



# Draft Genome Sequences of Three Clinical Isolates of the Pathogenic Yeast *Candida glabrata*

Pedro Pais,<sup>a,b</sup> Mónica Galocha,<sup>a,b</sup> Isabel M. Miranda,<sup>c,d</sup> Acácio G. Rodrigues,<sup>c,d</sup>  Miguel C. Teixeira<sup>a,b</sup>

<sup>a</sup>Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

<sup>b</sup>Institute for Bioengineering and Biosciences (IBB), Biological Sciences Research Group, Instituto Superior Técnico, Lisbon, Portugal

<sup>c</sup>Department of Pathology, Division of Microbiology, Faculty of Medicine, University of Porto, Porto, Portugal

<sup>d</sup>CINTESIS-Center for Health Technology and Services Research, Faculty of Medicine, University of Porto, Porto, Portugal

**ABSTRACT** Here, we report the draft genome sequences of three *Candida glabrata* clinical isolates, 040, 044, and OL152. The isolates were recovered from patients admitted to Centro Hospitalar de S. João (CHSJ) in Porto, Portugal. Isolates 040 and 044 were taken from blood samples, while isolate OL152 was collected from urine.

*Candida glabrata*, a yeast belonging to the *Nakaseomyces* genus from the Saccharomycetes class and Ascomycota phylum, is a major cause of candidemia worldwide (1). In this announcement, we present the draft genome sequences of three clinical isolates of *C. glabrata*. These draft genome sequences will contribute to our understanding of the genetic variability among *C. glabrata* isolates and provide clues on genomic evolution in this species. These isolates are part of a collection previously screened for azole resistance (2, 3). Isolate 044 was previously used to study azole resistance acquisition *in vitro* (4).

Isolates were grown overnight in liquid yeast extract-peptone-dextrose (YPD) medium at 30°C with orbital shaking (250 rpm) and reinoculated in fresh YPD medium until early log phase. Genomic DNA (gDNA) was extracted using the NZY microbial gDNA isolation kit (NZYTech), following the manufacturer's instructions. Genomic DNA was sequenced on an Illumina HiSeq X platform, producing 2 × 150-bp paired-end reads. Library preparation (Nextera XT) and sequencing were carried out by Admera Health, LLC.

Illumina sequencing produced 79,811,816, 93,096,166, and 81,058,166 raw paired-end reads for isolates 040, 044, and OL152, respectively. Low-quality bases and adapters were removed using Trimmomatic (v0.38) (5). Read duplicates were removed using PRINSEQ (v0.20.4) (6). Ultimately, 66,145,112, 76,825,456, and 67,787,160 high-quality reads were used for subsequent analysis for 040, 044, and OL152, respectively.

The draft genomes were assembled into scaffolds using SPAdes (v3.12.0) (7). Scaffolds smaller than 500 bp were filtered out, and the remaining sets of scaffolds were used as draft assemblies (BioProject no. [PRJNA525402](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA525402)). Assembly quality was analyzed using Quality Assessment Tool for Genome Assemblies (QUAST; v4.6.3) (8). The 040 assembly consists of 38.54% GC content, 202 scaffolds, a total length of 12.24 Mb, an  $N_{50}$  value of 400 kb, a longest scaffold of 854 kb, and an average coverage of 584×. For 044, the assembly consists of 38.51% GC content, 158 scaffolds, a total length of 12.24 Mb, an  $N_{50}$  value of 404 kb, a longest scaffold of 1.3 Mb, and an average coverage of 678×. For OL152, the assembly consists of 38.52% GC content, 207 scaffolds, a total length of 12.25 Mb, an  $N_{50}$  value of 272 kb, a longest scaffold of 1.14 Mb, and an average coverage of 583×. The assemblies align 97.6% to 98% against the reference genome (*C. glabrata* CBS138) (9) obtained from the *Candida* Genome Database (<http://www.candidagenome.org/>).

**Citation** Pais P, Galocha M, Miranda IM, Rodrigues AG, Teixeira MC. 2019. Draft genome sequences of three clinical isolates of the pathogenic yeast *Candida glabrata*. *Microbiol Resour Announc* 8:e00278-19. <https://doi.org/10.1128/MRA.00278-19>.

**Editor** Vincent Bruno, University of Maryland School of Medicine

**Copyright** © 2019 Pais 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 Miguel C. Teixeira, [mnpct@tecnico.ulisboa.pt](mailto:mnpct@tecnico.ulisboa.pt).

**Received** 9 March 2019

**Accepted** 26 April 2019

**Published** 29 August 2019

Genome annotation was performed by submitting the assemblies to the Yeast Genome Annotation Pipeline (YGAP), based on the Yeast Gene Order Browser (v7) (10). Annotation was performed specifying a post-whole-genome duplication (post-WGD) species and predicted 5,279 genes in 040, 5,278 genes in 044, and 5,287 genes in OL152.

