Mycopathologia (2023) 188:821–823
https://doi.org/10.1007/s11046-023-00772-8

Draft Genome Sequence of a Primate Isolate of *Kazachstania pintolopesii*

Received: 23 May 2023 / Accepted: 3 July 2023 / Published online: 17 August 2023

Abstract *Kazachstania pintolopesii* is an opportunistic mammalian pathobiont from the *K. telluris* species complex. No draft genomes of this species are currently available. Here, we report the first draft genome sequence of a primate isolate of *K. pintolopesii* (NCYC 4417).

Keywords *Kazachstania pintolopesii* · Draft genome · Non-human primate · Cynomolgus macaque · Fungal pathobiont · Gut mycobiota

Kazachstania pintolopesii is an ascomycetous yeast and relative of *K. bovina*, *K. heterogenica*, *K. slooffiae* and *K. telluris* [1]. Collectively, these species constitute the *K. telluris* complex, a phylogenetically distinct group of thermotolerant yeasts able to grow at 37 °C and above [1, 2]. They are widely distributed [1], but are found predominantly in the gastrointestinal (GI) tracts and nasal passages of birds and mammals [1–3]. Rodents are the principal hosts for *K. pintolopesii* [1, 2, 4] with a recent study extending its host range to include cynomolgus macaques [5]. These fungi are also considered pathobionts causing infections in rodents, primates, and less frequently in humans [2, 6, 7]. *K. pintolopesii* is in addition to being a gut commensal also associated with fatal infections in laboratory mice [2], and with ankylosing spondylitis in cynomolgus macaques [8]. Draft genomes for three members of the complex have been published [9–11], but despite its potential significance as a pathobiont no *K. pintolopesii* genomes have been published to date.

Here, we have combined short- and long-read sequencing to obtain the genome sequence of *K. pintolopesii* NCYC 4417, a faeces-derived isolate from a captive adult macaque. A faecal homogenate was prepared in sterile phosphate-buffered saline (PBS) and cultured onto Sabouraud dextrose (SD) agar plates containing penicillin (25 U/mL) and streptomycin (25 U/mL) at 37 °C. Species identity, from single colonies, was determined by PCR amplification and Sanger sequencing of the ribosomal DNA internal transcribed spacer 1 (ITS1) region of the ribosomal DNA locus using primers ITS1F [12] and ITS2 [13]. The ITS1 sequence of strain NCYC 4417 (GenBank accession number PRJEB63679) is 99.7% identical to that of the *K. pintolopesii* type strain CBS 2985 (GenBank accession number NR_155233).

For short- and long-read sequencing, total genomic DNA was extracted from a stationary phase SD culture using a MasterPure Yeast DNA Purification Kit (Cambio, Cambridge, UK). Short-read Illumina sequencing was performed using a modified 20-fold dilution of DNA Prep (Flex) reagent and run on a NextSeq 500 sequencer, producing 9,108,736 paired-end 150-bp reads (∼ 186 × coverage). Nanopore sequencing was performed using a MinION sequencer (Oxford Nanopore Technologies, ONT), ligation sequencing kit SQK-LSK109 (ONT) and flow cell FLO-MIN106 R9.4.1 (ONT). This produced a total of 437,446 reads with an average read length of 4,069 bases (∼ 127 × coverage).

