AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Genomics and Proteomics | Announcement

Draft genome sequence of 16 *Aspergillus flavus* isolated from cashew nuts from coastal Kenya

Pauline Wambui Gachanja,[1,2] Manase Aloo Onyango,[1,2] Colletah Rhoda Musangi,[1,2] Bicko Steve Juma,[2] Dennis Wamalabe Mukhebi,[1,2] Eugene Mwanza Muzami,[1,2] Kyalo Katua,[1,2] Wilton Mwema Mbinda[1,2]

AUTHOR AFFILIATIONS See affiliation list on p. 4.

ABSTRACT *Aspergillus flavus* is a soil-borne fungus known for its aflatoxin contamination of agricultural products. Here, we report the draft genome sequences of 16 predicted aflatoxin-producing *A. flavus* isolated from cashew nuts from coastal Kenya.

KEYWORDS *Aspergillus flavus*, cashew nut, draft genome, aflatoxigenic

Aspergillus flavus is a filamentous saprophytic fungus that contaminates important agricultural products, including cashew nuts (1). Under favorable conditions, such as drought and heat stress, *A. flavus* produces secondary metabolites, such as aflatoxins (2). Aflatoxins are carcinogenic, hepatotoxic, and can cause aspergillosis in immunocompromised individuals (2, 3). Aflatoxin contamination in crops poses significant threats to global food safety, particularly in sub-Saharan Africa, which has warm and humid climates that are conducive to aflatoxin production (4, 5). Aflatoxin contamination of crops is expected to increase due to climate change shocks (6). Here, we present draft genome sequences of 16 predicted aflatoxigenic *A. flavus* isolated from cashew nuts (Table 1), which is essential in understanding *A. flavus* phylogenetics and comparative and functional genomics.

Cashew nut samples were collected in coastal Kenya's production areas (Kilifi, Kwale, and Lamu) in May 2021 (7, 8). Whole cashew nuts were surface sterilized with 70% ethanol. Cashew shells were cut into four sections, and the kernels were cut into approximately 3 × 4-mm pieces. The pieces were directly plated on modified Rose Bengal Agar medium and incubated at 30°C for 7 days in darkness. Fungi growing were transferred to Water Agar medium and incubated for 3 days at 27°C in the light. The resulting hyphae were cultured on potato dextrose agar (PDA) and malt extract agar media at 25°C for 7 days under light to isolate pure cultures. Species identification was based on morphological determination, Sanger sequencing of PCR products of ITS (ITS1/ITS4), and 28s rRNA regions (NL1/NL4) (9), and calmodulin gene (Cmd5/Cmd6) (S1) (10) at Macrogen, Netherlands. Pure isolates were further cultured on PDA at 25°C for 7 days under darkness. Resulting mycelia were used for DNA extraction using ZR Fungal/Bacterial DNA Miniprep kit (Zymo, Irvine, USA). The sequences were queried using NCBI BLASTn v2.14.0 (11) (Table 1).

Sequencing library of the 16 *A. flavus* predicted to produce aflatoxin (12) was generated by TruSeq DNA PCR-Free kit (Illumina, San Diego, USA), and short reads paired-end genome sequencing was carried out using Illumina's NovaSeq-6000 at Macrogen, South Korea. FASTQC v0.12.1 was used to check the quality of the paired-end (PE) raw reads (13). PE raw reads were filtered using fastp v3.3.5 to eliminate reads with Q scores of <20 and adapters (14). The reads had an average length of 151 bp, and the total number of reads for each genome are listed in (Table 1). Clean reads were used for *de novo* genome assembly using SPAdes v3.15.4 with the "careful" option and k-mer sizes 21, 33, 55, 77, and 99 (15). SPAdes contigs and scaffolds are shown in Table 1.

Editor Leighton Pritchard, University of Strathclyde, Glasgow, United Kingdom

Address correspondence to Wilton Mwema Mbinda, wilton.mbinda@gmail.com, w.mbinda@pu.ac.ke.

The authors declare no conflict of interest.

See the funding table on p. 5.

