



Draft Genome Sequence of *Fusarium* sp. Strain DS 682, a Novel Fungal Isolate from the Grass Rhizosphere

Arunima Bhattacharjee,^a  Lindsey N. Anderson,^a Trinidad Alfaro,^a Andrea Porras-Alfaro,^c Ari Jumpponen,^d Kirsten S. Hofmockel,^{a,b} Janet K. Jansson,^a  Christopher R. Anderton,^a  William C. Nelson^a

^aEarth and Biological Sciences Directorate, Pacific Northwest National Laboratory, Richland, Washington, USA

^bDepartment of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa, USA

^cDepartment of Biological Sciences, Western Illinois University, Macomb, Illinois, USA

^dDivision of Biology, Kansas State University, Manhattan, Kansas, USA

ABSTRACT The novel fungal strain, *Fusarium* sp. strain DS 682, was isolated from the rhizosphere of the perennial grass, *Bouteloua gracilis*, at the Konza Prairie Biological Station in Kansas. This fungal strain is common across North American grasslands and is resilient to environmental fluctuations. The draft genome is estimated to be 97.2% complete.

Fusarium sp. strain DS 682 (DS 682) was isolated from the rhizosphere of *Bouteloua gracilis*, a blue grama perennial warm season (C4) grass, at the Konza Prairie Biological Station (KPBS) in Kansas. Specifically, the strain was isolated from field site 4A as part of a study to evaluate the importance of climatic, edaphic, and host plant traits as predictors of rhizobiome community structure (1). The isolation process included surface sterilization of excised root tissue of *B. gracilis* and plating onto malt extract agar (MEA; Difco Laboratories, MD) amended with 50 mg/liter streptomycin and tetracycline. After a 14-day incubation at room temperature, emerging colonies with different morphologies were transferred onto fresh MEA (supplemented with antibiotics) to establish pure cultures.

DS 682 was chosen for whole-genome sequencing because it is common to soil and grass rhizospheres across North America, demonstrating metabolic adaptability to environmental fluctuations. The fungal isolate was identified as a *Fusarium* sp. by amplicon sequencing of the internal transcribed spacer (ITS) region using the primers ITS1F (2) and ITS4 (3) and the Web resource FUSARIUM-ID (<http://isolate.fusariumdb.org/blast.php>). This identification was verified and refined using additional genomic markers (see below).

Genomic DNA was prepared from DS 682 mycelium grown on MEA. Cells were lysed via sonication (Branson SFX 150; Emerson Electric, MO) and extracted by using a DNeasy PowerSoil kit (Qiagen, Inc., Germany). DNA libraries were prepared by using a NEBNext Ultra DNA library prep kit (Illumina, San Diego, CA) and validated by using a DNA 1000 Chip on the Agilent TapeStation (Agilent Technologies, CA). Sequencing was performed on the Illumina HiSeq X platform (150 bp, paired end) by Genewiz, Inc. (South Plainfield, NJ).

Sequence reads were assembled using the ATLAS pipeline (4). Briefly, reads were deduplicated using Clumpify (BBTools v37.78), quality filtered using BBDukF, and decontaminated by searching against PhiX and rRNA reference data sets using BBSplit. Quality-filtered reads were normalized using KmerNormalize, error corrected using Tadpole, and merged using BBMerge. Assembly was performed by MEGAHIT (v1.1.2) (5), resulting in 13,284 contigs, totaling 56.5 Mbp and ranging in length from 200 to 215,770 bp (average, 4,253 bp; N_{50} , 24,701 bp). The average coverage was determined to

Citation Bhattacharjee A, Anderson LN, Alfaro T, Porras-Alfaro A, Jumpponen A, Hofmockel KS, Jansson JK, Anderton CR, Nelson WC. 2021. Draft genome sequence of *Fusarium* sp. strain DS 682, a novel fungal isolate from the grass rhizosphere. Microbiol Resour Announc 10:e00884-20. <https://doi.org/10.1128/MRA.00884-20>.

Editor Jason E. Stajich, University of California—Riverside

Copyright © 2021 Bhattacharjee et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to William C. Nelson, William.Nelson@pnsl.gov.

