



# Draft Whole-Genome Sequence of *Penicillium simplicissimum* A4, a Putative Endophyte from *Echium plantagineum*

Stacey Fisher,<sup>a</sup> Dewald De Villiers,<sup>b</sup> Morne Du Plessis,<sup>c</sup> Kacey Hattingh,<sup>b</sup> Crishe Saulse,<sup>b</sup> Gerhard Basson,<sup>d</sup> Adélé Barker,<sup>b</sup> Augustine Innalegwu Daniel,<sup>a</sup> Ali Al-Hashimi,<sup>a</sup> Arina Hitzeroth,<sup>e</sup>  Thulani Makhalanyane,<sup>f</sup> Vuyo Mavumengwana,<sup>g</sup> Arun Gokul,<sup>h</sup>  Marshall Keyster,<sup>b</sup>  Ashwil Klein<sup>a</sup>

<sup>a</sup>Plant Omics Laboratory, Department of Biotechnology, University of the Western Cape, Bellville, South Africa

<sup>b</sup>Environmental Biotechnology Laboratory, Department of Biotechnology, University of the Western Cape, Bellville, South Africa

<sup>c</sup>Sequencing Core Facility, National Institute for Communicable Diseases, Johannesburg, South Africa

<sup>d</sup>Plant Biotechnology Research Group, Department of Biotechnology, University of the Western Cape, Bellville, Western Cape, South Africa

<sup>e</sup>Institute for Microbial Biotechnology and Metagenomics, University of the Western Cape, Bellville, Western Cape, South Africa

<sup>f</sup>Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa

<sup>g</sup>Centre for Tuberculosis Research, South African Medical Research Council, Department of Biotechnology, University of the Western Cape, Bellville, Western Cape, South Africa

<sup>h</sup>Department of Plant Sciences, University of the Free State, Phuthadhithjaba, South Africa

**ABSTRACT** We report the draft whole-genome sequence of the putative endophytic fungus *Penicillium simplicissimum* A4, isolated from the roots of *Echium plantagineum* plants. The genome was sequenced using PacBio technology with an estimated genome size of 39 Mb.

*Penicillium simplicissimum* is an endophytic fungus belonging to the *Aspergillaceae* family (1, 2). Members of the *Aspergillaceae* family are widely distributed and have been isolated from diverse habitats including soil, vegetation, and food products (3–5).

*Echium plantagineum* plants were collected from Durbanville, South Africa (33°51'47.55S, 18°39'17.2E), in March 2019. Root material was surface sterilized with 70% ethanol for 3 min, 2.5% sodium perchlorate for 5 min, and 70% ethanol for 1 min and washed 5 times with reverse osmosis (RO) water (6). Sterilized roots were homogenized in a 0.9% saline (NaCl) solution using a sterile mortar and pestle. The homogenate was diluted (1:100), plated onto potato dextrose agar (PDA) containing 100 µg/mL chloramphenicol, and incubated at 25°C for 7 days. A single colony was picked, restreaked onto PDA, and incubated at 25°C for 7 days. Approximately 100 mg of biomass was collected, and total genomic DNA was extracted using the Zymo Research Quick-DNA Fungal-Bacterial Miniprep kit (Zymo Research, catalog number D6005). The purity and concentration of the DNA were measured using a NanoDrop 2000 spectrophotometer (Thermo Scientific). The isolate A4 was identified as *Penicillium simplicissimum* (99% identity to accession [KY315584.1](#)) based on the DNA sequence of the internal transcribed spacer (ITS) region (ITS1 and ITS4), which has been deposited in GenBank under the accession no. [OP512539](#).

Genomic DNA sequencing was performed at Inqaba Biotechnical Industries (Pretoria, South Africa) with the single-molecule real-time (SMRT) technology from PacBio. The genomic DNA (~1 µg) was size selected using the AMPure PB bead size selection kit (Pacific Biosciences, Menlo Park, CA). The HiFi SMRTbell library was constructed using the SMRTbell template preparation kit v3.0 and SMRTlink software v11 according to the manufacturer's instructions. The HiFi SMRTbell library was sequenced at Inqaba Biotechnical Industries (Pty) Ltd. (Pretoria, South Africa) using the PacBio Sequel IIe system in circular consensus sequencing (CCS) mode with 110× coverage per the manufacturer's instructions, resulting in a total read length of 531,767 bp and an  $N_{50}$  read of 5,000 kb.

