

ORIGINAL ARTICLE

Open Access



Degeneration of aflatoxin gene clusters in *Aspergillus flavus* from Africa and North America

Bishwo N. Adhikari¹, Ranajit Bandyopadhyay² and Peter J. Cotty^{1*}

Abstract

Aspergillus flavus is the most common causal agent of aflatoxin contamination of food and feed. However, aflatoxin-producing potential varies widely among *A. flavus* genotypes with many producing no aflatoxins. Some non-aflatoxigenic genotypes are used as biocontrol agents to prevent contamination. Aflatoxin biosynthesis genes are tightly clustered in a highly conserved order. Gene deletions and presence of single nucleotide polymorphisms (SNPs) in aflatoxin biosynthesis genes are often associated with *A. flavus* inability to produce aflatoxins. In order to identify mechanisms of non-aflatoxigenicity in non-aflatoxigenic genotypes of value in aflatoxin biocontrol, complete cluster sequences of 35 *A. flavus* genotypes from Africa and North America were analyzed. Inability of some genotypes to produce aflatoxin resulted from deletion of biosynthesis genes. In other genotypes, non-aflatoxigenicity originated from SNP formation. The process of degeneration differed across the gene cluster; genes involved in early biosynthesis stages were more likely to be deleted while genes involved in later stages displayed high frequencies of SNPs. Comparative analyses of aflatoxin gene clusters provides insight into the diversity of mechanisms of non-aflatoxigenicity in *A. flavus* genotypes used as biological control agents. The sequences provide resources for both diagnosis of non-aflatoxigenicity and monitoring of biocontrol genotypes during biopesticide manufacture and in the environment.

Keywords: *Aspergillus flavus*, Aflatoxin gene cluster, Non-aflatoxigenic, Cluster degeneration, Biocontrol, Evolution

Introduction

Aspergillus flavus, the primary causal agent of food and feed contamination with the toxic fungal metabolites, aflatoxins (Cotty et al. 1994; Klich 2007; Probst et al. 2010, 2011), is ubiquitous in the environment. This prolific saprophyte (Klich 2002) is an opportunistic pathogen of plants and animals (Leger et al. 2000) including humans (Hedayati et al. 2007; Sepahvand et al. 2011). A wide variety of crops including maize, cottonseed, peanuts, and tree nuts are susceptible to infection and subsequent aflatoxin contamination (Cotty et al. 1994; Doster and Michailides 1994). Aflatoxin B₁, the most toxic aflatoxin, is carcinogenic for both humans and

animals (McKean et al. 2006). Aflatoxins in foods and feed are limited through strict regulation which results in significant economic loss for producers and processors of contaminated crops (Robens and Cardwell 2003; van Egmond et al. 2007).

Populations of *A. flavus* in agricultural fields are complex communities that contain many Vegetative Compatibility Groups (VCGs) (Cotty et al. 1994; Ehrlich et al. 2007). Aflatoxin-producing potential is known to vary less among genotypes within a VCG than among genotypes from different VCGs (Ehrlich and Cotty 2004) and the average aflatoxin-producing potential of populations varies among fields, areas, and regions (Cotty 1997; Lisker et al. 1993; Schroeder and Boller 1973). Endemic *A. flavus* genotypes that don't produce aflatoxin, frequently called non-aflatoxigenic strains, have been used successfully as biocontrol agents to reduce aflatoxin contamination in cottonseed (Cotty 1989; Ehrlich et al. 2007),

*Correspondence: pjcotty@email.arizona.edu

¹ USDA-ARS, The University of Arizona, School of Plant Sciences, 303 Forbes Building, P.O. Box 210036, Tucson, AZ 85721, USA
Full list of author information is available at the end of the article

peanut (Dorner 2008, 2009; Dorner and Horn 2007), corn (Atehnkeng et al. 2008, 2014; Brown et al. 1999), and pistachio (Doster et al. 2014) where they competitively displace aflatoxin-producing fungi (Cotty and Mellon 2006). Biocontrol applications reshape the fungal community that grows in association with developing crops so that the non-aflatoxigenic genotypes dominate the fungal community structure and as a result the aflatoxin-producing potential of that fungal community is greatly reduced. Biological control products directed at reducing contamination and utilizing endemic well adapted, non-aflatoxigenic genotypes of *A. flavus* as active ingredients are registered for use in the US, Nigeria, and Kenya, and are under development in several other nations in Africa, the Americas, and Europe (Atehnkeng et al. 2014; Chulze et al. 2015; Cotty 2006; Mauro et al. 2013). The first non-aflatoxigenic genotype to be registered as a biopesticide active ingredient by a regulatory authority was *A. flavus* AF36 which, after more than a decade, is still used in commercial agriculture in the US on cottonseed, corn, pistachios, and figs (Grubisha and Cotty 2015).

Enzymes and regulatory proteins for aflatoxin synthesis in *A. flavus* are encoded by 25 clustered genes in a 70-kb region (Ehrlich et al. 2005; Yu et al. 2004b). The genes involved in aflatoxin biosynthesis are clustered together and the order of genes within the cluster is highly conserved within *Aspergillus* section *Flavi*. Presence and absence of a complete functional aflatoxin gene cluster is often associated with the ability to produce toxin by the members of the *Aspergillus* section *Flavi*. Deletion of portions of the aflatoxin biosynthetic gene cluster within non-aflatoxigenic *A. flavus* genotypes is common (Chang et al. 2005) and strains of *Aspergillus* section *Flavi* with large deletions in the aflatoxin gene cluster have been used to study the genetics of aflatoxin biosynthesis (Prieto et al. 1996). Both *A. sojae* and *A. oryzae*, close relatives of *A. flavus* which are used for food fermentation, don't produce aflatoxins (Wei and Jong 1986) even though homologues of aflatoxin biosynthesis genes are present in the genomes of both species (Chang et al. 1995; Klich et al. 1995; Yu et al. 2000). Similarly, a single nucleotide polymorphism (SNP) in the aflatoxin pathway polyketide synthase gene in *A. flavus* AF36 (Ehrlich and Cotty 2004; Ehrlich et al. 2007) and deletion of the entire aflatoxin gene cluster (Chang et al. 2005; Dorner 2004) in *A. flavus* NRRL21882 (active ingredient of afla-guard®) are sufficient to explain non-aflatoxigenicity in these active ingredients of commercially used biopesticides. Nevertheless, mechanisms responsible for non-aflatoxigenicity are diverse and for most non-aflatoxigenic *A. flavus* genotypes, the specific genetic changes leading to non-aflatoxigenicity are unknown.

Degradation of the aflatoxin gene cluster in *A. oryzae* included indel and SNP mutations. Deletion of a large portion of the aflatoxin biosynthesis gene cluster, including *aflR*, was detected in 40 % of the *A. oryzae* RIB strains (groups 2 and 3) in analyses including 39 *A. oryzae* genotypes (Kusumoto et al. 2000). A number of mutations relative to the *A. flavus* sequence were observed in the aflatoxin biosynthesis gene cluster in *A. oryzae* RIB 40 (Tominaga et al. 2006). Eight deletion patterns were detected among 38 non-aflatoxigenic *A. flavus* genotypes from the southern United States (Chang et al. 2005). When selective forces necessary to maintain the aflatoxin gene cluster are relaxed, genetic drift may lead to further mutations and deletions in genes related to aflatoxin synthesis (Chang et al. 2005). Although the loss of production of aflatoxin by some non-aflatoxigenic genotypes is attributed to point mutations (Ehrlich and Cotty 2004) or small deletions in genes essential for aflatoxin production (Calvo et al. 2004), the process of degeneration of aflatoxin gene cluster is not well understood.

Non-aflatoxigenic *A. flavus* genotypes have been identified from different regions of the world (Cotty 1997; Joffe 1969; Lisker et al. 1993; Schroeder and Boller 1973) and abilities of these genotypes to reduce aflatoxin have been demonstrated in many crops (Probst et al. 2014). Previous studies on non-aflatoxigenic *A. flavus* have primarily focused on analysis of few genotypes through analysis of selected genes in the aflatoxin gene cluster and no study has compared complete genome sequences of aflatoxin gene clusters. Studies have primarily relied upon PCR amplification to evaluate gene presence (Callicott and Cotty 2015; Chang et al. 2005; Kusumoto et al. 2000; Tominaga et al. 2006). Such studies provide an incomplete picture of cluster structure and may result in misleading conclusions. Though the diversity of *A. flavus* biocontrol agents is associated with variation in the aflatoxin gene cluster and deletion is a major cause of non-aflatoxigenicity, the process of acquisition and the evolution of non-aflatoxigenicity needs more detailed description to facilitate both biopesticide regulatory processes and DNA based monitoring of non-aflatoxigenicity. Molecular elucidation of aflatoxin gene clusters in non-aflatoxigenic *A. flavus* also has important implications for understanding mechanisms of non-aflatoxigenicity and the evolution of aflatoxin gene clusters in *A. flavus*. Effectiveness of non-aflatoxigenic genotypes varies by crop, location and nutrient environment (Mehl and Cotty 2010, 2013; Mehl et al. 2012). Thus it is important to identify new genotypes that are well adapted to both target crops and target regions in order to optimally reduce aflatoxin concentration in infected crops. In order to identify mechanisms of non-aflatoxigenicity

and understand the evolution of aflatoxin gene clusters in non-aflatoxigenic genotypes, we sequenced and analyzed aflatoxin gene clusters from 35 *A. flavus* genotypes from North America, West Africa and East Africa. By assembling the aflatoxin gene cluster and flanking regions, we analyzed the diversity of aflatoxin gene clusters of *A. flavus* genotypes currently incorporated as active ingredients in aflatoxin preventing biopesticides. The molecular characteristics of these factors will be useful for both understanding mechanisms of non-aflatoxigenicity and for monitoring non-aflatoxigenic genotypes on crops and in the environment.

Materials and methods

Fungal cultures

Thirty-five genotypes were chosen to represent the non-aflatoxigenic genotypes from Africa and North America. Non-aflatoxigenic genotypes of *A. flavus* from Kenya and Nigeria were collected and processed during previous studies (Donner et al. 2010; Probst et al. 2011). Genotypes from Burkina Faso and Senegal were genotyped and identified from samples transported to IITA's plant pathology laboratory in Ibadan, Nigeria. Maize and soil samples collected in the United States (Texas and Arizona) were transported to the United States Department of Agriculture (USDA), Agricultural Research Service (ARS), at the University of Arizona. Genotype recovery, identification and aflatoxin quantification was done as previously (Cotty 1997; Cotty and Cardwell 1999). Non-aflatoxigenicity was confirmed on maize for all fungi from cultures resulting from two serial single spore transfers. Geographic origins and sources of the 35 genotypes are shown in Table 1.

DNA isolation, library preparation and sequencing

Genomic DNA was isolated from conidia collected after 7 days culture (31 °C, dark) on 5 % V-8 vegetable juice, 2 % salt, and 2 % agar. The FastDNA SPIN Kit and the FastPrep Instrument were used following the manufacturer's instructions (MP Biomedicals LLC, Santa Ana, CA, USA). DNA from the FastDNA SPIN Kit was applied to a SPIN filter column following manufacturer's instructions to remove small DNA fragments and other contaminants <10,000 Da. Genomic DNA was quantified with both spectrophotometer (modelND-1000, NanoDrop) and with the Qubit dsDNA BR assay kit (Q32850) using the Qubit 1.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA) following the manufacturer's guidelines. Libraries were prepared according to Illumina's HiSeq 2000 library preparation protocol (Illumina, San Diego, CA, USA). Whole genome sequencing was performed at Arizona Genomics Institute (AGI) located at the University of Arizona's BIO5 institute using Illumina

HiSeq 2000. Libraries were sequenced with 100-bp paired-end reads and an insert size of 250-bp.

