




Draft Genome Sequence of the Birch Tree Fungal Pathogen *Taphrina betulina* UCD315

Padraic Heneghan,^a Adam P. Ryan,^a Darragh Nimmo,^a Claudine Duggan,^a Paurush Kumar,^a Peadar O'Gaora,^a Eoin Ó'Cinnéide,^b Kevin P. Byrne,^b Kenneth H. Wolfe,^b Caoimhe E. O'Brien,^a  Geraldine Butler^a

^aSchool of Biomedical and Biomolecular Sciences, Conway Institute, University College Dublin, Dublin, Ireland

^bSchool of Medicine, Conway Institute, University College Dublin, Dublin, Ireland

ABSTRACT *Taphrina betulina* is the ascomycete yeast that causes the formation of witches' brooms in birch trees. Here, we report the first draft genome sequence of *T. betulina*, from strain UCD315, isolated from soil in Ireland. The genome is haploid and 12.5 Mb long.

Taphrina species are plant pathogens in the subphylum Taphrinomycotina of the phylum Ascomycota (1). *Taphrina* species cause plant deformity diseases in a diversity of tree species, including *Prunus* (edible fruit trees and shrubs), *Cerasus* (sour cherry), and *Populus* (poplar) (2). *Taphrina betulina* was first described in Norway in 1883 (3). It infects *Betula pubescens* (downy birch), *Betula nana* (dwarf birch), crosses between the two species (*Betula intermedia*), and *Betula pendula* (silver birch) (1, 4). Infection results in host tissue deformities, such as nest-like tumors called "witches' broom" (5). Ultimately, infection affects the diameter, height, and life span of the tree (4).

T. betulina UCD315 was isolated from soil near Lough Corrib, County Galway, Ireland (global positioning system coordinates, 53.4344816, -9.1534624). The yeast was cultured at room temperature on yeast extract-peptone-dextrose (YPD) medium with chloramphenicol (3% [wt/vol]) and ampicillin (10% [wt/vol]). The species was identified by sequencing the internal transcribed spacer (ITS) of the ribosomal DNA (rDNA) gene locus (GenBank accession number [MN540705](https://www.ncbi.nlm.nih.gov/nuclseq/MN540705)). Genomic DNA was extracted and purified using Qiagen's QIAamp DNA minikit. Libraries with an insert size of 800 bp were made from genomic DNA and sequenced by BGI Tech Solutions using an Illumina HiSeq 4000 instrument with 150-bp paired-end reads (9.5 million spots). All parameters used for sequence assembly and analysis are available at <https://www.doi.org/10.6084/m9.figshare.9775517>.

Low-quality reads (1.69 million) were trimmed using Skewer v0.2.2 (6). The genome was assembled from all reads using SPAdes v3.11.1 with the "careful" parameter (7). Based on coverage-versus-length analysis (8), contigs below 72 \times coverage or 1-kb length were removed. The results were then analyzed using QUAST v4.6.1 (9). The genome size was 12.5 Mb with an N_{50} value of 321 kb, an L_{50} value of 16, an average coverage of 101 \times , and a GC content of 49.8%. The largest contig is 682,403 bp. This is similar in size and contiguity to other sequenced *Taphrina* genomes, which range from 11.9 to 15.7 Mb (2). Using BUSCO v3.0.1 (10), genome completeness was estimated at 91.5% (compared to the Ascomycota lineage data set). This is similar to the other eight *Taphrina* genomes (2, 11), which have an average completeness of 91.95% (ranging from 89.6% to 93.1%). The mitochondrial genome was assembled as a 42.9-kb contig (GenBank accession number [VWS10100053](https://www.ncbi.nlm.nih.gov/nuclseq/VWS10100053)).

