AMERICAN SOCIETY FOR MICROBIOLOGY
Microbiology Resource Announcements

AUTHOR REPLY

Lessons Learned: the Importance of Biological Curation

Jake C. Fountain,[a] Josh P. Clevenger,[b] Justin N. Vaughn,[c] Baozhu Guo[d]

[a]Department of Biochemistry, Molecular Biology, Entomology, and Plant Pathology, Mississippi State University, Mississippi State, Mississippi, USA
[b]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, USA
[c]USDA-ARS, Genomics and Bioinformatics Research Unit, Stoneville, Mississippi, USA
[d]USDA-ARS, Crop Genetics and Breeding Research Unit, Tifton, Georgia, USA

Proper species identification of sequenced fungal isolates or strains is imperative for the interpretation of genomics data. Through this letter (1), and through another recently published research note (2), it has come to our attention that one of the isolates presented in our previous publication (3) was misidentified. The presumptive *Aspergillus parasiticus* isolate NRRL 2999 has been shown to be an *Aspergillus flavus* isolate, a clonal derivative of *A. flavus* NRRL 3357 (2). In our recent publication presenting two new reference genomes for *A. flavus* (4), we performed a phylogenomics analysis of several publicly available *Aspergillus* assemblies and noted the sequence similarity between *A. flavus* NRRL 3357 and *A. parasiticus* NRRL 2999. We also noted a close relationship between *A. flavus* WRRL 1519 and *A. oryzae* RIB40 and that *A. oryzae* RIB40 grouped among *A. flavus* isolates, not as an outgroup. Based on these analyses, we concluded that the *A. flavus* species is polyphyletic. The revelation that NRRL 2999 is indeed an *A. flavus* isolate clonal to NRRL 3357 (2) does point toward *A. flavus* as a distinct species from *A. parasiticus* but does not change the conclusion of the potential polyphyletic nature of *A. flavus* based on the position of *A. oryzae* in the analysis (4). All additional conclusions were based on phylogenetic relationships defined within the scope of the manuscript (4) and were therefore independent of species labels.

As reported by Smith et al. (5) on "the early scientific literature," the issue of isolate identification is further complicated for aflatoxigenic fungal isolation and classification. Over the years, there have been occurrences of the same strain being assigned many different designations by different research groups. One interesting example is NRRL 2999, which was originally isolated from Ugandan peanuts in 1961 (6, 7). Over time, this isolate has received numerous other designations, such as Austwick strain V. 3734/10, Hodges M-3, SYS-4, ATCC 56775, ATCC 26692, CMI 91019b, NRRL 5862, ATCC 15517, SU-1, and SRRC 143, depending on the research groups the isolate was received from. It is now accessioned as IMI 91019b, NRRL 2999, and ATCC 15517, serving as the type strain for the American Type Culture Collection (2). There are two genome sequence data sets for *A. parasiticus* SU-1 (8) in NCBI. Since *A. parasiticus* SU-1 and NRRL 2999 are actually the same isolate, we plan to remove the current sequence data for the misidentified "NRRL 2999" from NCBI (GenBank accession no. CP051027 to CP051034) (3).

This situation does highlight the danger inherent in large-scale genome sequencing experiments, where the identity of isolates is presumed correct based on identification provided by the source of an isolate. Therefore, we must join with Houbraken et al. (1) in agreement that additional safeguards should be taken to ensure that the proper identification of isolates be determined before conclusions are made based on sequencing data.

Editor Antonis Rokas, Vanderbilt University

Address correspondence to Baozhu Guo, Baozhu.Guo@usda.gov.

This is a response to a letter by Houbraken et al. (https://doi.org/10.1128/MRA.01074-20).

Published 2 December 2021

REFERENCES

1. Houbraken J, Visagie CM, Frisvad JC. 2021. Recommendations to prevent taxonomic misidentification of genome-sequenced fungal strains. Microbiol Resour Announc 10:e01074-20. https://doi.org/10.1128/MRA.01074-20.
2. Chang PK. 2021. Authentication of *Aspergillus parasiticus* strains in the genome database of the National Center for Biotechnology Information. BMC Res Notes 14:111. https://doi.org/10.1186/s13104-021-05527-6.
3. Fountain JC, Clevenger JP, Nadon B, Wang H, Abbas HK, Kemerait RC, Scully BT, Vaughn JN, Guo B. 2020. Draft genome sequences of one *Aspergillus parasiticus* isolate and nine *Aspergillus flavus* isolates with varying stress tolerance and aflatoxin production. Microbiol Resour Announc 9:e00478-20. https://doi.org/10.1128/MRA.00478-20.
4. Fountain JC, Clevenger JP, Nadon B, Youngblood RC, Korani W, Chang P-K, Starr D, Wang H, Isett B, Johnston HR, Wiggins R, Agarwal G, Chu Y, Kemerait RC, Pandey MK, Bhatnagar D, Ozias-Akins P, Varshney RK, Scheffler BE, Vaughn JN, Guo B. 2020. Two new *Aspergillus flavus* reference genomes reveal a large insertion potentially contributing to isolate stress tolerance and aflatoxin production. G3 (Bethesda) 10:3515–3531. https://doi.org/10.1534/g3.120.401405.
5. Smith JS, William WP, Windham GL. 2019. Aflatoxin in maize: a review of the early literature from “moldy-corn toxicosis” to the genetics of aflatoxin accumulation resistance. Mycotoxin Res 35:111–128. https://doi.org/10.1007/s12550-018-00340-w.
6. Sargeant K, Sheridan A, O’Kelly J, Carnaghan RBA. 1961. Toxicity associated with certain samples of groundnuts. Nature 192:1096–1097. https://doi.org/10.1038/1921096a0.
7. Hesseltine CW, Shotwell OL, Ellis JJ, Stubblefield RD. 1966. Aflatoxin formation by *Aspergillus flavus*. Bacteriol Rev 30:795–805. https://doi.org/10.1128/br.30.4.795-805.1966.
8. Linz JE, Wee J, Roze LV. 2014. *Aspergillus parasiticus* SU-1 genome sequence, predicted chromosome structure, and comparative gene expression under aflatoxin-inducing conditions: evidence that differential expression contributes to species phenotype. Eukaryot Cell 13:1113–1123. https://doi.org/10.1128/EC.00108-14.