



# Draft Genome Sequence of the Yeast *Nadsonia starkeyi-henricii* UCD142, Isolated from Forest Soil in Ireland

Sinéad O'Boyle,<sup>a</sup> Sean A. Bergin,<sup>a</sup> Éabha E. Hussey,<sup>a</sup> Aaron D. McLaughlin,<sup>a</sup> Luan R. Riddell,<sup>a</sup> Kevin P. Byrne,<sup>b</sup> Kenneth H. Wolfe,<sup>b</sup> Caoimhe E. O'Brien,<sup>a</sup>  Geraldine Butler<sup>a</sup>

<sup>a</sup>School of Biomedical and Biomolecular Sciences, Conway Institute, University College Dublin, Dublin, Ireland

<sup>b</sup>School of Medicine, Conway Institute, University College Dublin, Dublin, Ireland

**ABSTRACT** We report a draft genome sequence of a strain of the nonfermentative yeast *Nadsonia starkeyi-henricii*, isolated from soil in a forest in Ireland. Comparison to *Nadsonia fulvescens* shows few rearrangements and a level of divergence similar to that of *Saccharomyces cerevisiae* versus *Saccharomyces paradoxus*. Its mitochondrial genome lacks *NAD* genes.

*Nadsonia starkeyi-henricii* (1) is a little-studied yeast. It is one of the four accepted species and varieties in the genus *Nadsonia* and was formerly called *Schizoblastosporion starkeyi-henricii* (2–4). It is unable to ferment sugars and is only able to grow by respiration (4).

*Nadsonia* species have lower maximum temperatures for growth than most other yeasts, which may have caused them to be overlooked in previous surveys of yeast diversity. Most previous isolates of *N. starkeyi-henricii* were found in soils from forests or bogs, mainly in Northern Europe (2, 3, 5). It will not grow at temperatures above 25 to 30°C, and it dies at 35°C (6). *Nadsonia* species are characterized by their unusual mode of mitotic cell division, with buds forming alternately from opposite poles of the cell. Sporulation has never been observed in *N. starkeyi-henricii*, unlike the other members of this genus. Phylogenetically, *Nadsonia* is in the *Yarrowia* clade, distantly related to *Saccharomyces cerevisiae* (7, 8). One genome sequence has been published from this genus, *Nadsonia fulvescens* var. *elongata* (7).

Strain UCD142 was isolated from soil in a coniferous forest in Dublin, Ireland (global positioning satellite [GPS] coordinates N53.241645, W6.294017). It was cultured at room temperature on yeast extract-peptone-dextrose (YPD) agar plates containing chloramphenicol (3% wt/vol) and carbenicillin (10% wt/vol). Genome sequencing (6.6 million paired-end reads of 150 bp) was performed by BGI Tech Solutions (China) with an Illumina HiSeq 4000. The genome was assembled into 318 contigs (60× coverage, >1 kb) using SPAdes v 3.11.1 (9).

The *N. starkeyi-henricii* genome assembly is 14.6 Mb, similar to *Nadsonia fulvescens* (13.7 Mb; reference 7). Half the data are present in the 40 scaffolds ( $L_{50}$ ) that are larger than 120 kb ( $N_{50}$ ), and the largest scaffold is 404 kb. BWA and SAMtools (10, 11) identified only 2,337 single nucleotide polymorphisms (SNPs), and the genome appears to be haploid. Annotation using YGAP (12) predicted 5,216 protein-coding genes, which compares to 5,657 in *N. fulvescens* (7). tRNAscan-SE (13) predicted 607 tRNA genes in the nuclear genome, an unusually high number for a yeast. Similar to *N. fulvescens*, the 5S rRNA genes are not located within the repeating 18S-5.8S-26S array but instead are dispersed at approximately 115 locations (7).