Analyses using SAMtools v1.1.19 (12), Burrows-Wheeler Aligner MEM (BWA-MEM) v0.7.12-r1039 (13), and Genome Analysis Toolkit (GATK) v4.0.1.2 (14) with default settings identified very small numbers of putative heterozygous single-nucleotide

Citation Heneghan P, Ryan AP, Nimmo D, Duggan C, Kumar P, O'Gaora P, Ó'Cinnéide E, Byrne KP, Wolfe KH, O'Brien CE, Butler G. 2019. Draft genome sequence of the birch tree fungal pathogen *Taphrina betulina* UCD315. *Microbiol Resour Announc* 8:e01255-19. <https://doi.org/10.1128/MRA.01255-19>.

Editor Antonis Rokas, Vanderbilt University

Copyright © 2019 Heneghan 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.

Received 4 October 2019

Accepted 1 November 2019

Published 27 November 2019

polymorphisms (1,385) and insertion/deletions (181), suggesting that the genome is haploid.

Data availability. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession number [VWSI00000000](https://www.ncbi.nlm.nih.gov/nuccore/VWSI00000000) and the raw reads under SRA number [SRX6812536](https://www.ncbi.nlm.nih.gov/sra/SRX6812536). These data are also available under BioProject number [PRJNA564291](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA564291). The ITS sequence is available at accession number [MN540705](https://www.ncbi.nlm.nih.gov/nuccore/MN540705), and the mitochondrial genome is available under accession number [VWSI01000053](https://www.ncbi.nlm.nih.gov/nuccore/VWSI01000053).

ACKNOWLEDGMENTS

This work was supported by undergraduate teaching resources from University College Dublin and Science Foundation Ireland (13/IA/1910). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

We thank James Mooney for collecting the soil sample.

REFERENCES

- Rodrigues MG, Fonseca A. 2003. Molecular systematics of the dimorphic ascomycete genus *Taphrina*. *Int J Syst Evol Microbiol* 53:607–616. <https://doi.org/10.1099/ijs.0.02437-0>.
- Tsai IJ, Tanaka E, Masuya H, Tanaka R, Hirooka Y, Endoh R, Sahashi N, Kikuchi T. 2014. Comparative genomics of *Taphrina* fungi causing varying degrees of tumorous deformity in plants. *Genome Biol Evol* 6:861–872. <https://doi.org/10.1093/gbe/evu067>.
- Rostrup FGE. 1883. Fortsatte Underogolser over Snyltesvampes Angreb paa Skovtraeorne. *Tidsskr Skogbruk* 6:199–300.
- Spanos YA, Woodward S. 1994. The effects of *Taphrina betulina* infection on growth of *Betula pubescens*. *Forest Pathol* 24:277–286. <https://doi.org/10.1111/j.1439-0329.1994.tb00997.x>.
- Jump BA, Woodward S. 1994. Histology of witches' brooms on *Betula pubescens*. *Forest Pathol* 24:229–237. <https://doi.org/10.1111/j.1439-0329.1994.tb00989.x>.
- 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>.
- 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>.
- Douglass AP, O'Brien CE, Offei B, Coughlan AY, Ortiz-Merino RA, Butler G, Byrne KP, Wolfe KH. 2019. Coverage-versus-length plots, a simple quality control step for *de novo* yeast genome sequence assemblies. *G3 (Bethesda)* 9:879–887. <https://doi.org/10.1534/g3.118.200745>.
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
- Waterhouse RM, Seppely M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva EV, Zdobnov EM. 2018. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol Biol Evol* 35:543–548. <https://doi.org/10.1093/molbev/msx319>.
- Cissé OH, Almeida JMGCF, Fonseca A, Kumar AA, Salojärvi J, Overmyer K, Hauser PM, Pagni M. 2013. Genome sequencing of the plant pathogen *Taphrina deformans*, the causal agent of peach leaf curl. *mBio* 4:e00055. <https://doi.org/10.1128/mBio.00055-13>.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>.
- 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>.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytzky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303. <https://doi.org/10.1101/gr.107524.110>.