

The University of Arizona Electronic Theses and Dissertations Reproduction and Distribution Rights Form

The UA Campus Repository supports the dissemination and preservation of scholarship produced by University of Arizona faculty, researchers, and students. The University Library, in collaboration with the Honors College, has established a collection in the UA Campus Repository to share, archive, and preserve undergraduate Honors theses.

Theses that are submitted to the UA Campus Repository are available for public view. Submission of your thesis to the Repository provides an opportunity for you to showcase your work to graduate schools and future employers. It also allows for your work to be accessed by others in your discipline, enabling you to contribute to the knowledge base in your field. Your signature on this consent form will determine whether your thesis is included in the repository.

Name (Last, First, Middle)
Almader, Angelica, Estella
Degree title (eg BA, BS, BSE, BSB, BFA):
BS
Honors area (eg Molecular and Cellular Biology, English, Studio Art):
Microbiology
Date thesis submitted to Honors College:
04/30/2013
Title of Honors thesis:
Mechanisms of Atoxigenicity: Characterizing Genetic Properties of <u>Atoxigenic Aspergillus Flavus Isolates</u>
The University of Arizona Library Release Agreement
I hereby grant to the University of Arizona Library the nonexclusive worldwide right to reproduce and distribute my dissertation or thesis and abstract (herein, the "licensed materials"), in whole or in part, in any and all media of distribution and in any format in existence now or developed in the future. I represent and warrant to the University of Arizona that the licensed materials are my original work, that I am the sole owner of all rights in and to the licensed materials, and that none of the licensed materials infringe or violate the rights of others. I further represent that I have obtained all necessary rights to permit the University of Arizona Library to reproduce and distribute any nonpublic third party software necessary to access, display, run or print my dissertation or thesis. I acknowledge that University of Arizona Library may elect not to distribute my dissertation or thesis in digital format if, in its reasonable judgment, it believes all such rights have not been secured.
<input checked="" type="checkbox"/> Yes, make my thesis available in the UA Campus Repository!
Student signature: <u>A. Almader</u> Date: <u>5/3/13</u> Thesis advisor signature: <u>[Signature]</u> Date: <u>3 MAY 2013</u>
<input type="checkbox"/> No, do not release my thesis to the UA Campus Repository.
Student signature: _____ Date: _____

MECHANISMS OF ATOXIGENICITY:
CHARACTERIZING GENETIC PROPERTIES OF ATOXIGENIC *ASPERGILLUS FLAVUS* ISOLATES

By

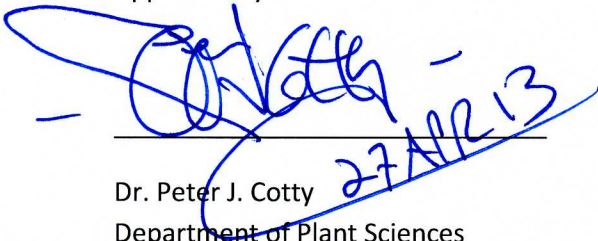
ANGELICA ESTELLA ALMADER

A Thesis Submitted to The Honors College
In Partial Fulfillment of the Bachelors degree
With Honors in
Microbiology

THE UNIVERSITY OF ARIZONA

MAY 2013

Approved by:



Dr. Peter J. Cotty
Department of Plant Sciences

ABSTRACT

Aspergillus flavus is the primary causal agent of aflatoxin contamination in crops. Aflatoxins are highly toxic carcinogens that threaten many agricultural systems throughout the world. Twenty-eight genes located in a single aflatoxin biosynthesis gene cluster are required for aflatoxin production. *A. flavus* isolates that are not capable of producing aflatoxins (atoxigenics) have been utilized commercially as biocontrol agents. Atoxigenicity results from the modification of genes within the aflatoxin biosynthesis cluster. Mechanisms by which *A. flavus* isolates become atoxigenic are unknown for many atoxigenics. Twelve atoxigenic isolates under development as biocontrol agents in the United States and Africa were subjected to several genetic analyses including multiplex PCR and sequencing of aflatoxin biosynthesis genes. Deletion patterns within the aflatoxin gene cluster varied among the atoxigenics, but certain deletion patterns were shared by multiple isolates. Sequence analysis revealed that either breakpoints or insertions exist in several isolates. The data obtained from this study will be used to develop assays to interrogate large populations of *A. flavus* for specific DNA features. This will aid the search for elite atoxigenic vegetative compatibility groups with superior value as biocontrol agents and allow for monitoring specific genotypes in the environment for detection of adaptive differences.

INTRODUCTION

There are several species of fungi belonging to the genus *Aspergillus* that produce aflatoxins. Aflatoxins are some of the most potent, naturally-occurring liver carcinogens currently identified (Probst et al., 2011). *Aspergillus flavus* is the most common causal fungus, producing aflatoxins B₁ and B₂ (Chang et al., 2005). Agricultural systems are subjected to contamination by these toxic chemicals, which then become a major threat to the health of individuals who consume the contaminated crops. Health effects range from immune suppression and growth retardation to cancer and, in cases of severe poisonings, liver failure followed by death (Williams et al., 2004). In most developed countries aflatoxin contamination is not a critical concern due to strict regulations implemented by governmental agencies that limit the concentrations of aflatoxins allowed in food for human consumption as well as feed for livestock (van Egmond et al., 2007). Maximum tolerable levels for aflatoxin range from < 1 µg/kg in some dairy products to 300 µg/kg for certain feeds (Van Egmond, 2002). However, not all individuals are fortunate enough to live in areas where such guidelines are followed. Consumers in underdeveloped nations, lacking an appropriate system of regulation, are most at risk for unsafe levels of aflatoxin exposure. There are many regions where the products from small-scale farms move from field to mouth without the opportunity for practical monitoring of aflatoxin content (Cotty, 2006). It is crucial that the spread of toxigenic *A. flavus* be significantly decreased in order to prevent the populations of developing countries from suffering the hazardous health effects caused by aflatoxin contamination.