Base calling was performed using Guppy v.6.1.2 (basecall_model_version_id = 2021–05-17_dna_r9.4.1_minion_384_d37a2ab9). Raw short- and long-read polishing, including the removal of adapters and low-quality bases, was performed using SeqFu 1.16 [14] and fastp 0.23 [15]. The genome was assembled using Flye 2.9.1 [16] and polished with one round of Pilon 1.24 [17]. The genome assembly comprised of 34 contigs with the largest being 1,738,127 bp in length. The total size of the genome was 13,992,981 bp, the N_{50} value was 947 kb, and the G + C content was 30.63%. In addition, a putative 8-nt telomeric repeat was identified (5'-WGTATGGG-3'), similar in sequence to the canonical eukaryotic telomere motif [18], and present in 50–100 tandem copies at one or both termini of 13 contigs. Augustus v.3.3.3 [19] predicted 4884 protein-coding genes using the *Saccharomyces cerevisiae* (S288C) training data set, and 196 tRNA genes were detected using tRNAscan-SE 2.0 [20]. Genome completeness was estimated as 91.0% using BUSCO v5.4.4 [21]. Dependencies and scripts are available at https://github.com/quadram-institute-bioscience/ont-candida.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by SAJ, AP, CP, AT, DB, RH, SH, SGPF and SRC. The first draft of the manuscript was written by SAJ and SRC, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This research was funded in part by a Biotechnology and Biological Sciences Research Council (BBSRC) awarded Core Capability Grant BB/CCG1860/1, project grant BB/N000250/1, and Institute Strategic Programme Grant Gut Microbes and Health (BB/R012490/1) and its constituent projects BBS/E/F/000PR10353, BBS/E/F/000PR10355.

Data availability This whole-genome shotgun project has been deposited at DDJB/ENA/Genbank (Biosample Number SAMEA112953918, BioProject Number PRJEB61520, and Assembly Accession Number GCA_950065675.1). The version described in this paper is version 1. The raw reads were deposited at SRA (Accession Numbers ERR11267923 and ERR11267961).This article was written according to MycopathologiaGENOMES checklist requirements [22].

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethics Approval This article does not include human participants, or any procedures carried out on animals.

Author names and affiliations

[1]Steve A. James
[1]Aimee Parker
[1]Catherine Purse
[1]Andrea Telatin
[1]David Baker
[1]Rhiannon Evans
[2]Sandy Holmes
[1,2]Simon G. P. Funnell
[1,3]Simon R. Carding

[1]Food, Microbiome and Health, Quadram Institute Bioscience, Norwich, UK
e-mail: steve.james@quadram.ac.uk

[2]UK Health Security Agency, Porton Down, Salisbury, UK

[3]Norwich Medical School, University of East Anglia, Norwich, UK

References

1. Vaughan-Martini A, Lachance M-A, Kurtzman CP. Kazachstania Zubkova (1971), p 439–470. *In* Kurtzman CP, Fell JW, Boekhout T (ed), The yeasts: A taxonomic study, 5 ed, vol 2. Elsevier, Amsterdam, 2021.
2. Kurtzman CP, Robnett CJ, Ward JM, Brayton C, Gorelick P, Walsh TJ. Multigene phylogenetic analysis of pathogenic Candida species in the Kazachstania (Arxiozyma) telluris complex and description of their ascosporic states as Kazachstania bovina sp nov., K-heterogenica sp nov., K-pintolopesii sp nov., and K-slooffiae sp nov. J Clin Microbiol. 2005;43:101–11. https://doi.org/10.1128/JCM.43.1.101-111.2005.
3. Arfken AM, Frey JF, Ramsay TG, Summers KL. Yeasts of burden: Exploring the Mycobiome-Bacteriome of the Piglet GI Tract. Front Microbiol. 2019;10:2286. https://doi.org/10.3389/fmicb.2019.02286.

