Mycopathologia (2006) 162: 167–177
DOI 10.1007/s11046-006-0051-8

Aflatoxigenicity in *Aspergillus*: molecular genetics, phylogenetic relationships and evolutionary implications

Jeffrey W. Cary & Kenneth C. Ehrlich
USDA, ARS, Southern Regional Research Center, New Orleans, LA, 70124, USA

Abstract

Aflatoxins (AFs) are toxic and carcinogenic secondary metabolites produced by isolates of *Aspergillus* section *Flavi* as well as a number of *Aspergillus* isolates that are classified outside of section *Flavi*. Characterization of the AF and sterigmatocystin (ST) gene clusters and analysis of factors governing regulation of their biosynthesis has resulted in these two mycotoxins being the most extensively studied of fungal secondary metabolites. This wealth of information has allowed the determination of the molecular basis for non-production of AF in natural isolates of *A. flavus* and domesticated strains of *A. oryzae*. This review provides an overview of the molecular analysis of the AF and ST gene clusters as well as new information on an AF gene cluster identified in the non-section *Flavi* isolate, *Aspergillus ochraceoroseus*. Additionally, molecular phylogenetic analysis using AF biosynthetic gene sequences as well as ribosomal DNA internal transcribed spacer (ITS) sequences between various section *Flavi* and non-section *Flavi* species has enabled determination of the probable evolutionary history of the AF and ST gene clusters. A model for the evolution of the AF and ST gene clusters as well as possible biological roles for AF are discussed.

Key words: aflatoxin, *Aspergillus flavus*, *Aspergillus ochraceoroseus*, gene cluster, secondary metabolite, sterigmatocystin

Introduction

The goal of this review is to summarize our current knowledge of the various genetic factors in *Aspergillus* section *Flavi* and non-section *Flavi* species that govern aflatoxigenicity. In this review we will: (i) compare the genetic composition and physical structure of the aflatoxin (AF) and sterigmatocystin (ST) biosynthetic gene clusters; (ii) analyze the phylogenetic relationships among AF- and ST-producing isolates; and (iii) provide a model for the evolution of the AF and ST gene clusters and their role in the biology of the fungus.

Both section *Flavi* isolates and isolates outside of section *Flavi* produce AF [1–4] (Table 1). Normally, there is no accumulation of ST by AF-producing isolates from *Aspergillus* section *Flavi* [2]. However, non-section *Flavi* isolates also accumulate ST. Two non-section *Flavi* isolates (an undescribed *Aspergillus* sp. SRRC 1468 and *E. venezuelensis* SRRC 2520) produced levels of AF comparable to levels of AF produced by some section *Flavi* isolates (Table 1). Different media were used to grow section *Flavi* and non-section *Flavi* isolates to achieve maximal AF production levels shown in Table 1. When grown under the same conditions as the non-section *Flavi* isolates, *A. parasiticus* NRRL2999 produces 200,000 ng/g AF, an amount only 2-fold more than that of *Aspergillus* SRRC 1468. Compared to section *Flavi* AF-producers, production of AF by most non-section *Flavi* isolates is relatively low [1–3].

The non-section *Flavi* isolates are in two different subgenera. *A. ochraceoroseus*, is in section *Circumdati* [5], and *Emericella astellata* and *E. venezuelensis*, like *A. nidulans*, which produces

Table 1. Biogeography and mycotoxin production

	Non-section *Flavi*				Section *Flavi*				
	Aspergillus sp. SRRC 1468	*A. ochraceoroseus* SRRC 1432	*E. venezuelensis* SRRC 2520	*E. astellata* SRRC 503	*A. nomius* NRRL13137	*A. parasiticus* NRRL2999	*A. flavus* AF13	*A. flavus* AF70	*A. pseudotamarii* NRRL443
Biogeography									
Location	Tai National Forest, Ivory Coast	Tai National Forest, Ivory Coast	Venezuela	Galapagos Islands	United States	Uganda	United States	United States	Japan
Habitat	Rain forest soil	Rain forest soil	Porifera on Mangrove	Dead leaf	Wheat	Peanuts	Cotton soil	Cotton soil	Tea field
Mycotoxins[a]									
Aflatoxin B_1	100000	40	51000	580	800000	4350000	270000	550000	11000
Aflatoxin G_1	None	None	None	None	800000	850000	None	None	None
Sterigmatocystin	36	20	60	60	ND[b]	ND	ND	ND	ND

[a]Determined by HPLC or TLC.
[b]ND indicates not detected.

the AF precursor, ST, are in section *Nidulantes*. These non-section *Flavi* species are associated with woody plant biomes located near coastal areas in the tropics. *E. venezuelensis* was isolated from sponge in mangrove surface water and is the first organism of marine origin reported to produce AF. None of the above mentioned non-section *Flavi* AF producers has been associated with agricultural crops or foods.

A. flavus, a member of the subgenera, section *Flavi*, is the most common species implicated in preharvest contamination of crops with AF [6]. *A. flavus* populations are genetically and phenotypically diverse [7] with some isolates producing abundant conidiospores, large (L) sclerotia, and variable amounts of AF, while another type produces abundant, small (S) sclerotia, fewer conidiospores and high levels of AF [6, 8]. A related type, *A. oryzae* is atoxigenic and has been used as a source of industrial enzymes and as a koji (starter) mold for Asian fermented foods, such as sake, miso, and soy sauce [9]. Other AF-producing section *Flavi* species, *A. nomius* [10], *A. bombycis*, *A. parasiticus*, an unnamed species related to *A. parasiticus* [6], and *A. pseudotamarii* [11] have rarely been associated with crop contaminations. These species are frequently found in soil and have been isolated from insect debris [12]. When soils from non-agricultural areas were examined, these species were found to be more abundant than they appear to be in agricultural soils (Ehrlich and Cotty, unpublished results) [13].