In order to minimize human exposure to aflatoxins, biological control strategies utilizing atoxigenic *A. flavus* strains have been developed (Chang et al., 2005). Specifically, two *A. flavus* isolates, AF36 and NRRL 21882, have been registered as biopesticides with the United States Environmental Protection Agency for the management of aflatoxin-producing fungi (Chang et al., 2005). Atoxigenic strains that are used to control aflatoxin levels in commercial crops utilize competitive exclusion as the mechanism to reduce contamination. All genes to produce aflatoxins are located in one cluster of 28 genes. Atoxigenic isolates have lost their ability to produce aflatoxins as a result of genetic changes in this cluster. For example, the biocontrol strain AF36 contains a single nucleotide polymorphism (SNP) in a polyketide synthase (PKS) gene, resulting in a premature stop codon and atoxigenicity (Ehrlich and Cotty 2004; Ehrlich et al. 2007). Not all atoxigenic isolates contain the same genetic alterations. Atoxigenicity can result from multiple events such as deletions, frame-shift mutations, base pair substitutions, DNA insertions, and SNPs. The molecular mechanisms responsible for the loss of aflatoxin production are diverse and for most atoxigenic *A. flavus* strains the specific genetic changes resulting in atoxigenicity are unknown (Donner et al., 2009).

Several atoxigenic strains of *A. flavus* are currently under development as agents for the prevention of aflatoxin contamination. Regulatory agencies frequently ask for mechanism of atoxigenicity during biopesticide registration (Cotty et al., 2007). Precise characterization of molecular lesions can facilitate development of rapid assays that will monitor genetic groups to which the atoxigenic strains belong (i.e. through vegetative

compatibility analyses) (Cotty, 1994) and atoxigenicity itself (Das et al., 2008). The current study sought to determine mechanisms of atoxigenicity for twelve atoxigenic strains currently under development as biocontrol agents and for which mechanisms have not previously been established. In the process, several mechanisms of atoxigenicity were

characterized and assays for monitoring the molecular lesions were developed. Ultimately, the current results will be used to develop methods to rapidly identify specific genetic alterations within *A. flavus* populations and facilitate the process of developing methods to displace aflatoxin producers from agroecosystems

MATERIALS AND METHODS

Table 1: *Aspergillus* section *Flavi* isolates utilized in the current study

ID	Isolate	Origin	Year	VCG	Source
B19	G018-2	Burkina Faso	2010	BF10 G018-2	Groundnut grain
K10	C6-E	Kenya	2006	KN00A	Maize
K11	C8-F	Kenya	2006	KN012	Maize
K39	E63-I	Kenya	2006	KN001	Maize
K49	R7-H	Kenya	2006	KN011	Maize
S01	M21-11	Senegal	2008	SN03	Maize
S16	Ms14-19	Senegal	2008	SN004	Maize soil
S25	Ss19-14	Senegal	2008	SN005	Sesame soil
T01	AT5-B	Texas	2008	TXAT043	Maize grain
T13	DO11-K	Texas	2008	TXAT016	Maize grain
T16	DO114-A	Texas	2008	TXAT039	Maize grain
T19	DO38-B	Texas	2008	TXAT005	Maize grain

ID = refers to the DNA code given to the isolate

Year = Year of isolation

VCG = Vegetative Compatibility Group

Twelve *A. flavus* isolates were chosen based on their atoxigenicity and availability from personal (Peter J. Cotty's Laboratory, School of Plant Sciences, University of Arizona, Tucson, AZ) and public (USDA-ARS) culture collections. Each of these isolates contains deletions in either the aflatoxin cluster or the CPA cluster. These deletions have not yet been characterized. The origins of the 12 isolates are shown in Table 1.

DNA Extraction. Isolates were cultured on a minimal nutrient media, 5-2 with 2% NaCl, and

incubated at 31° C for at least 3 days. This media ensures high sporulation, allowing for an appropriate amount of spores to be harvested for DNA extraction.

In a laminar flow hood, 1.5 mL 0.1% TWEEN-80 was pipeted onto each cultured plate. A cell spreader was utilized to suspend the spores in the buffer. 1.2 mL of the spore suspension in TWEEN-80 was transferred into a 1.5 mL microcentrifuge tube. The microcentrifuge tubes were spun in a centrifuge for 5 minutes at

8000 xg. The supernatant was carefully removed using a pipet tip.

450 µL of Lysis Buffer (30 mM Tris, 10mM EDTA, 1% SDS, pH 8.0) was added to each microcentrifuge tube followed by a brief vortex to suspend the spores. The tubes were then placed in a ThermoMixer at 60° C and 8,000 rpm for 1 hour.

The tubes were spun in a centrifuge at 14,000 xg for 30 minutes and 370 µL of the supernatant was transferred to a new 1.5 µL centrifuge tube. The tubes were placed on ice and 370 µL of 4M ammonium acetate (NH₄OAc) was added. The tubes were inverted several times to ensure proper mixing.

740 µL of ice-cold 100% ethanol was added to the tubes and mixed by inverting. The tubes were placed in the freezer at -20° C for 30 minutes. The tubes were then spun at maximum speed for 5 minutes and the supernatant was poured off.

The tubes were inverted and left upside down on a paper towel to air dry for 1 hour. The pellet was resuspended in 25 µL of sterile H₂O. The tubes were vortexed briefly to re-suspend the DNA.

The concentration of the isolated DNA was measured utilizing a spectrophotometer (model ND-1000 Nano Drop).

PCR Conditions

Primer sets targeting PCR products within the aflatoxin biosynthesis cluster (Table 3) were used to profile the 12 isolates and determine the presence or absence of aflatoxin biosynthesis genes. The same primers were also used to attempt amplifications across missing gene segments, as summarized in

Figure 1. Two primers that target the telomere repeat of *A. flavus* were used (Table 2) in order to attempt amplifications from the last gene present to the telomeric end of the chromosome.

Each PCR reaction contained 2 µL genomic DNA, 0.5 µL of 10 µM forward primer, 0.5 µL of 10 µM reverse primer, 17 µL deionized H₂O and AccuPower® HotStart PCR PreMix (Bioneer, Alameda, CA). Each PCR reaction had a 20 µL volume and was performed with a MyCycler thermocycler (Bio-Rad Laboratories). The PCR cycle conditions were as follows: 5 minutes at 94°C followed by 38 cycles of 94°C for 20 seconds, [the annealing temperature specific to the primers being used] for 30 seconds, 72°C for [the appropriate extension time depending on product size, ranging from 30 seconds to 6 minutes] and a final elongation period of 10 min at 72° C.

Gel electrophoresis was run using 1.0% agarose gel. Afterward, the amplicons were visualized with GelRed Nucleic Acid Stain. Products to be sent for sequencing were prepped for clean-up utilizing 1 µL of ExoSAP and 16 µL of the amplicon and run on the Exo-SAP 60 program in the thermocycler. The Exo-SAP program consists of 60 minute incubation at 37° C, followed by an inactivation step of 85° C for 15 minutes.