Aflatoxin gene cluster assembly and comparative analysis

To ensure the quality of the reads, the raw reads obtained from AGI were quality trimmed and cleaned using cutadapt v1.8.3 (Martin 2011). Cleaned reads were assembled using a de novo genome assembly program Velvet v1.2.10 (Zerbino and Birney 2008). The aflatoxin gene cluster from each assembled genome was extracted using the published aflatoxin gene cluster of *A. flavus* AF13 as a reference (Ehrlich et al. 2005) through BLAST alignment. Aflatoxin gene clusters were annotated using MAKER (Cantarel et al. 2008) genome annotation pipeline and multiple alignments were performed using CLUSTALW (Thompson et al. 1994) and MUMMER (Delcher et al. 2002). Putative functional annotation was assigned by searching the gene models against UniProt (Bateman et al. 2015) and previously annotated aflatoxin gene clusters from *A. flavus* using BLASTX (Altschul et al. 1990). To investigate the evolution of several portions of the aflatoxin-gene cluster, we divided the aflatoxin gene cluster pathway into two halves considering *ver-1* as a midpoint (Roze et al. 2007) and thus excluded from the analysis when comparing early and late portions of the gene cluster. The early part of the cluster includes *norB* through *norA* genes on the telomere side of the cluster and later part includes *verA* through *hypA* genes on the centromere side of the cluster. The assembled aflatoxin gene clusters are deposited at EMBL-EBI (ENA) database under the study accession number PRJEB11911.

Variant analysis

Single nucleotide polymorphisms (SNPs) in aflatoxin gene clusters were computationally detected using two SNP calling methods. First, reads from different genotypes were mapped to aflatoxin gene cluster from *A. flavus* AF13 using Bowtie v1.1.1 (Langmead et al. 2009) and the resulting BAM file was fed to SAMTools v0.1.16 (Li et al. 2009). SNPs positions were identified using mpileup function in SAMTools. SNPs were filtered using vcfutils.pl varFilter with minimum mapping quality (-Q) of 20 and minimum and maximum read coverage of 20 and 100 respectively. Second, to improve the confidence in variant calling, SNPs were also identified using MAQ v0.7.1 (Li et al. 2008) easyrun pipeline. SNPs were filtered using maq.pl SNPfilter with minimum read depth (-d) of 20, minimum consensus quality (-q) of 20, and minimum adjacent consensus quality (-n) of 20. Only SNPs called by both methods were considered for the analysis. In order to find the false positive rate of SNP called by two different methods used in this analysis, a random sample of 3–9 SNPs from six different genes were selected

Table 1 *Aspergillus flavus* genotypes used in the current study

	Country	Substrate	Group ^a	Genotype ^b	VCG	Culture accession/source ^c	
East Africa	Kenya	Maize	A	C6-E	KN00A	USDA-ARS, Tucson, USA	
			C	C8-F	KN012	USDA-ARS, Tucson, USA	
			B	E63-I	KN001	USDA-ARS, Tucson, USA	
			B	R7-H	KN011	USDA-ARS, Tucson, USA	
West Africa	Burkina Faso	Groundnut	D	GO18-2	BF018	IITA, Ibadan, Nigeria	
			A	GO67-10	BF067	IITA, Ibadan, Nigeria	
		Maize	A	M011-8	BF011	IITA, Ibadan, Nigeria	
			D	M092-15	BF092	IITA, Ibadan, Nigeria	
			A	M102-11	BF102	IITA, Ibadan, Nigeria	
			A	M109-2	BF109	IITA, Ibadan, Nigeria	
			F	M110-7	BF110	IITA, Ibadan, Nigeria	
			A	M129-5	BF129	IITA, Ibadan, Nigeria	
	Nigeria	Maize	A	Ka16127	AV16127	IITA, Ibadan, Nigeria	
			A	La3279	AV3279	IITA, Ibadan, Nigeria	
			A	La3304	AV3304	IITA, Ibadan, Nigeria	
			F	Og0222	AV0222	IITA, Ibadan, Nigeria	
		Senegal	Maize	A	M2-7	SN002	IITA, Ibadan, Nigeria
				E	M21-11	SN021	IITA, Ibadan, Nigeria
			A	Ms14-19	SN014	IITA, Ibadan, Nigeria	
			A	Ss19-14	SN019	IITA, Ibadan, Nigeria	
North America	USA	Sesame	A	Ss19-14	SN019	IITA, Ibadan, Nigeria	
		Cottonseed	A	AF36	YV36	NRRL 18543	
			F	AT21-A	TX021	USDA-ARS, Tucson, USA	
			F	AT4-C	TX004	USDA-ARS, Tucson, USA	
			D	AT5-B	TX005	USDA-ARS, Tucson, USA	
			A	BA16-F	TX016	FGSC A2220	
			F	BA35-C	IC001	FGSC A2223	
			A	BY18-A	EC11-D	USDA-ARS, Tucson, USA	
			F	BY19-D	IC001	USDA-ARS, Tucson, USA	
			A	DO107-L	TX107	USDA-ARS, Tucson, USA	
			D	DO114-A	TX114	USDA-ARS, Tucson, USA	
			C	DO38-B	TX038	FGSC A2226	
			F	DO46-G	TX046	FGSC A2229	
			F	EC19-B	TX019	USDA-ARS, Tucson, USA	
			A	EC69-E	EC69-E	USDA-ARS, Tucson, USA	
			F	NRRL21882		NRRL21882	
			Cottonseed	A	AF13	YV13	ATCC 96044
					EB01	EB01	USDA-ARS, Tucson, USA
	MR2-17	MR17		USDA-ARS, Tucson, USA			
	OD02	OD02		USDA-ARS, Tucson, USA			

^a Refers to the grouping based on aflatoxin gene cluster alignment in Fig. 1 and Neighbor-Net network in Fig. 2

^b Isolates from Kenya were first reported in Probst et al. (2011); isolates from Nigeria were first reported in Atehnkeng et al. (2008); AF13 and AF36 were first reported in Cotty (1989); EB01 and MR2-17 were first reported in Mehl and Cotty (2013); and OD02 was first reported in Grubisha and Cotty (2009). Rest of the isolates are first reported in this manuscript

^c Culture collection/source designation (IITA, The International Institute of Tropical Agriculture, Oyo Road, Ibadan, Nigeria; NRRL, ARS Culture Collection; FGSC, Fungal Genetics Stock Center)

for validation. SNP validation was done by manually checking the alignment and by PCR amplification and sequencing of the polymorphic sites in five genotypes. Deletions in aflatoxin gene clusters were predicted using

variant detection program DELLY (Rausch et al. 2012). By using paired-end mapping and split-read analysis, DELLY identifies deletions, duplications, inversions and translocations in the genomes. Deletion predictions with

supporting reads less than three and the mapping quality below 20 were rejected. Only deletions greater than 10 bp were considered for further analysis. Deletions predicted by DELLY were validated by manually checking the alignments and through PCR, by designing primers that bridge the putative gaps or that amplify from regions flanking the gap within the deleted region. PCR conditions used to validate the deletions were described previously (Mehl and Cotty 2010).

Neighbor-Net network

Genetic relationships among the 35 genotypes were examined by using simple sequence repeat (SSR) data for 17 loci identified in the previous study (Grubisha and Cotty 2009) scored manually for all included genotypes according to their amplified fragment size. Genetic distance was calculated across 17 loci with START2 program (Jolley et al. 2001). The distance matrix was analyzed with the Neighbor-Net algorithm in SplitsTree v4.13.1 (Huson and Bryant 2006). The tree was displayed using SplitsTree. Edges were color shaded according to the grouping of genotypes based on structure of aflatoxin gene cluster.

Identification of similarly evolving genes

To identify the genes evolving together, correlations between SNP densities from different genotypes were calculated. Pearson's correlation coefficient, r , was calculated for every gene combination in the combined aflatoxin biosynthesis and sugar clusters. Groups of four genes each from early and late portions of the aflatoxin gene cluster and from the sugar cluster were selected and the level of correlation of these gene groups with each gene in both clusters was estimated as the mean of the correlation coefficients of the four groups of genes with that gene. Two or more genes were considered evolving together if the average correlation coefficient between them was higher than 0.50 and considered highly evolving if correlation coefficient was higher than 0.70 (Fraser et al. 2004).

Ka/Ks analysis

To measure the rates of evolution of aflatoxin gene cluster genes, the ratio of non-synonymous to synonymous substitutions (Ka/Ks) between pairs of genes from five non-aflatoxigenic genotypes and toxigenic genotype *A. flavus* AF13 was calculated using the yn00 package in PAML (Yang 2007). Five non-aflatoxigenic genotypes were selected to represent the genotypes with full and partial aflatoxin gene clusters. Four to six genes representing early, middle and late segments of the aflatoxin gene cluster were examined from genotypes with complete as well

as partial gene clusters. Only orthologous genes were used for the analysis.

Statistical analysis

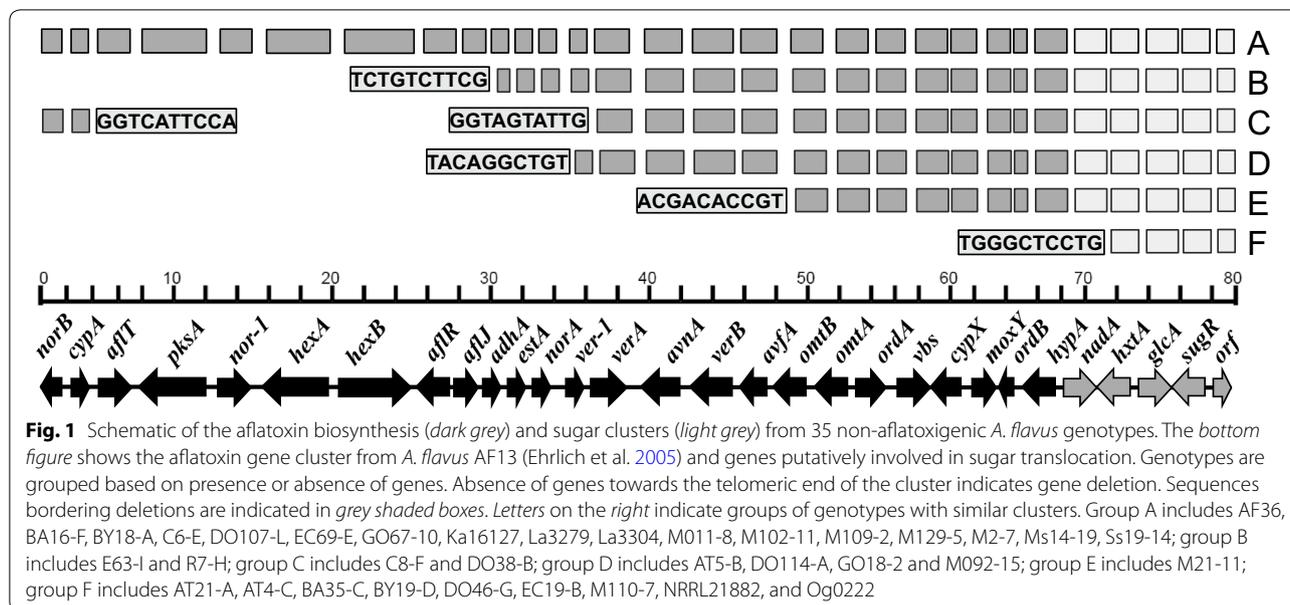
The number and density of SNPs were subjected to analysis of variance (ANOVA). Pearson's correlation coefficient was calculated on SNP densities. Tests of differences in means were performed following significant ANOVAs with Tukey's honest significant difference (HSD) test. Pair-wise comparisons of number of synonymous to non-synonymous SNPs, SNP densities and Ka/Ks values were done using Paired t test. Statistical analyses were performed with SAS (version 9.1, SAS Institute Inc., Cary, NC, USA).

