



# Whole-Genome Sequencing and *De Novo* Assembly of *Malassezia pachydermatis* Isolated from the Ear Canal of a Dog with Otitis

S. D'Andreano,<sup>a,b</sup>  J. Viñes,<sup>a,b</sup>  O. Francino<sup>a</sup>

<sup>a</sup>SVGM, Molecular Genetics Veterinary Service, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

<sup>b</sup>Vetgenomics, Edifici EUREKA, PRUAB, Bellaterra, Barcelona, Spain

**ABSTRACT** We have *de novo* assembled the genome sequence of *Malassezia pachydermatis* isolated from a canine otitis sample with Nanopore-only long reads. With 99× coverage and 8.23 Mbp, the genome sequence was assembled in 10 contigs, with 6 of them corresponding to chromosomes, improving the scaffolding of previous genome assemblies for the species.

*Malassezia* (Eukaryota; Fungi; Basidiomycetes) is a lipid-dependent yeast of the skin mycobiome in humans and other mammals (1–3). The genus includes 18 closely related species described to date (4), and *M. globosa*, *M. restricta*, and *M. sympodialis* are the most common human skin residents, whereas *M. pachydermatis* is the most common species related to skin disorders in veterinary medicine (3, 5).

As of February 2021, there are only two assembled genome sequences for *M. pachydermatis*, under NCBI accession no. [GCF\\_001278385.1](https://.ncbi.nlm.nih.gov/assembly/GCF_001278385.1) (6) and [GCA\\_001264975.1](https://.ncbi.nlm.nih.gov/assembly/GCA_001264975.1) (7), which are both from the collection strain CBS 1879. There is a need to gather genomic information from *Malassezia* isolates associated with skin disorders important in animal and human health. We have assembled a high-quality *M. pachydermatis* genome sequence isolated from a dog with otitis using a *de novo* assembly and polishing strategy with Nanopore-only long reads.

Leti Animal Health kindly provided an isolate identified as *M. pachydermatis* after microbiological culture of a swab sample collected from a dog's ear canal with otitis in Barcelona, Spain. The culture was grown at 28°C in dermatophyte test medium (DTM) agar and Saboraud-chloramphenicol agar. DNA was extracted with a Biomix DNA miniprep kit (Zymo Research, USA). DNA quality and quantity were determined using a Nanodrop 2000 spectrophotometer and Qubit double-stranded DNA (dsDNA) broad-range (BR) assay kit (Fisher Scientific SL, Spain). Before whole-genome sequencing, the isolate was confirmed as *Malassezia pachydermatis* by long-amplicon Nanopore sequencing of the fungal ribosomal operon and taxonomy assignment at the species level with the cloud-based WIMP application from the EPI2ME platform (Oxford Nanopore Technologies [ONT] Ltd., UK), as previously reported (8). The sequencing library was prepared with the 1D native barcoding genomic DNA kits (EXP-NBD103 and SQK-LSK109; ONT) and sequenced in a MinION FLO-MIN106 v9.4.1 flow cell in a MinION Mk1B instrument for 21 h, using the MinKNOW v18.12.9 software.

The fast5 files were base called with Guppy v2.3.7 (ONT), with the flip-flop algorithm for improved accuracy. Reads with a quality score lower than 7 were discarded. The fastq files were demultiplexed and trimmed with Deepbiner (9). We further *de novo* assembled the genome sequence with Flye 2.6 (10) and corrected the contigs with racon (<https://github.com/lbcb-sci/racon>) and Medaka (<https://github.com/nanoporetech/medaka>). We assessed the genome completeness with Benchmarking Universal Single-Copy Orthologs (BUSCO) v4.0.1 (<https://gitlab.com/ezlab/busco>) (11) in mode genome with basidiomycota\_odb10

**Citation** D'Andreano S, Viñes J, Francino O. 2021. Whole-genome sequencing and *de novo* assembly of *Malassezia pachydermatis* isolated from the ear canal of a dog with otitis. *Microbiol Resour Announc* 10:e00205-21. <https://doi.org/10.1128/MRA.00205-21>.

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

**Copyright** © 2021 D'Andreano 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 O. Francino, [olga.francino@uab.cat](mailto:olga.francino@uab.cat).