Amplicons were sent to The Genomic Analysis and Technology Core Facility (GATC) at the University of Arizona, Tucson. The product sizes of these amplicons ranged from 200 – 3000 basepairs.

Sequence Analysis

Product sequences were initially aligned to the aflatoxin biosynthesis gene cluster sequences of

3 genomes from Genbank. The reference sequences were *A. flavus* isolate 3357, *A. flavus* isolate AF13, and *A. flavus* isolate AF36. The

Primer Design

Primers were designed using the Primer3 program, version 0.4.0 (Rozen and Skaletsky, 2000). The sequences from PCR amplicons were used as the templates for creating the primers. The following conditions were used

alignments were analyzed specifically by searching for breakpoints and/or InDels (insertions/deletions).

when designing all primers: [divalent cations] = 1.5 mM, [monovalent cations] = 50.0 mM, [dNTPs] = 0.2 mM, Primer Tm: Min = 62° C, Opt = 64° C, Max = 68°C.

Table 2: Primers used for genetic analyses

Gene	Forward Primer	Reverse Primer	T ^a (°C)	Product (bp)	Use
Telomere repeat*		teloR: ATGTTGACCCCTAATGTTGACCCCTA	54	--	PCR, sequencing
Telomere repeat*		AgtloR: TGTTGACCCCTAATGTTGACCCCT	54	--	PCR, sequencing
Intergenic region*	3SA3-2F: GCCAATGCTTGTCCCATG	3SA3-2R: GCAATCAACATAACCCACG	50	924	PCR
<i>afIT</i>	afIT-F: ATGACATGCTAATCGACGAG	afIT-R: AGGCGCATGCTACGGATC	50	1,141	PCR
<i>avfA</i>	avfA-F: ATCAAAATCCCGTTCGGTCCG	avfA-R: TAGCCCGTGGTGTGTGTTCC	53	491	PCR
<i>avnA</i>	avnA-F: GCG ATAGAAGCTGACAAAGGCA	avnA-R: GAATGAGTCTCCAAAGGGCAG	53	541	PCR, sequencing
<i>dmaT</i>	cpaAF: GTC TCTGGATCGTCCGTCG	cpaAR: GTATAGCACAGCTCCGATGT	55	765	PCR, sequencing
<i>estA</i>	estA-F: CGATGGGACTGACGGTGATT	estA-R: ACCACGCCGCTGACTTTAT	53	530	PCR, sequencing
<i>norB</i>	norB-F: GTGAGGGATAGCAATAAGTGTG	norB-R: TCCTGGATTTCCGCATAC	50	452	PCR, sequencing
<i>omtB</i>	omtB1F: ACACATCCAGAGGCAGAAATG	omtB1R: GACCACATCCCAATCCGAGTA	53	681	PCR, sequencing
<i>omtB</i>	omtB2F: TTTACTCGGATGGGATGTTGGT	omtB2R: CGCAGTCTTGTAGAGGTGAT	55	554	PCR, sequencing
<i>omtB</i>	omtB3F: GAGTCCGTTACAGATGCTTTTGC	omtB3R: TCTGAATAGGCTGTGGCTCG	53	554	PCR
<i>omtB</i>	omtB4F: AACCTTCTGTCGCCCTTAT	omtB4R: GCGTTAAACCCATGCAAAATG	50	500	PCR
<i>ver1</i>	ver1-1F: GGAGCACCTCAAGGAGAACA	ver1-1R: GGAACCTTGCCTCGATTCTA	53	357	PCR
<i>ver1</i>	ver1-2F: GCCTGAGAATCTTTGCTTGG	ver1-2R: GCGGTAGTTCACCCACGACTT	53	753	PCR
<i>ver1</i>	ver1-3F: AGCCAAAAGTCGTGGTGAAT	ver1-3R: CCATCCACCCCAATGATCT	53	787	PCR, sequencing
<i>ver1</i>	ver1-4F: AAAGATCATTGGGGTGGATG	ver1-4R: CACGCCTTTACCCATATCCAG	53	738	PCR, sequencing
<i>verA</i>	verA-F: CCGCAACACCACAAGTAGCA	verA-R: AAACGCTCTCCAGGCACCTT	53	424	PCR, sequencing
<i>verB</i>	verB1-F: CCCAATACAGTCCGCAGTC	verB-1R: AGTGAAGAGTGCCGACGATAA	53	568	PCR, sequencing
<i>verB</i>	verB2-F: TTATCGTCGGCACTCTTCACT	verB-2R: AACCTGTTCTTCCATCCGCT	53	755	PCR, sequencing
Subtelomeric region*	sub1-F** : TCCGCGGTATGAAGGGCCCTCTA	sub1-R** : CGCAATGCCTTCTCCCGGGTT	58	800	PCR
Subtelomeric region*	sub2-F** : CGGCTTCAGACACCAGCCGCT	sub2-R** : TGAACAGCAGCCTCCCAACCCT	58	802	PCR
Subtelomeric region*	sub3-F** : CGTGACCACCCTAGTGGGGACC	sub3-R** : ACTCCAACCCGAGTGCAGACT	58	783	PCR
Subtelomeric region*	sub4-F** : ACGGTAGGTGCAGACGGCA	sub4-R** : ACTCTCAATCGTAAGGCCAACTTGCAAT	58	782	PCR
Subtelomeric region*	sub5-F** : GGCCAAACCCACCCTCTGCCA	sub5-R** : TTTCAGCGCTAGAACCTTCACATCGGA	58	770	PCR

T^a = annealing temperature;

(*) indicates that the primers do not target a specific gene;

(**) signifies primers that were designed specifically for this study;

(--) indicates that the product size resulting from these primers is dependent on the forward primer used in conjunction with the reverse primer listed