Results

Variation in the aflatoxin gene cluster

Whole genome sequence of the non-aflatoxigenic genotypes analyzed in this study (Table 1) was obtained by aligning assembled genomes with published aflatoxin gene cluster from *A. flavus* AF13 (Ehrlich et al. 2005). For all genotypes, the aflatoxin gene cluster was part of a large contig/scaffold, which mapped to the sub-telomeric region of chromosome 3 of *A. oryzae* RIB40 (Machida et al. 2005). Distribution of genes within the cluster in *A. oryzae* RIB40 and directional alignment of each are identical in *A. flavus* NRRL3357 (Ehrlich et al. 2005) and *A. parasiticus* SRRRC 2043 (Yu et al. 2004a).

Analyses of aflatoxin gene clusters from 35 genotypes showed high level of diversity in terms of number and size of deletions in the genes (Fig. 1). Based on the size and type of deletion, genotypes ranged from those with full aflatoxin gene cluster to those having no genes at all with most deletions happening at the left end of the cluster or towards the telomeric end (Fig. 1). Based on the structure and presence of genes, genotypes were divided into 6 different groups (Group A through F). Group A contains 17 genotypes (49 % of examined genotypes) including AF36, a commercially used biocontrol agent, that have all the aflatoxin cluster genes. Groups B through E contain partial aflatoxin gene clusters with varying level of deletions and Group F contains genotypes with none of the aflatoxin genes. Group B contains two genotypes (6 % of examined genotypes) with an identical deletion including 11 genes. In addition to those 11 genes, genotypes in this group have also lost part of *aflJ* gene. The terminal gene, *aflJ*, from these two genotypes has an identical deletion indicating a shared deletion event. Group C, includes two genotypes that share identical deletions at the left end of the cluster. Genotypes in this group contain two genes (*norB* and *cypA*) on the left end and lost 13 genes leaving partially deleted *verA*



at the middle of the cluster. Similarly, group D contains four genotypes (12 % of examined genotypes) with deletion of 14 genes including *norA* but with *ver-1* left intact. Group E contains only one genotype with a large deletion including 19 complete genes and partial deletion of *omtB* gene. Despite the presence of partial gene clusters in groups B, C, D and E each group has remnants of different deletion events as indicated by the sequenced ends (Fig. 2). The fact that groups B, C, D and E have varying level of deletions suggest each went through different deletion events. Finally, Group F contains a distinct set of nine biocontrol genotypes (27 % of examined genotypes) including NRRL21882 with a completely deleted aflatoxin gene cluster. In addition to genes from the aflatoxin gene cluster, *nadA* gene from sugar cluster is also deleted from this group.

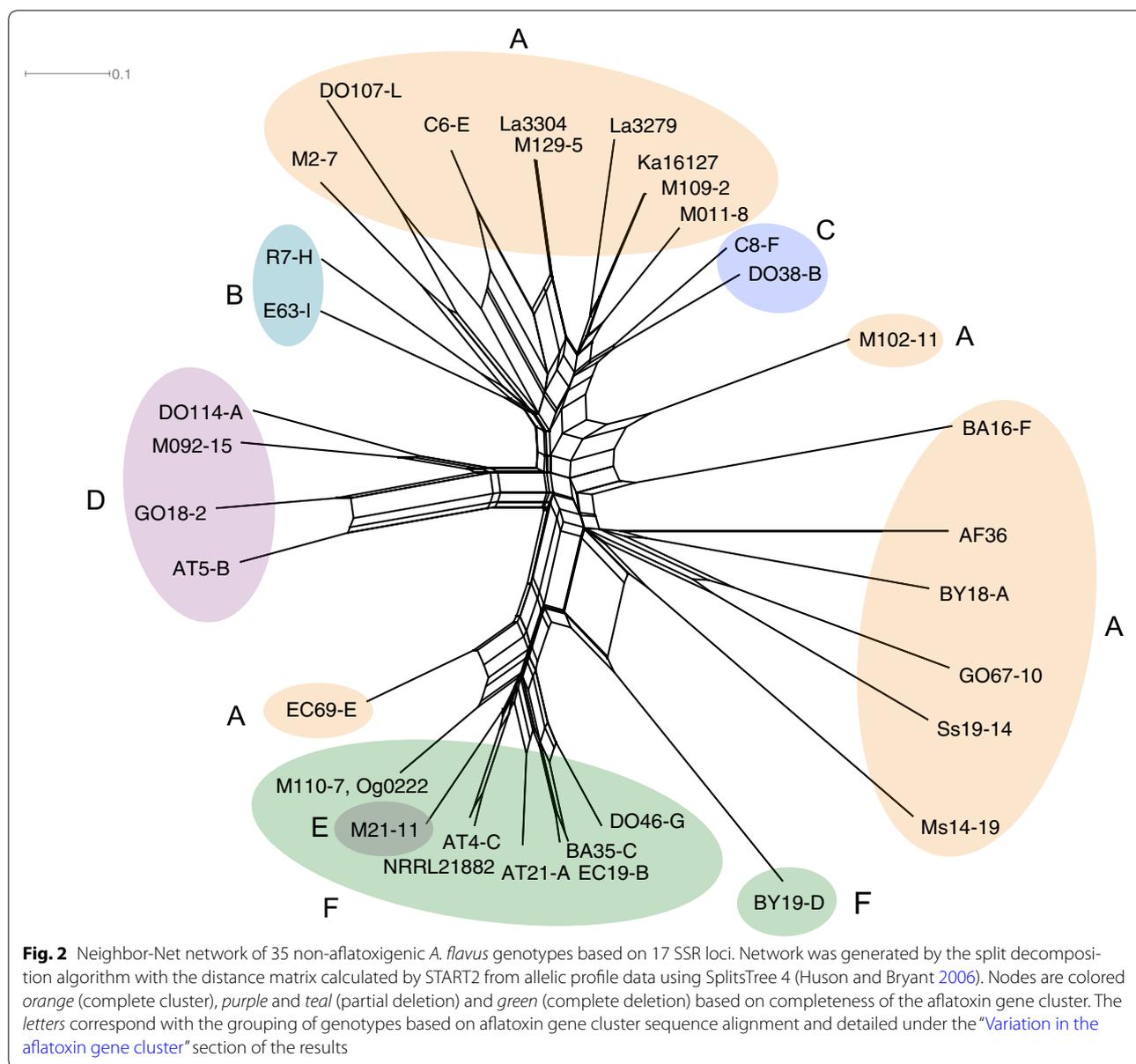
Evolutionary relationship between *A. flavus* genotypes

In order to examine the genetic relationships among non-aflatoxigenic genotypes analyzed in this study, a Neighbor-Net network (Fig. 2) was constructed using the genetic distance estimated from SSR data from 17 loci distributed throughout 8 chromosomes of *A. flavus* (Gru-bisha and Cotty 2009). The Neighbor-Net network has a high degree of congruence with the genotype graphs (group A–F) revealed by sequencing (Figs. 1, 2). The nine genotypes in group F (lacking the entire aflatoxin gene cluster) are clustered together in the same region of the Neighbor-Net network. Nested within group F is the sole genotype in group E with over 19 genes deleted. The Neighbor-Net network clustered all isolates by group for groups B, C, and D. The only group for which genotypes

were not consistently grouped together was group A, the ancestral genotype. Genotypes in group A contain all the aflatoxin biosynthesis genes, each with distinct aspects of degeneration partially described by the lack of shared deletion pattern (Table 2). These observations are consistent with groups B through F arising from group A, the ancestral state. The network represents a nontree-like structure that recapitulates diversity but fails to show clustering based on geographic origins. Despite having different levels of partial deletion of aflatoxin genes, groups B, C, D and E are not clustered together (Fig. 2). Among the genotypes, genotypes with similar sequence ends are in the same group suggesting distinct deletion events for each group.

Deletion in aflatoxin gene cluster

In addition to deletions in the *norB*–*cypA* region present in all *A. flavus* (Chang et al. 2005; Probst et al. 2014), several other large deletions in the aflatoxin-biosynthesis gene cluster have been described (Callicott and Cotty 2015; Chang et al. 2005; Donner et al. 2010). The current study revealed both large (>1-kb, Fig. 1) and small (Table 2; Additional file 1: Table S1) deletions through alignment of sequences of 35 non-aflatoxigenic genotypes with the sequence of *A. flavus* AF13 (Ehrlich et al. 2005). Deletions were widespread across the clusters (Table 2; Fig. 1). Group A has deletions ranging from 17–278 bp (Table 2; Additional file 1: Table S1). Although deletions were found in 12 of the 25 genes present in the cluster; *cypA*, *pksA*, *aflR*, *verA*, *verb* and *avfA* genes have deletions in at least ten genotypes analyzed. Genotypes in Group B share the same approximately 30-kb deletion



on the telomeric end of the cluster with partial *aflJ* gene remaining at the end of the cluster. Group C genotypes have unique approximately 40-kb deletion with *norB* and *cypA* genes at the telomeric end of the cluster. Genotypes in Group D have an identical approximately 35-kb deletion. Group E genotype has approximately 50-kb deletion in the aflatoxin gene cluster. Group F is composed of genotypes that lack all genes in the aflatoxin gene cluster. Genotypes in the same group (Group B through F) have identical sequence ends, indicating that the deletion event in each group probably occurred in a common ancestor. In addition to larger deletions, genotypes in

group A, B, C, D, and E have 1-9 smaller deletions of varying sizes scattered among 12 genes (Table 2; Additional file 1: Table S1). Of the 35 genotypes analyzed in this study, more than 49 % have larger deletions in between *norB* to *estA* (Table 3) whereas smaller deletions were more concentrated in *norB*, *cypA*, *pksA*, *aflR*, *verA* followed by *verB*. The results from validation clearly showed that the smaller deletions were real (Additional file 1: Table S1). In line with the deletion pattern, it is interesting to note that the genes in early portions of the cluster are present only in 49–69 % of genotypes while genes in late portions of the cluster are present in at least 71–74 %

Table 2 Small (<1 kb) deletions in aflatoxin biosynthesis genes sequenced from 26 genotypes of *Aspergillus flavus*

Genotype	Group	No of deletions	<i>cypA</i>	<i>aflT</i>	<i>pksA</i>	<i>hexA</i>	<i>aflR</i>	<i>verA</i>	<i>verB</i>	<i>avfA</i>	<i>omtB</i>	<i>vbs</i>	<i>cypX</i>	<i>ordB</i>
AF36*	A	4	17 ^a		61		29						21	
AT5-B	D	2							57					33
BA16-F*	A	6	17	253	61			20	57				21	
BY18-A*	A	6	17		61	278	29	20		29				
C6-E*	A	6	23		61		29	20				42	21	
DO107-L*	A	4	23		61		29			29				
DO114-A	D	2							57					33
DO38-B	C	2	17											33
C8-F	C	2	17											33
E63-I	B	3							57	29	25			
EC69-E*	A	4	17	42	61									33
GO18-2	D	2							57					33
GO67-10*	A	4	23		61			20		29				
Ka16127*	A	6	23		61	278	29	20		29				
La3279*	A	9	17		61	278	29	20	57	29		42	21	
La3304*	A	5	17		61	278	29			29				
M011-8*	A	8	17		61	278	29	20		29		42	21	
M092-15	D	1							57					
M109-2*	A	6	17		61	278	29			29				33
M129-5*	A	6	17		61	278	29	20		29				
M2-7	A	7	17		61	278	29	20		29			21	
Ms14-19*	A	7	17	253	61			20	57				21	33
R7-H	B	3							57	29	25			
Ss19-14*	A	7	17	253	61			20	57				21	33
M102-11	A	5	23		61		29	20		29				
M21-11	E	1											21	
Total deletions			19	4	17	8	12	12	10	13	2	3	9	9

^a Numbers in italics indicate deletion sizes in bp

*These genotypes have complete sets of genes in aflatoxin gene cluster

of the genotypes (Table 3). Four genes from the sugar cluster (*hxtA*, *glcA*, *sugR* and *orf*) were the only genes present in all genotypes.

