

8 Applied and Industrial Microbiology | Announcement

Draft genome sequence of *Aspergillus ochraceus* VKM-F4104D (L-1) producing peptidase activator of protein C

Anna Shestakova,^{1,2} Daria Surkova,¹ Artem Fatkulin,³ Alexander Osmolovskiy,¹ Elizaveta Popova¹

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT Aspergillus ochraceus VKM-F4104D (L-1) is a saprotrophic fungus isolated from buried soils of Phanagoria (Russia). This strain is known as a producer of the fibrinolytic peptidase-activating plasma protein C. We have sequenced and assembled its genome for a more detailed understanding of the fungus' physiology and encoded peptidases.

KEYWORDS Aspergillus, peptidases, protein C

A spergillus ochraceus VKM-F4104D (L-1) is a known producer of a peptidase possessing the unique property of activating protein C of human blood plasma (1). It was previously identified by ITS sequencing (available at GenBank under the accession number KC859009) and deposited in the Russian collection of industrially important microorganisms ("VKPM") (2). Initially, *A. ochraceus* was isolated on malt-extract agar (MEA) by a serial dilution method from samples of buried soils during excavations of Phanagoria, the largest ancient Greek city on the Taman peninsula (Russia, 44.47035, 38.97919).

A. ochraceus VKM-F4104D was cultivated on MEA slants at 28°C. After 7 days, it was used to prepare a spore suspension, which was then inoculated into 100 mL of the liquid medium (g/L: malt extract—67.0, glucose—20.0, and peptone—1.0) and cultivated for 48 hours at 28°C, 200 rpm (3). The biomass was subsequently harvested by centrifugation and used for DNA isolation.

The genomic DNA was isolated using GNOME DNA Isolation Kit (MP Biomedicals). Libraries were prepared with KAPA HyperPrep Kit (Roche) and sequenced on Illumina NextSeq with 2×150 bp reads. A total of 41,018,432 reads were generated. Quality control of raw reads was performed using FastQC 0.12.1 (4) and SPAdes 3.15.5 (5) was used to assemble reads into contigs. Assembly quality was evaluated with QUAST 5.2.0 (6). Genes were predicted with AUGUSTUS 3.5.0 using --species=aspergillus_fumigatus (*Aspergillus fumigatus*) and --genemodel=complete options (7). For other tools, the default parameters were used.

The assembled genome consists of 862 contigs longer than 200 bp, with a maximum contig length of 720,519 bp. The total length is 36,493,367 bp. The assembled genome is available in the NCBI database with the accession number JBAFXH000000000 and raw reads with SRR28732222. The GC content of the genome is 49.2%. *N*50 and *L*50 values equal 131 kb and 89, respectively. The average coverage depth for the assembled genome is 319x. A total of 11,038 putative proteins were predicted to be encoded on the genome.

A maximum-likelihood phylogenetic tree based on the alignment of 18 beta-tubulin coding sequences [*BenA, in silico* amplified using Pincer (8) and Bt2a and Bt2b primers (9)] done with muscle (version 3.8.31) (10) was constructed in MEGA 11 with automated model selection, confirming that strain VKM-F4104D is indeed *A. ochraceus* (Fig. 1).

Editor Vincent Michael Bruno, University of Maryland School of Medicine, Baltimore, Maryland, USA

Address correspondence to Anna Shestakova, shestakovaaa@my.msu.ru.

The authors declare no conflict of interest.

See the funding table on p. 3.

Received 19 April 2024 Accepted 12 September 2024 Published 30 September 2024

Copyright © 2024 Shestakova et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.







0.10

FIG 1 Phylogenetic tree illustrating the taxonomy of A. ochraceus VKM-F4104D based on the BenA coding sequence.

ACKNOWLEDGMENTS

This research was funded by RCF, grant no. 22-74-00015.

AUTHOR AFFILIATIONS

¹Faculty of Biology, Department of Microbiology, M.V. Lomonosov Moscow State University, Moscow, Russia ²Faculty of Biology and Biotechnology, HSE University, Moscow, Russia ³Laboratory of Molecular Physiology, HSE University, Moscow, Russia

AUTHOR ORCIDs

Anna Shestakova (b) http://orcid.org/0000-0001-6894-8461

FUNDING

Funder	Grant(s)	Author(s)
Russian Science Foundation (RSF)	№ 22-74-00015	Elizaveta Popova

AUTHOR CONTRIBUTIONS

Anna Shestakova, Investigation, Methodology, Project administration, Writing – original draft, Writing – review and editing | Daria Surkova, Investigation, Validation | Artem Fatkulin, Data curation, Formal analysis, Software | Alexander Osmolovskiy, Supervision, Validation | Elizaveta Popova, Funding acquisition, Project administration, Resources, Supervision, Visualization

DATA AVAILABILITY

The assembled genome was deposited at GenBank under the accession number JBAFXH00000000 and the BioProject accession number PRJNA1075430. The raw sequence reads have been submitted to the Sequence Read Archive (SRA) under the accession number SRR28732222. ITS sequence previously used for identification is available at GenBank under the accession number KC859009. mRNA sequence of the BenA is available under the accession number PP695246.

REFERENCES

- Komarevtsev SK, Evseev PV, Shneider MM, Popova EA, Tupikin AE, Stepanenko VN, Kabilov MR, Shabunin SV, Osmolovskiy AA, Miroshnikov KA. 2021. Gene analysis, cloning, and heterologous expression of protease from a micromycete *Aspergillus ochraceus* capable of activating protein C of blood plasma. Microorganisms 9:1936. https://doi.org/10. 3390/microorganisms9091936
- Osmolovskiy AA, Kreier VG, Kurakov AV, Baranova N, Egorov NS. 2012. Aspergillus ochraceus micromycetes—producers of extracellular proteinases—protein C activators of blood plasma. Appl Biochem Microbiol 48:488–492. https://doi.org/10.1134/S0003683812050109
- Osmolovskiy AA, Lukianova AA, Zvonareva ES, Kreyer VG, Baranova NA, Egorov NS. 2018. Combined microbiological approach to screening of producers of proteases with hemostasis system proteins activity among micromycetes. Biotechnol Rep (Amst) 19:e00265. https://doi.org/10. 1016/j.btre.2018.e00265
- Andrews S. 2021. FastQC: a quality control analysis tool for high throughput sequencing data. Github. Available from: https://github. com/s-andrews/FastQC. Retrieved 18 Apr 2024.

- Spades: SPAdes genome assembler. 2024. Github. Available from: https:/ /github.com/ablab/spades. Retrieved 18 Apr 2024.
- Quast: genome assembly evaluation tool. 2024. Github. Available from: https://github.com/ablab/quast. Retrieved 18 Apr 2024.
- Brůna T, Li H, Guhlin J, Honsel D, Herbold S, Stanke M, Nenasheva N, Ebel M, Gabriel L, Hoff KJ. 2023. Galba: genome annotation with miniprot and AUGUSTUS. BMC Bioinformatics 24:327. https://doi.org/10.1186/s12859-023-05449-z
- Tan A. 2017. Pincer: in silico PCR for Python3. Github. Available from: https://github.com/Andre-Tan/Pincer. Retrieved 19 Apr 2024.
- Glass NL, Donaldson GC. 1995. Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. Appl Environ Microbiol 61:1323–1330. https://doi.org/10. 1128/aem.61.4.1323-1330.1995
- BioWeb. 2020. Available from: https://bioweb.pasteur.fr/packages/ pack@muscle@3.8.31. Retrieved 18 Apr 2024.