Received 29 July 2020

Accepted 2 October 2020

Published 7 January 2021

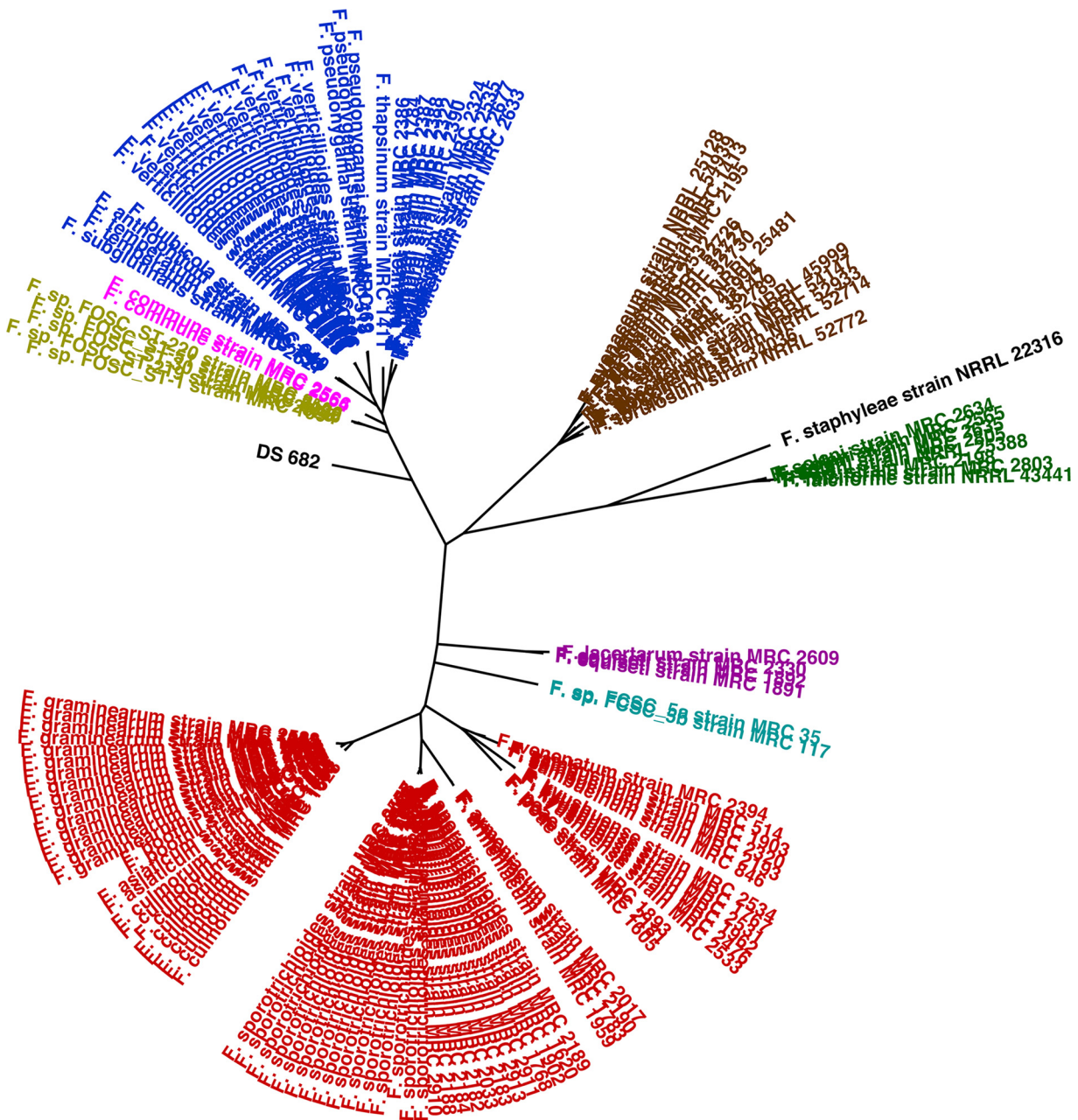


FIG 1 Molecular phylogeny of *Fusarium* sp. strain DS 682. Sequences of the RNA polymerase II (RPB2) and translation elongation factor 1 (TEF1) genes were concatenated and aligned using MUSCLE v3.8.425 (9). An approximate maximum-likelihood tree was generated using FastTree (10) and visualized using FigTree v1.43 (<http://tree.bio.ed.ac.uk/software/figtree/>). Label colors denote species complex memberships: red, *F. sambucinum* species complex (FSAMSC); purple, *F. incarnatum-equiseti* species complex (FIESC); cyan, *F. chlamydosporum* species complex (FCSC); mustard, *F. oxysporum* species complex (FOSC); blue, *F. fujikuroi* species complex (FFSC); brown, *F. tricinctum* species complex (FTSC); green, *F. solani* species complex (FSSC); magenta, *F. nisikadoi* species complex (FNSC); black, none assigned.

be 233-fold by QAST (v5.0.2) (6). Genemark-ES (v4.59) was used to predict coding genes, resulting in 19,949 predicted proteins, of which 8,849 were assigned to a KEGG Orthology (KO) family using kofamskan (v1.3). Genome completeness was estimated to be 97.2% with low duplication and low fragment count (C:97.2%[S:95.7%,D:1.5%],F:1.7%, M:1.1%,n:4494) by BUSCO (v4.0.6) using the *hypocreales_odb10* reference gene set.

