Inferring the presence of aflatoxin-producing *Aspergillus flavus* strains using RNA sequencing and electronic probes as a transcriptomic screening tool

Andres S. Espindola[1,2], William Schneider[3], Kitty F. Cardwell[1,2], Yisel Carrillo[1], Peter R. Hoyt[3], Stephen M. Marek[1], Hassan Melouk[1] and Carla D. Garzon[1]

[1] Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, Oklahoma, United States of America

[2] National Institute for Microbial Forensics and Food and Agricultural Biosecurity (NIMFFAB), Oklahoma State University, Stillwater, Oklahoma, United States of America

[3] Department of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, Oklahoma, United States of America

Corresponding Author

E-mail: andres.espindola@okstate.edu

ABSTRACT

E-probe Diagnostic for Nucleic acid Analysis (EDNA) is a bioinformatic tool originally developed to detect plant pathogens in metagenomic databases. However, enhancements made to EDNA increased its capacity to conduct hypothesis directed detection of specific gene targets present in transcriptomic databases. To target specific pathogenicity factors used by the pathogen to infect its host or other targets of interest, e-probes need to be developed for transcripts related to that function. In this study, EDNA transcriptomics (EDNAtran) was developed to detect the expression of genes related to aflatoxin production at the transcriptomic level. E-probes were designed from genes up-regulated during *A. flavus* aflatoxin production. EDNAtran detected gene transcripts related to aflatoxin production in a transcriptomic database from corn, where aflatoxin was produced. The results were significantly different from e-probes being used in the transcriptomic database where aflatoxin was not produced (atoxigenic AF36 strain and toxigenic AF70 in Potato Dextrose Broth).

INTRODUCTION

Maize [1], peanuts [2], tree nuts, dried spices [3] and cottonseed [4] are crops that can be infected during the pre-harvest, post-harvest and/or storage period with *Aspergillus flavus* Link. This fungus produces polyketide secondary metabolites named aflatoxins. Among the four known aflatoxins (B_1, B_2, G_1, G_2), B_1 has been of special interest to food biosecurity due to its toxicity and potent carcinogenic properties [5]. *A. flavus* is a ubiquitous saprophytic ascomycete fungus grouped in the *Aspergillus* section Flavi, species with aflatoxin-producing strains including *A. flavus, A. parasiticus* and *A. nomius* [6,7].

Aflatoxin is produced through the interaction of approximately 25 genes in a cluster cascade [8–10]. Regulatory genes for the cluster are *aflR* and *aflS* (*aflJ*), where *aflR* encodes for a transcriptional factor of the type $Zn(II)_2Cys_6$ which binds promoter regions of many aflatoxin genes [11–14]. In contrast, aflS (*aflJ*) regulates aflatoxin production through binding and activating aflR [15]. Some strains of *A. flavus* do not produce aflatoxin and these have been shown to have deletion mutations, identified by 32 separate PCR amplifications [16]. Callicott and Cotty [17] have begun to use cluster amplification patterns (CAPS) to evaluate *A. flavus* populations based on deletions genotypes.

Aflatoxin contamination in food is highly regulated in multiple countries, consequently increasing management costs and final product price [18–20]. In the United States alone, the maximum allowed concentration of aflatoxin in food for human consumption is 20 ppb, as dictated by the U.S. Food and Drug Administration (FDA). Appropriate and accurate aflatoxin testing is necessary to opportunely control *A. flavus* infected crops. Among the most used techniques for aflatoxin detection and quantification are thin layer chromatography (TLC), high-performance liquid chromatography (HPLC), enzyme-linked immunosorbent assay (ELISA) and fluorometry [18], however, there are limitations in all of these for rapid testing. Industry costs for testing crops for aflatoxins in the United States alone have ranged from $30 to $50 million per year at approximately at $10 to $20 per sample tested [20].

A promising management strategy for aflatoxin reduction is inundative biological control using atoxigenic strains of *A. flavus* (Aflaguard® & AF36) [21–24]. Location or region-specific indigenous atoxigenic strains are selected to take advantage of inherent fitness

in the environment and to avoid potential adaptation problems [25]. The indigenous isolated atoxigenic strains are multiplied *in vitro,* grown onto a carrier and inoculated on crop fields with the goal of competitive niche exclusion of the toxigenic strain, by the atoxigenic strains. The result is the almost complete elimination of aflatoxin contamination in the crop [21,26]. While the exact mechanism is not understood, a high prevalence of atoxigenic inoculum results in temporal and physical niche displacement of the toxigenic strains. Nevertheless, a single year release is not sufficient to render the field permanently atoxigenic. Thus a yearly application of the biological control agent is recommended for optimum efficacy [21,24,27].