Received 2 May 2024
Accepted 19 August 2024
Published 9 September 2024

TABLE 1 Genome sequences of *Aspergillus flavus* sp. isolated from cashew nuts from coastal Kenya

Parameter/ Sample No.	10B	11B	12A	13B	15A	16A	17A	18B	19B	1B	20B	22B	4B	5B	7B	9A
Genbank accession No.	JBA-WIZ00000000	JBA-WIU00000000	JBA-WIT00000000	JBA-WIS00000000	JBA-WIR00000000	JBA-WIQ00000000	JBA-WIP00000000	JBA-WIO00000000	JBA-WIN00000000	JBA-WIK00000000	JBA-WIM00000000	JBA-WIL00000000	JBA-WIX00000000	JBA-WIY00000000	JBA-WIV00000000	JBAWIW000000000
SRA accession No.	SRR28841496	SRR28841501	SRR28841488	SRR28841489	SRR28841490	SRR28841491	SRR28841492	SRR28841493	SRR28841494	SRR28841503	SRR28841495	SRR28841502	SRR28841498	SRR28841497	SRR28841500	SRR28841499
Number of reads	21,050,726	22,806,790	17,032,196	20,990,232	19,001,446	17,128,796	21,850,708	19,905,174	20,603,930	19,419,956	20,553,698	21,310,018	#########	20,196,650	22,126,146	19,341,246
Genome size (bp)	37,449,923	37,283,523	37,243,444	37,657,964	37,316,005	37,274,043	39,804,705	37,649,694	38,097,618	37,177,891	38,422,971	37,308,414	37,559,519	37,759,462	37,310,799	37,307,524
Genome coverage (×)	83.4	90.8	67.9	82.7	75.5	68.2	78.1	78.4	80.2	79.3	84.7	77.5	92.4	79.4	88	76.9
% G + C	47.5	47.5	47.5	47.5	47.5	47.5	47.5	47.5	47.5	48	47.5	47.5	47.5	47.5	47.5	47.5
No. of contigs	63	87	158	164	78	126	11,715	183	609	437	859	80	102	127	66	81
N50 contigs	1.4 Mb	1.2 Mb	993 kb	894.4 kb	1.1 Mb	988.7 kb	58.1 kb	894.4 kb	775 kb	1 Mb	384.5 kb	899.9 kb	903.4 kb	978.4 kb	931.3 kb	1.1 Mb
No. of scaffolds	25	58	122	122	38	80	9,715	135	560	399	760	36	65	84	29	48
N50 scaffolds	2.5 Mb	1.9 Mb	1.8 Mb	1.7 Mb	1.9 Mb	2.1 Mb	271.9 kb	2.1 Mb	983.3 kb	1.8 Mb	579.5 kb	2 Mb	1.6 Mb	2.1 Mb	2.1 Mb	2.4 Mb
Total no. of Busco orthologs	1,706	1,706	1,706	1,706	1,706	1,706	1,706	1,706	1,706	1,706	1,706	1,706	1,706	1,706	1,706	1,706
Complete single-copy, complete multicopy, fragmented, and missing orthologs (%)	98.6, 0.4, 0.0, 1.0	98.5, 0.4, 0.0, 1.0	98.5,0.4, 0.0, 1.1	98.6, 0.4, 0.0, 1.0	98.5,0.4, 0.0, 1.1	98.7, 0.4, 0.0, 0.9	93.1, 0.5, 3.2, 3.2	98.5,0.4, 0.0, 1.1	98.6, 0.4, 0.0, 1.0	98.5, 0.5, 0.0, 1.0	98.4, 0.5, 0.1, 1.0	98.7, 0.4, 0.0, 0.9	98.7, 0.4, 0.0, 0.9	98.5, 0.5, 0.0, 1.0	98.6, 0.4, 0.0, 1.0	98.6, 0.4, 0.0, 1.0
ITS BLAST similarity (%), reference match and accession	*Aspergillus flavus*, 95, MT573498.1.	*Aspergillus flavus*, 91, MT071404. 1	*Aspergillus niger*, 98, MK372989.1	*Aspergillus flavus*, 99, OK086056. 1	*Aspergillus flavus*, 95, MK992255.1		*Aspergillus flavus*, 85, MT994594.1	*Aspergillus flavus*, 94, MK992254. 1	*Aspergillus sp.*, OK210353. 1	*Aspergillus flavus*, 99, KX067886. 1	*Aspergillus terreus*, 98, KJ685810.1				*Aspergillus aculeatus*, 99, KP965728.1	