Table 3: Additional primers used for Multiplex PCR

Gene	Forward Primer	Reverse Primer	T _a (°C)	Product (bp)	Use
<i>glcA</i>	43T: ATACCTCATGATCTGGTGCACGG	43C: CTTGCAGCGACAATGATACGTC	54	883	PCR
<i>nadA</i>	48T: GTCCCCAGGTACGATAGGTCTCT	48C: GCTGGATATTCCAAGGAGTGGCT	54	742	PCR
<i>moxY</i>	53TGACTGCCACCCATCACTCTTCC	53C: TGGCTCGACTGGGTATGAAATCC	54	613	PCR
<i>vbs</i>	58T: GCATTGCCAGCATCGGTTTCATA	58C: AGCAGACCCGTACTAAGTGATGC	54	487	PCR
<i>omtA</i>	63T: CATGATGGAGCATGACATTCGGC	63C: GCGCCACCATAATCTTCAGTCT	54	387	PCR, sequencing
<i>verB</i>	68T: TTTAACCCTTCAYGCCCTCGAAT	68C: TGCRTARCTAATCTCATCGGGT	54	297	PCR
<i>verA</i>	73T: TGCTGAGCGAGTAGGTAGGT	73C: CCGGATCATCCCTCCAATCTGT	54	194	PCR
<i>norA</i>	78T: CCTGTGAGGGACACAAGACACT	78C: AAGAATAGCGGTGACATCCAGCA	54	1427	PCR
<i>afIJ</i>	83T: GAGGACAGGTTGTGTGCTGTTG	83C: GTTACGAGCTATCCTCAGCCAT	54	1092	PCR
<i>hexB</i>	88T: GAAC TGAGCCATTTCCATCAGCG	88C: GTC TTGTACAGGGAACGTGGTGA	54	897	PCR
<i>hexA</i>	93T: AACGCTTCAACGTGGAGGACATA	93C: AATAGCGTTGGCGTTGAAGTCAC	54	736	PCR
<i>hexA</i>	98T: CCCGCATTTTCTCGATCCCCTTG	98C: GCGACGACCAGTCAATTATGAAGC	54	633	PCR
<i>pkxA</i>	103T: GTCAGACCACAGTGAGTGCTTCT	103C: AAGCTGACTGGGAGAATGTTGCT	54	536	PCR
<i>pkxA</i>	108T: CCCCTCAACTTCTGTCTCCTAC	108C: GCTGGGTAGCGAACAATCCAATG	54	425	PCR
<i>cypA</i>	113T: GCACACAGCAGAGGCCATTTCTAC	113C: AATCTATCTAGCCATCGCCACCG	54	330	PCR, sequencing
Intergenic region*	118T: GCC TGCTAGGCTTGAACATATGT	118C: CGCAATGCTAGTATGCCCTTGTG	54	209	PCR
Intergenic region*	123T: GACACTCGTACCATCTATGCACC	123C: GATCCCTGATCCATTCCACCTTG	54	1219	PCR
<i>PKS</i>	128T: ACGATACGAGCTTTAGTGCAAGG	128C: GATATAGACCTCAGGGTGCAGCA	54	925	PCR
<i>PKS</i>	133T: TGAATGACACGGTCAATGATGCTT	133C: GACATTTCCCTGTCTATCGATGC	54	821	PCR
<i>PKS</i>	138T: ACCTCAACAATTACACCGGATGG	138C: GTTGTAGCTCAACGTCAC TAGCA	54	648	PCR
Subtelomeric region*	143T: TATCTATCTGGGATACGGGCTGG	143C: TATGCCGTTGCTATCCAATGAGG	54	521	PCR
Subtelomeric region*	148T: AAGTCAGATTCGCGGATGAAG	148C: TCATCGCATTAATCGAGGCAGTT	54	416	PCR, sequencing
Subtelomeric region*	153T: CCTCCTGCACAAAAATACTCCCA	153C: GATCAGATCTTTGAGCGTAGCGT	54	320	PCR, sequencing
Subtelomeric region*	158T: TCATGTTTCGGATCGGAGATTGG	158C: ACATTTCCAAGTGAGAGATGTGGC	54	234	PCR
<i>cifR2</i>	163T: ACTGGTGTGGATAGAGCTCAGA	163C: TGGAAGGTTCTCCGGATACTTGA	54	908	PCR
<i>ord1</i>	168T: TACTCCGTTGCTGTCAATGGATG	168C: CGAATTCCTGGTTGAGCAGCTTG	54	782	PCR
<i>ord3</i>	173T: TGCTGAATAACAACCTCGACCCAG	173C: CAGGCTGGTATAGCACCAATGTT	54	684	PCR
<i>apmA</i>	178T: GGTTTCGTCTTGCCCTTCTCTCA	178C: AGCAAAGTGATGCCGTTCAATG	54	584	PCR, sequencing
<i>mdr4</i>	183T: CGTACTTTTGTACGGCGTACATC	183C: GCTGTTTTCGCGTTAGTTGGTAAC	54	512	PCR, sequencing
Subtelomeric region	188T: GCCCGTAAATGAGGTGCAGATAA	188C: TTTGGGTGTGCTTCTTCATGCTA	54	404	PCR
<i>sidC</i>	193T: GGGGACTTAGTCGCGAATGGTTA	193C: TATGAAGGCCACCAACTGAGGAC	54	258	PCR
<i>sidC</i>	198T: AATGACGACACTTGAGGCCACAG	198C: TCGGCTCCGTGACACCATATTA	54	185	PCR

T_a = Annealing temperature;

(*) indicates that the primers do not target a specific gene;