Rapid assessment of determining changing profiles in atoxigenic to toxigenic population shifts have been lacking. Callicot and Cotty (2015) are attempting to create sensitive assays with cluster amplification patterns (CAPS), which involves some 32 known markers, however at this time; the CAP markers are primarily suggested to be a research technology for understanding dynamics of the atoxigenic populations. Absent that, current methods for assessing soil toxigenic to atoxigenic ratios of *A. flavus* are expensive, time, labor and skill intensive. Both, the evaluation of levels of soil *A. flavus* strain ratios and post-biocontrol screening could be facilitated by new methods. If there were a sensitive, simple assay for relative toxigenicity, inoculation of the field with additional atoxigenic strains could be recommended only when empirical tests showing higher relative toxigenicity reach an action threshold [28].

We propose that for practical application, the viability and presence of the toxigenic strains can be inferred by testing for upregulation of the gene cascade that leads to presence of aflatoxin. A fast and tentatively less expensive screening tool for toxigenic

A. flavus strains might be sequencing the whole transcriptome (metatranscriptome) of the pathogen niche and determining the presence of toxigenic gene up-regulation. Previously, various approaches have been developed to use metagenomes in ecology studies and determine the microbial profile of natural ecosystems [29,30]. Yet, few have focused on the development of tools to detect microbes at the species/isolate level in agricultural ecosystems [31–33]. However, none of them has addressed the detection of gene activity and upregulation in agroecosystems. E-probe Diagnostic for Nucleic acid Analysis (EDNA) was designed to detect viruses, bacteria, fungi and oomycete plant pathogens by using species-specific markers named e-probes [31,32,34]. Here we modified EDNA to be used as a gene functional analysis and detection tool to infer the presence of aflatoxin. EDNA transcriptomics — a modification of the original EDNA's bioinformatic pipeline — was designed to incorporate genome annotations on the e-probe design as well as on the detection pipelines. EDNAtran is a theoretical approach that is being tested for the first time with *A. flavus* and could be extended to detect metabolic functions associated to pathogenicity in other host-pathogen systems. Detecting metabolic functions that could potentially lead to plant disease is crucial to incorporate proper and timely management practices in agroecosystems.

MATERIALS AND METHODS

Fungal isolates and culture methods

A. flavus strains were obtained as freeze-dried (AF36; ATCC 96045; atoxigenic) and frozen (AF70; ATCC MYA-384; toxigenic) cultures from ATCC (Manassas, VA). AF36 was reactivated by rehydration, adding 500 µL of sterilized distilled water inside the vial. Subsequently, 100 µL of the re-suspended AF36 was plated on Malt extract agar

112 Blakeslee's formula (MEAbl) and incubated at 31℃ in darkness until mycelium was
113 developed (72 hours), according to ATCC instructions. AF70 was thawed for 5 minutes,
114 directly plated onto Malt extract agar, and incubated at 25 ℃ in darkness until mycelium
115 was developed (72 hours), according to ATCC instructions. Agar plugs with actively
116 growing mycelia were re-plated in MEAbl agar and incubated at their optimal
117 temperatures in the dark until extensive conidial development (5 days) was observed.
118 The cultures (AF36 and AF70) containing extensive conidia growth were used to
119 inoculate ground corn and Potato Dextrose Broth (PDB).

120 Corn substrate was prepared using dried corn kernels (*Zea mays*). Kernels were weight
121 (20g) and ground (using a coffee grinder Mr. Coffee Precision Coffee Grinder IDS77)
122 until obtaining pieces with the approximate texture of coarse sand (0.5-1mm in
123 diameter). The coarse grains were autoclaved (dry cycle) for 20 minutes in
124 polycarbonate containers (Magenta GA-7, Plantmedia, US) and its humidity was
125 adjusted to keep between 25 – 33% w/v (Modified from Woloshuk, Cavaletto, and
126 Cleveland 1997).