(*Continued on next page*)

TABLE 1 Genome sequences of *Aspergillus flavus* sp. isolated from cashew nuts from coastal Kenya (*Continued*)

Parameter/ Sample No.	10B	11B	12A	13B	15A	16A	17A	18B	19B	1B	20B	22B	4B	5B	7B	9A
28s rRNA BLAST similarity (%), reference match and accession	*Aspergillus novoparasiticus, 99,* NG069972. 1	*Aspergillus oryzae, 99,* KX958066. 1	*Aspergillus sp., 99,* MN515285. 1	*Aspergillus flavus, 98,* MT509808. 1	*Aspergillus flavus, 99,* MT252035. 1	*Aspergillus flavus, 95,* MT509808. 1	*Aspergillus flavus, 99,* MT252035. 1	*Aspergillus novoparasiticus, 99,* NG069972. 1		*Aspergillus flavus, 100,* MT509808. 1	*spergillus terreus, 99,* MH877949. 1	*Aspergillus aculeatus, 99,* MK518351. 1	*Aspergillus novoparasiticus, 99,* NG066672. 1	*Aspergillus novoparasiticus, 100,* MT252035.1	*Aspergillus aculeatus, 99,* MH870630. 1	*Aspergillus aculeatus, 99,* JQ301899.1
Calmodulin BLAST similarity (%), reference match and accession	*Aspergillus flavus, 99.48,* KY272751.1	*Aspergillus flavus, 98.81,* LS999591.1	*Aspergillus flavus, 98.48,* MK304471.1	*Aspergillus flavus, 99.48,* MN271387.1	*Aspergillus flavus, 99.14,* LC061194.1	*Aspergillus flavus, 99.66,* MG991523.1	*Aspergillus flavus,* 99.48, KY272751.1	*Aspergillus flavus,* 99.14, MG991523.1	*Aspergillus flavus,* 99.65, MN271387.1	*Aspergillus flavus,* 91.40, MN271387.1	*Aspergillus flavus,* 98.95, LC061194.1	*Aspergillus flavus,* 98.62, LS999591.1	*Aspergillus flavus,* 99.65, MN271387.1	*Aspergillus flavus,* 98.97, HF570041.1	*Aspergillus flavus,* 98.80, MG991523.1	*Aspergillus flavus,* 99.37, MN416023.1