Single nucleotide polymorphisms (SNPs) were identified against the reference genome using the Burrows-Wheeler Aligner (BWA; v0.7.17) (11) and the Genome Analysis Toolkit HaplotypeCaller (v4.0.8.1) (12). Low-quality variants were filtered out using BCFtools (v1.9) (13). The number of variants identified were 28,270, 37,158, and 37,743 for 040, 044, and OL152, respectively.

**Data availability.** This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession no. [SKBI000000000](https://www.ncbi.nlm.nih.gov/nuccore/SKBI000000000), [SKBJ000000000](https://www.ncbi.nlm.nih.gov/nuccore/SKBJ000000000), and [SKBK000000000](https://www.ncbi.nlm.nih.gov/nuccore/SKBK000000000). The versions described in this paper are the first versions, SKBI01000000, SKBJ01000000, and SKBK01000000. Raw reads are available at [SRR8953802](https://www.ncbi.nlm.nih.gov/sra/SRR8953802), [SRR8953803](https://www.ncbi.nlm.nih.gov/sra/SRR8953803), and [SRR8953804](https://www.ncbi.nlm.nih.gov/sra/SRR8953804).

## ACKNOWLEDGMENTS

This work was financially supported by Fundação para a Ciência e a Tecnologia (FCT) contracts PTDC/BBB-BIO/4004/2014 and PTDC/BII-BIO/28216/2017 and Ph.D. scholarships to P.P. and M.G., associated with the BIOTECNico and AEM FCT Ph.D. programs. Funding received from the FCT (grant UID/BIO/04565/2013) and Programa Operacional Regional de Lisboa 2020 (project no. 007317) is also acknowledged.

## REFERENCES

- Azie N, Neofytos D, Pfaller M, Meier-Kriesche H-U, Quan S-P, Horn D. 2012. The PATH (Prospective Antifungal Therapy) Alliance registry and invasive fungal infections: update 2012. *Diagn Microbiol Infect Dis* 73: 293–300. <https://doi.org/10.1016/j.diagmicrobio.2012.06.012>.
- Faria-Ramos I, Neves-Maia J, Ricardo E, Santos-Antunes J, Silva AT, Costa-de-Oliveira S, Cantón E, Rodrigues AG, Pina-Vaz C. 2014. Species distribution and in vitro antifungal susceptibility profiles of yeast isolates from invasive infections during a Portuguese multicenter survey. *Eur J Clin Microbiol Infect Dis* 33:2241–2247. <https://doi.org/10.1007/s10096-014-2194-8>.
- Costa C, Ribeiro J, Miranda IM, Silva-Dias A, Cavalheiro M, Costa-de-Oliveira S, Rodrigues AG, Teixeira MC. 2016. Clotrimazole drug resistance in *Candida glabrata* clinical isolates correlates with increased expression of the drug:H(+) antiporters CgAqr1, CgTpo1\_1, CgTpo3, and CgQdr2. *Front Microbiol* 7:526. <https://doi.org/10.3389/fmicb.2016.00526>.
- Cavalheiro M, Costa C, Silva-Dias A, Miranda IM, Wang C, Pais P, Pinto SN, Mil-Homens D, Sato-Okamoto M, Takahashi-Nakaguchi A, Silva RM, Mira NP, Fialho AM, Chibana H, Rodrigues AG, Butler G, Teixeira MC. 2019. A transcriptomics approach to unveiling the mechanisms of in vitro evolution towards fluconazole resistance of a *Candida glabrata* clinical isolate. *Antimicrob Agents Chemother* 63:e00995-18. <https://doi.org/10.1128/AAC.00995-18>.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
- 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>.
- 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>.
- Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, de Montigny J, Marck C, Neuvéglise C, Talla E, Goffard N, Frangeul L, Aigle M, Anthouard V, Babour A, Barbe V, Barnay S, Blanchin S, Beckerich J-M, Beyne E, Bleykasten C, Boisramé A, Boyer J, Cattoico L, Confanioleri F, de Daruvar A, Despons L, Fabre E, Fairhead C, Ferry-Dumazet H, Groppi A, Hantraye F, Hennequin C, Jauniaux N, Joyet P, Kachouri R, Kerrest A, Koszul R, Lemaire M, Lesur I, Ma L, Muller H, Nicaud J-M, Nikolski M, Oztas S, Ozier-Kalogeropoulos O, Pellenz S, Potier S, Richard G-F, Straub M-L, et al. 2004. Genome evolution in yeasts. *Nature* 430:35–44. <https://doi.org/10.1038/nature02579>.
- 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>.
- 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, Kernysky 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>.
- Narasimhan V, Danecsek P, Scally A, Xue Y, Tyler-Smith C, Durbin R. 2016. BCFtools/RoH: a hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics* 32: 1749–1751. <https://doi.org/10.1093/bioinformatics/btw044>.