**Received** 23 February 2021

**Accepted** 28 April 2021

**Published** 27 May 2021

**TABLE 1** *De novo* assembly of the *Malassezia pachydermatis* genome sequence isolated from a canine otitis sample<sup>a</sup>

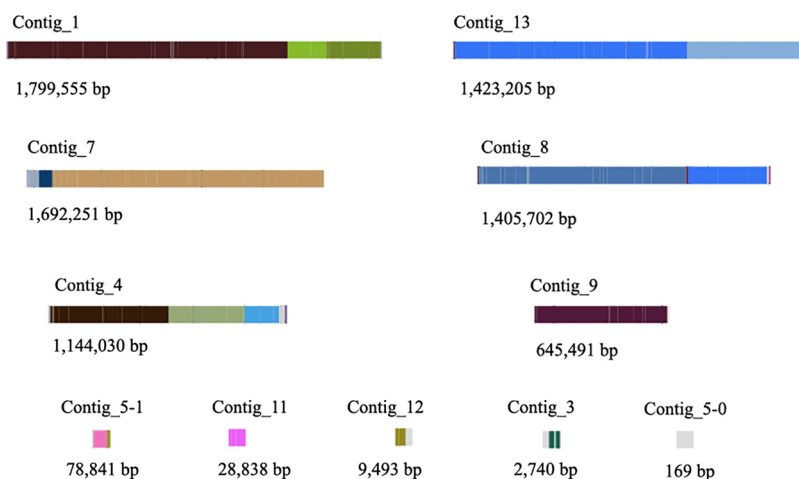
Sequence name	Sequence role	Assigned molecule	Assigned molecule location type	GenBank accession no.	Sequence length (bp)
Contig_1	Assembled molecule	1	Chromosome	CM025284.1	1,799,555
Contig_7	Assembled molecule	2	Chromosome	CM025285.1	1,692,251
Contig_13	Assembled molecule	3	Chromosome	CM025286.1	1,423,205
Contig_8	Assembled molecule	4	Chromosome	CM025287.1	1,405,702
Contig_4	Assembled molecule	5	Chromosome	CM025288.1	1,144,030
Contig_9	Assembled molecule	6	Chromosome	CM025289.1	645,491
Contig_11	Unplaced scaffold	NA	NA	JABNMX010000001.1	28,388
Contig_12	Unplaced scaffold	NA	NA	JABNMX010000002.1	9,493
Contig_3	Unplaced scaffold	NA	NA	JABNMX010000005.1	2,740
Contig_5	Unplaced scaffold	NA	NA	JABNMX010000007.1	74,481

<sup>a</sup>BioProject PRJNA631787, assembly GCA\_014493585.1, accession no. JABNMX01, and isolate M13-UAB2019. NA, not applicable.

(20 November 2019; no. of species, 133; no. of BUSCOs, 1,764). We used Bandage (12) for contig visualization and comparison with the CBS 1879 *M. pachydermatis* reference genome (GCF\_001278385.1) (6).

Nanopore sequencing generated 265,298 reads with an  $N_{50}$  value of 5,648 bp and allowed successful *de novo* assembly and polishing of the genome sequence of *M. pachydermatis* (Table 1). Coverage was 99 $\times$ , with a GC content of 54.54%. The genome was 8.23 Mbp in 10 contigs (longest contig, 1.799 Mbp;  $N_{50}$ , 1.423 Mbp), with the longest ones corresponding to the 6 expected chromosomes (13) (Fig. 1). The average nucleotide identity (ANI) (14) was 96.3% with the *M. pachydermatis* reference genome (GCF\_001278385.1), confirming the isolate as the same species (15, 16). The 76.6% BUSCO completeness (C,76.6% [S,76.5%; D,0.1%]; F,3.8%; M,19.6%; n,1,764) is similar to the *M. pachydermatis* reference genome with short reads (6, 7) and with the same duplication of the glutathione reductase gene as the reference genome.

**Data availability.** The standardized strain descriptions and accession numbers are presented in Table 1; the genome assemblies and genomic data are publicly available in DDBJ/ENA/GenBank under BioProject no. PRJNA631787 with the assembly accession no. GCA\_014493585.1 and genome accession no. JABNMX01. The version described in this paper is the first version. The raw data are available from the Sequence Read Archive (SRA) under the accession no. SRR13764144. Raw data (fast5) are available at Zenodo (<https://zenodo.org/record/3874381#.YJQ1rLVKi70>).



**FIG 1** Bandage visualization of the genome assembly for *Malassezia pachydermatis* (GenBank accession no. GCA\_014493585.1). Contig\_1, Contig\_13, Contig\_7, Contig\_8, Contig\_4, and Contig\_9 correspond to the six chromosomes of *M. pachydermatis*. Contigs are colored relative to the reference assembly for the CBS 1879 strain (GCF\_001278385.1) (6) and indicate the merge of contigs that are separate from the CBS reference genome assembly with short reads.