127 Ground corn kernels and PDB media were inoculated with conidial suspensions
128 obtained by washing *A. flavus* MEAbl plates with 2 mL of sterile distilled water. Conidia
129 collected (2 mL) were then added to a single vial containing 4mL of distilled water for a
130 final dilution of 3:1 v/v (Spore suspension was not quantified). Six mL of spore
131 suspension was used to inoculate each replicate (20 g of ground corn and PDB). The
132 ground grain was inoculated with the *A. flavus* suspension in polycarbonate containers
133 and homogeneously mixed by rolling the containers to allow uniform distribution of the
134 conidia. Similarly, 250 mL flasks containing 44mL of PDB were inoculated with 6 mL of

135 *A. flavus* spore suspension. The containers and PDB plates were incubated at 31℃ in
136 the dark for 10 days.

137 ***RNA extraction and sequencing***

138 Ground corn kernels inoculated with the AF70 and AF36 strains produced extensive
139 conidia, which were suspended by gently adding 10 mL of sterilized water to the
140 magenta containers. The containers were shaken gently to homogenize the spores and
141 then 1 mL of the spore suspension was obtained and added to a capped 2mL tube
142 containing silica beads. The conidia cell walls were disrupted by shaking the 2mL tubes
143 using a bead beater (2 cycles of 20 seconds). The lysate was then transferred (500 µL)
144 to a column of the Qiagen RNeasy Plant Mini Kit for RNA extraction. On the other hand,
145 for AF36 and AF70 growing on PDB, mycelia/spores were recovered by filtrating them
146 using Whatman paper. 100 mg of mycelium/spores were weight and added to a column
147 of the Qiagen RNeasy Plant Mini Kit to continue with the RNA extraction procedure. The
148 RNA quality and integrity were assessed using a 2100 Agilent Bioanalyzer (Agilent
149 Technologies) for 12 RNA extraction samples from AF36 and AF70. A sample (per
150 strain) having RIN numbers higher than eight were selected for RNA sequencing. After
151 quality control, RNA was sequenced using the Illumina HiSeq 2500 sequencer at the
152 Core Facility of the University of Illinois at Urbana-Champaign, IL. The mRNA
153 sequencing library was created with PolyA capture method per manufacturer's protocol
154 and the library was sequenced as single-end.

155 ***Gene expression analysis***

156 RNA sequencing reads from samples AF70-corn and AF70-PDB were mapped onto the
157 *A. flavus* AF70 genome using STAR software [36] and bam binary files were

158 transformed from sam files using SAMtools (http:// samtools.sourceforge.net). Gene
159 expression analysis was performed with DeSeq2 in R by comparing AF70 growing on
160 two substrates (corn and PDB). The control was AF70 in PDB (non-conducive for
161 aflatoxin production), and the treatment was AF70 in corn (conducive for aflatoxin
162 production) [37]. Positive fold change (up-regulated) genes were selected using the
163 log2 fold change metric obtained from the DeSeq2 analysis. Upregulated gene
164 sequences having log2 fold changes greater than five were retrieved by an in-house
165 Linux bash script and kept in a multi-fasta file for later e-probe design.

166 ***E-probe design***

167 The genomes from *A. flavus* AF70 (Accession: JZDT00000000.1) and NRRL3357
168 (Accession: AAIH00000000.2) [38] were obtained from Genbank. Sequences for the
169 aflatoxin gene cluster of AF70 (AY510453) and AF36 (AY510455) were also retrieved
170 from GenBank [39]. E-probes 80 nt long were generated using the e-probe pipeline for
171 EDNAtran (Espindola *et al.,* 2015; Stobbe *et al.,* 2013). The aflatoxin gene cluster of
172 AF70 was used as target sequence and the same gene cluster for AF36 was used as
173 near neighbor sequence. E-probe specificity was verified by local alignment of each e-
174 probes with the intended target genome (AF70) using an stringency of 100% identity
175 and query coverage. Metadata information about the gene where it belongs was also
176 retrieved by the EDNAtran e-probe design pipeline. E-probe annotation was utilized to
177 select e-probes that belonged only to the up-regulated genes previously identified with
178 DeSeq2.

Rapid assessment of active aflatoxin metabolic pathway using EDNAtran

EDNAtran was utilized with default parameters (percent identity and query coverage of 100%) to assess four transcriptomic databases (**Table 1**) which included toxigenic (AF70) and atoxigenic (AF36) *A. flavus* strains growing in conducive (ground corn) and non-conducive (PDB) environment for aflatoxin production. E-probes designed in up-regulated genes of the aflatoxin gene cluster of AF70 was utilized during this analysis. Hit frequencies of raw reads with e-probes were recorded for each of the four treatments. Data on hit frequencies were analyzed for variance by ANOVA. For analysis of significant differences between hit frequencies, Tukey's HSD test and T-test at P=0.05 were used.