Although studies have not convincingly shown that AF is required for growth and fungal survival, the genes for AF production have persisted for more than 100 million years (My) (see below). There are many theories explaining why fungi make AFs [14, 15]. Among these are: AF and its precursors are a defense response by fungi to stress; they provide protection from UV damage; they are by-products of primary metabolism; they are virulence factors; they increase fungal fitness [16, 17]; and they provide protection from predators for reproductive structures such as conidia and sclerotia.

Toxin synthesis gene clusters

Homology

AF and ST biosynthesis require as many as 25 enzymes and two regulatory proteins encoded by

contiguous genes in a 70-kb cluster (Figure 1). The AF gene clusters of S- and L-type *A. flavus* are >99% identical. The *A. flavus* cluster is 96% identical to that of *A. parasiticus* and 91% identical to that of *A. nomius*. Coding regions generally have 4–10% higher sequence identity than intergenic regions. Among protein homologs encoded by the different gene clusters, the predicted proteins differ somewhat from one another in their degree of sequence conservation. For example, AflR, AflJ, AvfA, OrdB, and HypA in *A. nomius* have 5–10% lower amino acid sequence identity than Vbs, AflT, NorA, and Ver-1. The relative differences in amino acid identity of proteins encoded by genes in the AF cluster to proteins encoded by genes in the ST cluster was even greater. OmtB, Vbs, VerB, and Ver-1 have almost 25% higher amino acid identity than their ST cluster-encoded homologs compared to the proteins AflR, AflJ, HexA, HexB and EstA. The closest *A. nidulans* homologs for proteins required for conversion of ST to AF, CypA, OmtA and OrdA, were proteins encoded by genes outside of the ST cluster locus. These homologs have lower than 41% amino acid identities.

The introns in AF and ST cluster genes showed typical intron–exon boundary regions for fungal splice sites [18] with nucleotides GTR being at the first three positions of the intron and YAG at the last. The average intron size is 58 bp. Intron numbers and positions in genes in the AF clusters of section *Flavi* isolates are completely conserved, but differ considerable from homologous genes in section *Nidulantes* isolates [19]. For the eight genes analyzed in *A. ochraceoroseus*, the number of introns matched that found in the *A. flavus* and *A. nidulans* gene homologs with the exception of *aflJ, estA*/*stcI*, and *vbs*/*stcN* (Table 2).

Select genes from the *A. ochraceoroseus* AF gene cluster were compared to their homologs in the *A. flavus* and *A. nidulans* AF and ST gene clusters respectively (Table 2). The *A. ochraceoroseus aflR*, *aflJ*, *nor-1*/*stcE*, and *estA*/*stcI* genes had a lower level of nucleotide and amino acid identity to their *A. flavus* and *A. nidulans* homologs than did the genes, *ver-1*/*stcU*, *verb*/*stcL*, *omtB*/*stcP*, and *vbs*/*stcN*. This result is similar to that found for the comparison of nucleotide and deduced amino acid identities for these genes in section *Flavi* species and *A. nidulans* [20]. For *aflR*, *aflJ*, *nor-1*/*stcE*, and *estA*/*stcI*, identity between *A. ochraceoroseus* and *A. nidulans* was higher than that observed between *A. ochraceoroseus* and *A. flavus*. This is in agreement with molecular phylogenetic analyses that showed a closer relationship of *A. ochraceoroseus* to *A. nidulans* than to *A. parasiticus* (see below) [1, 4]. As observed for sequence identity between the *A. flavus* and *A. nidulans ver-1*/*stcU*, *verb*/*stcL*, *omtB*/*stcP*, and *vbs*/*stcN* genes, a higher level of identity was seen between *A. ochraceoroseus* and *A. flavus* and *A. nidulans* compared to the four low-identity genes. This is especially true for the deduced protein sequences in which amino acid

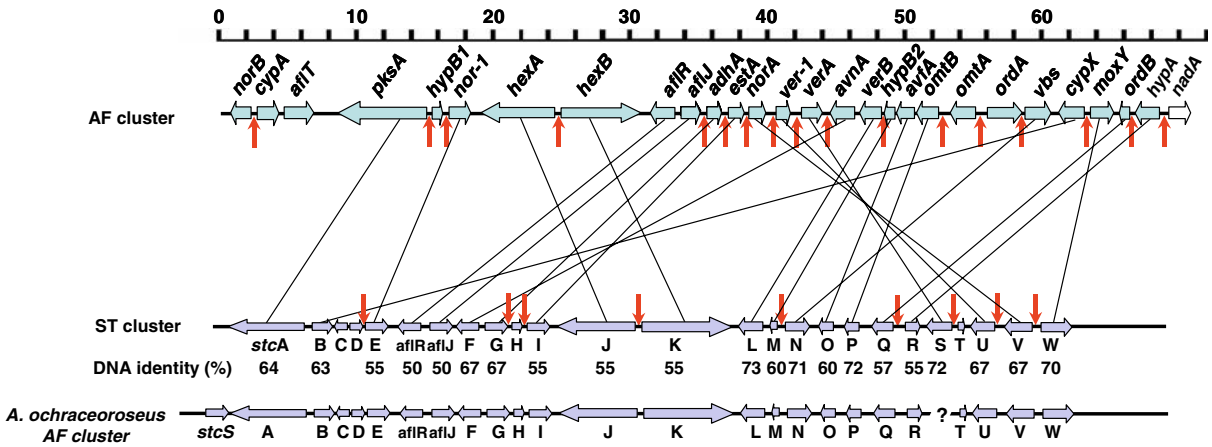

Figure 1. Comparison of section *Flavi* and non-section *Flavi* biosynthetic gene clusters. Horizontal arrows show the direction of transcription. Vertical arrows show promoter regions that contain AflR-binding sites. Lines connect gene homologs. The DNA sequence identity between the *A. nidulans* and *A. flavus* homologs is given below the ST cluster. The scale is length in kb.