RESULTS

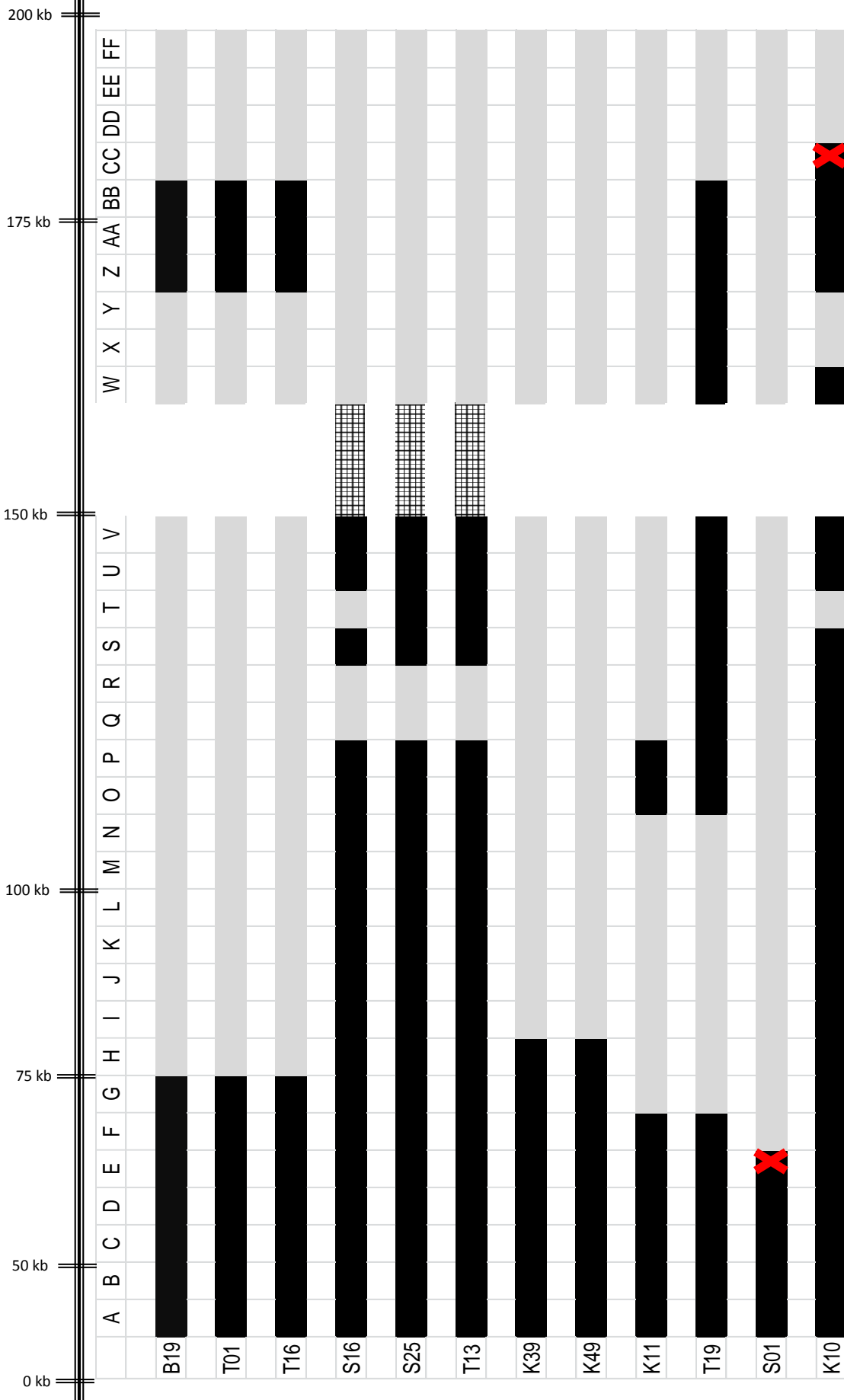


Figure 1 Schematic representation of the aflatoxin biosynthesis gene presence/absence in each of the isolates. The letters across the top refer to the genes that were targeted using the primers listed in Table 2. (See key below for additional gene information). The isolate ID names are listed down the left side. The presence of a gene is indicated by a black solid fill, while the absence of a gene is indicated by a grey solid fill. The checker-filled areas are DNA insertions (see Figure 3). The red "x" signifies a breakpoint (see Figures 2 & 3).

Figure 1 Key: Detailed positions of amplifications attempted

- A: 743 bp of the telomeric side of *glcA*, 140 bp of intergenic space
- B: 196 bp of the telomeric side of *nadA*, 40 bp of intergenic space, 82 bp of the centromeric side of *hypA*
- C: 573 bp of the telomeric side of *moxY*, 40 bp of intergenic space
- D: 350 bp of the telomeric side of *vbs*, 137 bp of intergenic space
- E: 387 bp of the telomeric side of *omtA*
- F: 54 bp of intergenic space, 243 bp of the centromeric side of *verB*
- G: 194 bp of the middle region of *verA*
- H: 574 bp of the centromeric side of *norA*, 59 bp of intergenic space, 569 bp telomeric side of *norA*, 225 bp of intergenic space
- I: 1071 bp of telomeric side of *afII*, 21 bp of intergenic space
- J: 897 bp of the centromeric side of *hexB*
- K: 153 bp of intergenic space, 583 bp of the centromeric side of *hexA*
- L: 619 bp of the telomeric side of *hexA*, 14 bp of intergenic space
- M: 536 bp of the centromeric side of *pksA*
- N: 425 bp of the telomeric side of *pksA*
- O: 175 bp of intergenic space, 155 bp of the centromeric side of *cypA*
- P: 209 bp of intergenic space
- Q: 1219 bp of intergenic space
- R: 286 bp of intergenic space, 639 bp of the centromeric side of *PKS*
- S: 770 bp of the middle region of *PKS*
- T: 648 bp of the telomeric side of *PKS*
- U: 541 bp of the subtelomeric region
- V: 416 bp of the subtelomeric region
- W: 320 bp of the subtelomeric region
- X: 234 bp of the subtelomeric region
- Y: 908 bp of the centromeric side of *ctfR2*
- Z: 724 bp of the subtelomeric region, 28 bp of the centromeric side of *ord1*
- AA: 650 bp of the telomeric side of *ord3*, 34 bp of the subtelomeric region
- BB: 528 bp of the subtelomeric region, 56 bp of the centromeric side of *apmA*
- CC: 447 bp of the centromeric side of *mdr4*, 22 bp of intergenic space, 1 bp of the telomeric side of *mdr4*
- DD: 404 bp of the subtelomeric region
- EE: 285 bp of the centromeric side of *sidC*
- FF: 76 bp of the telomeric side of *sidC*, 109 bp of the subtelomeric region

