Title

Draft genome sequence of the Tremellomycetes yeast Papiliotrema laurentii 5307AH, isolated from aircraft.

Authors

Roman, Victor
Haridas, Sajeet
Hung, Chia
et al.

Publication Date

2024-07-05

DOI

10.1128/mra.00141-24

Peer reviewed

Microbiology Resource Announcements

| Environmental Microbiology | Announcement

Draft genome sequence of the Tremellomycetes yeast *Papiliotrema laurentii* 5307AH, isolated from aircraft

Victor A. Roman,[1,2] Sajeet Haridas,[3] Chia S. Hung,[1] Jasmyn Pangilinan,[3] Anna Lipzen,[3] Hyunsoo Na,[3] Mi Yan,[3] Vivian Ng,[3] Igor V. Grigoriev,[3,4] Justin Biffinger,[5] Daniel Barlow,[6] Nancy Kelley-Loughnane,[1] Wendy J. Crookes-Goodson,[7] Vanessa A. Varaljay,[8] Blake W. Stamps[1]

AUTHOR AFFILIATIONS See affiliation list on p. 2.

ABSTRACT *Papiliotrema laurentii* 5307AH was isolated from an aircraft polymer-coated surface. The genome size is 19,510,785 bp with a G + C content of 56%. The genome harbors genes encoding oxygenases, cutinases, lipases, and enzymes for styrene degradation, all of which could play a critical role in survival on xenobiotic surfaces.

KEYWORDS biodegradation, aircraft, *Papiliotrema laurentii*, polymers

Papiliotrema laurentii strain 5307AH was isolated from an environmental consortium collected inside an aircraft (1). *P. laurentii* species have been found inhabiting a variety of habitats such as natural, agricultural, and human-made environments (2). For instance, *P. laurentii* has been found in hydrocarbon-contaminated soils (3), inside aircrafts (1), and on surfaces within the International Space Station (4). The aircraft-derived *P. laurentii* strain 5307AH was maintained in glycerol stocks at −80°C, and overnight cultures were inoculated from single colonies on tryptic soy agar plates. Overnight broth cultures were grown on tryptic soy broth at 27°C with shaking (200 rpm). A 1 mL aliquot of overnight culture was centrifuged, and genomic DNA was extracted from the pellet using the Qiagen PowerMicrobiome kit. The DNA was quantified using the Qubit dsDNA Quantification Assay Kit (Life Technologies) following the manufacturer's instructions (5).

The *P. laurentii* 5307AH genome was sequenced using PacBio technology. Five micrograms of genomic DNA was sheared to >10 kb using Covaris g-Tubes. The sheared DNA was treated with exonuclease to remove single-stranded ends and DNA damage repair mix, followed by end repair and ligation of blunt adapters using SMRTbell Template Prep Kit 1.0 (Pacific Biosciences). The library was purified with AMPure PB beads. PacBio Sequencing primer was then annealed to the SMRTbell template library, and sequencing polymerase was bound using Sequel Binding kit 2.0. The prepared SMRTbell template libraries were then sequenced on a Pacific Biosciences Sequel sequencer using v3 sequencing primer, 1M v2 SMRT cells, and version 2.1 sequencing chemistry with 1 × 360 and 1 × 600 sequencing movie run times. The resulting 2,143,392 reads totaling 8.04 Gb were assembled with Falcon version 1.8.8 (6), improved with finisherSC version 2.0 (7), and polished with Arrow version SMRTLink version 5.1.0.26412, yielding 24 scaffolds, 364.38× sequencing read coverage depth, and an N_{50} value of 1.59 Mb with a G + C content of 56%. The assembled genome was annotated using the JGI Annotation pipeline (8). There were 7,537 predicted protein-coding genes with a median length of 422 amino acids, of which 5,474 (72.63%) had Pfam domains, 252 (3.35%) had CAZyme annotations, 1,417 (18.80%) had signal peptides, and 1,497 (19.86%) had transmembrane domains. Mitochondria were assembled separately from the Falcon pre-assembled reads (preads) using an in-house tool (assemblemito.sh), used to filter the preads, and polished with Arrow version SMRTLink version 5.1.0.26412 (9).

Editor Jason E. Stajich, University of California Riverside, Riverside, California, USA

Address correspondence to Blake W. Stamps, blake.stamps.1@us.af.mil, or Vanessa A. Varaljay, varaljay.1@osu.edu.

The authors declare no conflict of interest.

See the funding table on p. 2.

See the companion article at https://doi.org/10.1128/aem.01694-23.

Received 12 February 2024
Accepted 18 June 2024
Published 5 July 2024

A secondary Falcon assembly was generated using the mitochondria-filtered preads, improved with finisherSC version 2.0 (10), and polished with Arrow version SMRTLink version 5.1.0.26412 (9). Contigs less than 1,000 bp were excluded.

Furthermore, a search of the KEGG Metabolic Pathway database showed that the *P. laurentii* 5307AH genome encodes enzymes from the family of cutinases, lipases, hydroxylases, oxygenases, among others, which could provide an advantage to this fungal strain for survival in xenobiotic environments.

