7-23-2020

Sixteen Draft Genome Sequences Representing the Genetic Diversity of Aspergillus flavus and Aspergillus parasiticus Colonizing Peanut Seeds in Ethiopia

Renee Arias
United States Department of Agriculture, renee.arias@usda.gov

Abdi Mohammed
Haramaya University, Dire Dawa

Valerie A. Orner
United States Department of Agriculture

Paola C. Faustinelli
United States Department of Agriculture

Marshall C. Lamb
United States Department of Agriculture, marshall.lamb@usda.gov

See next page for additional authors

AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Sixteen Draft Genome Sequences Representing the Genetic Diversity of *Aspergillus flavus* and *Aspergillus parasiticus* Colonizing Peanut Seeds in Ethiopia

Renee S. Arias,[a] Abdi Mohammed,[b] Valerie A. Orner,[a] Paola C. Faustinelli,[a]* Marshall C. Lamb,[a] Victor S. Sobolev[a]

[a]U.S. Department of Agriculture, Agricultural Research Service, National Peanut Research Laboratory, Dawson, Georgia, USA
[b]School of Plant Science, College of Agriculture and Environmental Sciences, Haramaya University, Dire Dawa, Ethiopia

ABSTRACT Draft genomes of 16 isolates of *Aspergillus flavus* Link and *Aspergillus parasiticus* Speare, identified as the predominant genotypes colonizing peanuts in four farming regions in Ethiopia, are reported. These data will allow mining for sequences that could be targeted by RNA interference to prevent aflatoxin accumulation in peanut seeds.

Aspergillus flavus and *Aspergillus parasiticus* are commonly found in staple food grains such as maize, peanut, and many other crops (1). In seeds, these fungi can accumulate carcinogenic compounds called aflatoxins (2), as well as the neurotoxin cyclopiazonic acid (3). Aflatoxin accumulation in seeds can be controlled by plant host RNA interference-mediated silencing of fungal aflatoxin biosynthesis genes (4, 5); this requires knowing gene sequences of the invading fungus. *A. flavus* and *A. parasiticus* strains were collected from four peanut-growing districts in Ethiopia, namely, Fedis, Babile, Darolabu, and Gursum. The fungi were isolated on modified dichloran-rose bengal (MDRB) agar medium (6), restreaked onto MDRB agar, from which hyphal tips of single colonies were transferred to slants of Czapek's medium (7), and identified following a previously described protocol (8). Cluster analysis of the genetic fingerprints of the isolates using the 25 insertion/deletion markers within the aflatoxin biosynthesis pathway and a protocol published by Faustinelli et al. (9) revealed several clades, from which 16 representative isolates (4 *A. parasiticus* strains, 5 *A. flavus* strains, 2 *A. flavus* L strains [10], and 5 *A. flavus* S strains [10]) were chosen for whole-genome sequencing, as reported in Table 1. DNA was extracted from spores and sclerotia using the DNeasy plant minikit (Qiagen, Valencia, CA) after growth of the isolates on MDRB agar for 3 days in the dark at 30°C.

Barcode-indexed sequencing libraries were generated from genomic DNA sheared on an E220 focused ultrasonicator (Covaris, Woburn, MA) and size selected with a double-sided solid-phase reversible immobilization (SPRI) protocol with AMPure XP beads (Beckman Coulter, Brea, CA) using bead-to-sample ratios of 0.6 and 0.73. The size-selected DNA samples were converted to sequencing libraries using a KAPA HyperPrep library preparation kit (Kapa Biosystems-Roche, Basel, Switzerland). The libraries were amplified with six PCR cycles, analyzed with a Bioanalyzer 2100 instrument (Agilent, Santa Clara, CA), quantified by fluorometry with a Qubit instrument (Life Technologies, Carlsbad, CA), and combined in a pool at equimolar ratios. The pool was quantified with a KAPA library quantification kit (Kapa Biosystems-Roche) on a QuantStudio 5 real-time PCR system (Applied Biosystems, Foster City, CA) and sequenced on a HiSeq 4000 system (Illumina, San Diego, CA) with paired-end 150-bp reads. Reads were trimmed for removal of adapters using the Phred quality score in the modified Mott trimming algorithm for this purpose (average length after trimming was 148.4 nucleotides) and then were *de novo* assembled with CLC Genomics Workbench version

Citation Arias RS, Mohammed A, Orner VA, Faustinelli PC, Lamb MC, Sobolev VS. 2020. Sixteen draft genome sequences representing the genetic diversity of *Aspergillus flavus* and *Aspergillus parasiticus* colonizing peanut seeds in Ethiopia. Microbiol Resour Announc 9:e00591-20. https://doi.org/10.1128/MRA.00591-20.