Polymorphism in aflatoxin gene cluster

Single nucleotide polymorphisms (SNPs) were identified by mapping Illumina sequencing reads from different genotypes to the *A. flavus* AF13 aflatoxin gene cluster (Ehrlich et al. 2005). Analysis included 26 non-aflatoxigenics with complete (17 genotypes) and partial (9 genotypes) sets of aflatoxin genes and 3 toxigenic genotypes. The average SNP density (SNPs/kb of gene) for non-aflatoxigenic and toxigenic genotypes varied from 2–6 and 0–2 respectively (Fig. 3). Although the average SNP density for all genotypes was 4 SNPs/kb, non-aflatoxigenic genotypes have significantly higher SNP density (4) as compared to toxigenic genotypes (2) ($P < 0.05$). Among 25 genes in aflatoxin gene cluster, *nor-1* has the lowest average SNP density (less than 1) while *omtA* has

the highest (8) followed by *vbs* (7) and *cypX* (7) (Fig. 3). For the genotypes with partial aflatoxin gene cluster, the average SNP density for genes ranged from 0–6 SNPs/kb but average SNP density for genotypes ranged from 1–7 SNPs/kb. Of all SNPs, we found transitions to be over-represented: 64–74 % of all single base pair substitutions were transitions. Examination of polymorphic sites from aflatoxin biosynthetic gene cluster showed all analyzed sites with true polymorphisms. The results from validation clearly showed that the SNP calling methods are robust and selected SNPs, which were called by two methods, have very low false positive rate (Additional file 1: Table S2). Genotypes from the same geographical region were not necessarily similar in terms of number and density of SNPs. For example, the four genotypes from Senegal were no more similar to each other than to the genotypes from North America; nevertheless the genotypes with similar deletion patterns have more shared SNPs than genotypes with different deletions (Fig. 3).

Table 3 Frequencies of aflatoxin biosynthesis and sugar cluster genes among 35 atoxigenic genotypes from Africa and North America

Cluster ^a	Gene ^b	Genotype (%) ^c	Synonymous SNPs ^d	Non-synonymous SNPs ^d	Nonsense SNPs ^e	
Early genes	<i>norB</i>	19 (54)	0–3 (43)	0–2 (49)	0	
	<i>cypA</i>	19 (54)	0–3 (43)	0–5 (43)	0–1 (49)	
	<i>aflT</i>	17 (49)	0–3 (43)	0–4 (43)	0	
	<i>pksA</i>	17 (49)	6–18 (40)	3–16 (40)	0–1 (3) ^f	
	<i>nor-1</i>	17 (49)	0–1 (6)	0–1 (3)	0	
	<i>hexA</i>	17 (49)	5–18 (40)	2–15 (40)	0	
	<i>hexB</i>	17 (49)	9–20 (49)	7–15 (49)	0	
	<i>aflR</i>	17 (49)	1–3 (49)	1–3 (49)	0	
	<i>aflJ</i>	19 (54)	2–9 (51)	0–2 (51)	0–1 (3)	
	<i>adhA</i>	19 (54)	1–2 (54)	1–4 (49)	0	
	<i>estA</i>	19 (54)	2–6 (54)	1–3 (54)	0	
	<i>norA</i>	24 (69)	0–5 (46)	0–3 (46)	0–1 (3)	
		<i>ver-1</i>	25 (71)	0–5 (63)	0–2 (57)	0
	Late genes	<i>verA</i>	25 (71)	0–37 (69)	1–19 (71)	0
		<i>avnA</i>	27 (77)	2–5 (77)	2–5 (77)	0
<i>verB</i>		25 (71)	0–5 (69)	1–3 (71)	0	
<i>avfA</i>		25 (71)	0–4 (57)	0–4 (63)	0	
<i>omtB</i>		26 (74)	0–6 (57)	1–3 (63)	0	
<i>omtA</i>		26 (74)	0–7 (66)	0–4 (71)	0–1 (3)	
<i>ordA</i>		26 (74)	0–9 (63)	0–6 (57)	0	
<i>vbs</i>		26 (74)	1–12 (74)	1–6 (74)	0–1 (3)	
<i>cypX</i>		26 (74)	1–11 (74)	1–3 (74)	0	
<i>moxY</i>		26 (74)	0–7 (69)	1–8 (74)	0	
<i>ordB</i>		26 (74)	0–2 (69)	0–2 (66)	0–1 (14)	
<i>hypA</i>		26 (74)	2–5 (74)	1–6 (74)	0–1 (14)	
Sugar cluster		<i>nada</i>	26 (74)	1–2 (74)	1–6 (74)	0–3 (63)
	<i>hxtA</i>	35 (100)	0–5 (89)	1–2 (100)	0	
	<i>glcA</i>	35 (100)	1–3 (100)	2–3 (100)	0	
	<i>sugR</i>	35 (100)	1–5 (100)	0–4 (77)	0–1 (14)	
	<i>orf</i>	35 (100)	0–3 (83)	1–2 (100)	0	

^a Classification is based on the enzymes encoded by these genes, which are involved in early and late portions of aflatoxin biosynthesis, considering *ver-1* (bold italics) as the middle gene

^b Genes are listed in the same order as in aflatoxin gene cluster

^c Percentages were calculated as (number of genotypes with the gene or SNP/total number of genotypes) × 100. The total number of genotypes was 35

^d Number of synonymous or non-synonymous SNPs present. Numbers in the parenthesis include percentage of isolates with at least one synonymous or non-synonymous SNP

^e Number of nonsense SNPs and percentage of isolate with at least one nonsense SNP

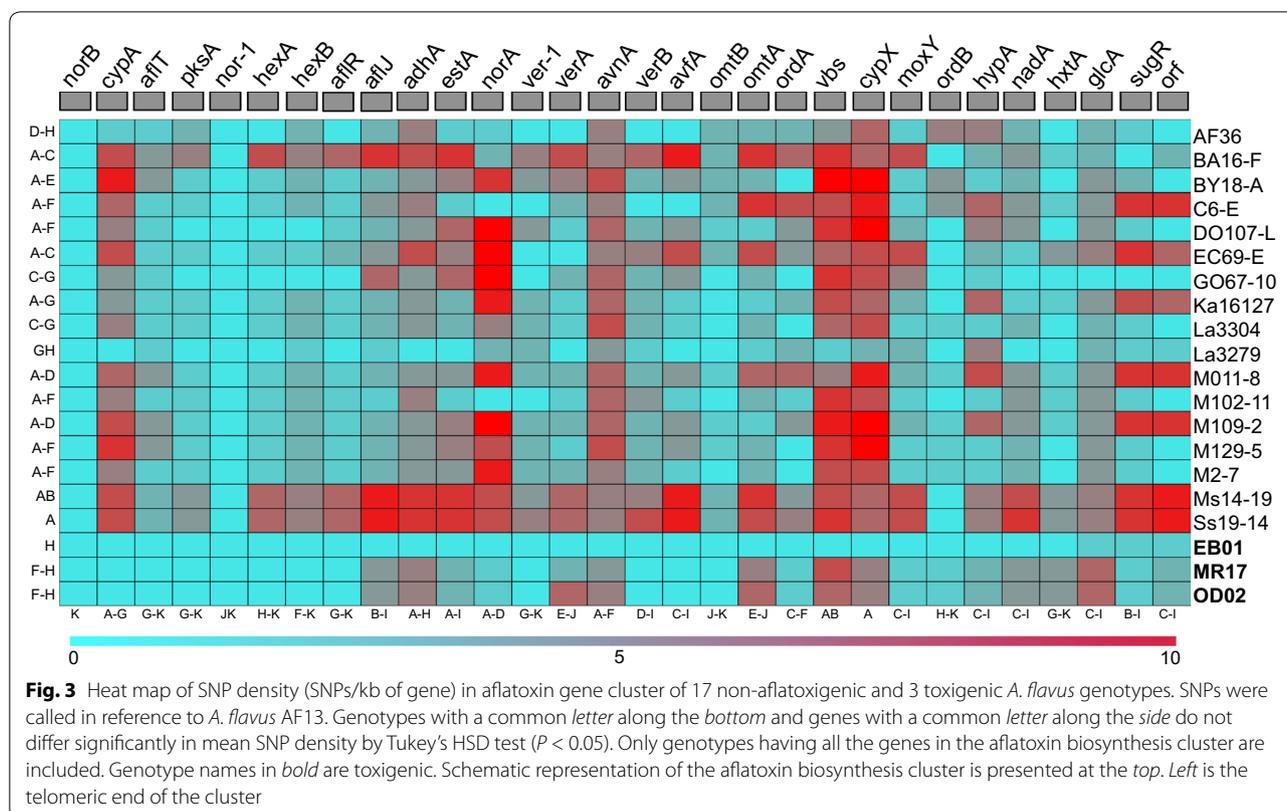
^f Stop codon present in *A. flavus* AF36 as reported in Ehrlich and Cotty (2004)

To further investigate SNPs across the cluster, annotation was performed with Annovar (Wang et al. 2010). The effect of SNP mutation varied among genotypes. Synonymous SNPs were significantly more abundant ($P < 0.005$) than non-synonymous SNPs for all genotypes analyzed. Of the 25 aflatoxin cluster genes, eight genes (*cypA*, *pksA*, *aflJ*, *norA*, *omtA*, *vbs*, *ordB*, *hypA*) have stop-gain SNPs, four (*norB*, *cypA*, *ordB*, *glcA*) have stop-lost SNPs, and one gene (*hexA*) has a start-lost SNP eliminating a start codon (Table 3; Additional file 1: Table S3). To better

understand SNP distribution, we analyzed SNP density in early and late portions of the gene cluster. The later portions had significantly higher SNP density ($P < 0.05$) (Fig. 3).