RESULTS

Assessing appropriate growing conditions for the production of aflatoxin

The isolates of *A. flavus* AF70 and AF36 showed different growth patterns and morphology on the different media (PDB and ground corn). Aflatoxin production is directly correlated with concentrations of free saccharides [40–43]; therefore AF70 and AF36 were grown in a toxin-inducing substrate (ground corn) and non-toxin inducing media (PDB). Free saccharides (glucose, sucrose, raffinose, etc) are localized in the germ of the corn grain when matured non-germinated seeds; however, in germinating corn, the endosperm is used as a source of saccharides [44]. Increased sclerotia production was observed in AF70, in contrast, AF36 produced conidia in all media. However, AF70 produced conidia 10 days post inoculation in corn.

200 ***RNA sequencing and Gene expression analysis***

201 RNA extracted from AF36 and AF70 strains grown on PDB and ground corn yielded
202 from 20 to 24,9 million reads per sequencing run (**Table 1**). The sequenced reads then
203 were mapped to the *A. flavus* AF70 strain genome to retrieve information about
204 potential up-regulation and down-regulation of genes by using STAR [36] and DESeq2.
205 In total, 44 genes were identified as up-regulated and 129 genes were down-regulated
206 in the *A. flavus* genome (*Error! Reference source not found.* **Table**). Identified as part of
207 the aflatoxin gene cluster, only 17 out of 44 upregulated genes. From two to six gene
208 fold changes were plotted in a hierarchical clustering heat map as well as in a MAPlot
209 (**Fig 1 Fig 2**).

210 **Fig 1. Mean Average Plot for RNA sequencing gene expression analysis. Red line**
211 **shows zero change in gene expression.** Blue dashed lines show no change in gene
212 expression and green dashed lines show a five-fold change in gene expression. Red
213 dots are genes that have been either up-regulated or down-regulated in *A. flavus* AF70
214 infecting ground corn. Gray dots depict genes that have not had enough statistical
215 evidence to be assigned a gene expression fold change.

216 **Fig 2. Hierarchical clustering map depicting *A. flavus* AF70 growing on PDA and**
217 **ground corn.** Gene expression fold change is differentiated by a color palette ranging
218 from red (most up-regulated genes have plus six-fold changes) to blue (most down-
219 regulated genes have minus six-fold change). Genes are clustered based on their
220 gene expression fold change to facilitate gene co-expression analysis.

221

E-probe generation for aflatoxin detection

In total, 231 highly specific e-probes were generated to detect the production of aflatoxin specifically for AF70. AF36 genome-wide e-probes were not generated because there is not a genome sequence available yet for that specific strain.

Detecting aflatoxin production using EDNAtran in A. flavus

As expected, 231 e-probes had hits creating High-Quality Matches (HQMs) in AF70-corn transcriptome datasets; meanwhile, AF70-PDB had only 39 HQMs (**Fig 3**). AF36-corn had only two HQMs and AF36-PDB had 12 HQMs (**Table 1**). EDNAtran discriminated between the transcriptomic databases with abundant aflatoxin production and the transcriptomes from low-toxin production based on EDNA eukaryotic metrics [32] (**Table 1**). However, to infer the presence of aflatoxin we can only use frequencies of hits as an indirect measure. In this case, the number of times a read was mapped to an e-probe was recorded and counted without any limits. A dot plot of alignment length vs. percent identity with marginal hit frequencies facilitates visualizing hit frequencies. Specifically for *A. flavus* AF70 in corn, it was observable that the hit frequencies were very high — around 9,000 hits per e-probe — when the alignments are above 90% identity and the alignment length was approaching to the total length of the e-probe (**Fig 3A**). Conversely, for AF70 in PDB and AF36, the marginal plots show a low frequency of hits when alignment lengths and percent identities were above the threshold of 35nt and 90% respectively (**Fig 3B-3D**). Frequencies of hit values were square root converted and statistically analyzed with ANOVA to compare all the samples/treatments (**Fig 4**).

Fig 3. EDNA transcriptomics hits distribution and frequencies for *A. flavus* aflatoxin detection. (A and C) RNA sequencing of *A. flavus* AF70 and AF36 respectively growing on corn identified with 80-mer AF70 aflatoxin-specific e-probes. (B and D). RNA sequencing of *A. flavus* AF70 and AF36 respectively growing on PDB.