ACKNOWLEDGMENTS

The research reported in this publication has been cleared for public release under reference number AFRL-2024-0727 and was supported by funding from the Air Force Office of Scientific Research (AFRL grant 12RX14 COR and NRL grant F4FGA05338G001) and the Strategic Environmental Research and Development Program (SERDP, grant WP-1381). The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

AUTHOR AFFILIATIONS

[1]Biomaterials Branch, Materials and Manufacturing Directorate, Air Force Research Laboratory, Wright-Patterson AFB, Dayton, Ohio, USA
[2]Advanced Research and Development Division, UES Inc, a BlueHalo company, Dayton, OH, USA
[3]U.S. Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, California, USA
[4]Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, California, USA
[5]Chemistry Department, University of Dayton, Dayton, Ohio, USA
[6]Chemistry Division, Naval Research Laboratory, Washington DC, USA
[7]Microbial Cell Solutions, Ginkgo Bioworks, Boston, Massachusetts, USA
[8]Infectious Disease Institute, The Ohio State University, Ohio, Columbus, USA

AUTHOR ORCIDs

Victor A. Roman http://orcid.org/0009-0005-4442-3755
Vanessa A. Varaljay http://orcid.org/0000-0001-8815-2583
Blake W. Stamps http://orcid.org/0000-0001-9713-9014

FUNDING

Funder	Grant(s)	Author(s)
DOD | USAF | AMC | Air Force Office of Scientific Research (AFOSR)	12RX14COR	Victor A. Roman Chia S. Hung Justin Biffinger Nancy Kelley-Loughnane Vanessa A. Varaljay Blake W. Stamps
DOD | USAF | AMC | Air Force Office of Scientific Research (AFOSR)	F4FGA05338G001	Daniel Barlow
DOD | Strategic Environmental Research and Development Program (SERDP)	WP-1381	Victor A. Roman Justin Biffinger Daniel Barlow
Joint Genome Institute (JGI)	DE-AC02-05CH11231	Sajeet Haridas Jasmyn Pangilinan

Funder	Grant(s)	Author(s)
		Anna Lipzen
		Hyunsoo Na
		Mi Yan
		Vivian Ng
		Igor Grigoriev

AUTHOR CONTRIBUTIONS

Sajeet Haridas, Data curation, Formal analysis, Methodology, Software | Chia S. Hung, Conceptualization, Investigation, Methodology | Jasmyn Pangilinan, Formal analysis, Methodology, Software | Anna Lipzen, Formal analysis, Methodology, Software | Hyunsoo Na, Formal analysis, Methodology, Software | Mi Yan, Formal analysis, Methodology, Software | Vivian Ng, Formal analysis, Methodology, Software | Igor V. Grigoriev, Funding acquisition, Software, Supervision | Justin Biffinger, Funding acquisition, Project administration, Writing – review and editing | Daniel Barlow, Funding acquisition, Project administration | Nancy Kelley-Loughnane, Funding acquisition, Project administration, Resources | Wendy J. Crookes-Goodson, Conceptualization, Funding acquisition, Project administration, Resources | Vanessa A. Varaljay, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Supervision, Writing – review and editing.

DATA AVAILABILITY

The raw sequencing reads and draft genome with annotation have been deposited in GenBank under BioProject accession number PRJNA500119. The Sequence Read Archive (SRA) accession number for the raw sequencing reads is SRS4101967. The draft annotated genome assembly accession number is JAODAN000000000. Mitochondrial reads can be accessed using accession number PP706695.

REFERENCES

1. Hung C-S, Barlow DE, Varaljay VA, Drake CA, Crouch AL, Russell JN, Nadeau LJ, Crookes-Goodson WJ, Biffinger JC. 2019. The biodegradation of polyester and polyester polyurethane coatings using *Papiliotrema laurentii*. International Biodeterioration & Biodegradation 139:34–43. https://doi.org/10.1016/j.ibiod.2019.02.002
2. de Almeida ELM, Ventorim RZ, de Moura Ferreira MA, da Silveira WB. 2022. *Papiliotrema laurentii*: general features and biotechnological applications. Appl Microbiol Biotechnol 106:6963–6976. https://doi.org/10.1007/s00253-022-12208-2
3. Chandran P, Das N. 2012. Role of plasmid in diesel oil degradation by yeast species isolated from petroleum hydrocarbon-contaminated soil. Environ Technol 33:645–652. https://doi.org/10.1080/09593330.2011.587024
4. Satoh K, Alshahni MM, Umeda Y, Komori A, Tamura T, Nishiyama Y, Yamazaki T, Makimura K. 2021. Seven years of progress in determining fungal diversity and characterization of fungi isolated from the Japanese experiment module KIBO, International Space Station. Microbiol Immunol 65:463–471. https://doi.org/10.1111/1348-0421.12931
5. Wagner DN, Varaljay VA, Lyon WJ, Crouch AL, Drake CA, Allex-Buckner C, Biffinger JC, Kelley-Loughnane N, Crookes-Goodson WJ. 2022. Draft genome sequence of the nonmotile *Tremellomycetes* yeast *Naganishia albida*, isolated from aircraft. Microbiol Resour Announc 11:e0024222. https://doi.org/10.1128/mra.00242-22
6. Pacific BioSciences. 2018. Https://Github.com/Pacificbiosciences/FALCON.
7. Lam K-K, Khalak A, Tse D. 2014. Near-optimal assembly for shotgun sequencing with noisy reads. BMC Bioinformatics 15 Suppl 9:S4. https://doi.org/10.1186/1471-2105-15-S9-S4
8. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I. 2014. Mycocosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res 42:D699–704. https://doi.org/10.1093/nar/gkt1183
9. Pacific BioSciences. 2019. Pacbio variant and consensus caller V5.1.0.26412 on Pacbio. https://web.archive.org/web/20220313151841/https://github.com/PacificBiosciences/GenomicConsensus.
10. Lam KK, LaButti K, Khalak A, Tse D. 2015. FinisherSC: a repeat-aware tool for upgrading *de novo* assembly using long reads. Bioinformatics 31:3207–3209. https://doi.org/10.1093/bioinformatics/btv280