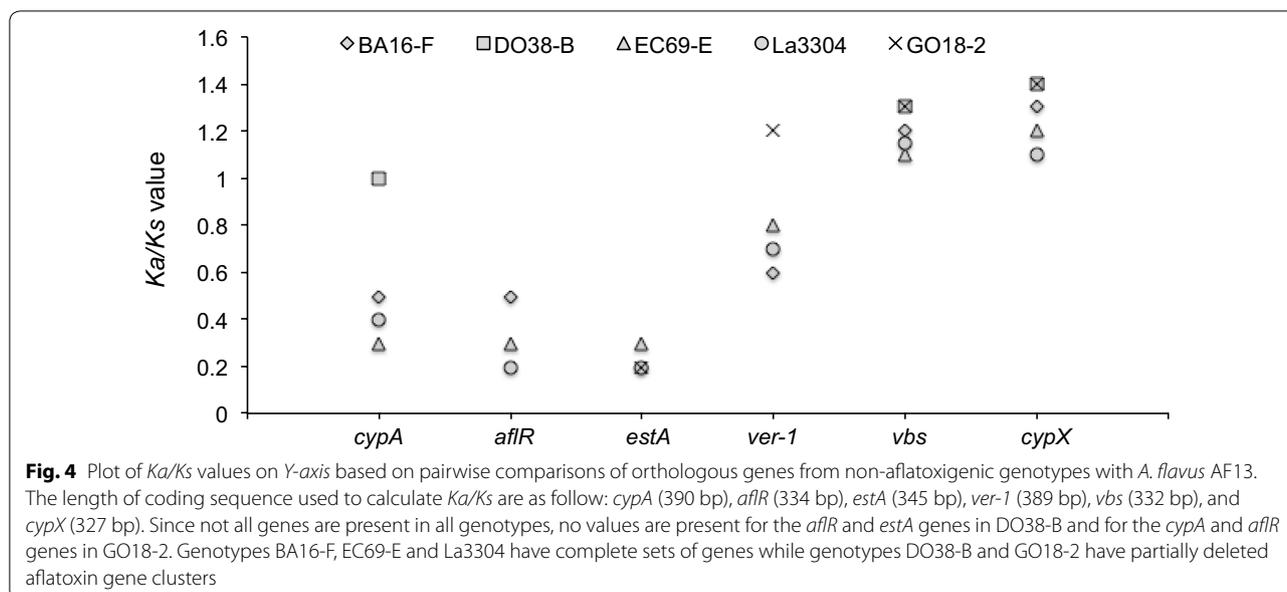
Selective pressure in aflatoxin gene cluster

To study the evolution of the genes within the aflatoxin gene cluster, *Ka/Ks* ratio was calculated for pairs of orthologous genes between non-aflatoxigenic genotypes and toxigenic *A. flavus* AF13. Six genes representing



early, middle and late portions of the cluster were selected from five genotypes either with full or partial sets of aflatoxin genes (Fig. 4). Of the six genes analyzed in this study, *cypX* and *vbs* have $Ka/Ks > 1$ for all genotypes, possibly indicating positive selection acting on these genes. The two genes, which are not present in all

genotypes, with lowest Ka/Ks ratio were *aflR* and *estA*. Out of six genes, *aflR*, *cypX*, *estA* and *vbs* have Ka/Ks ratio either higher or lower than 1 for all genotypes analyzed. On the contrary, *cypA* and *ver-1* have a higher Ka/Ks ratio only for genotypes with partial gene cluster and lower for all genotypes with full gene clusters, suggesting



that these two genes are evolving differently in different genotypes. Results showed that *cypX* and *vbs* have significantly higher *Ka/Ks* values (*Ka/Ks* > 1, *P* < 0.05, *t test*) than *aflR* and *estA* while *cypA* and *ver-1* are intermediate to those two groups of genes.

Differential evolution of genes

In order to identify the genes that are evolving together, correlation coefficients were calculated for SNP densities from 25 aflatoxin genes and 5 genes in sugar cluster from 17 non-aflatoxigenic genotypes with complete gene sets. In order to represent the whole cluster, two groups of four genes from the aflatoxin gene cluster, representing early and late portions, and one group from the sugar cluster were selected. The two groups from the aflatoxin gene cluster were designated as the *pksA* group (containing *pksA*, *nor-1*, *hexA* and *hexB* genes) and the *cypX* group (containing *ordA*, *vbs*, *cypX*, *moxY* genes) while the group from the sugar cluster was designated as the *sugR* group (containing *hxtA*, *glcA*, *sugR*, *orf* genes) (Table 4). Because it has been shown that one of the genes from sugar cluster (*hxtA*) is concurrently expressed with aflatoxin genes (Yu et al. 2000), we reasoned that genes within these clusters might be evolving together. Considering Pearson’s correlation coefficient threshold of 0.5, all four genes within *pksA* group were highly evolving together (Pearson’s *r* = 0.85) but no such correlation was found with the *cypX* (Pearson’s *r* = 0.21) and *sugR* groups (Pearson’s *r* = 0.19) (Table 4). In contrast, genes from the *cypX* group were not evolving with genes from

either *cypX* (average *r* = 0.35), *pksA* (average *r* = 0.21) or *sugR* (Pearson’s *r* = 0.15) groups. Four out of five genes from the sugar cluster, which are included in the *sugR* group, were evolving together (Pearson’s *r* = 0.66) while there was complete lack of such relationship with genes from the *pksA* (Pearson’s *r* = 0.19) and *cypX* (Pearson’s *r* = 0.15) groups.

Discussion

Genotypes of *A. flavus* that do not produce aflatoxins exist in all warm agricultural areas (Cotty et al. 1994; Probst et al. 2011). These non-aflatoxigenics are used to reduce aflatoxin contamination of maize, cottonseed, groundnut, pistachio, and fig (Bandyopadhyay and Cotty 2013; Dorner 2008; Doster et al. 2008, 2014; Probst et al. 2014). Applications of non-aflatoxigenics reduce aflatoxins by altering compositions of fungal populations associated with crops (Antilla and Cotty 2002; Dorner 2004). As frequencies of non-aflatoxigenics increase, quantities of aflatoxins in the crop decreases (Cotty 1994). Non-aflatoxigenicity in *A. flavus* genotypes originate from deletions of aflatoxin-biosynthesis genes, including deletion of the entire cluster as in NRRL 21882 the active ingredient of afla-guard™ (NRRL 21882), or from a single inactivating SNP, as with the SNP induced stop codon in the *pks* gene of the biopesticide *A. flavus* AF36 (Ehrlich and Cotty 2004; Chang et al. 2005). However, processes through which aflatoxin-biosynthesis gene clusters of non-aflatoxigenics evolve and mechanisms behind cluster degeneration have not been adequately described for

Table 4 Correlations among SNP densities in three regions of the aflatoxin-biosynthesis and sugar clusters

Group	Gene	SNP density	Significance	<i>pksA</i> group ^a	Significance ^b	<i>cypX</i> group	Significance	<i>sugR</i> group	Significance
<i>sugR</i> group	<i>hxtA</i>	2		0.19	e	0.35	cde	0.7	ab
	<i>glcA</i>	4		0.15	e	0.01	e	0.55	a
	<i>sugR</i>	5		0.22	e	0.05	e	0.7	ab
	<i>Orf</i>	4		0.22	e	0.19	e	0.71	ab
	Average	4	AB	0.19	CD	0.15	D	0.66	B
<i>cypX</i> group	<i>ordA</i>	3		0.19	e	0.31	cde	0.35	cde
	<i>vbs</i>	7		0.27	de	0.34	cde	0.01	e
	<i>cypX</i>	8		0.27	de	0.41	bcd	0.05	e
	<i>moxY</i>	4		0.12	e	0.36	cde	0.19	e
	Average	6	A	0.21	CD	0.35	C	0.15	D
<i>pksA</i> group	<i>pksA</i>	2		0.87	a	0.19	e	0.19	e
	<i>nor-1</i>	0		0.83	a	0.27	de	0.15	e
	<i>hexA</i>	2		0.9	a	0.27	de	0.22	e
	<i>hexB</i>	3		0.82	a	0.12	e	0.22	e
	Average	2	B	0.85	A	0.21	CD	0.19	CD

^a Values are means of either 3 (within group comparisons) or 4 (between group comparisons) correlation coefficients. Self-correlations (only occur in within groups comparisons) are not included to avoid bias

^b Means followed by the same letter do not differ significantly (*P* < 0.05) by Tukey’s HSD test

many biopesticide active ingredients. By sequencing and comparatively analyzing complete aflatoxin biosynthesis gene clusters of 35 non-aflatoxigenic genotypes of interest as active ingredients of biopesticides for aflatoxin mitigation in Africa and North America, the current work sheds light on the process of cluster degeneration. The results suggest the cluster is not evolving as one unit but that portions are responding divergently to selective forces and the degeneration process is not similar for all non-aflatoxigenic genotypes.

Diversity of non-aflatoxigenic genotypes

Analyses of complete aflatoxin gene cluster sequences from 35 non-aflatoxigenic *A. flavus* genotypes from Africa and North America suggest 6 genotype groups based on cluster structure (Figs. 1, 2). Group A genotypes retain all aflatoxin biosynthesis genes and include the active ingredient of the USEPA registered biocontrol product *A. flavus* AF36 (AF36, Antilla and Cotty 2002). Group A genotypes have small (<1 kb) deletions (Table 2) and SNPs with potential to cause non-aflatoxigenicity. Frequently single genotypes have multiple genomic lesions sufficient to individually cause non-aflatoxigenicity suggesting phenotypic stability and long histories of the non-aflatoxigenicity. Groups B, C, D, and E consist of genotypes with varying levels of deletion, ranging from a few bases to most of the cluster. Group F is distinguished by absence of all aflatoxin biosynthesis genes resulting from a single group-wide deletion. Although deletions in aflatoxin gene clusters are known for *A. flavus* (Chang et al. 2005) and *A. oryzae* (Kusumoto et al. 2000), in the current study 35 genotypes were assembled against a reference sequence from an aflatoxin producer to reveal previously unknown diversity in both deletion pattern and mechanism of non-aflatoxigenicity. Genotypes within the same group have identical deletion patterns, including identical sequences bordering the deletions, suggesting common ancestry. Common ancestry is supported by the Neighbor-Net network based on 17 SSR loci (Grubisha and Cotty 2009) (Fig. 2). Clustering occurs among genotypes with similar deletion patterns but no clustering is seen based on genotype geographic origin reflecting both shared ancestry within each group and migration of groups across vast landscapes during intra-group divergence and after arisal of atoxigenicity.

Degeneration of aflatoxin gene cluster

A variety of mutations shape genes and genomes. Protein coding genes may be eliminated by large deletion and smaller deletions may render genes inactive. In agreement with previous reports (Callicott and Cotty 2015; Chang et al. 2005; Criseo et al. 2001; Mauro et al. 2013; Yin et al. 2009), our results provide strong evidence

that partial or complete gene deletion is an important mechanism causing non-aflatoxigenicity in *A. flavus* populations. The current work suggests deletion of aflatoxin-biosynthesis genes is an ongoing process with genotypes retaining both varying percentages of the aflatoxin cluster and multiple genetic defects sufficient to cause non-aflatoxigenicity. Shared deletion patterns (Table 2; Fig. 2) suggest different deletion events lead to non-aflatoxigenicity in founding genotypes for the various lineages. A general pattern among genotypes was high frequencies of deletions (small and large) at the telomeric end of the cluster. This is likely a reflection of the neighboring sub-telomeric region which has a high rate of genetic flux compared to other regions (Chang et al. 2005; Liti and Louis 2005; Maciaszczyk et al. 2004).