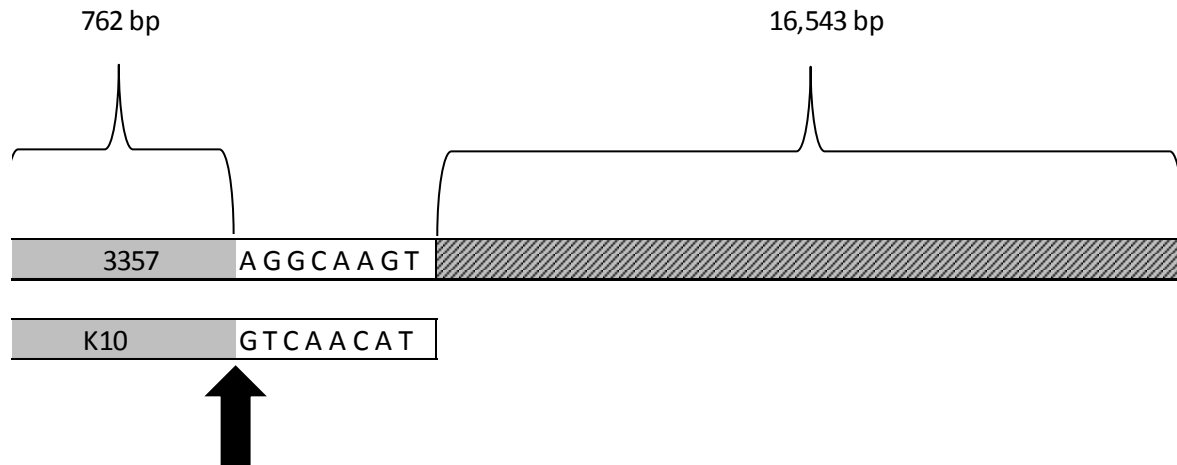


Figure 2 Schematic of the breakpoint found in isolate K10. *A. flavus* isolate 3357 (top) was aligned to the K10 183T/teloR amplicon sequence (bottom) and a breakpoint, represented by the arrow, was detected. The initial 762 bp of the amplicon sequence was identical to 3357, with the exception of 2 SNPs. The last 8 bases following the breakpoint arrow on isolate K10 form the reverse complement of the reverse primer utilized, teloR. The breakpoint is located 1,176 bases after the transcription start site of the *mdr4* gene. The solid grey fill indicates consensus among the sequences, while the dashed grey fill in 3357 indicates the remainder of the aflatoxin biosynthesis genes that are not present in isolate K10.

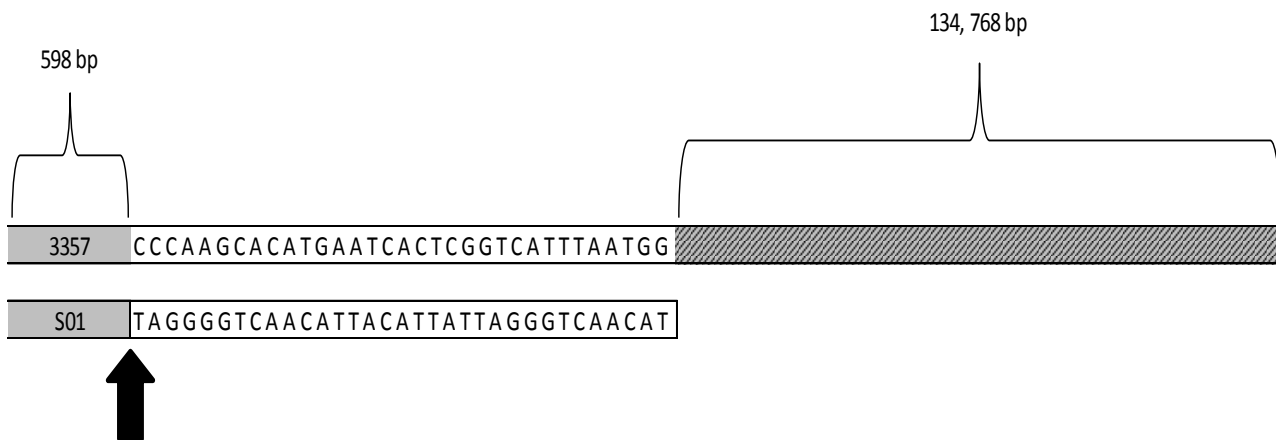


Figure 3 Schematic of the breakpoint found in isolate S01. *A. flavus* isolate 3357 (top) and the S01 omtB2F/teloR amplicon sequence (bottom) were aligned and revealed the breakpoint presented by the arrow. The amplicon size was 631 bp. The initial 598 bp align with 100% identity to 3357, while the last 33 bases begin to form the telomeric repeat. The breakpoint is located 725 bases after the transcription start site of the *omtB* gene. The solid grey fill represents the consensus between isolate 3357 and isolate S01, while the dashed grey fill represents the remainder of the aflatoxin biosynthesis genes that are not present in isolate S01.

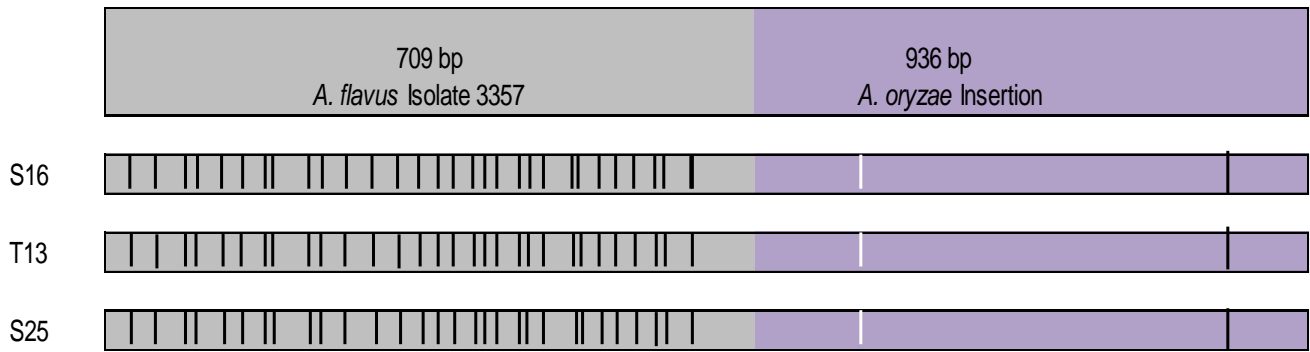


Figure 3 Schematic of the insertion found in isolates S16, T13, and S25. This diagram represents the alignment between *A. flavus* isolate 3357 with the S16, T13, and S25 148T/teloR amplicon sequences. The insertion was found through BLAST and is part of an *Aspergillus oryzae* isolate sequence (Accession number: AB506492.1), represented with the purple fill. The solid grey fill indicates consensus among the sequences, while the solid black lines represent SNPs that were found. There are 30 SNPs throughout the first 709 basepairs that aligned with 3357 and only one SNP throughout the insertion sequence. The solid white lines represent a single basepair insertion.