## ACKNOWLEDGMENTS

Funding was provided by Industrial Doctorate Program grants 2015DI044 to S.D. and 2017DI037 to J.V., AGAUR, Generalitat de Catalunya, Spain.

O.F. has received unrelated honoraria for lecturing from Oxford Nanopore Technologies.

## REFERENCES

- Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, Schoenfeld D, Nomicos E, Park M, Kong HH, Segre JA, NIH Intramural Sequencing Center Comparative Sequencing Program. 2013. Topographic diversity of fungal and bacterial communities in human skin. *Nature* 498:367–370. <https://doi.org/10.1038/nature12171>.
- Oh J, Byrd A, Deming C, Conlan S, Kong H, Segre J, NISC Comparative Sequencing Program. 2014. Biogeography and individuality shape function in the human skin metagenome. *Nature* 514:59–64. <https://doi.org/10.1038/nature13786>.
- Cabañas FJ. 2014. *Malassezia* yeasts: how many species infect humans and animals? *PLoS Pathog* 10:e1003892. <https://doi.org/10.1371/journal.ppat.1003892>.
- Theelen B, Cafarchia C, Gaitanis G, Bassukas ID, Boekhout T, Dawson TL. 2018. *Malassezia* ecology, pathophysiology, and treatment. *Med Mycol* 56:S10–S25. <https://doi.org/10.1093/mmy/myx134>.
- Coelho MA, Sampaio JP, Gonçalves P. 2013. Living and thriving on the skin: *Malassezia* genomes tell the story. *mBio* 4:e00117-13. <https://doi.org/10.1128/mBio.00117-13>.
- Triana S, González A, Ohm RA, Wösten HA, de Cock H, Restrepo S, Celis A. 2015. Draft genome sequence of the animal and human pathogen *Malassezia pachydermatis* strain CBS 1879. *Genome Announc* 3:e01197-15. <https://doi.org/10.1128/genomeA.01197-15>.
- Wu G, Zhao H, Li C, Rajapakse MP, Wong WC, Xu J, Saunders CW, Reeder NL, Reilman RA, Scheynius A, Sun S, Billmyre BR, Li W, Averette AF, Mieczkowski P, Heitman J, Theelen B, Schröder MS, De Sessions PF, Butler G, Maurer-Stroh S, Boekhout T, Nagarajan N, Dawson TL, Jr. 2015. Genus-wide comparative genomics of *Malassezia* delineates its phylogeny, physiology, and niche adaptation on human skin. *PLoS Genet* 11:e1005614. <https://doi.org/10.1371/journal.pgen.1005614>.
- D'Andreano S, Cusco A, Francino O. 2020. Rapid and real-time identification of fungi up to species level with long amplicon nanopore sequencing from clinical samples. *Biol Methods Protoc* 6:bpaa026. <https://doi.org/10.1093/biomethods/bpaa026>.
- Wick RR, Judd LM, Holt KE. 2018. Deepbinner: demultiplexing barcoded Oxford Nanopore reads with deep convolutional neural networks. *PLoS Comput Biol* 14:e1006583. <https://doi.org/10.1371/journal.pcbi.1006583>.
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–549. <https://doi.org/10.1038/s41587-019-0072-8>.
- Seppy M, Manni M, Zdobnov EM. 2019. BUSCO: assessing genome assembly and annotation completeness. In Kollmar M (ed), *Gene prediction: methods in molecular biology*, vol 1962. Humana, New York, NY.
- Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 31:3350–3352. <https://doi.org/10.1093/bioinformatics/btv383>.
- Sankaranarayanan SR, Ianiri G, Coelho MA, Reza MH, Thimmappa BC, Ganguly P, Vadnala RN, Sun S, Siddharthan R, Tellgren-Roth C, Dawson TL, Heitman J, Sanyal K. 2020. Loss of centromere function drives karyotype evolution in closely related *Malassezia* species. *Elife* 9:e53944. <https://doi.org/10.7554/eLife.53944>.
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>.
- Lachance MA, Lee DK, Hsiang T. 2020. Delineating yeast species with genome average nucleotide identity: a calibration of ANI with haplontic, heterothallic *Metschnikowia* species. *Antonie Van Leeuwenhoek* 113:2097–2106. <https://doi.org/10.1007/s10482-020-01480-9>.
- Gostinčar C. 2020. Towards genomic criteria for delineating fungal species. *J Fungi* 6:246. <https://doi.org/10.3390/jof6040246>.