Table 2. Comparison of AF and ST cluster genes

AF/ST cluster gene[a]	Introns			% Nucleotide identity[b]			% Amino acid identity[b]		
	A.o.	A.n.	A.f.	A.o./A.n.	A.o./A.f.	A.n./A.f.	A.o./A.n.	A.o./A.f.	A.n./A.f.
aflR	0	0	0	68	51	50	64	34	36
aflJ	2	1	2	67	52	50	62	40	35
nor-1/*stcE*	3	3	3	66	62	56	68	63	56
estA/*stcI*	1	0	1	64	57	55	67	53	50
ver-1/*stcU*	2	2	2	74	74	72	90	92	90
verB/*stcL*	1	1	1	74	74	74	83	83	81
omtB/*stcP*	3	3	3	70	71	72	78	83	75
vbs/*stcN*	2	2	1	76	72	71	87	81	80

[a]Sequence data for the A. *nidulans* ST cluster was obtained from GenBank accession no. U34740 or AACD01000132; AY510453 for *A. flavus*.
[b]Alignments were performed using the full alignment tool in DNAMAN software.
Abbreviations: A.o., *Aspergillus ochraceoroseus*; A.n., *Aspergillus nidulans*; A.f., *Aspergillus flavus*.

identities are between 78 and 92%. Amino acid identity was equal to or higher between *A. ochraceoroseus* and *A. flavus* (with the exception of *vbs*/*stcN*) than to *A. nidulans*.

Gene order

Gene order in the ST and AF clusters and direction of transcription are conserved for most of the genes and their homologs, with the exception of four AF genes, *cypX*, *moxY* , *avnA* and *vbs* (Figure 1). The gene order is inverted for the genes *norA*, *ver-1* and *verA* and their ST cluster homologs. If it is assumed that the *A. nidulans* cluster is the ancestral cluster type (see below for the rationale for this assumption), then at least two and possibly three rearrangements must have occurred to change the gene order for these genes (Figure 1, lines connecting genes). The rearrangements appear to have occurred in blocks related to putative function. For example *cypX* and *moxY* are jointly involved in averufin conversion to versicolorin A and *adhA* and *avnA* encode enzymes for averufin formation. DNA sequence analysis of the *A. ochraceoroseus* AF gene cluster demonstrated that it was very similar to that of the *A. nidulans* ST gene cluster with respect to gene order and direction of gene transcription (Figure 1). The only major difference was that the *A. ochraceoroseus stcS* gene homolog was located directly upstream of the *stcA* gene coding region instead of between the *stcR* and *stcT* gene regions as defined by the *A. nidulans* ST gene cluster. In the place of the *stcS* coding region between *stcR* and *stcT* in the *A. ochraceoroseus* AF gene cluster, there is an insertion of greater than 4 kb in which no genes expected to play a role in AF biosynthesis were found. It is possible that the *A. ochraceoroseus* AF gene cluster is fragmented into two clusters; one spanning the *stcS* to *stcR* region and the other spanning the *stcW* to *stcT* region. Homologs of *omtA* and *ordA*, genes required for conversion of ST to AF in section *Flavi* isolates, have not been identified in the *A. ochraceoroseus* gene cluster. Searches of databases containing *A. nidulans* genomic sequence for such homologs revealed the presence of several high scoring candidates in *A. nidulans* located outside of the ST cluster. It is probable that homologs for these genes are also located outside of the gene cluster in *A. ochraceoroseus*.

Genes adjacent to the AF or ST clusters

The ORFs distal (closest to the chromosomal end) to the cluster are not conserved in section *Flavi* isolates and the promoter regions for these putative genes do not contain AflR-binding sites. A gene predicted to encode a protein with high sequence identity to a xylanase occurs within 2 kb of the distal end of the AF cluster of *A. nomius*, but is not found in *A. flavus*, which contains, in this region, genes encoding a possible MFS efflux transporter and a dehydrogenase. Proximal to the cluster in all of the section *Flavi* isolates is a four-gene sugar utilization cluster [21]. A similar sugar cluster is not found adjacent to the ST cluster in *A. nidulans*. The two immediate ORFs in

A. nidulans at the distal end of the ST cluster in *A. nidulans* are predicted to encode a transporter protein (AN7826, *E* value = $4e^{-37}$) and glucuronyl hydrolase (COG4225). The closest ORF at the other end of the ST cluster is predicted to encode an MFS transporter protein. Different predicted genes flanked the ends of the *A. ochraceoroseus* cluster. One of these is predicted to encode a protein with high sequence identity to microbial cyanovirin-N (AN8062.2, *E* value = $1e^{-30}$). Sequence variation in genes outside of the proximal end of the AF cluster suggests that this region is not under the same evolutionary constraints as are the genes in the AF and sugar clusters. As observed for the *A. nidulans* ST cluster genes predicted to encode homologs to *ordA* and *omtA*, genes necessary for conversion of ST to AF were not in the proximity of the *A. ochraceoroseus* ST biosynthesis gene cluster.

Putative binding sites for transcriptional regulatory proteins

Thirteen of the genes in the AF cluster contain consensus TATA sites [22], and 3 others have AT-rich domains within 150 bp of the translational start site, that may serve a similar function. In the *A. nidulans* ST cluster 11 genes have TATA sites, but only 4 of these are in the promoter regions of their AF cluster homologs (*stcF*, *stcG*, *stcE*, and *stcU*). In a few of the AF and ST clusters, the promoter regions have possible CCAAT-box domains [23, 24]. Except for the promoter regions of *aflR*, *avfA*, *aflT*, and *aflJ*, AF cluster genes have one or two AflR-binding sites with the minimal consensus sequence TCGN5CGR. In the ST cluster, these genes and the homolog of *pksA* lack AflR-binding sites (Figure 1).