Fig 4. Hit frequencies of AF70 e-probes in RNA sequencing databases of *A. flavus*. Atoxigenic strain of *A. flavus* AF36 growing on both PDB and Corn. Similarly, the toxigenic strain of *A. flavus* AF70 growing on PDB and Corn. Differences in gene expression levels are directly correlated to hit frequencies.

Table 1. EDNA transcriptomics output table for the inference of aflatoxin in *A. flavus*.

Culture	Total reads	ARL	LRL	MRL	Probe length	TNP	HQM		
							Ev	NoEv	HSGM
AF70[a]	20657024	98.9	100	35	80	231	39	40	39
AF36[b]	22495368	99.0	100	35	80	231	2	2	2
AF36[a]	24134226	98.7	100	35	80	231	12	12	12
AF70[b]	24902500	98.9	100	35	80	231	231	231	231

ARL, average read length; LRL, largest read length; MRL, minimum read length; TNP, total number of probes; HQM, high-quality matches; Ev, matches include e-value in the scoring method; NoEv, matches do not include e-value in the scoring method; HSGM, high scoring general matches.

[a]*A. flavus* strain growing on PDB.

[b]*A. flavus* strain growing on ground corn.

The ANOVA in the *A. flavus* experiment had a p-value lower than 0.05 which rejects the null hypotheses (all hit frequencies are equal); therefore, a post-hoc analysis was

263 automatically performed using the Tukey HSD function in R. The post-hoc analysis and
264 T-test for *A. flavus* showed that e-probes hitting on RNA sequencing databases
265 obtained from *A. flavus* AF70 growing on ground corn were different from those of AF70
266 growing on PDB, and AF36 on corn and PDB (**Fig. 5** and **Table 2**). In conclusion,
267 EDNAtran was able to find statistically significant differences between the transcriptomic
268 data set of the highly toxigenic sample, from the non-toxigenic samples, using 231 e-
269 probes generated in this study.

270 **Fig. 5.** Post-hoc analysis of ANOVA using Tukey HSD with 95% of confidence for the
271 inference of aflatoxin transcriptional activation. Lines close to zero are interactions that
272 had no difference in their hit frequency means while lines closer to 30 are interactions
273 that had different hit frequency means.

274 **Table 2.** Pairwise T-test p-values comparing e-probe hit frequencies for *A. flavus* toxin
275 detection analysis using the aflatoxin e-probes.

	AF36-Corn	AF36-PDB	AF70-Corn
AF36-PDB	1	NA	NA
AF70-Corn	**3.72E-222***	**8.38E-221***	NA
AF70-PDB	1	1	**1.04E-220***

276 *Comparisons having p-values lower than 0.05 are considered significant.

277 **DISCUSSION**

278 EDNA has previously been proven to successfully detect a variety of plant pathogens
279 from raw metagenomic databases [32,34,45]. DNA as the main source of identification
280 has always been the gold standard for detecting organisms in a sample, although
281 viability is not assessed. Therefore, the question about “dead or alive” is left
282 undetermined unless the organism is isolated and cultured, or, transcriptome analysis is

283 used as a complementary detection tool or, complementary molecular viability analysis
284 is included [46].

285 Inspectors at international ports require a rapid detection method when decisions need
286 to be done on site. EDNA has been considered a good candidate to be used as a
287 diagnostic tool in ports of entry, due to its multiplexing capacity and rapidness. Yet,
288 EDNA does not include an analysis of pathogen viability. If DNA-based detection
289 (metagenomic analysis) is positive and viability needs to be addressed, the use of
290 additional tests is not a viable approach for perishable or time-sensitive shipments.
291 Using RNA sequencing and relative quantification of active genes is ideal to infer the
292 viability of plant pathogens. The use of EDNA transcriptomics to infer the production of
293 aflatoxin is a first attempt to introduce a novel strategy by using new sequencing
294 technologies to identify viable plant pathogens. The use of e-probes that are designed
295 on up-regulated genes incorporates an advantage to EDNA transcriptomics over other
296 tools that use RNA sequencing to assess gene expression [37]. The advantage of
297 EDNA transcriptomics over other methods of transcript frequency inference and
298 calculation is that the time-consuming map against the reference genome is not
299 necessary. Instead, we align the sample reads to the highly specific e-probes, which are
300 designed for known up-regulated genes. Directing the analysis to genes that are known
301 to be up-regulated reduces the analysis time tremendously since a mapping against a
302 whole genome is no longer necessary. Where needed total nucleic acids (DNA and
303 RNA) can be extracted from the sample of interest to perform both pathogen detection
304 and gene activity.