Unlike the genotypes with partial gene clusters where most deletions were predominantly >1-kb, genotypes with complete gene sets have several smaller deletions (<1-kb) (Table 2). The smaller deletions were dispersed throughout the cluster with *norB*, *cypA*, *pksA*, *aflR*, *verA* and *verB* having higher frequencies of deletions. It is not surprising that *pksA* and *aflR* are two of the six genes, which are present in only 49 % of the genotypes analyzed in this study (Table 3). High frequencies of smaller deletions and higher probabilities of deletions in early portions of the gene cluster suggest that these genes may have gone through a combined process of inactivation by small deletions followed by complete gene loss from large deletions. Portions of the gene cluster are not only evolving differently but also going through different processes of degeneration (e.g. deletion vs SNPs) further suggesting multiple events are driving the degeneration of aflatoxin gene clusters.

Differential evolution of aflatoxin gene cluster

Deletions in the gene cluster may not be sufficient to explain the variation among non-aflatoxigenic genotypes and mechanisms of non-aflatoxigenicity since genotypes with complete gene sets have no large deletions. SNP formation is an important mechanism through which non-aflatoxigenicity forms (Ehrlich and Cotty 2004). Three patterns are apparent in SNP distribution. First, with the exception of few genes (*norB*, *nor-1*), SNPs are common (Fig. 3; Additional file 1: Figure S1). Second, SNPs are not distributed randomly among genotypes. Third, late genes in the aflatoxin cluster have elevated SNP density as compared to early genes (Fig. 3). Not surprisingly, SNPs are also more abundant in non-aflatoxigenic genotypes than toxigenic genotypes as expected if purifying selection is maintaining toxigenicity. The observed variation in mutation in early and late portions of the cluster suggests the possibility of a differential evolution. Results from the current study reveal multiple forces driving evolution of

aflatoxin genes and diverse mechanisms for non-aflatoxigenicity. This adds to recent observations (Grubisha and Cotty 2015) that non-aflatoxigenic phenotypes can be highly stable and retained for long-periods.

The physical order of the genes in aflatoxin gene cluster is similar to the order of enzyme reactions catalyzed by their gene products (Roze et al. 2007; Trail et al. 1995). However, expression of genes required for early stages of aflatoxin biosynthesis is modulated differently than those from the later steps (Ehrlich et al. 1999; Roze et al. 2007; Schmidt-Heydt et al. 2009). In agreement with these findings, are the different rates of evolution between early and late portions of the cluster (Fig. 3). This may result from selection driving divergent changes across the cluster. This hypothesis is consistent with the observation that two genes (*cypA* and *ver-1*) in the late portion are under positive selection (Fig. 4). This suggests that in non-aflatoxigenic phenotypes, certain genes in the late portion of the cluster are evolving new functions allowed by relaxation of selection for aflatoxin production. At the same time, lack of functional constraints might have allowed deletion of early genes in many genotypes. Occurrence of deletions and other mutations at different rates in different portions of the cluster results in a complex process of aflatoxin biosynthesis loss.

Evolution of clustered genes

Genes for biosynthesis of secondary metabolites are typically clustered together (Ehrlich et al. 2005; Keller and Hohn 1997; Walton 2000; Yu et al. 2004b). Genes encoding aflatoxin pathway enzymes are highly coexpressed. However, strong correlation coefficients between genomic changes suggest that certain genes involved in early steps of aflatoxin biosynthesis and neighboring *pksA* (Table 4) are evolving together and separately from the remainder of the cluster. The three genes (*hexA*, *hexB* and *pksA*) in the *pksA* group are involved in the conversion of acetate to norsolorinic acid (*nor-1*) (Yu et al. 2004a). The *pksA* gene, which is divergently transcribed from a 1.5 kb intergenic region with *nor-1*, along with *hexA* and *hexB* is involved in the synthesis of a polyketide from the primary metabolite acetate (Brown et al. 1996). Similar to the *pksA* group, genes in *sugR* group are also evolving together. None of the genes in the sugar cluster are evolving with any of the aflatoxin genes suggesting differential evolution and adaptation of the two clusters (Table 4). Clusters of genes that work together to produce a product (e.g., a secondary metabolite) may show coordinated changes in evolutionary rates because of increased or decreased utilization of those genes over evolutionary time.

Implications for biological control of aflatoxin

Phenotypic variation among *A. flavus* genotypes confers differential adaptation to hosts, soils, and environmental

conditions (Garber and Cotty 2014; Mehl and Cotty 2013; Mehl et al. 2012). Biocontrol methods that decrease human exposure to aflatoxins by selection and application of native, well adapted, non-aflatoxigenic genotypes of *A. flavus* have been successfully used in diverse locations and on several crops (Atehnkeng et al. 2008, 2014; Cotty 1994; Cotty et al. 2007; Probst et al. 2011, 2014). This study provides insight into diversity of mechanisms through which non-aflatoxigenicity has evolved. Twenty-six of the thirty-five non-aflatoxigenic genotypes of *A. flavus* included in the current analyses are active ingredients in biopesticides developed for aflatoxin management. Ten of the genotypes are active ingredients in biopesticides already registered for use (AF36 in the USA for *A. flavus* AF36 and AF36 Prevail™; NRRL 21882 in the USA for afla-guard®, Ka16127, La3279, La3304, and Og0222 in Nigeria for Aflasafe, and C6-E, C8-F, E63-I, and R7-H in Kenya for Aflasafe KE01) in aflatoxin management programs. Sixteen genotypes (GO18-2, GO67-10, MO11-8, MO92-15, M102-11, M109-2, M110-7, M129-5, M2-7, M21-11, Ms14-19, Ss19-14 AT21-A, BA35-C, DO114-A, EC69-E) are active ingredients in different biopesticides at various stages of the development cycle. The current report is the first report of the mechanisms of non-aflatoxigenicity for all the included genotypes except AF36 (Ehrlich and Cotty 2004) and NRRL 21882 (Chang et al. 2005). Molecular characterization of mechanisms of non-aflatoxigenicity is sometimes required for full biopesticide registration. The rationale for this may be to verify the specific nature of the failure to produce aflatoxins or as a required tool for assessing the stability of non-aflatoxigenicity (Ehrlich and Cotty 2004; Grubisha and Cotty 2015). Sequences for aflatoxin biosynthesis gene clusters associated with the current report are also available to support tools for monitoring proportions of the specific non-aflatoxigenic genotypes in fungal populations associated with crops or in the environment, as well as, incidences of the specific mechanisms of non-aflatoxigenicity (Das et al. 2008; Mehl and Cotty 2010, 2013). Ten of the non-aflatoxigenic genotypes analyzed here are already being applied widely to commercial crops in the target nations (USA, Nigeria, or Kenya) and another sixteen are being applied to farmer's fields on a smaller scale during biopesticide development in Senegal, Burkina Faso, and the USA. Thus, needs for the reported sequences already exist.

The structure of the aflatoxin biosynthesis cluster differs significantly among non-aflatoxigenic genotypes (Fig. 1). Closely related genotypes belonging to different VCGs have identical deletion patterns suggesting clonal derivation from a common ancestor. Deletion patterns within each of the 6 groups (group A to F; Tables 1, 2; Fig. 1) likely formed before separation into different VCGs. Multiple deletions in the aflatoxin gene cluster

along with high frequencies of other mutations indicate origins of non-aflatoxigenicity in many lineages are old, sufficiently old to have allowed time, in some genes, for divergent adaptive evolution. In natural systems, multiple and diverse mechanisms of non-aflatoxigenicity in individual non-aflatoxigenic genotypes indicate long-term persistence and stability of non-aflatoxigenicity that extends over many thousands of years (Grubisha and Cotty 2010, 2015). In each of the non-aflatoxigenic gene clusters examined in the current study, once established, cluster non-aflatoxigenicity was maintained over sufficiently long periods for additional degeneration in the aflatoxin gene cluster to form. The multiple mechanisms of non-aflatoxigenicity suggest the clusters are not vulnerable to conversion to aflatoxin-producing phenotype and the *A. flavus* genotypes harboring these clusters cannot convert to aflatoxin-producers short of acquisition of an intact aflatoxin-biosynthesis gene cluster through either horizontal gene transfer or independent assortment of chromosomes during meiosis. Loci on all 8 chromosomes of *A. flavus* are in linkage disequilibrium with each other indicating that independent assortment of chromosomes does not occur among either aflatoxin producers (Grubisha and Cotty 2010) or among mixed populations of aflatoxin-producing and non-aflatoxigenic *A. flavus* (Grubisha and Cotty 2010, 2015; Ortega-Beltran et al. 2016). Although, simple mutations can cause non-aflatoxigenicity at any time in otherwise aflatoxin-producing lineages and such non-aflatoxigenics may revert back to aflatoxin-producer with a similar mutation, this is not the case for the non-aflatoxigenic genotypes used as active ingredients in biocontrol products. Non-aflatoxigenic active ingredients are selected from VCGs and/or SSR haplotypes that are relatively common and widely distributed in the target agroecosystem and which contain no aflatoxin-producing members (Atehnkeng et al. 2014; Bandyopadhyay and Cotty 2013; Cotty 2006; Probst et al. 2011). As revealed in the current report, these genotypes have numerous lesions in the aflatoxin biosynthesis gene cluster that have formed over long periods of non-aflatoxigenicity and are not vulnerable to conversion to aflatoxin-producing phenotype.

Additional file

Additional file 1. Table S1. Validation of small deletions in aflatoxin gene clusters. **Table S2.** Validation of SNPs from aflatoxin biosynthesis gene clusters. **Table S3.** Annotation of SNPs from aflatoxin biosynthesis gene clusters. **Figure S1.** Heat map of SNP density in partial sets of genes in aflatoxin biosynthesis gene clusters.

Abbreviations

SNP: Single Nucleotide Polymorphism; VCG: Vegetative Compatibility Group; USDA: United States Department of Agriculture; ARS: Agricultural Research

Service; AGI: Arizona Genomics Institute; SSR: Simple Sequence Repeat; USEPA: United States Environmental Protection Agency; ANOVA: Analysis of Variance.

Authors' contributions

BA, RB and PC contributed to the conception and design of the experiments. BA performed the experiments, analyzed the data and drafted the manuscript. All authors read and approved the final manuscript.

Author details

¹ USDA-ARS, The University of Arizona, School of Plant Sciences, 303 Forbes Building, P.O. Box 210036, Tucson, AZ 85721, USA. ² International Institute of Tropical Agriculture (IITA), PMB 5320, Oyo Road, Ibadan, Nigeria.

Acknowledgements

This work was supported by Agricultural Research Service, US Department of Agriculture, CRIS project 5347-42000-019-00D. 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 US Department of Agriculture.

Competing interests

The authors declare that they have no competing interests.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Funding

The funding for this work was provided by Agricultural Research Service, US Department of Agriculture, CRIS project 5347-42000-019-00D.