Image 1: Gel Electrophoresis Products using Telomere Primers

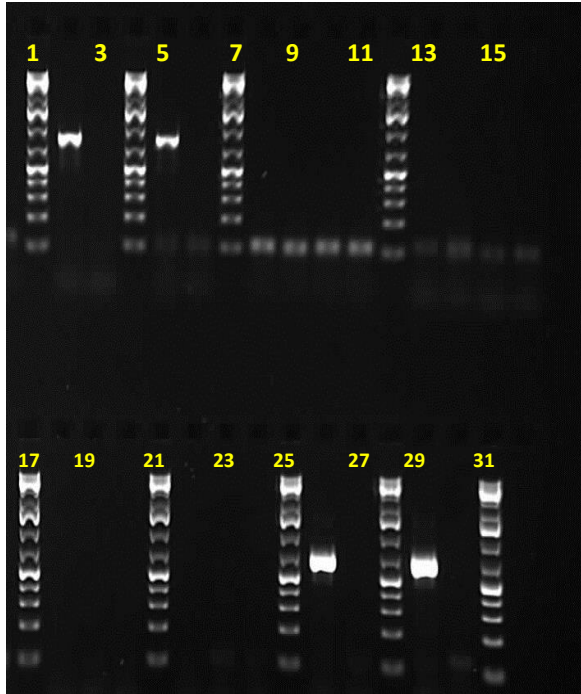


Image 1 shows the products obtained from amplifying from the last gene segment present in the aflatoxin gene cluster to the telomere using the primers AgteloR and teloR (Table 2). Lanes 2 and 5 are amplicons of the isolate K10; lane 2 contains the 183T/teloR amplicon, while lane 5 contains the

183T/AgteloR amplicon. Lanes 26 and 29 are amplicons of the isolate S01; lane contains the omtB2-F/teloR amplicon, while lane 29 contains the omtB2-F/AgteloR amplicon. Lanes 3, 6, 11, 16, 20, 24, 27 and 30 are all negative controls that contain deionized H₂O as the template. Lanes 1, 4, 7, 12, 17, 21, 25, 28 and 31 are all 1kb ladders.

Image 2: Multiplex Gel Electrophoresis Results

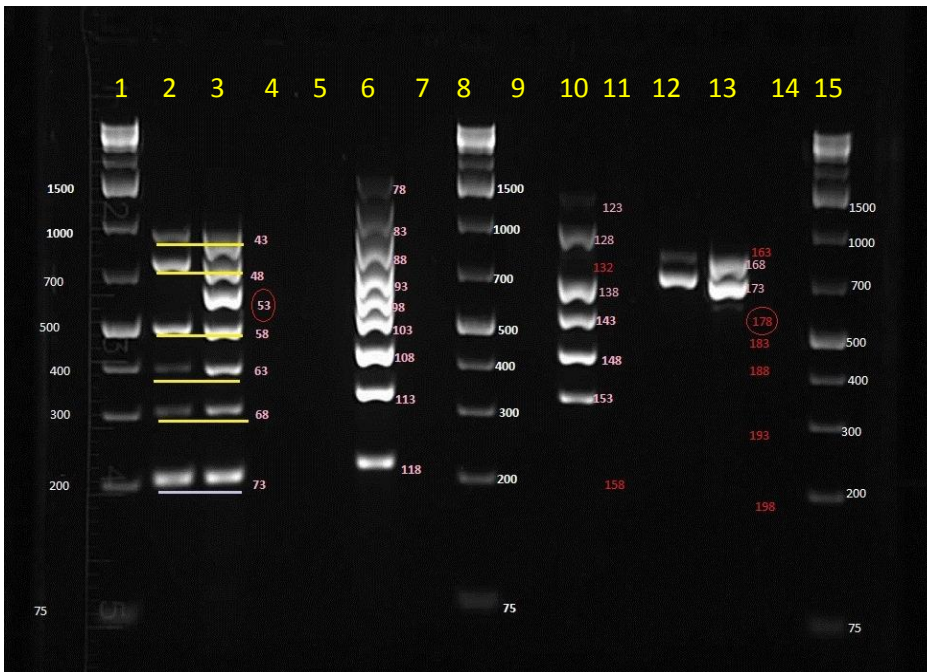


Image 2 shows the products from a Multiplex PCR assay. Lanes 1, 8, and 15 contain the 1 kb ladder and lanes 4, 7, 11 and 14 contain the negative control – deionized H₂O. Lanes 2, 5, 9 and 12 contain the isolate being assayed, T01, and lanes 3, 6, 10 and 13 contain a positive control. The numbers of the primer pairs corresponding to each band are listed in pink and red (pink representing amplicons that are present and red representing products that did not amplify). The sizes of the ladder bands are listed in white. The bands in the 2nd and 3rd lanes are underlined in order to ease visualization of the presence of amplicons in isolate T01 that align properly with the positive control amplicons.

DISCUSSION

Previously the mechanisms causing atoxigenicity among *A. flavus* isolates were not completely understood. The current study unveiled some of the genetic changes responsible for the atoxigenicity in 12 isolates under development as biocontrol agents in the United States and Africa. Each of these isolates has been analyzed with a Multiplex PCR assay, which allowed for a brief characterization of the genes present, or lack thereof, within the aflatoxin biosynthesis cluster. The isolates can be divided into six groups according to their deletion patterns shown in Figure 1.

The first group consists of isolates B19, T01, and T16. These isolates all share identical deletion patterns. The next group, consisting of isolates S16, S25, and T13, shares a common insertion sequence (see Figure 3). The third group consists of isolates K39 and K49, each of which share identical deletion patterns. The fourth group, isolates K11 and T19, share the same deletion pattern from the beginning of the aflatoxin biosynthesis gene cluster to the area surrounding the *norB* gene. This area includes the last known genes present in isolate K11, while isolate T19 has an additional 8,271 basepairs present that are not found in isolate K11. The fifth and sixth groups consist of one isolate each, due to their unique deletion patterns. Isolates S01 and K10 have deletion patterns unlike any of the other isolates in this study.

The assembly of these isolates into separate groups based on their genetic properties provides valuable information that will later be useful in the development of genetic assays that screen for potential biocontrol isolates. It is interesting to note the variability within each group. The first group

contains isolates from Burkina Faso and Texas. Similarly, the second group contains isolates from Senegal and Texas. The third group only has isolates from Kenya, while the fourth group has isolates from Kenya and Texas. Evidently, the origin of the isolate does not always have a direct correlation with the genetic changes that occur within the aflatoxin biosynthesis cluster. Phylogenetic analyses would need to be performed in order to trace the divergence of these isolates originating from different areas of the world.

The genetic profiles created from the Multiplex PCR data served as the basis for choosing which areas to amplify with the telomere-repeat sequenced derived primers. PCRs were performed from the last genes present in the cluster to the telomeric end of the chromosome. Although primers *teloR* and *AgteloR* (Table 2) are essentially the same primer sequence, differing only slightly in the anchor region, they were both used in each attempted amplification to increase accuracy. Upon sequence analysis there were two breakpoints identified.

The first breakpoint was found in isolate K10 (Figure 2). This was discovered by aligning the K10 183T/*teloR* product sequence with *A. flavus* isolate 3357 from the GenBank database. Isolate K10 aligns with 3357 uninterrupted, with the exception of 2 SNPs present on the *mdr4* gene. The breakpoint occurs on the *mdr4* gene and is immediately followed by the reverse complement of the *teloR* primer sequence. This discovery is a key element that will aid in the creation of a PCR protocol that can rapidly identify atoxigenic isolates with similar genetic alterations.