Editor Christina A. Cuomo, Broad Institute

Address correspondence to Renee S. Arias, renee.arias@usda.gov.

* Present address: Paola C. Faustinelli, JLA International, Albany, Georgia, USA.

Received 11 June 2020
Accepted 1 July 2020
Published 23 July 2020

TABLE 1 Data for 16 *Aspergillus* isolates from peanuts in Ethiopia

Sample name[a]	District of origin	NCBI accession no. for assembled genome	NCBI accession no. for raw data	No. of scaffolds	No. of contigs	ABC contig location (nucleotide position)	Approx ABC position (kbp)	No. of raw reads	No. of reads after trimming	N_{50} (bp), including scaffolded regions	Genome size (bp)	Coverage (×)
E1201, *A. flavus*	Babile	SKBW00000000	PRJNA509212	271	482	37	91.8–171	31,638,796	31,603,343	1,237,922	37,716,145	124
E1236, *A. flavus*	Babile	SKBX00000000	PRJNA522284	76	496	29	180–259	29,293,138	29,269,867	1,326,349	37,575,901	116
E1275, *A. flavus*	Babile	SKBY00000000	PRJNA522289	68	421	2	69–148	30,051,746	30,015,148	1,138,433	37,567,828	119
E1288, *A. flavus* (S)	Babile	SKBZ00000000	PRJNA522452	136	561	58	41.7–123	33,291,634	33,246,727	1,149,099	38,045,091	130
E1293, *A. flavus* (S)	Babile	SJEZ00000000	PRJNA522457	84	419	44	586–666	32,918,814	32,844,669	995,785	37,882,296	128
E1316, *A. flavus* (S)	Darolabu	SJFA00000000	PRJNA522783	127	580	17	131–213.5	30,963,252	30,897,449	1,093,467	38,074,444	120
E1319, *A. parasiticus*	Darolabu	SJFB00000000	PRJNA522792	143	666	1	718–807	30,698,482	30,648,087	673,791	37,947,312	120
E1337, *A. parasiticus*	Darolabu	SJFC00000000	PRJNA522796	233	695	9	25–114	31,448,004	31,378,441	696,941	39,150,789	119
E1345, *A. flavus*	Darolabu	SJFD00000000	PRJNA523165	82	434	17	75–153	29,777,742	29,716,598	1,360,789	37,430,661	118
E1348, *A. parasiticus*	Darolabu	SJFE00000000	PRJNA523169	194	863	43	135–224	27,589,882	27,550,162	704,672	39,351,768	104
E1376, *A. flavus* (S)	Fedis	SJFG00000000	PRJNA523208	124	647	21	1,423–1,502	31,129,974	31,046,692	1,027,715	37,914,373	121
E1402, *A. flavus* (L)	Fedis	SJFH00000000	PRJNA523216	86	485	3	1,044–1,124	28,333,840	28,275,793	1.037,074	37,867,696	111
E1404, *A. flavus* (L)	Fedis	SJFI00000000	PRJNA523224	100	571	22	81–161	29,343,244	29,303,931	979,662	37,695,620	115
E1406, *A. flavus* (S)	Gursum	SJFJ00000000	PRJNA523229	96	532	14	1,428–1,507	28,985,738	28,914,704	928,368	37,990,949	113
E1443, *A. parasiticus*	Gursum	SJFK00000000	PRJNA523233	663	1,243	45	65–154	26,190,370	26,116,731	740,392	41,459,518	94
E1445, *A. flavus*	Gursum	SJFL00000000	PRJNA523235	76	491	34	180–259	30,424,294	30,380,530	1,120,665	37,603,622	120

[a] L and S strains are described in reference 10.

