High-Quality Draft Genome Sequence of the Microcolonial Black Fungus *Aeminium ludgeri* DSM 106916

João Trovão,[a] Igor Tiago,[a] Fabiana Soares,[a] Diana Sofia Paiva,[b] Nuno Mesquita,[a] Catarina Coelho,[a] Lídia Catarino,[c] Francisco Gil,[d] António Portugal[a]

[a]Centre for Functional Ecology–Science for People & the Planet, Life Sciences Department, University of Coimbra, Coimbra, Portugal
[b]Fitolab–Laboratory for Plant Health, Instituto Pedro Nunes, Coimbra, Portugal
[c]Geosciences Center, Earth Sciences Department, University of Coimbra, Coimbra, Portugal
[d]CFisUC–Center for Physics of the University of Coimbra, Coimbra, Portugal

ABSTRACT *Aeminium ludgeri* is an extremotolerant microcolonial black fungus isolated from a biodeteriorated limestone art piece in the Old Cathedral of Coimbra, Portugal (a UNESCO World Heritage Site). The high-quality draft genome sequence of *Aeminium ludgeri* presented here represents the first sequenced genome for both the recently described fungal family *Aeminiaceae* and the genus *Aeminium*.

Microcolonial black fungi are a diverse group of slow-growing fungi that pose three major problems when they colonize historical stone monuments, esthetic, biophysical, and biochemical biodeterioration (1–3). Microcolonial black fungi usually display high resistance to various extreme environmental conditions, being considered one of most resistant groups of eukaryotic organisms and a serious challenge in the field of biodeterioration of cultural heritage materials (4). Recently, the family *Aeminiaceae* (*Capnodiales*) was described based on the typification of the species *Aeminium ludgeri* and the genus *Aeminium*. *Aeminium ludgeri* DSM 106916 (strain E14) was isolated from a biodeteriorated limestone art piece in the Old Cathedral of Coimbra, Portugal, and it is characterized by slow growth, late melanization, and extremotolerance (halotolerance, xerophilia, and facultative alkaliphilia) (5). In this study, we were able to produce a high-quality draft genome sequence of *Aeminium ludgeri* DSM 106916 that constitutes valuable data for genomic studies regarding the extremotolerance pathways of microcolonial black fungi and to further understand their contribution to biodeterioration of stone monuments.

Fresh cultures of *Aeminium ludgeri* DSM 106916 were grown in potato dextrose agar (Difco, USA) and incubated aerobically in the dark at room temperature (28 ± 1°C), until full melanization of the cultures could be observed (6 months). Genomic DNA was extracted with a DNeasy PowerSoil extraction kit (Qiagen), and 200 ng of high-quality genomic DNA was used for DNA library preparation with the TruSeq Nano DNA library kit (Illumina, USA) and sequenced using paired-end (PE) 2 × 150-bp technology on the NextSeq 550 Illumina platform at Genoinseq (Cantanhede, Portugal). All procedures were performed according to the manufacturers' protocols.

Sequenced reads were demultiplexed using Bcl2fastq v.2.20 (Illumina, USA) and quality filtered with Trimmomatic v.0.30 (6) using the following parameters: ILLUMINACLIP, TruSeq3-PE.fa:2:30:10, SLIDINGWINDOW:5:25, and MINLEN:50. High-quality, adapter-free reads were assembled with dipSPAdes v.3.11.1 (7) with the parameters –careful and –cov-cutoff auto. Assembled scaffolds with a size of <1,000 bp were removed from the assemblies. Assembly metrics were calculated with Quast v.4.6.1 (8). Coding gene predictions were performed with AUGUSTUS v.2.5.5 (9) with the option –especies=*Botrytis cinerea* for training, rRNA genes were detected using Barrnap v.0.8 (https://github.com/

Address correspondence to João Trovão, jtrovaosb@gmail.com, or Igor Tiago, itiago@uc.pt.

tseemann/barrnap), and tRNA genes were identified with ARAGORN v.1.2 (10). Coding gene annotation was carried out with DIAMOND v.0.9.22 (11) against the Swiss-Prot database (12) and with HMMER 3.1b2 (13) against the High-quality Automated and Manual Annotation of Proteins (HAMAP) (14), TIGRFAMs (15), and Pfam (16) repositories and used to determine the existing pathways using the Kyoto Encyclopedia of Genes and Genomes (KEGG) (17).