The second breakpoint was identified in isolate S01 (Figure 3). This breakpoint was recognized in a similar manner to the previous one found in isolate K10. The alignment of the S01 omtB2F/teloR product sequence to *A. flavus* isolate 3357 demonstrates 100% identity with no disruptions. The breakpoint occurs on the *omtB* gene and, similar to the first breakpoint, is followed by the reverse complement of the telomeric repeat. The telomeric repeat associated with this breakpoint differs slightly from the known *A. flavus* telomere repeat. However, the differences in sequence are not drastic enough to doubt the validity of the finding; instead, they provide a new insight to variability within this repeat region. Further investigations could be conducted that determine whether this change in the telomere repeat sequence is involved at any level with the atoxigenicity of the isolate.

Another finding from the sequence analysis of products amplified with the telomere-repeat sequence derived primers was the presence of an insertion sequence in 3 of the isolates: T13, S16, and S25 (Figure 3). Upon alignment with the reference isolate, 3357, the 148T/teloR product sequences of isolates T13, S16, and S25 displayed a considerably large area of mis-consensus. The sequences of these disagreements within the alignment were put into the BLAST search engine and the top hit was from an *Aspergillus oryzae* isolate (Accession number: AB506482.1). The alignments of the sequences from isolates T13, S16, and S25 can be seen in Figure 3. The 148T/teloR amplicons for all three of the isolates were 1,645 basepairs long and when aligned together displayed 100% consensus. The initial 709 basepairs of the amplicon sequences aligned with 3357 consists of 30

SNPs. The insertion sequence only contains 1 SNP and one single basepair insertion when aligned to the *A. oryzae* isolate obtained from BLAST. From these findings, it can be inferred that *A. oryzae* diverged from *A. flavus*. The divergence of the two species could potentially be involved in the causation of atoxigenicity. It is interesting to note the origins of these 3 isolates. T13 originates from Texas, while S16 and S25 originate from Senegal. Further research would need to be conducted to explain the evolution of this genetic change that occurs in these isolates.

The discovery of this insertion led to the creation of 5 primer sets in the subtelomeric region of the *A. oryzae* sequence from the GenBank database (Table 2). The primers were specifically made to target areas closer to the centromeric region of the chromosome in order to determine whether the insertion sequence continued further into this region. Upon analysis of the PCR products of these primers with gel electrophoresis, it was evident that amplicons were present; however, these amplicons have not yet been sent for sequencing. More primers can be designed following a similar approach to verify the exact size of the insertion present in the isolates. The amplicons would need to be sent for sequencing to analyze new genes that may be encoded by the insertion.

Although internal amplifications across missing gene segments were attempted, there were not any that were successful. Future studies involving these 12 isolates will consist of designing primers that target specific sequences within the aflatoxin biosynthesis cluster. These primers will allow for internal amplifications to be done across missing genes. The potential amplicons that will be produced from these

amplifications will be sent for sequencing and subsequently analyzed. The analysis should provide further insight of what is really occurring within the aflatoxin biosynthesis cluster of atoxigenic isolates. Resultantly, we should be able to identify additional genetic modifications that will allow us to detect

potential bio-control agents more efficiently and also help us to better understand evolution and divergence among atoxigenic strains. The ultimate goal is to facilitate regulatory approval of additional atoxigenic strains as active ingredients in biopesticides and to do so in a quick and effective manner.

References

- Cotty PJ, Bayman P, Egel DS, Elias KS. 1994. Agriculture, Aflatoxins, and *Aspergillus*. p 1-27 in: "The Genus *Aspergillus*: From Taxonomy and Genetics to Industrial Applications" FEMS Symposium No. 69. Powell KA, Renwick A, Peberdy JF, editors. Plenum Press.
- Cotty PJ. 2006. Biocompetitive exclusion of toxigenic fungi. p. 179-197 in: "The Mycotoxin Factbook food and feed topics". Barug D, Bhatnagar D, van Egmond HP, van der Kamp JW, van Osenbruggen WA, Visconti A, editors. Wageningen Academic Publishers, the Netherlands.
- Cotty PJ, Antilla L, Wakelyn PJ. 2007. Competitive Exclusion of Aflatoxin Producers: Farmer Driven Research and Development. p. 241-253 in: "Biological Control: A Global Perspective". Vincent C, Goettel N, Lazarovits G, editors. CAB International, Oxfordshire, UK.
- Das MK, Ehrlich EC, Cotty PJ. 2008. Use of pyrosequencing to quantify incidence of a specific *Aspergillus flavus* strain within complex fungal communities associated with commercial cotton crops. *Phytopathology* 98: 282-288
- Donner M, Atehnkeng J, Sikora RA, Bandyopadhyay R, Cotty PJ. 2010. Molecular characterization of atoxigenic strains for biological control of aflatoxins in Nigeria. *Food Addit Contam* 27: 576-590
- Ehrlich KC, Cotty PJ. 2004. An isolate of *Aspergillus flavus* used to reduce aflatoxin contamination in cottonseed has a defective polyketide synthase gene. *Appl Microbiol Biotechnol*. 65:473-478.
- Ehrlich KC, Montalbano BG, Cotty PJ. 2007. Analysis of single nucleotide polymorphisms in three genes shows evidence for genetic isolation of certain *Aspergillus flavus* vegetative compatibility groups. *FEMS Microbiol Lett*. 268:231-236.
- Probst, C., Callicott, K.A. & Cotty, P. J. (2012). Deadly strains of Kenyan *Aspergillus* are distinct from other aflatoxin producers. *European Journal of Plant Pathology* 132 (3):419-429.
- Steve Rozen and Helen J. Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386
- Van Egmond. H.P., 2002. Worldwide regulations for mycotoxins. *Advances in Experimental Medicine and Biology* 504:257-269.
- Van Egmond HP, Schothorst RC, Jonker MA. 2007. Regulations relating to mycotoxins in food: perspectives in a global and European context. *Anal Bioanal Chem*. 389:147-157.
- Williams, J. H., Phillips, T. D., Jolly, P. E., Stiles, J. K., Jolly, C. M., & Aggarwal, D. (2004). Human aflatoxicosis in developing countries: a review of toxicology, exposure, potential health consequences, and interventions. *American Journal of Clinical Nutrition*, 80, 1106-22.