multiple tandem repeats of the sequence (GTGTAGGATGGT)_n at both ends, so they are inferred to be complete chromosomes with telomeres (8). This telomeric repeat was found at 16 scaffold ends, indicating that this

Editor Jason E. Stajich, University of California
Riverside, USA

Address correspondence to Kevin P. Byrne,
kevin.byrne@ucd.ie.

The authors declare no conflict of interest.

See the funding table on p.2.

Received 8 November 2023

Accepted 23 January 2024

Published 5 February 2024

Copyright © 2024 Heneghan et al. This is an
open-access article distributed under the terms of
the [Creative Commons Attribution 4.0 International
license](#).

yeast strain has 8 chromosomes. BLASTN searches detected rDNA arrays at the ends of three scaffolds ([JAVSEF010000003](#), [JAVSEF010000012](#), and [JAVSEF010000013](#)).

The UCD805 genome assembly has average nucleotide identity (9) of 99.75% to the *S. capriottii* type strain NRRL Y-7423 (accession no. [JAKTYZ010000000](#)) (10). Using BUSCO version 5.1.2, genome completeness was estimated as 96.83% compared to the Ascomycota lineage data set. G + C content is 34.84%.

ACKNOWLEDGMENTS

This work was supported by undergraduate teaching resources from University College Dublin and by the European Research Council (789341).

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

AUTHOR AFFILIATIONS

¹School of Medicine, Conway Institute, University College Dublin, Belfield, Ireland

²School of Biomolecular and Biomedical Science, Conway Institute, University College Dublin, Belfield, Ireland

AUTHOR ORCID

Kenneth H. Wolfe  <http://orcid.org/0000-0003-4992-4979>

Geraldine Butler  <http://orcid.org/0000-0002-1770-5301>

Kevin P. Byrne  <http://orcid.org/0000-0002-6368-154X>

FUNDING

| Funder | Grant(s) | Author(s) |
|---------------------------------|----------|------------------|
| European Research Council | 789341 | Kenneth H. Wolfe |
| University College Dublin (UCD) | | Geraldine Butler |

DATA AVAILABILITY

This whole-genome shotgun project has been deposited at DDBJ/ ENA/GenBank accession no. [JAVSEF000000000](#) (BioProject no. [PRJNA1018912](#)). The version described in this paper is version 1. The raw reads were deposited at SRA (accession no. [SRR26109203](#) and [SRR26109698](#)). The ITS sequence is at [OR578565](#) and the D1/D2 region sequence at [OR578569](#).

REFERENCES

- Kurtzman CP, Robnett CJ. 1995. Molecular relationships among hyphal ascomycetous yeasts and yeastlike taxa. *Can J Bot* 73:824–830. <https://doi.org/10.1139/b95-328>
- Kurtzman CP, Suzuki M. 2010. Phylogenetic analysis of ascomycete yeasts that form coenzyme Q-9 and the proposal of the new genera *Babjeviella*, *Meyerozyma*, *Milleroyzyma*, *Priceomyces*, and *Scheffersomyces*. *Mycoscience* 51:2–14. <https://doi.org/10.1007/S10267-009-0011-5>
- Capriotti A. 1958. *Debaryomyces castellii* nov. spec. *Arch Microbiol* 28:344.
- Bergin SA, Allen S, Hession C, Ó Cinnéide E, Ryan A, Byrne KP, Ó Cróinín T, Wolfe KH, Butler G. 2022. Identification of European isolates of the lager yeast parent *Saccharomyces eubayanus*. *FEMS Yeast Res* 22:foac053. <https://doi.org/10.1093/femsyr/foac053>
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. Nanopack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive K-MER weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>
- Chen Z, Erickson DL, Meng J. 2021. Polishing the Oxford nanopore long-read assemblies of bacterial pathogens with Illumina short reads to improve genomic analyses. *Genomics* 113:1366–1377. <https://doi.org/10.1016/j.ygeno.2021.03.018>
- Cohn M, McEachern MJ, Blackburn EH. 1998. Telomeric sequence diversity within the genus *Saccharomyces*. *Curr Genet* 33:83–91. <https://doi.org/10.1007/s002940050312>
- Rodriguez-R LM, Konstantinidis KT. 2014. Bypassing cultivation to identify bacterial species. *Microbe Magazine* 9:111–118. <https://doi.org/10.1128/microbe.9.111.1>
- Opulente DA, Leavitt LaBella A, Harrison M-C, Wolters JF, Liu C, Li Y, Kominek J, Steenwyk JL, Stoneman HR, VanDenAvond J, et al. 2023. Genomic and ecological factors shaping specialism and generalism across an entire subphylum. *bioRxiv*:2023.06.19.545611. <https://doi.org/10.1101/2023.06.19.545611>