The primary genome assembly was performed with Canu v2.2 (7) with default

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

**Copyright** © 2022 Fisher et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

Address correspondence to Ashwil Klein, [aklein@uwc.ac.za](mailto:aklein@uwc.ac.za).

The authors declare no conflict of interest.

**Received** 18 August 2022

**Accepted** 9 October 2022

**Published** 26 October 2022

parameters and a genome size setting at 35 Mb. The assembly resulted in a total genome size of 39 Mb with 50.61% GC content, 110× sequencing coverage, 196 contigs,  $N_{50}$  contig length of 3,965.722 kb, and a maximum contig size of 6,478,589 bp. The genome assembly was validated using BUSCO v. 5.4.2 (8), with default parameters, which showed a 96% hit against 4,191 fungal genes. Gene prediction was performed using Augustus 3.4.0 (9), with default parameters against *Aspergillus nidulans* as reference, and produced 11,862 protein-coding genes.

**Data availability.** The assembled genome sequence of *Penicillium simplicissimum* isolate A4 is available at GenBank under the accession number [JAMZTW000000000](https://www.ncbi.nlm.nih.gov/nuccore/JAMZTW000000000) with BioProject number [PRJNA855613](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA855613), BioSample number [SAMN29498895](https://www.ncbi.nlm.nih.gov/biosample/SAMN29498895), and SRA accession number [SRR20281321](https://www.ncbi.nlm.nih.gov/sra/SRR20281321).

## ACKNOWLEDGMENTS

This study was supported by the University of the Western Cape. Funding was provided by the National Research Foundation (NRF), the Department of Science and Innovation (DSI), and the Technology Innovation Agency (TIA).

## REFERENCES

1. Nguyen TT, Kwan Noh KJ, Lee HB. 2021. New species and eight undescribed species belonging to the families *Aspergillaceae* and *Trichocomaceae* in Korea. *Mycobiology* 49:534–550. <https://doi.org/10.1080/12298093.2021.1997461>.
2. Wakelin SA, Warren RA, Harvey PR, Ryder MH. 2004. Phosphate solubilization by *Penicillium* spp. closely associated with wheat roots. *Biol Fertil Soils* 40:36–43. <https://doi.org/10.1007/s00374-004-0750-6>.
3. Samson RA, Visagie CM, Houbraeken J, Hong SB, Hubka V, Klaassen CH, Perrone G, Seifert KA, Susca A, Tanney JB, Varga J, Kocsube S, Szigeti G, Yaguchi T, Frisvad J. 2014. Phylogeny, identification and nomenclature of the genus *Aspergillus*. *Stud Mycol* 78:141–173. <https://doi.org/10.1016/j.simyco.2014.07.004>.
4. Visagie CM, Hirooka Y, Tanney JB, Whitfield E, Mwange K, Meijer M, Amend AS, Seifert KA, Samson RA. 2014. *Aspergillus*, *Penicillium* and *Talaromyces* isolated from house dust samples collected around the world. *Stud Mycol* 78:63–139. <https://doi.org/10.1016/j.simyco.2014.07.002>.
5. Yilmaz N, Visagie CM, Houbraeken J, Frisvad JC, Samson RA. 2014. Polyphasic taxonomy of the genus *Talaromyces*. *Stud Mycol* 78:175–341. <https://doi.org/10.1016/j.simyco.2014.08.001>.
6. Istain B, Mokgokong PS, Wu R, Mia J, Gokul A, Mendoza-Cozatl D, Du Plessis M, Keyster M. 2021. Draft genome sequence of the putative endophytic bacterium *Pantoea agglomerans* R6, associated with *Lactuca serriola* from South Africa. *Microbiol Resour Announc* 10:e00023-21. <https://doi.org/10.1128/MRA.00023-21>.
7. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res* 27:722–736. <https://doi.org/10.1101/gr.215087.116>.
8. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* 38:4647–4654. <https://doi.org/10.1093/molbev/msab199>.
9. Stanke M, Steinkamp R, Waack S, Morgenstern B. 2004. AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res* 32:W309–W312. <https://doi.org/10.1093/nar/gkh379>.