4. Bendova B, Pialek J, Dureje L, Schmiedova L, Cizkova D, Martin JF, Kreisinger J. How being synanthropic affects the gut bacteriome and mycobiome: comparison of two mouse species with contrasting ecologies. BMC Microbiol. 2020;20:194. https://doi.org/10.1186/s12866-020-01859-8.
5. James SA, Parker A, Purse C, Telatin A, Baker D, Holmes S, Durham J, Funnell SGP, Carding SR. The Cynomolgus Macaque intestinal mycobiome is dominated by the Kazachstania Genus and K. pintolopesii species. J Fungi. 2022; 8:1054. https://doi.org/10.3390/jof8101054.
6. Alvarez-Perez S, Mateos A, Dominguez L, Martinez-Nevado E, Rodriguez-Bertos A, Blanco JL, Garcia ME. First isolation of the anamorph of Kazachstania heterogenica from a fatal infection in a primate host. Med Mycol. 2012;50:193–6. https://doi.org/10.3109/13693786.2011.578155.
7. Brunet K, Minoza A, Rammaert B, Portet-Sulla V, Hubert F, Lorenzo JC, Rodier MH, Cateau E. Invasive Candida bovina infection, France. Emerging Infect Dis. 2020;26:626–7. https://doi.org/10.3201/eid2603.191371.
8. Zhang HT, Wei Y, Jia HH, Chen DL, Tang XC, Wang J, Chen ML, Guo YR. Immune activation of characteristic gut mycobiota Kazachstania pintolopesii on IL-23/IL-17R signaling in ankylosing spondylitis. Front Cellular Infection Microbiol. 2022;12:1035366. https://doi.org/10.3389/fcimb.2022.1035366.
9. Morio F, O'Brien CE, Butler G. Draft genome sequence of the yeast Kazachstania telluris CBS 16338 isolated from forest soil in Ireland. Mycopathologia. 2020;185:587–90. https://doi.org/10.1007/s11046-020-00449-6.
10. Davies CP, Arfken AM, Frey JF, Summers KL. Draft genome sequence of Kazachstania slooffiae, isolated from post-weaning piglet feces. Microbiol Resource Announce. 2021;10:e00198-e221. https://doi.org/10.1128/MRA.00198-21.
11. Deroche L, Buyck J, Cateau E, Rammaert B, Marchand S, Brunet K. Draft genome sequence of Kazachstania bovina yeast isolated from human infection. Mycopathologia. 2022;187:413–5. https://doi.org/10.1007/s11046-022-00639-4.
12. Gardes M, Bruns TD. ITS primers with enhanced specificity for basidiomycetes–application to the identification of mycorrhizae and rusts. Mol Ecol. 1993;2:113–8. https://doi.org/10.1111/j.1365-294x.1993.tb00005.x.
13. White TJ, Bruns TD, Lee SL, Taylor JW. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, editors. PCR protocols: a guide to methods and applications. San Diego: Academic Press; 1990. p. 315–22.
14. Telatin A, Fariselli P, Birolo G. SeqFu: A Suite of Utilities for the Robust and Reproducible Manipulation of Sequence Files. Bioeng-Basel 2021;8:59. https://doi.org/10.3390/bioengineering8050059.
15. Chen SF, Zhou YQ, Chen YR, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90. https://doi.org/10.1093/bioinformatics/bty560.
16. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol. 2019;37:540–6. https://doi.org/10.1038/s41587-019-0072-8.
17. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng QD, Wortman J, Young SK, Earl AM. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. Plos One. 2014;9:e112963. https://doi.org/10.1371/journal.pone.0112963.
18. Cervenak F, Sepsiova R, Nosek J, Tomaska L. Step-by-step evolution of telomeres: Lessons from yeasts. Genome Biol Evolut. 2021;13:evaa268. https://doi.org/10.1093/gbe/evaa268.
19. Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics. 2008;24:637–44. https://doi.org/10.1093/bioinformatics/btn013.
20. Lowe TM, Eddy SR. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 1997;25:955–64. https://doi.org/10.1093/nar/25.5.955.
21. Manni M, Berkeley MR, Seppey M, Simao FA, Zdobnov EM. BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021;38:4647–54. https://doi.org/10.1093/molbev/msab199.
22. Mac Aogain M, Chaturvedi V, Chotirmall SH. MycopathologiaGENOMES: The New 'Home' for the publication of fungal genomes. Mycopathologia. 2019;184:551–4. https://doi.org/10.1007/s11046-019-00366-3.