AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

| Applied and Industrial Microbiology | Announcement

Draft genome assemblies of 38 *Aspergillus parasiticus* isolates collected from South Georgia crop fields

Sunil S. Gangurde,[1,2,3] Sejal Asija,[1,3] Prasad Bajaj,[4] Jake C. Fountain,[5] Hamed K. Abbas,[6] C. Corley Holbrook,[1] Robert C. Kemerait,[3] Baozhu Guo[1]

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT *Aspergillus parasiticus* is a fungus recognized for producing highly carcinogenic mycotoxins. In this study, we collected 38 isolates of *A. parasiticus* from fields in South Georgia. We performed whole genome re-sequencing and developed 38 draft genome assemblies of *A. parasiticus*. The average genome size was 38.7 Mb, with larger genomes (~40 Mb) found in peanut fields in Turner County. Scaffold N50 was recorded highest for isolates collected from the corn fields of Tifton. The average BUSCO completeness score for these assemblies was 99.1%. The genome sequences generated for these 38 isolates will serve as a valuable genomic resource for the community working on aflatoxin mitigation strategies in crops.

KEYWORDS *Aspergillus parasiticus*, aflatoxin, genome sequencing, carcinogenic, mycotoxins

Aspergillus parasiticus is a saprophytic mold, classified as a subspecies of *Aspergillus flavus* (1). The *A. parasiticus* was first reported in the Sugarcane field of Hawaii in 1912 (2). *A. parasiticus* is a highly toxigenic species that belongs to section "*Flavi*" of the *Aspergillus* genus. Recent advancements in sequencing technologies generated massive sequencing data for many isolates, which helped uncover the regulation of various pathways associated with secondary metabolite production (3, 4). We recently developed a pangenome for *Aspergillus flavus* called "Aflapan" using 346 isolates collected from 11 states of the USA and identified novel gene clusters associated with aflatoxin production (4). In this study, we report 38 draft genome assemblies of *Aspergillus parasiticus* (Table 1). The isolates in this study were collected from the soils of three counties in Georgia. Specifically, three isolates were obtained from Irwin County fields where cotton and peanuts were grown. Eight isolates were collected from Turner County fields where peanuts, corn, and cotton were cultivated. The majority, 27 isolates, were collected from Tift County fields with sunflower, corn, cotton, and peanuts. The *A. parasiticus* isolates were cultured as per the protocol (5) with slight modifications. DNA from each isolate was isolated using the normal cetyltrimethylammonium bromide (CTAB) method (6). The DNA quality was checked using fluorometer Qubit 3.0 and 0.8% agarose gel. The sequencing and initial data filtering were conducted at Novogene Corp. (Sacramento, CA). Pair-end sequencing data were generated using Illumina HiSeq 4000 platform with read length PE150. The raw data generated for 38 *A. parasiticus* isolates were submitted to NCBI Sequence Read Archive Bio-project ID PRJNA1191183. Adapters were trimmed using Trimmomatic v0.40 (7), and trimmed reads were assembled using SPAdes v.3.15.4 (8). Pilon v.1.22 improved the assemblies (9). The assembly stats were generated using the tool QUAST v5.0.2 (10). The average genome size for *A. parasiticus* isolates was ~38.7 Mb (Table 1). The average BUSCO completeness score for these assemblies was 99.1% (Table 1). The BUSCO scores were calculated using the BUSCO v. 6.0.0 (11). Some isolates showed the largest genome size of ~40 Mb, which were

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

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

The authors declare no conflict of interest.