This analysis produced 43,665,202 paired reads with an average length of 150 bases for each pair (400× sequencing depth coverage and 58.57% average G+C content). Trimming of low-quality bases and removal of reads shorter than 50 bp yielded 34,772,376 high-quality paired-end reads and 5,971,206 unpaired reads. The *de novo* read assembly produced 228 scaffolds, and the longest scaffold had 704,538 bases. In total, 8,128 genes with 8,103 coding genes were identified, along with 25 RNA genes, including 23 tRNA genes and 2 rRNA genes.

Data availability. The raw reads and draft genome sequence of *Aeminium ludgeri* DSM 106916 have been deposited in the NCBI Sequence Read Archive (SRA) and GenBank databases under the accession numbers SRR8580644 and SGQK00000000, respectively, and BioProject number PRJNA520871.

ACKNOWLEDGMENTS

REFERENCES

1. Sterflinger K, Krumbein WE. 1997. Dematiaceous fungi as a major agent for biopitting on Mediterranean marbles and limestones. Geomicrobiol J 14:219–230. https://doi.org/10.1080/01490459709378045.
2. Lombardozzi V, Castrignanò T, D'Antonio M, Casanova Municchia A, Caneva G. 2012. An interactive database for an ecological analysis of stone biopitting. Int Biodeterior Biodegradation 73:8–15. https://doi.org/10.1016/j.ibiod.2012.04.016.
3. Breitenbach R, Silbernagl D, Toepel J, Sturm H, Broughton WJ, Sassaki GL, Gorbushina AA. 2018. Corrosive extracellular polysaccharides of the rock-inhabiting model fungus *Knufia petricola*. Extremophiles 22:165–175. https://doi.org/10.1007/s00792-017-0984-5.
4. Isola D, Selbmann L, Meloni P, Maracci E, Onofri S, Zucconi L. 2013. Detrimental rock black fungi and biocides: a study on the Monumental Cemetery of Cagliari, p 83–86. *In* Rogerio-Candelera MA, Lazzari M, Cano E (ed), Science and technology for the conservation of cultural heritage. CRC Press, London, United Kingdom.
5. Trovão J, Tiago I, Soares F, Paiva DS, Mesquita N, Coelho C, Catarino L, Gil F, Portugal A. 2019. Description of *Aeminiaceae* fam. nov., *Aeminium* gen. nov. and *Aeminium ludgeri* sp. nov. (*Capnodiales*), isolated from a biodeteriorated art-piece in the Old Cathedral of Coimbra, Portugal. MycoKeys 45:57–73. https://doi.org/10.3897/mycokeys.45.31799.
6. 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.
7. Safonova Y, Bankevich A, Pevzner PA. 2015. dipSPAdes: assembler for highly polymorphic diploid genomes. J Comput Biol 22:528–545. https://doi.org/10.1089/cmb.2014.0153.
8. 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.
9. Stanke M, Morgenstern B. 2005. AUGUSTUS: a Web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res 33:W465–W467. https://doi.org/10.1093/nar/gki458.
10. Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res 32:11–16. https://doi.org/10.1093/nar/gkh152.
11. Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods 12:59–60. https://doi.org/10.1038/nmeth.3176.
12. UniProt Consortium. 2017. UniProt: the universal protein knowledgebase. Nucleic Acids Res 45:D158–D169. https://doi.org/10.1093/nar/gkw1099.
13. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res 41:e121. https://doi.org/10.1093/nar/gkt263.
14. Pedruzzi I, Rivoire C, Auchincloss AH, Coudert E, Keller G, de Castro E, Baratin D, Cuche BA, Bougueleret L, Poux S, Redaschi N, Xenarios I, Bridge A. 2015. HAMAP in 2015: updates to the protein family classifi-

cation and annotation system. Nucleic Acids Res 43:D1064–D1070. https://doi.org/10.1093/nar/gku1002.

15. Haft DH, Selengut JD, White O. 2003. The TIGRFAMs database of protein families. Nucleic Acids Res 31:371–373. https://doi.org/10.1093/nar/gkg128.

16. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 44:D279–D285. https://doi.org/10.1093/nar/gkv1344.

17. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res 45:D353–D361. https://doi.org/10.1093/nar/gkw1092.