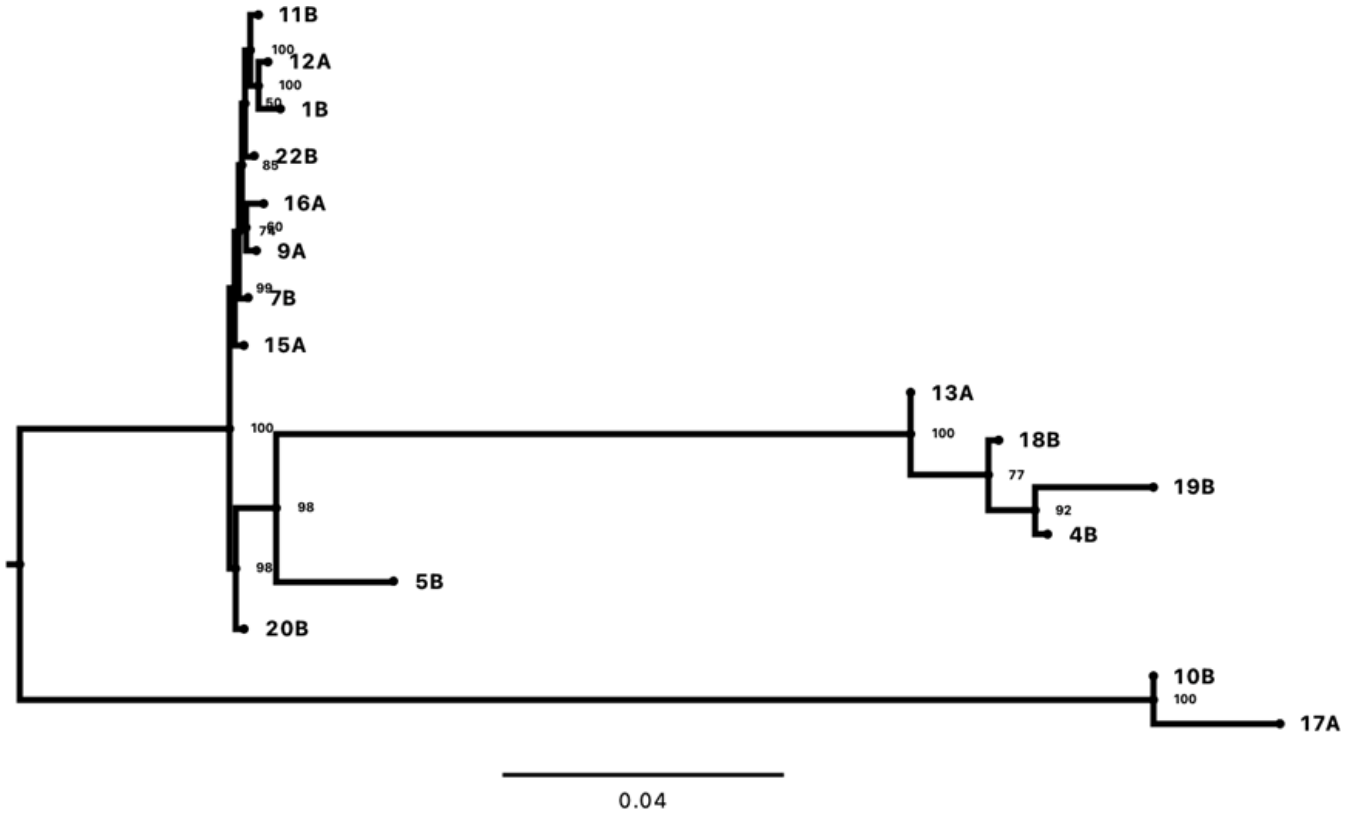

FIG 1 Phylogenetic tree of 16 predicted aflatoxin-producing *Aspergillus flavus de novo* genomes isolated from cashew nuts from coastal Kenya. The tree was constructed using the maximum likelihood algorithm in IQ-TREE v2.2.2.7 with the TVM + F + I + G4 model.

Completeness of the genome assemblies (Table 1) was determined using BUSCO v5.7.1 with the lineage database ascomycota_odb10 (16). The diversity of the *A. flavus* genomes was illustrated with phylogenetic analysis using maximum likelihood algorithm in IQTREE v2.2.2.7 with TVM + F + I + G4 model (17). The phylogenetic tree and midpoint rooting (Fig. 1) was generated using Figtree v1.4.4 (18). Default parameters were used for all software unless otherwise specified.

ACKNOWLEDGMENTS

We are grateful to Pwani University for providing the laboratory space at the Pwani University Bioscience Research Center (PUBReC) to perform this work. The authors acknowledge the National Research Fund for providing research grants.

This research was supported by the National Research Fund, Kenya (NRF/2/MMC/158).

P.W.G., M.A.O., E.M.M., K.K., and C.R.M. performed all the experiments, analyzed the data and wrote the manuscript, D.W.M., and B.S.J. assisted in some of the investigations and data analysis, W.M.M. supervised the work, contributed with experimental design and coordination, and reviewed the manuscript, W.M.M. conceptualized the idea. All authors have read and agreed to the published version of the manuscript.

The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

AUTHOR AFFILIATIONS

[1]Department of Biochemistry and Biotechnology, Pwani University, Kilifi, Kenya
[2]Pwani University Bioscience Research Centre PUBReC, Pwani University, Kilifi, Kenya

AUTHOR ORCIDs

Wilton Mwema Mbinda http://orcid.org/0000-0003-0586-0176

FUNDING

Funder	Grant(s)	Author(s)
National Research Fund, Kenya (NRF)	NRF/2/MMC/158	Wilton Mwema Mbinda