Some of the promoter regions of the AF and ST clusters contain consensus binding sites for PacC (GCCARG), involved in ambient pH regulation [25, 26], BrlA (MRAGGGR) and AbaA (CATTCY) involved in developmental regulation [27, 28], and AreA (HGATAR), involved in nitrogen regulation [29]. In the AF-type clusters, the genes *aflJ*, *omtB*, and *verB* and in *A. nidulans* clusters, the genes *stcU* and *stcP* have possible AreA sites in their promoter regions. Three BrlA sites surround the AflR site in the promoter region of *hypB1* from *A. parasiticus* and *A. flavus*, but in *A. nomius* there is only one BrlA site. For some of the species the intergenic regions of *aflR*/*aflJ*, *hexA*/*hexB*, and *omtA*/*ordA* have two or more sites for BrlA and AbaA. The *aflT* promoter of all of the AF clusters has four AbaA and two BrlA sites. The promoter regions of *aflJ*, *omtB*, *verA*, *verB*, and *hexB* contain putative PacC sites, but only in some of the species. One of the two PacC sites in the *A. parasiticus omtB* promoter is 8 bp from the AflR-binding site. In the ST cluster of *A. nidulans* the genes *stcF* and *stcG* have promoter region BrlA sites, *stcP*, *stcS*, and *stcU* have AreA sites, and *stcI*, *stcS*, and *stcU* have PacC sites. Therefore, except for AflR binding sites, there is little conservation of regulatory sites between the two types of gene clusters. Analysis of *A. ochraceoroseus* gene promoter regions for putative transcription factor binding sites has yet to be performed.

Phylogenetic relationships between AF and ST-producing isolates

Divergence of Eurotiomycete (*Aspergillus* sp.) and Sordariomycete (*Neurospora crassa*) fungi (ES) is estimated to have occurred between 300 and 700 million years ago (Mya) [30, 31]. Using the lower value for this estimate and the branch lengths in the phylogenetic trees in Figure 2A, B, we can estimate the divergence time between species with the AF-type and ST-type gene clusters (see [32]). Based on this estimate, *A. nidulans* and *A. nomius* diverged about 120 Mya and *A. nidulans* and either *E. astellata* or *A. ochraceoroseus* about 63 Mya. Figure 2B shows that *A. ochraceoroseus* and *A. parasiticus* diverged from a common ancestor about 75 Mya. Therefore the AF-type cluster evolved after this divergence and the ST-type gene cluster is the ancestral cluster type. However, it is likely that the ancestral species, like *A. ochraceoroseus*, was capable of AF production, even though the necessary genes for AF production were not yet part of the gene cluster, as they are in species containing the AF-type cluster.

The phylogenetic relationships in Figure 2A are based on the ribosomal DNA internal transcribed spacer region (ITS) while the studies in Figure 2B are based on the combined tree from sequence alignments of three genes, *aflR*, *nor-1*/*stcE*, and beta tubulin (*benA*). Bootstrap support for the cladal separations was 100 % except for the

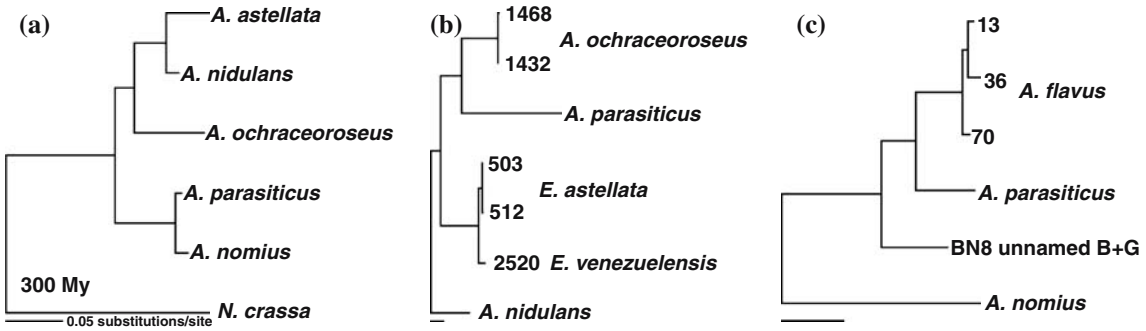

Figure 2. Phylogenetic relationships. Trees were based on Maximum likelihood method in PAUP*. (A) Cladogram based on ITS sequence of sections *Flavi*, *Circumdati*, and *Nidulantes* isolates. The tree is rooted with the ITS sequence for the Sordariomycete fungus, *Neurospora crassa*. The estimated time of divergence is based on the fossil record for divergence of Sordariomycete from Eurotiomycete fingi of 300 million years ago (Mya). (B) Cladogram based on the tree for the combined sequences of *stcE*, *aflR*, and *benA*. The tree is rooted with *A. nidulans*. The SRRC accession numbers for isolates are given next to the branches. (C) Cladogram based on concatenated intergenic regions of the genes in the AF clusters of three *A. flavus* isolates (AF70, AF13, and AF36), *A. parasiticus* NRRL2999, unnamed taxon BN008R, and *A. nomius* NRRL13137.