Comparison of the 10 largest *N. starkeyi-henricii* scaffolds to *N. fulvescens* by BLAST and dot-matrix plots showed that the genomes are mostly collinear. Protein sequence divergence between the two *Nadsonia* species is approximately equal to that between

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

**Citation** O'Boyle S, Bergin SA, Hussey EE, McLaughlin AD, Riddell LR, Byrne KP, Wolfe KH, O'Brien CE, Butler G. 2018. Draft genome sequence of the yeast *Nadsonia starkeyi-henricii* UCD142, isolated from forest soil in Ireland. *Genome Announc* 6:e00549-18. <https://doi.org/10.1128/genomeA.00549-18>.

**Copyright** © 2018 O'Boyle 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 Geraldine Butler, [geraldine.butler@ucd.ie](mailto:geraldine.butler@ucd.ie).

*S. cerevisiae* and *S. paradoxus* (e.g., 89% amino acid sequence identity in *MDN1*, the largest gene in the genome).

The mitochondrial genome is a 22,069-bp circle containing 8 protein-coding genes with no introns. It does not contain any *NAD* genes for NADH dehydrogenase subunits, in contrast to *Yarrowia lipolytica* (14), indicating an absence of mitochondrial respiratory complex I. Complex I is present in most yeasts but has been lost in three clades centered on *Saccharomyces*, *Schizosaccharomyces*, and *Nadsonia* (7).

**Accession number(s).** This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under accession no. [QBLK0000000](https://doi.org/10.1093/nar/gkz000). The version described in this paper is the first version, QBLK01000000.

## 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.

## REFERENCES

1. Starkey RL, Henrici AT. 1927. The occurrence of yeasts in soil. *Soil Science* 23:33–46. <https://doi.org/10.1097/00010694-192701000-00004>.
2. Smith MT. 2011. *Nadsonia* Sydow (1912), p 629–632. In Kurtzman CP, Fell JW, Boekhout T (ed), *The yeasts, a taxonomic study*. Elsevier, Amsterdam, The Netherlands.
3. Smith MT. 2011. *Schizoblastosporion* Ciferri (1930), p 1329–1330. In Kurtzman CP, Fell JW, Boekhout T (ed), *The yeasts, a taxonomic study*. Elsevier, Amsterdam, The Netherlands.
4. Ciferri R. 1930. Contribuzioni alla sistematica delle torulopsidaceae. II-XIV. *Arch Protistenkd* 71:405–452.
5. di Menna ME. 1965. *Schizoblastosporion starkeyi-henricii* Ciferri. *Mycopathol Mycol Appl* 25:205–212. <https://doi.org/10.1007/BF02049913>.
6. Lund A. 1954. *Studies on the ecology of yeasts*. Munksgaard, Copenhagen, Denmark.
7. Riley R, Haridas S, Wolfe KH, Lopes MR, Hittinger CT, Göker M, Salamov AA, Wisecaver JH, Long TM, Calvey CH, Aerts AL, Barry KW, Choi C, Clum A, Coughlan AY, Deshpande S, Douglass AP, Hanson SJ, Klenk H-P, LaButti KM, Lapidus A, Lindquist EA, Lipzen AM, Meier-Kolthoff JP, Ohm RA, Otiillar RP, Pangilinan JL, Peng Y, Rokas A, Rosa CA, Scheuner C, Sibirny AA, Slot JC, Stielow JB, Sun H, Kurtzman CP, Blackwell M, Grigoriev IV, Jeffries TW. 2016. Comparative genomics of biotechnologically important yeasts. *Proc Natl Acad Sci U S A* 113:9882–9887. <https://doi.org/10.1073/pnas.1603941113>.
8. Shen X-X, Zhou X, Kominek J, Kurtzman CP, Hittinger CT, Rokas A. 2016. Reconstructing the backbone of the *Saccharomycotina* yeast phylogeny using genome-scale data. *G3* 6:3927–3939. <https://doi.org/10.1534/g3.116.034744>.
9. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
10. 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>.
11. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The Sequence Alignment/Map Format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
12. 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 Bioinformatics* 13:237. <https://doi.org/10.1186/1471-2105-13-237>.
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. Kerscher S, Durstewitz G, Casaregola S, Gaillardin C, Brandt U. 2001. The complete mitochondrial genome of *Yarrowia lipolytica*. *Comp Funct Genomics* 2:80–90. <https://doi.org/10.1002/cfg.72>.