Received 27 January 2025
Accepted 14 August 2025
Published 12 September 2025

TABLE 1 Assembly statistics of the 38 *Aspergillus parasiticus* isolates

SN	Isolate name	Source	County	Number of scaffolds	No of contigs	Total genome length (bp)	Contig N50	Scaffold N50	BUSCO completeness (%)	NCBI accession
1	GA8_2	GA Soil: Cotton, Peanut	Irwin	193	210	38,678,779	647,683	721,253	99.2	JBPPNO0000000000
2	GA8_7	GA Soil: Cotton, Peanut	Irwin	205	214	38,651,458	490,355	604,993	99.2	JBPPNP000000000
3	GA8_1	GA Soil: Cotton, Peanut	Irwin	204	220	38,604,774	676,608	704,779	99.2	JBPPOU000000000
4	GA15_2	GA Soil: Peanut	Turner	696	754	40,445,389	170,002	201,733	99.2	JBPPOR000000000
5	GA15_5	GA Soil: Peanut	Turner	683	728	40,567,069	193,058	211,409	99.2	JBPPOS000000000
6	GA15_6	GA Soil: Peanut	Turner	954	1034	40653439	108,271	121,057	99.1	JBPPOT000000000
7	GA2_2	GA Soil: Corn, Cotton	Turner	216	220	38,849,244	1,011,652	1,017,690	99.2	JBPPOA000000000
8	GA15_1	GA Soil: Peanut	Turner	243	259	38930194	539,141	556,828	99.1	JBPPOL000000000
9	GA1_2	GA Soil: Cotton, Peanut	Turner	274	300	38,599,977	439,556	562,276	99.1	JBPPOV000000000
10	GA15_4	GA Soil: Peanut	Turner	315	337	38941268	298,107	342,207	99.8	JBPPOZ000000000
11	GA1_7	GA Soil: Cotton, Peanut	Turner	250	272	38,583,048	481,691	651,329	99.2	JBPPOW000000000
12	GA11_16	GA Soil: Sunflower	Tifton	220	244	38675675	560,226	634,552	99.2	JBPPOX000000000
13	GA1_14	GA Soil: Cotton, Peanut	Tifton	264	271	38,748,395	678,058	682,400	99.2	JBPPOY000000000
14	GA10_18	GA Soil: Corn, Cotton, Peanut	Tifton	223	245	38,569,654	645,191	756,912	99.2	JBPPNQ000000000
15	GA10_11	GA Soil: Corn, Cotton, Peanut	Tifton	255	275	38,478,180	405,183	501,646	99.2	JBPPNR000000000
16	GA10_16	GA Soil: Corn, Cotton, Peanut	Tifton	248	268	38,606,275	468,484	544,209	99.2	JBPPNS000000000
17	GA10_26	GA Soil: Corn, Cotton, Peanut	Tifton	169	180	38,554,143	1,104,656	1,104,656	99.2	JBPPNT000000000
18	GA10_14	GA Soil: Corn, Cotton, Peanut	Tifton	208	212	38,631,056	1,040,575	1,046,984	99.1	JBPPNU000000000
19	GA10_19	GA Soil: Corn, Cotton, Peanut	Tifton	171	190	38,538,164	663,647	802,154	99.2	JBPPNV000000000
20	GA10_21	GA Soil: Corn, Cotton, Peanut	Tifton	191	203	38,481,385	594,581	754,536	99.3	JBPPNW000000000
21	GA10_2	GA Soil: Corn, Cotton, Peanut	Tifton	200	223	38,462,387	485,004	584,674	99.2	JBPPNX000000000
22	GA10_3	GA Soil: Corn, Cotton, Peanut	Tifton	235	266	38,460,154	360,340	442,919	99.2	JBPPNY000000000
23	GA10_23	GA Soil: Corn, Cotton, Peanut	Tifton	164	183	38,515,256	707,862	995,569	99.2	JBPPNZ000000000
24	GA10_8	GA Soil: Corn, Cotton, Peanut	Tifton	202	207	38,450,145	657,000	607,000	99.2	JBPPOC000000000
25	GA10_24	GA Soil: Corn, Cotton, Peanut	Tifton	188	198	38,586,374	807,901	834,830	99.2	JBPPOD000000000
26	GA10_7	GA Soil: Corn, Cotton, Peanut	Tifton	185	203	38,482,017	628,815	731,947	99.2	JBPPOE000000000
27	GA10_6	GA Soil: Corn, Cotton, Peanut	Tifton	233	274	38,528,282	378,594	515,229	99.2	JBPPOF000000000
28	GA10_27	GA Soil: Corn, Cotton, Peanut	Tifton	194	220	38,523,427	500,262	641,900	99.1	JBPPOG000000000
29	GA10_22	GA Soil: Corn, Cotton, Peanut	Tifton	244	262	38,638,746	605,242	702,787	99.2	JBPPOH000000000
30	GA10_12	GA Soil: Corn, Cotton, Peanut	Tifton	244	255	38,453,068	445,104	476,244	99.2	JBPPOI000000000
31	GA10_9	GA Soil: Corn, Cotton, Peanut	Tifton	215	225	38,492,334	600,039	622,440	99.2	JBPPOJ000000000
32	GA10_20	GA Soil: Corn, Cotton, Peanut	Tifton	196	212	38,524,463	561,306	689,191	99.2	JBPPOK000000000
33	GA10_1	GA Soil: Corn, Cotton, Peanut	Tifton	203	216	38,538,030	503,419	602,793	99.1	JBPPOM000000000
34	GA10_13	GA Soil: Corn, Cotton, Peanut	Tifton	351	377	38,411,443	261,552	279,413	99.1	JBPPON000000000
35	GA10_15	GA Soil: Corn, Cotton, Peanut	Tifton	203	213	38,596,072	857,294	891,422	99.2	JBPPOO000000000
36	GA10_4	GA Soil: Corn, Cotton, Peanut	Tifton	236	271	38,442,523	421,903	518,109	99.2	JBPPOP000000000
37	GA10_10	GA Soil: Corn, Cotton, Peanut	Tifton	213	219	38,471,889	590,989	590,989	99.2	JBPPOQ000000000
38	GA10_17	GA Soil: Corn, Cotton, Peanut	Tifton	208	239	38,610,978	702,885	978,473	99.1	JBPPOB000000000