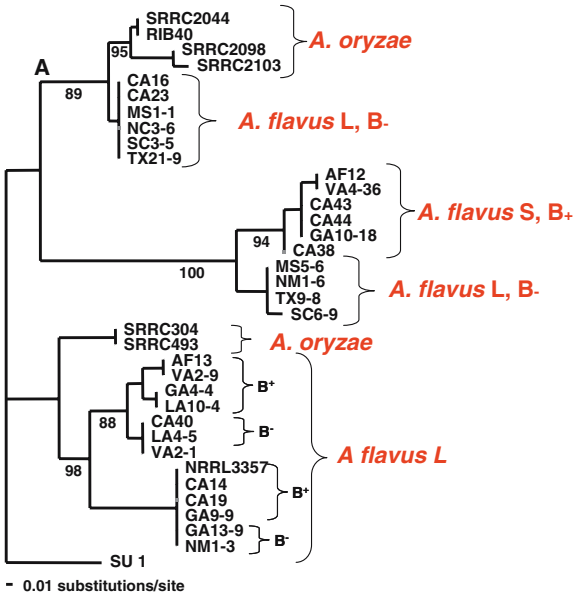

Figure 3. Comparison of non-aflatoxigenic (B−) and aflatoxin-producing (B+) *A. flavus* isolates. Phylogenetic relationships used the Distance method in PAUP*. Bootstrap values are given on branches. Isolates designated CA, LA, MS, SC, TX GA, VA, NM, SC were collected from soil of mainly peanut-growing fields in the United States. AF13 and AF12 were from cotton-growing fields. S (small) and L (large) indicate sclerotial size. Other designations refer to the culture collections of the Southern Regional Research Center (SRRC) and the Northern Center for Agricultural Utilization Research (NRRL). The designation *A. oryzae* is based on morphological comparison of isolates to known *A. oryzae* isolates used in commercial fermentations.

separations of *A. flavus* isolates in Figure 2C. The AF-producing, non-section *Flavi* isolates were from approximately the same latitude, suggesting that the west coast of Africa may be the origin of the ancestral species that radiated to the Galapagos Islands and Venezuela.

The relationships among species in the phylogenetic tree shown in Figure 2C suggests that within section *Flavi* isolates, the ability to produce both B and G aflatoxins is the ancestral trait. *A. flavus* diverged from *A. parasiticus* by loss of a portion of the gene, *cypA*, necessary for G-type AF production. *A. flavus* forms a polyphyletic assemblage containing isolates of different sclerotial morphotype and ability to produce AF (Figure 3). Included in this assemblage is

A. *oryzae*, a non-aflatoxigenic *A. flavus* species. Comparison of single nucleotide polymorphisms (SNPs) in the *omtA* coding sequence and the deletions at the distal end of the cluster, including *cypA* and *norB*, genes needed for G-type AF formation, showed that many of the non-aflatoxigenic L-isolates, including the *A. oryzae* isolates, more closely resemble S isolates, than they do L isolates (Figure 3). Sclerotial morphotype was found to be a poor indicator of phylogenetic relationships among *A. flavus* isolates.

Evolution of the ST and AF gene cluster

The basal AF gene cluster

The AF cluster contains a typical non-reducing PKS, whose members produce precursors for aromatic condensed ring metabolites [33]. We speculate that the basal AF gene cluster in ancestral *Aspergillus* species consisted of genes encoding the PKS, a few enzymes capable of stabilizing the nascent polyketide, and possibly *aflR* and *aflJ* for regulation of transcription. This basal cluster probably also contained genes for production of medium-chain FASs, which, as in yeast, may have evolved as separate entities from primary metabolism FASs to allow the fungi to survive in anaerobic environments [34]. The presumed basal cluster genes, *pksA*, *aflR*, *aflJ*, *hexA, hexB*, and *nor-1* and their ST-type cluster homologs form a discreet "mini-cluster" at the distal end of the biosynthetic gene cluster. The lower sequence identity of most of these genes compared to that of the genes encoding polyketide modifying enzymes suggests that they are subject to different selection pressures (Figure 1). Such a basal gene cluster would have allowed these ancestral fungi to synthesize colorful anthraquinones that may have helped to foster spore dispersal.

Gain of genes encoding modifying enzymes

Phylogenetic evidence suggests that the AF and ST clusters were formed mainly by a process of gene duplication and gene recruitment followed by purifying selection [20]. Evidence for gene duplication is seen in the AF cluster. The AF-type cluster contains duplicated genes encoding dehydrogenases, *norA* and *norB*, *O*-methyltransferases, *omtA* and *omtB*, NADH-dependent dehydrogenases, *aflX* and *avfA*, and proteins related to anthrone dioxygenases, *hypB1* and *hypB2* [35]. Gene recruitment may have involved adaptive translocation events similar to those found in other organisms [36–39]. Homologs of *omtA* and *ordA* are present elsewhere in the *A. nidulans* genome and the cluster genes *stcM* (*hypB1*) and *stcR* (*hypA*) show significant homology to the non-cluster genes, AN0151.2 and AN0152.2. Most of the genes in the AF and ST clusters have the typical small introns of fungi which precludes their having a bacterial origin.

Loss of AF cluster genes

At the distal end of the AF cluster the genes are not as well conserved as at the proximal end, suggesting that these genes may be under weaker selection pressure. *A. flavus* isolates are missing the genes necessary for production of AFG1 due to a large deletion of portions of *cypA* and *norB* (Figure 1). This loss is estimated to have occurred about 5 Mya after *A. flavus* and *A. parasiticus* diverged. *A. flavus* isolates also show considerable variability in AFB1 production. About 40% of the naturally occurring isolates of *A. flavus* lack the ability to produce AF and different isolates show large differences in the levels of AF produced. The non-aflatoxigenic *A. flavus* isolates in many cases have large deletions of portions of the distal end of the AF cluster. Relatively recent loss of genes for production of AF in *A. flavus* and the related A. *oryzae* is evidence that in certain environments AF-producing ability no longer serves an adaptive function [8, 40]. Several putative genes in the ST-type clusters, *stcC*, *stcD*, and *stcT*, are not present in the AF-type clusters. The role of these genes, if any, in ST production is not yet known.

