AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Whole-Genome Sequence and Draft Assembly of the Biocontrol Fungal Pathogen *Albifimbria verrucaria* CABI-IMI 368023

Mark A. Weaver,[a] Nahreen Mirza,[b,c] Jennifer R. Mandel,[b,c] C. Douglas Boyette,[a] Shawn P. Brown[b,c]

[a]USDA ARS, Biological Control of Pests Research Unit, National Biological Control Laboratory, Stoneville, Mississippi, USA
[b]Department of Biological Sciences, University of Memphis, Memphis, Tennessee, USA
[c]Center for Biodiversity Research, University of Memphis, Memphis, Tennessee, USA

ABSTRACT We report here the whole-genome sequence and draft assembly for a bioherbicidal strain of *Albifimbria verrucaria*, CABI-IMI 368023, which was formerly identified as *Myrothecium verrucaria*. This isolate has been well studied for the biological control of important weeds, including kudzu and giant salvinia.

Albifimbria verrucaria (formerly *Myrothecium verrucaria*; Ascomycota, Sordariomycetes, Hypocreales, Stachybotriaceae) (1) is a candidate biocontrol phytopathogen that has been demonstrated to aid in the control of many problematic and invasive weed species, including kudzu (*Pueraria montana* var. *lobata* Willd. [Ohwii]) and giant salvinia (*Salvinia molesta* [Mitchell]). The most studied strain of *A. verrucaria* for biocontrol purposes is CABI-IMI 368023, which was isolated from sicklepod (*Senna obtusfolia* [L.] Irwin & Barneby) in DeSoto County, Mississippi, United States, but was only recently confirmed to be *A. verrucaria* (2).

Here, we report a full genome assembly of *Albifimbria verrucaria* strain CABI-IMI 368023. This strain has been maintained and studied since 1983 at the USDA ARS in Stoneville, Mississippi, United States. This genome sequence could inform strategies for biocontrol implementation for unwanted weeds and plants.

Strain CABI-IMI 368023 was grown on potato dextrose agar (PDA) for 5 days at 28°C. Genomic DNA from ~10^8 conidia was extracted using the Zymo fungal miniprep kit (number D6005). Extracted DNA was quantified using a Qubit 3.0 fluorometer (double-stranded DNA [dsDNA] high-sensitivity [HS] assay kit; Thermo Fisher Scientific, Waltham, MA, USA), and 1.0 μg DNA was sequenced using one SpotON flow cell (R9) with the Oxford Nanopore Technologies (ONT) MinION platform following library generation (ligation kit SQK-KSK109) and enrichment of fragments of >3 kb using the long fragment buffer (LFB) according to the ligation kit protocol. Sequencing was conducted for 72 h, and base-calling data were base called using the Fast Basecalling implementation with the MinKNOW interface, with a minimum Q score of 7 for sequence inclusion. This resulted in 8.5 $\times$ 10^9 called bases, 1.3 $\times$ 10^6 sequences, an average sequence length of 6,154 bp, a longest read length of 75,816 bp, and a raw-data N_{50} value of 8,438 bp. The genome was assembled using the de Bruijn-graph-based long-read assembler Flye v.2.7-b1585 (https://github.com/fenderglass/Flye) (3). Each contig/scaffold obtained was queried against GenBank using BLASTn to ensure that contigs were from this pure culture, and mitochondrial reads were removed; this resulted in 61 contigs and 1 scaffold for a total assembled length of 48.5 $\times$ 10^6 bp, with an N_{50} value of 5,352,524 bp and mean coverage of 171$\times$. The assembled genome has a GC content of 49.3% and a total of 321,596 predicted open reading frames (ORFs) (predicted using Geneious v.11.1.5 with a minimum size of 150 bp). AUGUSTUS (4) (Geneious plugin v.0.1.1) was used to identify predicted genes using a *Fusarium* reference genome, and 12,893 predicted genes were identified.

Editor Jason E. Stajich, University of California, Riverside

Address correspondence to Mark A. Weaver, mark.weaver@usda.gov.

The authors declare no conflict of interest.

Received 21 September 2021
Accepted 3 December 2021
Published 6 January 2022

To investigate gene collinearity and genomic architecture within this *A. verrucaria* genome, we used the MCScanX toolkit (5) to classify all predicted genes as singletons, dispersed duplicates, proximal duplicates, tandem duplicates, or whole-genome duplication/segmental duplicates. Briefly, an all-versus-all BLASTp search (6) was performed using an E value of $\leq 1e^{-5}$ and the m8 output format. The BLASTp results along with a GFF file (formatted following author guidelines at https://github.com/wyp1125/MCScanX) were used to execute the MCScanX algorithm along with the downstream analysis tool, i.e., duplicate gene classifier. Output files, including the collinearity file, the tandem duplicate file, and HTML files showing multiple alignments, are available at https://doi.org/10.6084/m9.figshare.16587368. The duplicate gene classifier tool reported 8,794 singleton genes, 3,698 dispersed duplicates, 42 proximal duplicates, 35 tandem duplicates, and 324 whole-genome duplication/segmental duplicates.

Data availability. This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number JAHREC000000000. The version described in this paper is version JAHREC010000000. Raw reads have been deposited under the BioProject accession number PRJNA736674 and at the Sequence Read Archive (SRA) under BioSample accession number SAMN19655718.

ACKNOWLEDGMENTS

Partial support for this work was provided by the Department of Biological Sciences at the University of Memphis, and N.M. is supported by NSF CBET grant 2029764 (awarded to S.P.B.).

REFERENCES

1. Lombard L, Houbraken J, Decock C, Samson RA, Meijer M, Réblová M, Groenewald JZ, Crous PW. 2016. Generic hyper-diversity in *Stachybotriaceae*. Persoonia 36:156–246. https://doi.org/10.3767/003158516X691582.
2. Weaver M, Hoagland RE, Boyette CD, Brown SP. 2021. Taxonomic evaluation of a bioherbicidal isolate of *Albifimbria verrucaria*, formerly *Myrothecium verrucaria*. J Fungi (Basel) 7:694. https://doi.org/10.3390/jof7090694.
3. Kolmogorov M, Yuan H, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nat Biotechnol 37:540–546. https://doi.org/10.1038/s41587-019-0072-8.
4. 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.
5. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee T-h, Jin H, Marler B, Guo H, Kissinger JC, Paterson AH. 2012. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res 40:e49. https://doi.org/10.1093/nar/gkr1293.
6. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.