AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Mycology | Announcement

Genome sequence of the yeast *Cyberlindnera sylvatica* UCD1060, isolated from soil in Ireland

Padraic G. Heneghan,[1] Adam P. Ryan,[2] Sean A. Bergin,[2] Margarita Caka,[2] Aimee Coughlan,[2] Julianna Decuseara,[2] Michalina Gora,[2] Daniil Hudov,[2] Aoife McLoughlin,[2] Sheila Osadolor,[2] Isabel Saeed Maguire,[2] Weronika Swistowska,[2] Kenneth H. Wolfe,[1] Geraldine Butler,[2] Kevin P. Byrne[1,2]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT *Cyberlindnera sylvatica* is a member of the Cyberlindnera clade in the order Phaffomycetales. We present the genome sequence of *C. sylvatica*. The sequenced strain, UCD1060, was isolated from soil in Wicklow, Ireland. This genome is 13.5 Mb and was assembled into seven chromosome-sized contigs plus a mitochondrial genome contig.

KEYWORDS yeasts, genome analysis

Cyberlindnera sylvatica is a recently characterized (1) species in the budding yeast order Phaffomycetales. It was first isolated in Hungary and Germany from forests. It has no known food or biotechnological applications. As it grows at 37°C, it could potentially be associated with mammals (1).

C. sylvatica UCD1060 was isolated from soil collected from the base of a large oak in the Devil's Glen, County Wicklow, Ireland (GPS coordinates 53.0163025,–6.1524822) as part of an undergraduate research module (2).

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 PQ438548 and PQ438569), using primers ITS1 (TCCGTAGGT GAACCTGCGG) and ITS4 (TCCTCCGCTTATTGATATGC) (3), and for the D1/D2 region NL1 (G CATATCAATAAGCGGAGGAA) and NL4 (GGTCCGTGTTTCAAGACGG) (4). Sequence identity was 98.36% (540/549 bp) in the ITS and 99.96% (555/557 bp) in the D1/D2 region to the type strain of *C. sylvatica* (accession numbers MT305876 and MT316316; note MT305876 is misnamed in GenBank as "*Hyphopichia lachancei*" but comes from CBS 16335, the type strain of *C. sylvatica* [1]).

DNA was isolated by phenol/chloroform extraction from liquid YPD cultures grown at 20°C. Short-read library construction (Illumina DNA-Prep(M) Ref:20060060) and sequencing (UCD Conway Core Facility, Dublin, Ireland) used a NextSeq2000 and P1 flowcell, yielding 5.7 million read pairs (2 × 150 bp; 127× coverage). Adapters and low-quality reads were removed (Skewer version 0.2.2) (5). For long-read sequencing, after DNA extraction (Biosearch Technology Masterpure yeast DNA purification kit MPY80010), library preparation (Native Barcoding Kit SQK-NBD112-24), end repair (NEB-M6630), and ligation (NEB-E6056), we selected for DNA >3 kb with Long Fragment Buffer before Oxford Nanopore sequencing (MinION MK1C, flowcell FLO-MIN112 R10.4). Default fast basecalling was done on the instrument (MinKNOW version 23.07.12; default demultiplexing; barcode trimming; reads ≥500 bp kept; no quality cutoff; 74× coverage). NanoFilt (version 2.3.0) (6) retained 70,996 reads with quality ≥10 and length ≥1,000 bp (reads N_{50} = 21,004 bp), which were then assembled using Canu (version 2.2) (7),

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

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

The authors declare no conflict of interest.

Received 29 October 2024
Accepted 18 February 2025
Published 25 March 2025

followed by five rounds of error correction using NextPolish (version 1.4.1) with the Illumina reads (8). Default parameters were used, except where otherwise noted.

The assembly consisted of seven nuclear contigs (total 13.5 Mb; N_{50} = 2,385,092 bp) and the mitochondrial genome (33,689 bp circular unit; accession number CAXWVZ010000008). Six nuclear contigs terminate with tandem repeats of $(GGGTGTCT)_n$ at both ends, so they are inferred to be complete chromosomes with telomeres (9). The seventh nuclear contig (CAXWVZ010000007) has this telomeric repeat at one end and an rDNA array at the other end; this is the only rDNA locus in the genome.

Using BUSCO version 5.1.2 (10), genome completeness was estimated as 97.36% compared to the Ascomycota lineage data set. G + C content is 45.92%.

ACKNOWLEDGMENTS

This work was supported by undergraduate teaching resources from University College Dublin and Science Foundation Ireland (19/FFP/6668).

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

AUTHOR AFFILIATIONS

[1]School of Medicine, Conway Institute, University College Dublin, Belfield, Ireland
[2]School of Biomolecular and Biomedical Science, Conway Institute, University College Dublin, Belfield, Ireland

AUTHOR ORCIDs

Kenneth H. Wolfe http://orcid.org/0000-0003-4992-4979
Geraldine Butler https://orcid.org/0000-0002-1770-5301
Kevin P. Byrne http://orcid.org/0000-0002-6368-154X

DATA AVAILABILITY

This whole-genome shotgun project has been deposited as DDBJ/ENA/GenBank accession number CAXWVZ010000000 (BioProject no. PRJEB79408). The version described in this paper is version 1. The reads were deposited at ENA (accessions ERR13731774 and ERR13731842). The ITS sequence is PQ438548. The D1/D2 sequence is PQ438569.

Isolate UCD1060 has been deposited in the CBS and PYCC culture collections as CBS 18652 and PYCC 10019.

REFERENCES

1. Brysch-Herzberg M, Dlauchy D, Seidel M, Péter G. 2021. *Cyberlindnera sylvatica* sp. nov., a yeast species isolated from forest habitats. Int J Syst Evol Microbiol 71. https://doi.org/10.1099/ijsem.0.004477
2. 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. https://doi.org/10.1093/femsyr/foac053
3. Kurtzman CP, Robnett CJ. 1998. Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. Antonie Van Leeuwenhoek 73:331–371. https://doi.org/10.1023/a:1001761008817
4. Leaw SN, Chang HC, Sun HF, Barton R, Bouchara J-P, Chang TC. 2006. Identification of medically important yeast species by sequence analysis of the internal transcribed spacer regions. J Clin Microbiol 44:693–699. https://doi.org/10.1128/JCM.44.3.693-699.2006
5. Jiang H, Lei R, Ding SW, Zhu S. 2014. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. BMC Bioinformatics 15:182. https://doi.org/10.1186/1471-2105-15-182
6. 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
7. 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
8. 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
9. 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
10. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351