Evidence for purifying selection

Purifying selection and gene duplications have been invoked to explain the formation of gene families and gene clusters in other eukaryotes [33, 41–43]. Higher than expected nucleotide identity of some ST and AF cluster genes suggests that purifying selection occurred after divergence of these two clusters [20]. The ratio of non-synonymous to synonymous nucleotide polymorphisms is a good measure of purifying selection. A low ratio

indicates that adaptive selection has occurred, whereas a high ratio indicates a relative lack of selection. We found that the ratio of non-synonymous to synonymous substitutions between *A. nomius* and *A. flavus* was markedly lower for homologs with highest sequence identities in the ST and AF clusters. This result suggests that purifying selection affected the accumulation of cluster gene polymorphisms during divergence of *Aspergillus* section *Flavi* from non-section *Flavi* species. The lowest ratio of non-synonymous to synonomous substitutions for genes in the gene cluster was found for *ver-1*. The proteins encoded by this gene have 64% identity to a melanin biosynthesis protein T4HN of *Magnaporthe grisea* [44]. These results suggest a link between melanin and AF and ST biosynthesis [45].

Model for AF cluster evolution

A model for the staged evolution of the AF and ST gene cluster is shown in Figure 4. In this model we propose that a basal cluster gradually recruited blocks of biosynthetic genes, thereby allowing further metabolism of the initially formed anthraquinones. Colorful anthraquinone-derived metabolites are common to many fungal species and plants. We hypothesize that the function of the basal cluster was to produce colorful metabolites, perhaps to foster spore protection and dispersal, probably by enlisting insects or arthropods as carriers.

Dispersal may have forced these otherwise saprophytic fungi to adapt to increasingly hostile niches or to be exposed to novel food sources. *Dothistroma septosporum* (previously called *D. pini*) produces a compound related to versicolorin A called dothistromin that may be a causative agent in pine blight [46] (see Bradshaw and Zhang in this issue). Some of the genes responsible for production of dothistromin are clustered and have significant homology to AF and ST biosynthetic genes. The gene organization of this cluster deviates somewhat from that of the AF and ST clusters and could represent an ancestral form. Besides the apparent phytotoxicity of dothistromin, AF and ST are somewhat insecticidal [47]. Species such as *A. nomius* and *A. bombycis* which make AFs have been associated with insect debris, suggesting that AFs may be effective insect toxins under certain circumstances [10, 48, 49].

Homologs to genes such as *aflX* and *ver-1* in the *D. septosporum* and AF/ST type clusters encode proteins necessary for conversion of versicolorin A to ST. These proteins are similar to proteins involved in appessorium development, a hardened mycelial structure needed for fungal virulence [50, 51]. In the ST-type cluster these genes are adjacent to one another, suggesting that they may have indeed been "captured" by a putative basal cluster as a discreet unit. This observation provides support for the hypothesis that genes were recruited by the basal cluster to allow conversion of anthraquinones to increasingly toxic compounds.

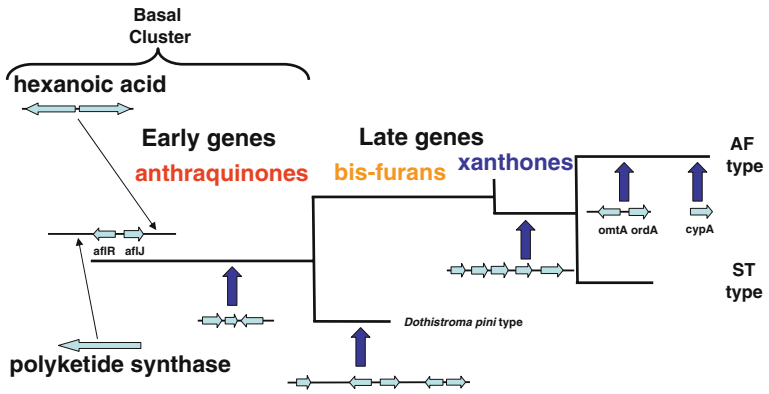

Figure 4. Model for steps in the evolution of the AF and ST gene clusters. The basal cluster includes the polyketide synthase gene, the hexanoic acid synthase genes and the "early genes required to stabilize the polyketide as an anthraquinone. Separate stages for acquisition of "later" genes are indicated by vertical arrows. Divergence of the *D. septosporum*-type cluster is indicated as ancestral to divergence of the AF and ST-type clusters.

In both types of cluster the genes necessary for conversion of versicolorin A to ST and ST to AF are at the ends of the cluster suggesting that these regions were most receptive to introduction of new genes. The increased toxicity that these more hydrophobic metabolites provided may have been necessary to allow the *Aspergillus* isolates to adapt to the increasingly hostile niches to which they had been dispersed.

Conclusions

The retention of an intact AF and ST gene cluster for over 100 My and the presence of AF/ST biosynthetic genes in other genera (e.g. *D. septosporum*) suggests that the metabolites produced by the gene cluster impart an adaptive advantage to the host organism in certain niches. Additional studies need to be performed to determine if the ability to produce AF or ST enhances the fitness and survival of the producing strain over that of strains that have lost the ability to produce AF or ST. No comparison has been made of the effects of different soil conditions and environmental factors on the survival of AF-producing strains to isogenic isolates that are AF non-producing due to mutations in AF biosynthetic genes. Using isogenic ST pathway mutants of *A. nidulans*, Wilkinson et al. [17] were able to show a stepwise increase in spore production with each progressive step in the ST pathway suggesting that there is a selective advantage to production of ST and its intermediates.

It has been estimated that only about 25% of fungal species have been identified; therefore it is probable that many fungi have not been discovered that possess some or all of the genes necessary for producing AF. In addition, some fungal taxa in various repositories throughout the world may be capable of producing AF and its precursors but have not been grown in the laboratory under suitable culture conditions for metabolite production. Further isolation and characterization of AF biosynthetic genes in non-section *Flavi* species will aid in determining the importance of AF and ST production for fungal adaptation to natural environments. As these novel cultures are identified they can be grown under a number of 'adverse' conditions to determine if AF-producing ability enhances survival.