12.0 (Qiagen, Aarhus, Denmark) using default parameters and no minimum contig length. The average G+C content observed across isolates was 47.4% ± 0.07%, and a summary of the geographic origins and genome sequencing statistics of the isolates is shown in Table 1. Assembled contigs of each isolate were converted to BLAST (11) databases in CLC Genomics Workbench, and then the aflatoxin biosynthesis cluster (ABC) of *A. flavus* or *A. parasiticus* was searched by BLAST analysis within each genome. In all 16 isolates, the ABC was observed in single contigs; ABC locations were estimated by alignment to contigs (Table 1).

Data availability. All 16 genomes have been deposited in NCBI GenBank (Table 1). Raw data are available in the NCBI SRA database (Table 1). Fungal isolates were obtained from foreign samples and are hosted at the USDA ARS National Peanut Research Laboratory culture collection; their access requires proper APHIS shipping/receiving permits (contact renee.arias@usda.gov).

ACKNOWLEDGMENTS

We thank Lutz Froenike for providing detailed information on library construction.

The sequencing was carried out at the UC Davis Genome Center DNA Technologies and Expression Analysis Core, supported by NIH shared instrumentation grant 1S10OD010786-01; the project was partly funded by grants from the USAID (agreement 58-0210-3-0012) and the USAID-funded Peanut Innovation Laboratory (award AIDECGA000700001). The overall project was supported by NP301 project 6044-21000-005-00D. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

REFERENCES

1. Perrone G, Susca A, Cozzi G, Ehrlich K, Varga J, Frisvad JC, Meijer M, Noonim P, Mahakarnchanakul W, Samson RA. 2007. Biodiversity of *Aspergillus* species in some important agricultural products. Stud Mycol 59:53–66. https://doi.org/10.3114/sim.2007.59.07.
2. IARC Working Group. 1980. An evaluation of chemicals and industrial processes associated with cancer in humans based on human and animal data: IARC Monographs Volumes 1 to 20. Cancer Res 40:1–12. https://cancerres.aacrjournals.org/content/40/1/1.
3. Lansden JA. 1986. Determination of cyclopiazonic acid in peanuts and corn by thin layer chromatography. J Assoc Off Anal Chem 69:964–966. https://doi.org/10.1093/jaoac/69.6.964.
4. Arias RS, Dang PM, Sobolev VS. 2015. RNAi-mediated control of aflatoxins in peanut: method to analyze mycotoxin production and transgene expression in the peanut/*Aspergillus* pathosystem. J Vis Exp 106:e53398. https://doi.org/10.3791/53398.
5. Thakare D, Zhang J, Wing RA, Cotty PJ, Schmidt MA. 2017. Aflatoxin-free transgenic maize using host-induced gene silencing. Sci Adv 3:e1602382. https://doi.org/10.1126/sciadv.1602382.
6. Horn BW, Dorner JW. 1998. Soil populations of *Aspergillus* species from section *Flavi* along a transect through peanut-growing regions of the United States. Mycologia 90:767–776. https://doi.org/10.1080/00275514.1998.12026969.
7. Horn BW, Dorner JW. 1999. Regional differences in production of aflatoxin B_1 and cyclopiazonic acid by soil isolates of *Aspergillus flavus* along a transect within the United States. J Appl Environ Microbiol 65:1444–1449. https://doi.org/10.1128/AEM.65.4.1444-1449.1999.
8. Mohammed A, Chala A, Dejene M, Fininsa C, Hoisington DA, Sobolev VS, Arias RS. 2016. *Aspergillus* and aflatoxin in groundnut (*Arachis hypogaea* L.) and groundnut cake in eastern Ethiopia. Food Addit Contam Part B Surveill 9:290–298. https://doi.org/10.1080/19393210.2016.1216468.
9. Faustinelli PC, Palencia ER, Sobolev VS, Horn BW, Sheppard HT, Lamb MC, Wang XM, Scheffler BE, Martinez-Castillo J, Arias RS. 2017. Study of the genetic diversity of the aflatoxin biosynthesis cluster in *Aspergillus* section *Flavi* using insertion/deletion markers in peanut seeds from Georgia, USA. Mycologia 109:200–209. https://doi.org/10.1080/00275514.2017.1307095.
10. Moore GG, Olarte RA, Horn BW, Elliott JL, Singh R, O'Neal CJ, Carbone I. 2017. Global population structure and adaptive evolution of aflatoxin-producing fungi. Ecol Evol 7:9179–9191. https://doi.org/10.1002/ece3.3464.
11. 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.