Although most of the potential controlled inputs must be maintained constant, the sample matrix could contain fungal biomass, spores or sclerotia, depending on the organism and its life cycle stage. Yet, the source of relative quantitation become irrelevant because gene-expression analysis tools (including DeSeq.) are developed to analyze bulk populations — containing millions of cells —. Consequently, cell number differences between the treatment (Corn+AF70) and control (PDB+AF70) are small-uncontrolled inputs when equivalent sequencing depth has been achieved. Different cell counts between the treatment and the control in gene expression studies is therefore not a factor. We intend to use EDNA transcriptomics in metatranscriptomic analyses, where cell counts of organisms is difficult (or impossible i.e. unculturable-unknown organisms). Therefore, EDNAtran relies on a good quality sequencing data and equivalent sequencing depth to be able to differentiate between high and low-frequency hits.

The use of replicates in gene expression analyses using NGS are crucial, yet, this study used one replicate for all four RNA sequencing samples. In a real case scenario — where soil samples potentially containing *A. flavus* are collected — obtaining high-quality RNA libraries is more important than replicates. This study produced twelve RNA extraction samples from which four RNA extractions having the highest RNA integrity and quality (RIN≥8) where selected to be sequenced. Replication may be needed for statistical hypothesis-driven research, but it is not required for presence/absence queries or for the development of e-probes. Sequencing depth and sequencing quality equivalence are the most important metric for diagnosis.

Future studies need to include multiple blind samples to assess the usefulness of the new EDNAtran protocol to indicate the presence of aflatoxin-producing *A. flavus*. In this study, we have shown that in a known positive transcriptomic database, EDNAtran is capable of discriminating between production and no-production of aflatoxin. However, blind samples will provide a realistic assessment of the tool.

ACKNOWLEDGMENTS

This work was supported by the USDA-CSREES Plant Biosecurity Program, grant number 2010-85605-20542. Portions of the computing for this project was performed at the OSU High-Performance Computing Center at Oklahoma State University supported in part through the National Science Foundation grant OCI-1126330.

LITERATURE CITED

1. Marsh SF, Payne G a. Preharvest Infection of Corn Silks and Kernels by Aspergillus flavus [Internet]. Vol. 74, Phytopathology. 1984. p. 1284–9.

2. Hill RA, Blankenship PD, Cole RJ, Sanders TH. Effects of soil moisture and temperature on preharvest invasion of peanuts by the Aspergillus flavus group and subsequent aflatoxin development. Appl Environ Microbiol. 1983;45(2):628–33.

3. Llewellyn GC, Mooney RL, Cheatle TF, Flannigan B. Mycotoxin contamination of spices : An update. Int Biodeterior Biodegradation. 1992;29(2):111–21.

4. Cotty PJ. Aflatoxin-producing potential of communities of Aspergillus section Flavi from cotton producing areas in the United States. Mycol Res. 1997;101(6):698–704.

5. Squire RA. Ranking animal carcinogens: a proposed regulatory approach. Sci . 1981 Nov 20;214(4523):877–80.

6. Klich M a., Pitt JI. Differentiation of Aspergillus flavus from A. parasiticus and other closely related species. Trans Br Mycol Soc. 1988;91(1):99–108.

7. Kurtzman CP, Horn BW, Hesseltine CW. Aspergillus nomius, a new aflatoxin-producing species related to Aspergillus flavus and Aspergillus tamarii. Antonie Van Leeuwenhoek. 1987;53(3):147–58.

8. Trail F, Mahanti N, Rarick M, Mehigh R, Liang SH, Zhou R, et al. Physical and transcriptional map of an aflatoxin gene cluster in Aspergillus parasiticus and functional disruption of a gene involved early in the aflatoxin pathway. Appl Environ Microbiol. 1995;61(7):2665–73.

9. Yu JJ, Chang PK, Cary JW, Wright M, Bhatnagar D, Cleveland TE, et al. Comparative Mapping of Aflatoxin Pathway Gene Clusters in Aspergillus-Parasiticus and Aspergillus-Flavus. Appl Environ Microbiol. 1995;61(6):2365–71.

10. Minto RE, Townsend CA. Enzymology and Molecular Biology of Aflatoxin Biosynthesis. Chem Rev. 1997 Nov 1;97(7):2537–56.