Much is known about AflR and its role as a positive regulator of AF biosynthesis, and the effects of plant metabolites, and environmental and nutritional factors on fungal development and AF biosynthesis. It is becoming apparent that AflR-like regulatory proteins are also functioning in regulation of AF production in isolates outside section *Flavi* such as *A. ochraceoroseus* and possibly in dothistromin biosynthesis in *D. septosporum*. However, there is still much that we do not know about the mechanisms by which global regulatory genes and signaling pathways control AF gene expression and fungal development. Comparison and contrast of putative transcription factor binding sites in the promoter regions of genes from section *Flavi* and non-section *Flavi* isolates may shed additional light on the role of environmental factors on AF and related metabolite production and fungal development, and virulence.

References

1. Klich MA, Cary JW, Beltz SB, Bennett CA. Phylogenetic and morphological analysis of *Aspergillus ochraceoroseus*. Mycologia 2003; 95: 1252–1260.
2. Frisvad JC, Samson RA. *Emericella venezuelensis*, a new species with stellate ascospores producing sterigmatocystin and aflatoxin B1. Syst Appl Microbiol 2004; 27: 672–680.
3. Frisvad JC, Samson RA, Smedsgaard J. *Emericella astellata*, a new producer of aflatoxin B, B and sterigmatocystin. Lett Appl Microbiol 2004; 38: 440–445.
4. Cary JW, Klich MA, Beltz SB. Characterization of aflatoxin-producing fungi outside of *Aspergillus* section *Flavi*. Mycologia 2005; 97: 425–432.
5. Bartoli A, Maggi O. Four new species of *Aspergillus* from Ivory Coast soil. Trans Br Mycol Soc 1978; 71: 383–394.
6. Cotty PJ, Bayman DS, Egel DS, Elias KS. Agriculture, aflatoxins and *Aspergillus*. In: Powell K, ed. The Genus *Aspergillus*, Plenum Press, New York, 1994: 1–27.
7. Geiser DM, Dorner JW, Horn BW, Taylor JW. The phylogenetics of mycotoxin and sclerotium production in *Aspergillus flavus* and *Aspergillus oryzae*. Fungal Genet Biol 2000; 31: 169–179.
8. Cotty PJ. Virulence and cultural characteristics of two *Aspergillus flavus* strains pathogenic on cotton. Phytopathology 1989; 79: 808–814.
9. Broek Pvan den, Pittet A, Hajjaj H. Aflatoxin genes and the aflatoxigenic potential of Koji moulds. Appl Microbiol Biotechnol 2001; 57: 192–199.
10. Kurtzman CP, Horn BW, Hesseltine CW. *Aspergillus nomius*, a new aflatoxin-producing species related to *Aspergillus flavus* and *Aspergillus tamarii*. Antonie Leeuwenhoek 1987; 53: 147–158.

