AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

EUKARYOTES

Draft Genome Sequences of Two Natural Isolates of the Yeast *Barnettozyma californica* from Ireland, UCD09 and UCD89

Alisha A. Mullen,[a] Ciara D. Lynch,[a] Shannon M. Hill,[a] Cian P. Holohan,[a] Tadhg Ó Cróinín,[a] Kenneth H. Wolfe,[b] Caoimhe E. O'Brien,[a] Geraldine Butler[a]

[a]School of Biomedical and Biomolecular Sciences, Conway Institute, University College Dublin, Dublin, Ireland
[b]School of Medicine, Conway Institute, University College Dublin, Dublin, Ireland

ABSTRACT No genome sequence of a species from *Barnettozyma*, a yeast genus in the family Phaffomycetaceae, is currently available. We isolated two *B. californica* strains from soils in Ireland and generated draft sequences of their 11.7-Mb genomes. Single nucleotide polymorphism (SNP) analysis showed 20,490 differences between the strains and suggests that *B. californica* is haploid.

We isolated yeasts from soil samples from Ireland following the protocol of Sylvester et al. (1). Samples were incubated at room temperature in yeast extract-peptone-dextrose (YPD) broth containing carbenicillin (10% [wt/vol]) and chloramphenicol (3% [wt/vol]). Following isolation of single colonies, species were identified by sequencing the internal transcribed spacer (ITS) region of rDNA. Two isolates, UCD09 and UCD89, were identified as *Barnettozyma californica* by searches against the YeastIP database (2). UCD09 was isolated from a wooded area at the Special Ops Paintball Range in Roundwood, County Wicklow, Ireland, and UCD89 from a patch of grass at Balally Park, Dublin, Ireland, about 25 km apart.

B. californica has a worldwide distribution and has previously been isolated from substrates including soil, water, animal dung, tree fluxes, and rotting wood (3, 4). The type strain, CBS 252, was isolated from soil in California in 1931 (5). *B. californica* has been found on the surfaces of olives and secretes lipase in sufficient amounts that it can degrade the quality of olive oil in storage (6). It is one of the six known species in the genus *Barnettozyma* and was formerly called *Williopsis californica* (3, 7). No genome sequence has previously been reported for any species in this genus. *Barnettozyma* is in the family Phaffomycetaceae, which also includes the genera *Wickerhamomyces*, *Cyberlindnera*, *Starmera*, and *Phaffomyces* (8).

Genomic DNAs were isolated by phenol-chloroform extraction and cleanup (DNA Clean & Concentrator-25 kit, Zymo Research). Sequencing was done by BGI Tech Solutions using an Illumina HiSeq 4000 instrument. Libraries with 170- to 800-bp inserts were generated, and 150 bases were sequenced from each end. Low-quality reads were removed using Skewer (9). The reads were then assembled into scaffolds using SPAdes version 3.11.1 (10). To remove contaminants, scaffolds with coverages below 10× or lengths below 1 kb were discarded. The final sets of scaffolds were quality analyzed using Quality Assessment Tool for Genome Assemblies (QUAST) (11).

The UCD09 assembly totaled 11.67 Mb, the N_{50} value was 164 kb, the longest scaffold was 415 kb, average coverage was 76×, and there were 164 contigs of >1 kb. For UCD89, the assembly totaled 11.68 Mb, the N_{50} value was 169 kb, the longest scaffold was 757 kb, average coverage was 69×, and there were 169 contigs of >1 kb. Protein-coding genes were annotated using AUGUSTUS (12) version 2.5.5 (with a *Saccharomyces cerevisiae* training set), which predicted 5,803 genes in UCD09 and 5,802 genes in UCD89. tRNAscan-SE version 1.3.1 (13) predicted 200 (UCD09) and 199 (UCD89) tRNA genes.

Received 14 May 2018 **Accepted** 15 May 2018 **Published** 21 June 2018

Citation Mullen AA, Lynch CD, Hill SM, Holohan CP, Ó Cróinín T, Wolfe KH, O'Brien CE, Butler G. 2018. Draft genome sequences of two natural isolates of the yeast *Barnettozyma californica* from Ireland, UCD09 and UCD89. Genome Announc 6:e00548-18. https://doi.org/10.1128/genomeA.00548-18.

Address correspondence to Geraldine Butler, geraldine.butler@ucd.ie.