11. Fernandes M, Keller NP, Adams TH. Sequence-specific binding by Aspergillus nidulans AflR, a C6 zinc cluster protein regulating mycotoxin biosynthesis. Mol Microbiol. 1998;28(6):1355–65.

12. Woloshuk CP, Foutz KR, Brewer JF, Bhatnagar D, Cleveland TE, Payne GA. Molecular characterization of aflR, a regulatory locus for aflatoxin biosynthesis.

Appl Env Microbiol. 1994;60(7):2408–14.

13. Chang PK, Cary JW, Bhatnagar D, Cleveland TE, Bennett JW, Linz JE, et al. Cloning of the Aspergillus parasiticus apa-2 gene associated with the regulation of aflatoxin biosynthesis. Appl Environ Microbiol. 1993;59(10):3273–9.

14. Yu JH, Butchko R a E, Fernandes M, Keller NP, Leonard TJ, Adams TH. Conservation of structure and function of the aflatoxin regulatory gene aflR from Aspergillus nidulans and A. flavus. Curr Genet. 1996;29(6):549–55.

15. Chang PK. Lack of interaction between AFLR and AFLJ contributes to nonaflatoxigenicity of Aspergillus sojae. J Biotechnol. 2004;107(3):245–53.

16. Chang PK, Horn BW, Dorner JW. Sequence breakpoints in the aflatoxin biosynthesis gene cluster and flanking regions in nonaflatoxigenic Aspergillus flavus isolates. Fungal Genet Biol. 2005;42(11):914–23.

17. Callicott K a., Cotty PJ. Method for monitoring deletions in the aflatoxin biosynthesis gene cluster of Aspergillus flavus with multiplex PCR. Lett Appl Microbiol. 2015;60(1):60–5.

18. Van Egmond HP, Schothorst RC, Jonker M a. Regulations relating to mycotoxins in food : Perspectives in a global and European context. Anal Bioanal Chem. 2007;389(1):147–57.

19. Wu F, Liu Y, Bhatnagar D. Cost-Effectiveness of aflatoxin control methods: economic incentives. 2008;9543(November):203–25.

20. Robens J, Cardwell K. The Costs of Mycotoxin Management to the USA:

391 Management of Aflatoxins in the United States. Toxin Rev. 2003;22(2–3):139–52.

392 21. Atehnkeng J, Ojiambo PS, Donner M, Ikotun T, Sikora RA, Cotty PJ, et al.
393 Distribution and toxigenicity of Aspergillus species isolated from maize kernels
394 from three agro-ecological zones in Nigeria. Int J Food Microbiol. 2008;122(1):74–
395 84.

396 22. Garber RK, Cotty PJ. Formation of Sclerotia and Aflatoxins in Developing Cotton
397 Bolls Infected by the S Strain of Aspergillus flavus and Potential for Biocontrol
398 with an Atoxigenic Strain. Phytopathology. 1997;87(9):940–5.

399 23. Probst C, Bandyopadhyay R, Price LE, Cotty PJ. Identification of Atoxigenic
400 Aspergillus flavus Isolates to Reduce Aflatoxin Contamination of maize in Kenya.
401 Plant Dis. 2011;95(2):212–8.

402 24. Bandyopadhyay R, Ortega-Beltran A, Akande A, Mutegi C, Atehnkeng J, Kaptoge
403 L, et al. Biological control of aflatoxins in Africa: current status and potential
404 challenges in the face of climate change. World Mycotoxin J. 2016;9(5):771–89.

405 25. Mehl HL, Jaime R, Callicott K a., Probst C, Garber NP, Ortega-Beltran A, et al.
406 Aspergillus flavus diversity on crops and in the environment can be exploited to
407 reduce aflatoxin exposure and improve health. Ann N Y Acad Sci.
408 2012;1273(1):7–17.

409 26. Cotty PJ. Influence of field application of an atoxigenic strains of Aspergillus
410 flavus on the populations of A. flavus infection cotton balls and on the aflatoxin
411 content of cottonseed. Vol. 84, Phytopathology. 1994. p. 1270–7.

412 27. Cotty PJ, Mellon JE. Ecology of aflatoxin producing fungi and biocontrol of
413 aflatoxin contamination. Mycotoxin Res. 2006;22(2):110–7.