11. Goto T, Wicklow DT, Ito Y. Aflatoxin and cyclopiazonic acid production by a sclerotium-producing *Aspergillus tamarii* strain. Appl Environ Microbiol 1996; 62: 4036–4038.
12. Dowd PF. Involvement of arthropods in the establishment of Mycotoxigenic fungi under field conditions. BibEditorName > KK Sinha, In: Bhatnagar D), ed. Mycotoxins in Agriculture and Food Safety, Marcel Dekker, New York, 1998: 307–350.
13. Boyd ML, Cotty PJ. *Aspergillus flavus* and aflatoxin contamination of leguminous trees of the Sonoran Desert in Arizona. Phytopathol 1997; 79: 808–814.
14. Bhatnagar D, Ehrlich KC, Chang P-K. Mycotoxins. In: Encyclopedia of Life Sciences, Vol. 12 London: Nature Publishing Company, 2000: 564–573.
15. Lillehoj EB. Aflatoxin: an ecologically elicited genetic/activation signal. In: Smith JE, Henderson RS, eds. Mycotoxins and Animal Foods, CRC Press, Boca Raton, FL, 1991: 2–30.
16. Pringle A, Taylor J. The fitness of filamentous fungi. Trends Microbiol 2002; 10: 474–481.
17. Wilkinson H, Ramaswamy A, Sim SC, Keller NP. Increased conidiation associated with progression along the sterigmatocystin biosynthetic pathway. Mycologia 2004; 96: 1190–1198.
18. Proctor RH, Brown DW, Plattner RD, Desjardins AE. Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in *Gibberella moniliformis*. Fungal Genet Biol 2003; 38: 237–249.
19. Yu J, Chang P-K, 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 2004; 70: 1253–1262.
20. Ehrlich KC, Yu J, Cotty PJ. Aflatoxin biosynthesis gene clusters and flanking regions. J Appl Microbiol 2005; 99: 518–527.
21. Yu J, Chang P, Bhatnagar D, Cleveland TE. Cloning of a sugar utilization gene cluster in *Aspergillus parasiticus*. Biochim Biophys Acta 2000; 1493: 211–214.
22. Bucher P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoters. J Mol Biol 1990; 212: 563–578.
23. Narendja FM, Davis MA, Hynes MJ, Kato M, Aoyama A, Naruse F, Kobayashi T, Tsukagoshi N, Heeswijck Rvan, Gomez D, Garcia I, Scazzocchio C, Cubero B, Kucharski R, Bartnik E. AnCF, the CCAAT binding complex of *Aspergillus nidulans*, is essential for the formation of a DNase I-hypersensitive site in the 5′ region of the *amdS* gene. Mol Cell Biol 1999; 19: 6523–6531.
24. Steidl S, Hynes MJ, Brakhage AA. The *Aspergillus nidulans* multimeric CCAAT binding complex AnCF is negatively autoregulated via its *hapB* subunit gene. J Mol Biol 2001; 306: 643–653.
25. Tilburn J, Sarkar S, Widdick DA, Espeso EA, Orejas M, Mungroo J, Penalva MA, Arst HN Jr. The *Aspergillus* PacC zinc finger transcription factor mediates regulation of both acid- and alkaline expressed genes by ambient pH. EMBO J 1995; 14: 779–790.
26. Then Bergh K, Brakhage AA. Regulation of the *Aspergillus nidulans* penicillin biosynthesis gene *acvA* (pcbAB) by amino acids: implication for involvement of the transcription factor PACC. Appl Environ Microbiol 1998; 64: 843–849.
27. Andrianopoulos A, Timberlake WE. The *Aspergillus nidulans abaA* gene encodes a transcriptional activator that acts as a genetic switch to control development. Mol Cell Biol 1994; 14: 2503–2515.
28. Han S, Navarro J, Greve RA, Adams TH. Translational repression of *brlA* expression prevents premature development in *Aspergillus*. Embo J 1993; 12: 2449–2457.
29. Muro-Pasteur MI, Gonzalez R, Strauss J, Narendja F, Scazzocchio C. The GATA factor AreA is essential for chromatin remodelling in a eukaryotic bidirectional promoter. EMBO J 1999; 18: 1584–1597.
30. Berbee ML, Taylor JW. Fungal molecular evolution: gene trees and geologic time. In: McLaughlin DJ, McLaughlin E, Lemke PA, eds. The Mycota, Springer, Berlin, 2001: 229–246.
31. Heckman DS, Geiser DM, Eidell BR, Stauffer RL, Kardos NL, Hedges SB. Molecular evidence for the early colonization of land by fungi and plants. Science 2001; 293: 1129–1133.
32. Kasuga T, White TJ, Taylor JW. Estimation of nucleotide substitution rates in eurotiomycete fungi. Mol Biol Evol 2002; 19: 2318–2324.
33. Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. Proc Natl Acad Sci USA 2003; 100: 15670–15675.
34. Bardi L, Cocito C, Marzona M. *Saccharomyces cerevisiae* cell fatty acid composition and release during fermentation without aeration and in absence of exogenous lipids. Int J Food Microbiol 1999; 47: 133–140.
35. Fujii I, Chen ZG, Ebizuka Y, Sankawa U. Identification of emodinanthrone oxygenase in fungus *Aspergillus terreus*. Biochem Int 1991; 25: 1043–1049.
36. Foster PL. Adaptive mutation: implications for evolution. Bioessays 2000; 22: 1067–1074.
37. Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, Botstein D. Characteristic genome rearrangements in experimental evolution of *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 2002; 99: 16144–16149.
38. Schmidt S, Sunyaev S, Bork P, Dandekar T. Metabolites: a helping hand for pathway evolution?. Trends Biochem Sci 2003; 28: 336–341.
39. Zuniga M, Perez G, Gonzalez-Candelas F. Evolution of arginine deiminase (ADI) pathway genes. Mol Phylogenet Evol 2002; 25: 429–444.
40. 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.
41. Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. Science 2000; 290: 1151–1155.
42. Ober D, Hartmann T. Phylogenetic origin of a secondary pathway: the case of pyrrolizidine alkaloids. Plant Mol Biol 2000; 44: 445–450.
43. Wheeler D, Hope R, Cooper SB, Dolman G, Webb GC, Bottema CD, Gooley AA, Goodman M, Holland RA. An orphaned mammalian beta-globin gene of ancient evolutionary origin. Proc Natl Acad Sci USA 2001; 98: 1101–1106.

44. Vidal-Cros A, Viviani F, Labesse G, Boccara M, Gaudry M. Polyhydroxynaphthalene reductase involved in melanin biosynthesis in *Magnaporthe grisea*. Purification, cDNA cloning and sequencing. Eur J Biochem 1994; 219: 985–992.
45. Gomez BL, Nosanchuk JD. Melanin and fungi. Curr Opin Infect Dis 2003; 16: 91–96.
46. Bradshaw RE, Bhatnagar D, Ganley RJ, Gillman CJ, Monahan BJ, Seconi JM. *Dothistroma pini*, a forest pathogen, contains homologs of aflatoxin biosynthetic pathway genes. Appl Environ Microbiol 2002; 68: 2885–2892.
47. Dowd PF. Insect interaction with mycotoxin-producing fungi and their hosts. In: Bhatnagar D, Lillehoj EB, Arora DK, eds. Mycotoxins in Ecological SystemsVol. 5, Marcel Dekker, New York, 1992: 137–155.
48. Llewellyn GC, Gee CL, Sherertz PC. Toxic responses of developing fifth instar milkweed bugs, *Oncopeltus fasciatus* (Hemiptera), to aflatoxin B1. Bull Environ Contam Toxicol 1988; 40: 332–338.
49. Peterson SW, Ito Y, Horn BW, Goto T. *Aspergillus bombycis*, a new aflatoxigenic species and genetic variation in its sibling species, *A. nomius*. Mycologia 2001; 93: 689–703.
50. Henson JM, Butler MJ, Day AW. The dark side of the mycelium: melanins of phytopathogenic fungi. Annu Rev Phytopathol 1999; 37: 447–471.
51. Inagaki A, Takano Y, Kubo Y, Mise K, Furusawa I. Construction of an equalized cDNA library from *Colletotrichum lagenarium* and its application to the isolation of differentially expressed genes. Can J Microbiol 2000; 46: 150–158.

Address for correspondence: Jeffrey W. Cary, USDA, ARS, Southern Regional Research Center, 1100 Robert E. Lee Blvd., New Orleans, LA 70124, USA
Phone: + 1-504-286-4264; Fax: + 1-504-286-4419,
E-mail: jcary@srrc.ars.usda.gov