AUTHOR CONTRIBUTIONS

Pauline Wambui Gachanja, Formal analysis, Investigation, Methodology, Writing – original draft | Manase Aloo Onyango, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft | Colletah Rhoda Musangi, Formal analysis, Investigation, Methodology, Writing – original draft | Bicko Steve Juma, Formal analysis, Investigation, Methodology, Writing – original draft | Dennis Wamalabe Mukhebi, Formal analysis, Investigation, Methodology, Writing – original draft | Eugene Mwanza Muzami, Formal analysis, Investigation, Methodology, Writing – original draft | Kyalo Katua, Formal analysis, Investigation, Methodology, Writing – original draft | Wilton Mwema Mbinda, Conceptualization, Data curation, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review and editing

DATA AVAILABILITY

The whole genome sequences of the 16 *A. flavus* were deposited in the NCBI Genbank project under Bioproject number PRJNA1051575. The GenBank and Sequence Read Archive (SRA) accession numbers are listed in Table 1. This is the first announcement of *A. flavus* draft genome sequences from cashew nuts from coastal Kenya.

REFERENCES

1. Amaike S, Keller NP. 2011. *Aspergillus flavus*. Annu Rev Phytopathol 49:107–133. https://doi.org/10.1146/annurev-phyto-072910-095221
2. Wang Y, Zhou Y, Qin Y, Wang L. 2022. Effect of environmental factors on the aflatoxin production by *Aspergillus flavus* during storage in upland rice seed using response surface methodology. LWT 169:113977. https://doi.org/10.1016/j.lwt.2022.113977
3. Weaver MA, Scheffler BE, Duke M, Ballard L, Abbas HK, Grodowitz MJ. 2017. Genome sequences of three strains of *Aspergillus flavus* for the biological control of aflatoxin. Genome Announc 5:e01204-17. https://doi.org/10.1128/genomeA.01204-17
4. Nji QN, Babalola OO, Mwanza M. 2022. Aflatoxins in maize: can their occurrence be effectively managed in Africa in the face of climate change and food insecurity? Toxins (Basel) 14:574. https://doi.org/10.3390/toxins14080574
5. Stepman F. 2018. Scalingup the impact of aflatoxin research in Africa. The role of social sciences. Toxins (Basel) 10:136. https://doi.org/10.3390/toxins10040136
6. Kos J, Anić M, Radić B, Zadravec M, Janić Hajnal E, Pleadin J. 2023. Climate change-a global threat resulting in increasing mycotoxin occurrence. Foods 12:2704. https://doi.org/10.3390/foods12142704
7. Whitaker T, Slate A, Doko B, Maestroni B, Cannavan A, eds. 2010. Sampling procedures to detect mycotoxins in agricultural commodities. Springer, Dordrecht, The Netherlands.
8. EC (Commission Regulation) No 401/2006 of 23 February (2006). Laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs. Official Journal of the European Union L 70:12–34.
9. White TJ, Bruns T, Lee SJ, Taylor J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics, p 315–322. In PCR protocols: a guide to methods and applications. Vol. 18.
10. Lamboni Y, Nielsen KF, Linnemann AR, Gezgin Y, Hell K, Nout MJR, Smid EJ, Tamo M, van Boekel MAJS, Hoof JB, Frisvad JC. 2016. Diversity in secondary metabolites including mycotoxins from strains of *Aspergillus* section *Nigri* isolated from raw cashew nuts from Benin, West Africa. PLoS One 11:e0164310. https://doi.org/10.1371/journal.pone.0164310
11. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. 2008. NCBI BLAST: a better web interface. Nucleic Acids Res 36:W5–W9. https://doi.org/10.1093/nar/gkn201
12. Musangi CR, Juma BS, Mukhebi DW, Isoe EM, Kibiti CM, Mbinda WM. 2024. *Aspergillus* population diversity and its role in aflatoxin contamination of cashew nuts from coastal Kenya. PLoS One 19:e0292519. https://doi.org/10.1371/journal.pone.0292519
13. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
14. Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560
15. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes *de novo* assembler. Curr Protoc Bioinformatics 70:e102. https://doi.org/10.1002/cpbi.102
16. Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol 38:4647–4654. https://doi.org/10.1093/molbev/msab199
17. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274. https://doi.org/10.1093/molbev/msu300
18. Rambaut A. Figtree 1.4.4. Available from: http://tree.bio.ed.ac.uk/software/figtree/. Accessed July 4, 2024