414 28. Mauro A, Battilani P, Cotty PJ. Atoxigenic Aspergillus flavus endemic to Italy for
415 biocontrol of aflatoxins in maize. BioControl. 2015;60(1):125–34.

416 29. Huson DH, Mitra S, Ruscheweyh H-J, Weber N, Schuster SC. Integrative analysis
417 of environmental sequences using MEGAN4. Genome Res. 2011;21(9):1552–60.

418 30. Truong DT, Franzosa E a, Tickle TL, Scholz M, Weingart G, Pasolli E, et al.
419 MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods.
420 2015;12(10):902–3.

421 31. Stobbe AH, Daniels J, Espindola AS, Verma R, Melcher U, Ochoa-Corona F, et
422 al. E-probe Diagnostic Nucleic acid Analysis (EDNA): A theoretical approach for
423 handling of next generation sequencing data for diagnostics. Microbiol Methods.
424 2013;94:356–66.

425 32. Espindola AS, Garzon CD, Schneider WL, Hoyt PR, Marek SM, Garzon CD. A
426 new approach for detecting Fungal and Oomycete plant pathogens in Next
427 Generation Sequencing metagenome data utilizing Electronic Probes. Int J Data
428 Min Bioinfomatics. 2015 May 18;12(2):115–28.

429 33. Visser M, Burger JT, Maree HJ. Targeted virus detection in next-generation
430 sequencing data using an automated e-probe based approach. Virology.
431 2016;495:122–8.

432 34. Blagden T, Schneider W, Melcher U, Daniels J, Fletcher J. Adaptation and

433 Validation of E-Probe Diagnostic Nucleic Acid Analysis for Detection of
434 Escherichia coli O157:H7 in Metagenomic Data from Complex Food Matrices. J
435 Food Prot. 2016;79(4):574–81.

436 35. Woloshuk CP, Cavaletto JR, Cleveland TE. Inducers of Aflatoxin Biosynthesis
437 from Colonized Maize Kernels Are Generated by an Amylase Activity from
438 Aspergillus flavus. Phytopathology. 1997;87(2):164–9.

439 36. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR:
440 ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.

441 37. Anders S, Huber W. Differential expression analysis for sequence count data.
442 Genome Biol. 2010;11(10):1–12.

443 38. Nierman WC, Yu J, Fedorova-Abrams ND, Losada L, Cleveland TE, Bhatnagar D,
444 et al. Genome Sequence of Aspergillus flavus NRRL 3357, a Strain That Causes
445 Aflatoxin Contamination of Food and Feed. Genome Announc. 2015;3(2):e00168-
446 15.

447 39. Ehrlich KC, Yu J, Cotty PJ. Aflatoxin biosynthesis gene clusters and flanking
448 regions. J Appl Microbiol. 2005;99(3):518–27.

449 40. Davis ND, Diener UL, Eldridge DW. Production of aflatoxins B1 and G1 by
450 Aspergillus flavus in a semisynthetic medium. Appl Microbiol. 1966;14(3):378–80.

451 41. Mateles RI, Adye JC. Production of Aflatoxins in Submerged Culture. Appl
452 Microbiol. 1965;13(2):208–11.

453 42. Mellon JE, Dowd MK, Cotty PJ. Substrate utilization by Aspergillus flavus in

454 inoculated whole corn kernels and isolated tissues. J Agric Food Chem.
455 2005;53:2351–7.

456 43. Probst C, Cotty PJ. Relationships between in vivo and in vitro aflatoxin
457 production: Reliable prediction of fungal ability to contaminate maize with
458 aflatoxins. Fungal Biol. 2012;116(4):503–10.

459 44. Keller N, Butchko AE, Sarr B, Phillips T. A visual pattern of mycotoxin production
460 in maize kernels by Aspergillus spp. Phytopathology. 1994;84:483–8.

461 45. Stobbe AH, Schneider WL, Hoyt PR, Melcher U. Screening Metagenomic Data for
462 Viruses Using the E-Probe Diagnostic Nucleic Acid Assay. Phytopathology. 2014
463 Sep 10;104(10):1125–9.

464 46. Cangelosi GA, Meschke JS. Dead or Alive: Molecular Assessment of Microbial
465 Viability. Appl Environ Microbiol. 2014;80(19):5884–91.

466

467 **SUPPORTING INFORMATION**

468 **S1 Table. Expression values (log2 fold Change) for AF70 *A. flavus* strain for two**
469 **culture conditions (ground corn and PDB).** For each Gene ID, expression levels are
470 listed along with p-values.