Received: 27 July 2016 Accepted: 12 August 2016

Published online: 31 August 2016

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol.* 1990;215(3):403–10. doi:10.1006/jmbi.1990.9999.
- Antilla L, Cotty PJ. The ARS-ACRPC partnership to control aflatoxin in Arizona cotton: current status. *Mycopathologia.* 2002;155:64.
- Atehnkeng J, Ojiambo PS, Cotty PJ, Bandyopadhyay R. Field efficacy of a mixture of atoxigenic *Aspergillus flavus* Link: Fr vegetative compatibility groups in preventing aflatoxin contamination in maize (*Zea mays* L.). *Biol Control.* 2014;72:62–70. doi:10.1016/j.biocontrol.2014.02.009.
- Atehnkeng J, Ojiambo PS, Ikotun T, Sikora RA, Cotty PJ, Bandyopadhyay R. Evaluation of atoxigenic isolates of *Aspergillus flavus* as potential biocontrol agents for aflatoxin in maize. *Food Addit Contam A.* 2008;25(10):1264–71. doi:10.1080/02652030802112635.
- Bandyopadhyay R, Cotty PJ. Biological controls for aflatoxin reduction, vol 2020 Vision Focus 20(16). Washington DC: International Food Policy Research Institute (IFPRI); 2013.
- Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, Antunes R, Ar-Ganiska J, Bely B, Bingley M, Bonilla C, Britto R, Bursteinas B, Chavali G, Cibrán-Uhalte E, Da Silva A, De Giorgi M, Dogan T, Fazzini F, Gane P, Cas-Tro LG, Garmiri P, Hatton-Ellis E, Hieta R, Huntley R, Legge D, Liu WD, Luo J, MacDougall A, Mutowo P, Nightin-Gale A, Orchard S, Pichler K, Poggioli D, Pundir S, Pura L, Qi GY, Rosanoff S, Saidi R, Sawford T, Shypitsyna A, Turner E, Volynkin V, Wardell T, Watkins X, Watkins Cowley A, Figueira L, Li WZ, McWilliam H, Lopez R, Xenarios I, Bougueleret L, Bridge A, Poux S, Redaschi N, Aimo L, Argoud-Puy G, Auchincloss A, Axelsen K, Bansal P, Baratin D, Blatter MC, Boeckmann B, Bolleman J, Boutet E, Breuza L, Casal-Casas C, De Castro E, Coudert E, Cucho B, Doche M, Dornevil D, Duvaud S, Estreicher A, Famiglietti L, Feuermann M, Gasteiger E, Gehant S, Gerritsen V, Gos A, Gruaz-Gumowski N, Hinz U, Hulo C, Junco F, Keller G, Lara V, Lemercier P, Lieberherr D, Lombardot T, Martin X, Masson P, Morgat A, Neto T, Noupikel N, Paesano S, Pedruzzi I, Pilbout S, Pozzato M, Pruess M, Rivoire C, Roehert B, Schneider M, Sigrist C, Sonesson K, Staehli S, Stutz A, Sundaram S, Tognolli M, Verbregue L, Veuthey AL, Wu CH, Arighi CN, Arminski L, Chen CM, Chen YX, Garavelli JS, Huang HZ, Laiho KT,

- McGarvey P, Natale DA, Suzek BE, Vinayaka CR, Wang QH, Wang YQ, Yeh LS, Yerramalla MS, Zhang J, Consortium U. UniProt: a hub for protein information. *Nucleic Acids Res.* 2015;43(D1):D204–12. doi:10.1093/nar/gku989.
- Brown DW, Adams TH, Keller NP. *Aspergillus* has distinct fatty acid syntheses for primary and secondary metabolism. *Proc Natl Acad Sci USA.* 1996;93(25):14873–7.
- Brown RL, Chen ZY, Cleveland TE, Russin JS. Advances in the Development of Host Resistance in Corn to Aflatoxin Contamination by *Aspergillus flavus*. *Phytopathology.* 1999;89(2):113–7. doi:10.1094/PHYTO.1999.89.2.113.
- Callicott KA, 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. doi:10.1111/lam.12337.
- Calvo AM, Bok J, Brooks W, Keller NP. veA is required for toxin and sclerotial production in *Aspergillus parasiticus*. *Appl Environ Microbiol.* 2004;70(8):4733–9. doi:10.1128/AEM.70.8.4733-4739.2004.
- Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Alvarado AS, Yandell M. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 2008;18(1):188–96. doi:10.1101/gr.6743907.
- Chang PK, Cary JW, Yu JJ, Bhatnagar D, Cleveland TE. The *Aspergillus parasiticus* polyketide synthase gene Pksa, a homolog of *Aspergillus nidulans* Wa, is required for aflatoxin B-1 biosynthesis. *Mol Gen Genet.* 1995;248(3):270–7. doi:10.1007/Bf02191593.
- Chang PK, Horn BW, Dörner 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. doi:10.1016/j.fgb.2005.07.004.
- Chulze SN, Palazzini JM, Torres AM, Barros G, Ponsone ML, Geisen R, Schmidt-Heydt M, Kohl J. Biological control as a strategy to reduce the impact of mycotoxins in peanuts, grapes and cereals in Argentina. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* 2015;32(4):471–9. doi:10.1080/19440049.2014.984245.
- Cotty PJ. Virulence and cultural characteristics of 2 *Aspergillus flavus* strains pathogenic on cotton. *Phytopathology.* 1989;79(7):808–14. doi:10.1094/Phyto-79-808.
- Cotty PJ. Influence of field application of an atoxigenic strain of *Aspergillus flavus* on the populations of *Aspergillus flavus* infecting cotton bolls and on the aflatoxin content of cottonseed. *Phytopathology.* 1994;84(11):1270–7. doi:10.1094/Phyto-84-1270.
- Cotty PJ. Aflatoxin-producing potential of communities of *Aspergillus* section Flavi from cotton producing areas in the United States. *Mycol Res.* 1997;101:698–704. doi:10.1017/S0953756296003139.
- Cotty PJ. Biocompetitive exclusion of toxigenic fungi. Wageningen: Wageningen Academic Publishers; 2006.
- Cotty PJ, Antilla L, Wakelyn PJ. Competitive exclusion of aflatoxin producers: farmer-driven research and development. *Biological control: a global perspective*; 2007. p. 241–53. doi:10.1079/9781845932657.0241.
- Cotty PJ, Bayman P, Egel DS, Elias KS. Agriculture, aflatoxins and *Aspergillus*. In: Keith AP, editor. *The genus Aspergillus*. New York: Plenum Press; 1994. p. 1–27.
- Cotty PJ, Cardwell KF. Divergence of West African and North American communities of *Aspergillus* section Flavi. *Appl Environ Microbiol.* 1999;65(5):2264–6.
- Cotty PJ, Mellon JE. Ecology of aflatoxin producing fungi and biocontrol of aflatoxin contamination. *Mycotoxin Res.* 2006;22(2):110–7. doi:10.1007/BF02956774.
- Criseo G, Bagnara A, Bisignano G. Differentiation of aflatoxin-producing and non-producing strains of *Aspergillus flavus* group. *Lett Appl Microbiol.* 2001;33(4):291–5.
- Das MK, Ehrlich KC, Cotty PJ. Use of pyrosequencing to quantify incidence of a specific *Aspergillus flavus* strain within complex fungal communities associated with commercial cotton crops. *Phytopathology.* 2008;98(3):282–8. doi:10.1094/PHYTO-98-3-0282.
- Delcher AL, Phillippy A, Carlton J, Salzberg SL. Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res.* 2002;30(11):2478–83. doi:10.1093/nar/30.11.2478.
- Donner M, Atehnkeng J, Sikora RA, Bandyopadhyay R, Cotty PJ. Molecular characterization of atoxigenic strains for biological control of aflatoxins in Nigeria. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* 2010;27(5):576–90. doi:10.1080/19440040903551954.
- Dörner JW. Biological control of aflatoxin contamination of crops. *J Toxicol-Toxin Rev.* 2004;23(2–3):425–50. doi:10.1081/Txr-200027877.
- Dörner JW. Management and prevention of mycotoxins in peanuts. *Food Addit Contam.* 2008;25(2):203–8. doi:10.1080/02652030701658357.
- Dörner JW. Biological control of aflatoxin contamination in corn using a non-toxicogenic strain of *Aspergillus flavus*. *J Food Prot.* 2009;72(4):801–4.
- Dörner JW, Horn BW. Separate and combined applications of nontoxicogenic *Aspergillus flavus* and *A. parasiticus* for biocontrol of aflatoxin in peanuts. *Mycopathologia.* 2007;163(4):215–23. doi:10.1007/s11046-007-9004-0.
- Doster MA, Cotty PJ, Michailides TJ. Evaluation of the atoxigenic *Aspergillus flavus* Strain AF36 in pistachio orchards. *Plant Dis.* 2014;98(7):948–56. doi:10.1094/Pdis-10-13-1053-Re.
- Doster MA, Michailides TJ. *Aspergillus* molds and aflatoxins in pistachio nuts in California. *Phytopathology.* 1994;84:583–90.
- Doster MA, Michailides TJ, Doster L, Cotty PJ. Biocontrol of aflatoxin in Figs. *Acta Hort.* 2008;798:223–6. doi:10.17660/ActaHortic.798.31.
- Ehrlich KC, Chang PK, Yu J, Cotty PJ. Aflatoxin biosynthesis cluster gene cypA is required for G aflatoxin formation. *Appl Environ Microbiol.* 2004;70(11):6518–24. doi:10.1128/AEM.70.11.6518-6524.2004.
- Ehrlich KC, Cotty PJ. An isolate of *Aspergillus flavus* used to reduce aflatoxin contamination in cottonseed has a defective polyketide synthase gene. *Appl Microbiol Biotechnol.* 2004;65(4):473–8. doi:10.1007/s00253-004-1670-y.
- Ehrlich KC, Montalbano BG, Cary JW. Binding of the C6-zinc cluster protein, AFLR, to the promoters of aflatoxin pathway biosynthesis genes in *Aspergillus parasiticus*. *Gene.* 1999;230(2):249–57.
- Ehrlich KC, Montalbano BG, Cotty PJ. Analysis of single nucleotide polymorphisms in three genes shows evidence for genetic isolation of certain *Aspergillus flavus* vegetative compatibility groups. *FEMS Microbiol Lett.* 2007;268(2):231–6. doi:10.1111/j.1574-6968.2006.00588.x.
- Ehrlich KC, Yu J, Cotty PJ. Aflatoxin biosynthesis gene clusters and flanking regions. *J Appl Microbiol.* 2005;99(3):518–27. doi:10.1111/j.1365-2672.2005.02637.x.
- Fraser HB, Hirsh AE, Wall DP, Eisen MB. Coevolution of gene expression among interacting proteins. *P Natl Acad Sci USA.* 2004;101(24):9033–8. doi:10.1073/pnas.0402591101.
- Garber NP, Cotty PJ. *Aspergillus parasiticus* communities associated with sugarcane in the Rio Grande Valley of Texas: implications of global transport and host association within *Aspergillus* section Flavi. *Phytopathology.* 2014;104(5):462–71. doi:10.1094/PHYTO-04-13-0108-R.
- Grubisha LC, Cotty PJ. Twenty-four microsatellite markers for the aflatoxin-producing fungus *Aspergillus flavus*. *Mol Ecol Resour.* 2009;9(1):264–7. doi:10.1111/j.1755-0998.2008.02378.x.
- Grubisha LC, Cotty PJ. Evolutionary relationships among *Aspergillus flavus* vegetative compatibility groups. *Phytopathology.* 2010;100(6):S43–4.
- Grubisha LC, Cotty PJ. Genetic analysis of the *Aspergillus flavus* Vegetative Compatibility Group to which a biological control agent that limits aflatoxin contamination in USA crops belongs. *Appl Environ Microbiol.* 2015;81(17):5889–99. doi:10.1128/AEM.00738-15.
- Hedayati MT, Pasqualotto AC, Warn PA, Bowyer P, Denning DW. *Aspergillus flavus*: human pathogen, allergen and mycotoxin producer. *Microbiology.* 2007;153(Pt 6):1677–92. doi:10.1099/mic.0.2007/007641-0.
- Huson DH, Bryant D. Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol.* 2006;23(2):254–67. doi:10.1093/molbev/msj030.
- Joffe AZ. Aflatoxin produced by 1626 isolates of *Aspergillus flavus* from groundnut kernels and soils in Israel. *Nature.* 1969;221(5179):492. doi:10.1038/221492a0.
- Jolley KA, Feil EJ, Chan MS, Maiden MCJ. Sequence type analysis and recombinational tests (START). *Bioinformatics.* 2001;17(12):1230–1. doi:10.1093/bioinformatics/17.12.1230.
- Keller NP, Hohn TM. Metabolic pathway gene clusters in filamentous fungi. *Fungal Genet Biol.* 1997;21(1):17–29.
- Klich MA. Biogeography of *Aspergillus* species in soil and litter. *Mycologia.* 2002;94(1):21–7. doi:10.2307/3761842.
- Klich MA. *Aspergillus flavus*: the major producer of aflatoxin. *Mol Plant Pathol.* 2007;8(6):713–22. doi:10.1111/j.1364-3703.2007.00436.x.
- Klich MA, Yu J, Chang PK, Mullaney EJ, Bhatnagar D, Cleveland TE. Hybridization of genes involved in aflatoxin biosynthesis to DNA of aflatoxigenic and non-aflatoxigenic *Aspergillus*. *Appl Microbiol Biotechnol.* 1995;44(3–4):439–43.