Single nucleotide polymorphisms were identified using the Burrows-Wheeler Aligner (BWA) (14) and SAMtools (15). The numbers of SNPs within each strain were low (1,482 in UCD09; 1,714 in UCD89), and the distribution of allele frequencies at SNP sites was flat, suggesting that both strains are haploid. Mapping the reads from UCD09 to the UCD89 assembly identified 20,490 SNPs between these two natural isolates. For comparison, strains of *Saccharomyces paradoxus* from Europe typically differ by 10,000 SNPs (16).

Accession number(s). These whole-genome shotgun projects have been deposited in DDBJ/ENA/GenBank under the accession no. QBLI00000000 and QBLJ00000000. The versions described in this paper are the first versions, QBLI01000000 and QBLJ01000000.

ACKNOWLEDGMENTS

This work was supported by an undergraduate teaching award from University College Dublin. The funders had no role in study design, data collection, and interpretation or the decision to submit the work for publication.

We thank Chris Todd Hittinger for help with the yeast isolation protocol.

REFERENCES

1. Sylvester K, Wang Q-M, James B, Mendez R, Hulfachor AB, Hittinger CT. 2015. Temperature and host preferences drive the diversification of *Saccharomyces* and other yeasts: a survey and the discovery of eight new yeast species. FEMS Yeast Res 15. https://doi.org/10.1093/femsyr/fov002.
2. Weiss S, Samson F, Navarro D, Casaregola S. 2013. YeastIP: a database for identification and phylogeny of *Saccharomycotina* yeasts. FEMS Yeast Res 13:117–125. https://doi.org/10.1111/1567-1364.12017.
3. Kurtzman CP. 2011. *Barnettozyma* Kurtzman, Robnett & Basehoar-Powers (2008), p 333–339. *In* Kurtzman CP, Fell JW, Boekhout T (ed), The yeasts, a taxonomic study, vol 2. Elsevier, Amsterdam, The Netherlands.
4. Morais CG, Cadete RM, Uetanabaro APT, Rosa LH, Lachance M-A, Rosa CA. 2013. D-Xylose-fermenting and xylanase-producing yeast species from rotting wood of two Atlantic Rainforest habitats in Brazil. Fungal Genet Biol 60:19–28. https://doi.org/10.1016/j.fgb.2013.07.003.
5. Lodder J. 1932. Über einige durch das "Centraalbureau voor Schimmelcultures" neuerworbene sporogene Hefearten. Zentralbl Bakteriol Parasitenkd Abt II 86:227–253.
6. Ciafardini G, Zullo BA, Cioccia G, Iride A. 2006. Lipolytic activity of *Williopsis californica* and *Saccharomyces cerevisiae* in extra virgin olive oil. Int J Food Microbiol 107:27–32. https://doi.org/10.1016/j.ijfoodmicro.2005.08.008.
7. Kurtzman CP, Robnett CJ, Basehoar-Powers E. 2008. Phylogenetic relationships among species of *Pichia*, *Issatchenkia* and *Williopsis* determined from multigene sequence analysis, and the proposal of *Barnettozyma* gen. nov., *Lindnera* gen. nov. and *Wickerhamomyces* gen. nov. FEMS Yeast Res 8:939–954. https://doi.org/10.1111/j.1567-1364.2008.00419.x.
8. Kurtzman CP, Robnett CJ. 2013. Relationships among genera of the *Saccharomycotina* (*Ascomycota*) from multigene phylogenetic analysis of type species. FEMS Yeast Res 13:23–33. https://doi.org/10.1111/1567-1364.12006.
9. Jiang H, Lei R, Ding S-W, 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.
10. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.
11. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
12. Stanke M, Diekhans M, Baertsch R, Haussler D. 2008. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. Bioinformatics 24:637–644. https://doi.org/10.1093/bioinformatics/btn013.
13. 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.
14. 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.
15. Hwang S, Kim E, Lee I, Marcotte EM. 2015. Systematic comparison of variant calling pipelines using gold standard personal exome variants. Sci Rep 5:17875. https://doi.org/10.1038/srep17875.
16. Bergström A, Simpson JT, Salinas F, Barré B, Parts L, Zia A, Nguyen Ba AN, Moses AM, Louis EJ, Mustonen V, Warringer J, Durbin R, Liti G. 2014. A high-definition view of functional genetic variation from natural yeast genomes. Mol Biol Evol 31:872–888. https://doi.org/10.1093/molbev/msu037.