Phylogenetic placement was determined by analyzing the genes for RNA polymerase II (RPB2) and translation elongation factor I (TEF1). The data set published by O'Donnell et al. (7) was used as a reference database. Sequences were imported into

the AliView (v1.19) (8) interface, aligned using MUSCLE v3.8.425 (9), and clustered using FastTree v2.2.11 (10) with the “nt” and “gtr” options. This analysis suggests that DS 682 represents a novel lineage within the *Fusaria* basal to the *F. fujikuroi*, *F. nisikadoi*, and *F. oxysporum* species complexes (Fig. 1).

Comparison with *Fusarium oxysporum* revealed 4,934 genes (~25% of the genome) in DS 682 that did not have close similarity (amino acid similarity > 60%, E value ≤ 10⁻¹⁰). Of these, only 664 were assigned to 355 KO families with diverse activities.

Data availability. Sequence data have been deposited at the PNNL DataHUB repository (<https://data.pnnl.gov/about>) and are available for download under project data DOI accession number [10.25584/KS4AAlsoGFspDS682/1635527](https://doi.org/10.25584/KS4AAlsoGFspDS682/1635527). The version described here is the first version. The data package contains raw reads, assembly, functional annotations, MIGS.eu.soil.5.0 metadata information, and a package “Read Me” file. The genome sequence information has also been deposited under BioProject number [PRJNA664411](https://doi.org/10.25584/PRJNA664411) and under BioSample accession number [SAMN16213216](https://doi.org/10.25584/SAMN16213216), and this whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under accession number [JACYFE000000000](https://doi.org/10.25584/JACYFE000000000), where the version described here is version [JACYFE010000000](https://doi.org/10.25584/JACYFE010000000). Gene models are available at https://www.ncbi.nlm.nih.gov/protein?LinkName=nucore_protein_wgs&from_uid=1910285265.

ACKNOWLEDGMENTS

This research was supported by the Department of Energy (DOE) Office of Biological and Environmental Research (BER) and is a contribution of the Scientific Focus Area “Phenotypic response of the soil microbiome to environmental perturbations” (number 70880). Pacific Northwest National Laboratory (PNNL) is operated for the DOE by Battelle Memorial Institute under Contract DE-AC05-76RLO1830. The ITS sequencing effort for taxonomy identification was supported by NSF-DEB 1457002 to A.P.-A. The material is based upon work supported by the National Science Foundation. Any opinion, findings, and conclusions or recommendations are those of the authors and do not necessarily reflect the views of the National Science Foundation.

We thank Terri Billingsley Tobias, Cedic Ndinga Muniania, Paris Hamm, Terry Torres-Cruz, and Adeyemi A. Olanrewaju for sample processing, culture isolation, and amplicon sequencing.

REFERENCES

- Jumpponen A, Herrera J, Porras-Alfaro A, Rudgers J. 2017. Biogeography of root-associated fungal endophytes, p 195–222. *In* Tedersoo L (ed), *Biogeography of mycorrhizal symbiosis*. Springer International Publishing, Cham, Switzerland.
- Gardes M, Bruns TD. 1993. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Mol Ecol* 2:113–118. <https://doi.org/10.1111/j.1365-294x.1993.tb00005.x>.
- White TJ, Bruns T, Lee S, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, p 315–322. *In* Innis MA, Gelfand DH, Sninsky JJ, White TJ (ed), *PCR protocols*. Academic Press, San Diego, CA.
- Kieser S, Brown J, Zdobnov EM, Trajkovski M, McCue LA. 2020. ATLAS: a Snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data. *BMC Bioinformatics* 21:257. <https://doi.org/10.1186/s12859-020-03585-4>.
- Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, Sadakane K, Yamashita H, Lam T-W. 2016. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102:3–11. <https://doi.org/10.1016/j.ymeth.2016.02.020>.
- 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>.
- O'Donnell K, McCormick SP, Busman M, Proctor RH, Ward TJ, Doehring G, Geiser DM, Alberts JF, Rheeder JP. 2018. Marasas et al. 1984 “Toxicogenic *Fusarium* Species: Identity and Mycotoxicology” revisited. *Mycologia* 110:1058–1080. <https://doi.org/10.1080/00275514.2018.1519773>.
- Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30:3276–3278. <https://doi.org/10.1093/bioinformatics/btu531>.
- Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5:113. <https://doi.org/10.1186/1471-2105-5-113>.
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2: approximate maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.