- Kusumoto K, Nogata Y, Ohta H. Directed deletions in the aflatoxin biosynthesis gene homolog cluster of *Aspergillus oryzae*. *Curr Genet*. 2000;37(2):104–11. doi:10.1007/s002940050016.
- Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):1. doi:10.1186/gb-2009-10-3-r25.
- Leger RJ, Screen SE, Shams-Pirzadeh B. Lack of host specialization in *Aspergillus flavus*. *Appl Environ Microbiol*. 2000;66(1):320–4.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc GPD. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9. doi:10.1093/bioinformatics/btp352.
- Li H, Ruan J, Durbin R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res*. 2008;18(11):1851–8. doi:10.1101/gr.078212.108.
- Lisker N, Michaeli R, Frank ZR. Mycotoxigenic potential of *Aspergillus flavus* strains isolated from groundnuts growing in Israel. *Mycopathologia*. 1993;122(3):177–83.
- Liti G, Louis EJ. Yeast evolution and comparative genomics. *Annu Rev Microbiol*. 2005;59:135–53. doi:10.1146/annurev.micro.59.030804.121400.
- Machida M, Asai K, Sano M, Tanaka T, Kumagai T, Terai G, Kusumoto K, Arima T, Akita O, Kashiwagi Y, Abe K, Gomi K, Horiuchi H, Kitamoto K, Kobayashi T, Takeuchi M, Denning DW, Galagan JE, Nierman WC, Yu J, Archer DB, Bennett JW, Bhatnagar D, Cleveland TE, Fedorova ND, Gotoh O, Horikawa H, Hosoyama A, Ichinomiya M, Igarashi R, Iwashita K, Juvvadi PR, Kato M, Kato Y, Kin T, Kokubun A, Maeda H, Maeyama N, Maruyama J, Nagasaki H, Nakajima T, Oda K, Okada K, Paulsen I, Sakamoto K, Sawano T, Takahashi M, Takase K, Terabayashi Y, Wortman JR, Yamada O, Yamagata Y, Anazawa H, Hata Y, Koide Y, Komori T, Koyama Y, Minetoki T, Suharnan S, Tanaka A, Isono K, Kuhara S, Ogasawara N, Kikuchi H. Genome sequencing and analysis of *Aspergillus oryzae*. *Nature*. 2005;438(7071):1157–61. doi:10.1038/nature04300.
- Maciaszczyk E, Wysocki R, Golik P, Lazowska J, Ulaszewski S. Arsenical resistance genes in *Saccharomyces douglasii* and other yeast species undergo rapid evolution involving genomic rearrangements and duplications. *FEMS Yeast Res*. 2004;4(8):821–32. doi:10.1016/j.femsyr.2004.03.002.
- Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;17:10–2. doi:10.14806/ej.17.1.200.
- Mauro A, Battilani P, Callicott KA, Giorni P, Pietri A, Cotty PJ. Structure of an *Aspergillus flavus* population from maize kernels in northern Italy. *Int J Food Microbiol*. 2013;162(1):1–7. doi:10.1016/j.ijfoodmicro.2012.12.021.
- McKean C, Tang L, Billam M, Tang M, Theodorakis CW, Kendall RJ, Wang JS. Comparative acute and combinative toxicity of aflatoxin B1 and T-2 toxin in animals and immortalized human cell lines. *J Appl Toxicol*. 2006;26(2):139–47. doi:10.1002/jat.1117.
- Mehl HL, Cotty PJ. Variation in competitive ability among isolates of *Aspergillus flavus* from different vegetative compatibility groups during maize infection. *Phytopathology*. 2010;100(2):150–9. doi:10.1094/Phyto-100-2-0150.
- Mehl HL, Cotty PJ. Nutrient environments influence competition among *Aspergillus flavus* genotypes. *Appl Environ Microbiol*. 2013;79(5):1473–80. doi:10.1128/AEM.02970-12.
- Mehl HL, Jaime R, Callicott KA, Probst C, Garber NP, Ortega-Beltran A, Grubisha LC, Cotty PJ. *Aspergillus flavus* diversity on crops and in the environment can be exploited to reduce aflatoxin exposure and improve health. *Ann NY Acad Sci*. 2012;1273:7–17. doi:10.1111/j.1749-6632.2012.06800.x.
- Ortega-Beltran A, Grubisha LC, Callicott KA, Cotty PJ. The vegetative compatibility group to which the US biocontrol agent *Aspergillus flavus* AF36 belongs is also endemic to Mexico. *J Appl Microbiol*. 2016;120:986–98.
- Prieto R, Yousibova GL, Woloshuk CP. Identification of aflatoxin biosynthesis genes by genetic complementation in an *Aspergillus flavus* mutant lacking the aflatoxin gene cluster. *Appl Environ Microbiol*. 1996;62(10):3567–71.
- Probst C, Bandyopadhyay R, Cotty PJ. Diversity of aflatoxin-producing fungi and their impact on food safety in sub-Saharan Africa. *Int J Food Microbiol*. 2014;174:113–22. doi:10.1016/j.ijfoodmicro.2013.12.010.
- Probst C, Bandyopadhyay R, Price LE, Cotty PJ. Identification of atoxigenic *Aspergillus flavus* isolates to reduce aflatoxin contamination of maize in Kenya. *Plant Dis*. 2011;95(2):212–8. doi:10.1094/Pdis-06-10-0438.
- Probst C, Schulthess F, Cotty PJ. Impact of *Aspergillus* section Flavi community structure on the development of lethal levels of aflatoxins in Kenyan maize (*Zea mays*). *J App Microbiol*. 2010;108(2):600–10. doi:10.1111/j.1365-2672.2009.04458.x.
- Rausch T, Zichner T, Schlattl A, Stutz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*. 2012;28(18):1333–9. doi:10.1093/bioinformatics/bts378.
- Robens J, Cardwell K. The costs of mycotoxin management to the USA: management of aflatoxins in the United States. *J Toxicol Toxin Rev*. 2003;22(2–3):139–52. doi:10.1081/Txr-120024089.
- Roze LV, Arthur AE, Hong SY, Chanda A, Linz JE. The initiation and pattern of spread of histone H4 acetylation parallel the order of transcriptional activation of genes in the aflatoxin cluster. *Mol Microbiol*. 2007;66(3):713–26. doi:10.1111/j.1365-2958.2007.05952.x.
- Schmidt-Heydt M, Abdel-Hadi A, Magan N, Geisen R. Complex regulation of the aflatoxin biosynthesis gene cluster of *Aspergillus flavus* in relation to various combinations of water activity and temperature. *Int J Food Microbiol*. 2009;135(3):231–7. doi:10.1016/j.ijfoodmicro.2009.07.026.
- Schroeder HW, Boller RA. Aflatoxin production of species and strains of the *Aspergillus flavus* group isolated from field crops. *Appl Microbiol*. 1973;25(6):885–9.
- Sepahvand A, Shams-Ghahfarokhi M, Allameh A, Jahanshiri Z, Jamali M, Razzaghi-Abyaneh M. A survey on distribution and toxigenicity of *Aspergillus flavus* from indoor and outdoor hospital environments. *Folia Microbiol (Praha)*. 2011;56(6):527–34. doi:10.1007/s12223-011-0078-1.
- Thompson JD, Higgins DG, Gibson TJ. Clustal-W—improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994;22(22):4673–80. doi:10.1093/nar/22.22.4673.
- Tominaga M, Lee YH, Hayashi R, Suzuki Y, Yamada O, Sakamoto K, Gotoh K, Akita O. Molecular analysis of an inactive aflatoxin biosynthesis gene cluster in *Aspergillus oryzae* RIB strains. *Appl Environ Microbiol*. 2006;72(1):484–90. doi:10.1128/Aem.72.1.484-490.2006.
- Trail F, Mahanti N, Rarick M, Mehig R, Liang SH, Zhou R, Linz JE. 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.
- van Egmond HP, Schothorst RC, Jonker MA. Regulations relating to mycotoxins in food: perspectives in a global and European context. *Anal Bioanal Chem*. 2007;389(1):147–57. doi:10.1007/s00216-007-1317-9.
- Walton JD. Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis. *Fungal Genet Biol*. 2000;30(3):167–71. doi:10.1006/fgbi.2000.1224.
- Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164. doi:10.1093/nar/gkq603.
- Wei DL, Jong SC. Production of aflatoxins by strains of the *Aspergillus flavus* group maintained in ATCC. *Mycopathologia*. 1986;93(1):19–24. doi:10.1007/Bf00437010.
- Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007;24(8):1586–91. doi:10.1093/molbev/msm088.
- Yin Y, Lou T, Yan L, Michailides TJ, Ma Z. Molecular characterization of toxigenic and atoxigenic *Aspergillus flavus* isolates, collected from peanut fields in China. *J Appl Microbiol*. 2009;107(6):1857–65. doi:10.1111/j.1365-2672.2009.04356.x.
- Yu J, Bhatnagar D, Cleveland TE. Completed sequence of aflatoxin pathway gene cluster in *Aspergillus parasiticus*. *FEBS Lett*. 2004a;564(1–2):126–30. doi:10.1016/S0014-5793(04)00327-8.
- Yu J, Chang PK, Ehrlich KC, Cary JW, Bhatnagar D, Cleveland TE, Payne GA, Linz JE, Woloshuk CP, Bennett JW. Clustered pathway genes in aflatoxin biosynthesis. *Appl Environ Microbiol*. 2004b;70(3):1253–62.
- Yu JJ, Woloshuk CP, Bhatnagar D, Cleveland TE. Cloning and characterization of avfA and omtB genes involved in aflatoxin biosynthesis in three *Aspergillus* species. *Gene*. 2000;248(1–2):157–67. doi:10.1016/S0378-1119(00)00126-8.
- Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 2008;18(5):821–9. doi:10.1101/gr.074492.107.