collected from peanut fields of Turner County. Scaffold N50 was recorded highest for GA10_26 collected from the corn field of Tifton. The lowest scaffold N50 was recorded for GA15_6 collected from the peanut field of Turner County. Maximum scaffold size was recorded for GA2_2 collected from the corn field of Turner County. The genome assemblies developed in this study will be an important resource for the community to understand the structural variations associated with aflatoxin production and secondary metabolite production.

ACKNOWLEDGMENTS

We thank Billy Wilson and Sheron Simpson for technical assistance in the laboratory and fields. This work is partially supported by the USDA-ARS, Georgia Agricultural Commodity Commission for Corn, National Corn Growers Association Aflatoxin Mitigation Center of

Excellence, Georgia Peanut Commission, National Peanut Board, and Peanut Research Foundation. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the USDA.

AUTHOR AFFILIATIONS

[1]USDA Agricultural Research Service, Crop Genetics and Breeding Research Unit, Tifton, Georgia, USA
[2]CIMMYT-China Shandong Wheat and Maize Research Center, College of Agronomy, Shandong Agricultural University, Tai'an, China
[3]Department of Plant Pathology, University of Georgia, Tifton, Georgia, USA
[4]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India
[5]Department of Plant Pathology, University of Georgia, Griffin, Georgia, USA
[6]USDA Agricultural Research Service, Biological Control of Pests Research Unit, Stoneville, Mississippi, USA

AUTHOR ORCIDs

Baozhu Guo http://orcid.org/0000-0002-9079-7126

AUTHOR CONTRIBUTIONS

Sunil S. Gangurde, Data curation, Writing – original draft | Sejal Asija, Data curation, Writing – review and editing | Prasad Bajaj, Data curation | Jake C. Fountain, Data curation, Investigation, Resources, Writing – review and editing | Hamed K. Abbas, Data curation, Writing – review and editing | C. Corley Holbrook, Investigation, Writing – review and editing | Robert C. Kemerait, Investigation, Writing – review and editing | Baozhu Guo, Conceptualization, Funding acquisition, Investigation, Resources, Supervision, Writing – review and editing

DATA AVAILABILITY

The raw sequencing data and genome assemblies are deposited in NCBI with Bioproject ID PRJNA1191183.

REFERENCES

1. 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
2. Speare AT. 1912. Fungi parasitic upon insects injurious to sugar cane. Pathol Physiol Series 12
3. Drott MT, Rush TA, Satterlee TR, Giannone RJ, Abraham PE, Greco C, Venkatesh N, Skerker JM, Glass NL, Labbé JL, Milgroom MG, Keller NP. 2021. Microevolution in the pansecondary metabolome of *Aspergillus flavus* and its potential macroevolutionary implications for filamentous fungi. Proc Natl Acad Sci USA 118:e2021683118. https://doi.org/10.1073/pnas.2021683118
4. Gangurde SS, Korani W, Bajaj P, Wang H, Fountain JC, Agarwal G, Pandey MK, Abbas HK, Chang P-K, Holbrook CC, Kemerait RC, Varshney RK, Dutta B, Clevenger JP, Guo B. 2024. *Aspergillus flavus* pangenome (AflaPan) uncovers novel aflatoxin and secondary metabolite associated gene clusters. BMC Plant Biol 24:354. https://doi.org/10.1186/s12870-024-04950-8
5. Abbas HK, Weaver MA, Zablotowicz RM, Horn BW, Shier WT. 2005. Relationships between aflatoxin production and sclerotia formation among isolates of *Aspergillus* section Flavi from the Mississippi Delta. Eur J Plant Pathol 112:283–287. https://doi.org/10.1007/s10658-004-4888-8
6. 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 9:10–1128. https://doi.org/10.1128/MRA.00478-20
7. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
8. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021
9. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9:e112963. https://doi.org/10.1371/journal.pone.0112963
10. Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 34:i142–i150. https://doi.org/10.1093/bioinformatics/bty266
11. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351