Recently available genome sequences for *Aspergillus flavus* and *A. oryzae* were used to gain insight into the biosynthesis of secondary metabolites and to identify species-specific characters for these fungi. Transcriptome analyses, comparative genome hybridizations, and bottom-up proteomics were used to study two interrelated aspects of the ecology of *A. flavus*: 1) Regulation of secondary metabolism, with emphasis on the carcinogenic mycotoxin, aflatoxin, and 2) Discovery of key differences between *A. flavus* and the closely related domesticated species *A. oryzae*.

Filamentous fungi such as *A. flavus* produce an abundance of diverse secondary metabolites, the most well-studied being aflatoxin. A defining feature of secondary metabolites is their production by clusters of genes. In this thesis I provide a comprehensive review of the current status of aflatoxin biosynthesis and regulation, including emerging genomic studies. I was the first to employ SILAC (stable isotope labeling by amino acids in cell culture), a technique to enable relative protein quantification by mass spectrometry, in a multicellular free-living prototroph. This technique allowed me to quantify 381 proteins during growth of *A. flavus* under conditions conducive (28°C) and non-conducive (37°C) for aflatoxin biosynthesis. From these studies I showed that enzymes needed for aflatoxin biosynthesis were lacking at 37°C. Additionally, I observed that protein concentration and transcript accumulation did not correlate well, with transcripts and proteins from genes within the aflatoxin cluster being a notable exception. I also showed through use of reporter constructs that the aflatoxin pathway specific transcription factor AflR is localized to the
nucleus and active at 37°C even though aflatoxin is not produced and most genes in the pathway are not expressed.

I have also studied the regulation of all predicted secondary metabolite gene clusters for *A. flavus* to better characterize their expression under environmental conditions. Of the predicted 55 secondary metabolite gene clusters in *A. flavus*, only three metabolites have been associated with a respective cluster. These are aflatoxin, cyclopiazonic acid, and aflatrem. Aside from aflatoxin, little knowledge is available about the regulation of other secondary metabolites in *A. flavus*. Transcriptional analysis of these secondary metabolite gene clusters over 28 experimental conditions showed the clusters to group into classes with similar profiles of expression. To further explore the correlations found by gene expression analysis, aflatoxin and CPA production were quantified under six cell culture environments known to be conducive or non-conducive for aflatoxin biosynthesis and in infected maize seeds. We found that CPA was not as tightly regulated as aflatoxin in response to cell culture environment however CPA and aflatoxin both accumulated similarly in developing maize seeds.

I compared the genome of *A. flavus* with *A. oryzae*, a non-aflatoxigenic and domesticated species related to *A. flavus*. I hypothesized that insights gained from knowledge of the differences between these species would reveal new information on regulation of secondary metabolism, and possibly pathogenicity. I used comparative genome hybridization to characterize genomic content among three strains of each species. These results revealed that *A. flavus* and *A. oryzae* are strikingly similar with regard to DNA sequence. In addition to defining core sequence variations between these species I investigated the gene expression differences on substrates unique to each species’ ecological
niche. *Aspergillus oryzae* has been used in fermentations through cultivation for thousand of years on wheat bran. *Aspergillus flavus* is an opportunistic pathogen of maize causing loss of crops through contamination with aflatoxin. From the expression analyses it was clear that despite similar genomes, *A. flavus* and *A. oryzae* use their genomes in vastly different ways. Among the most interesting of these differences was that *A. flavus* appears to be a much more capable producer of secondary metabolites.
Functional and Comparative Genomics of *Aspergillus flavus* to Characterize Secondary Metabolism

by

David Ryan Georgianna

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Functional Genomics

Raleigh, North Carolina

2009

APPROVED BY:

_______________________________  ______________________________
Gary A. Payne      James A. Alspaugh
Committee Chair

_______________________________  ______________________________
David C. Muddiman     Dahlia M. Nielsen
David Ryan Georgianna was born June 1, 1982 to parents David and Barbara Georgianna in Richmond, VA. He later moved to Williamsburg, VA at the age of 7 and lived there year-round until graduating from Jamestown High School in 2000. While in Williamsburg Ryan earned his Eagle Scout from Troop 103, was a Fifer in the Colonial Williamsburg Fifes and Drum Corp., played the saxophone in the high school band, ran on the high school track team, and enjoyed spending time sailing with his family.

Ryan attended Virginia Tech starting in 2000 and earned a B.S. in Biology in 2004. Ryan was a student-athlete while at Virginia Tech, rowing on the crew team all four years. During the summers Ryan spent time working as a sailing instructor and as a nursing assistant/orderly at the Williamsburg Community hospital. After realizing he did not want to be a medical doctor Ryan got more involved with undergraduate research, working in the tissue engineering lab of Dr. Aaron Goldstein while at Virginia Tech and working in the microarray division of the Chemical Engineering lab of Dr. Anthony Guiseppi-Elie at Virginia Commonwealth University during the summers of 1999 and 2000.

After graduating from Virginia Tech, Ryan joined the Functional Genomics Program at NC State University where he chose to work in the Plant Pathology laboratory of Dr. Gary Payne. Ryan is currently engaged to be married to Wesleigh Edwards, whom he met while at Virginia Tech. She is also currently attending NC State University working towards a doctorate in Chemistry. Upon completing his PhD, Ryan has accepted a post-doctoral research position at Duke University where he will work on dissecting quantitative trait-loci in yeast under the direction of Dr. John McCusker.
ACKNOWLEDGEMENTS

I express my sincerest appreciation to all of the people that made this work possible. Each one of my advisory committee members played an integral role in the completion of my dissertation research. Thanks to Dr. Gary A. Payne for taking me on as a Functional Genomics student in your Plant Pathology lab and for providing excellent research advice and conversation. Thanks to Dr. David C. Muddiman for developing a collaboration that enabled me access to technology and expertise for proteomics work. Thanks to Dr. James A. Alspaugh for providing a wealth of knowledge on fungal signaling and for allowing me to pursue some preliminary medical mycology work in his lab. Thanks to Dr. Dahlia M. Nielsen for all of the advice on statistics needed to create the results I have obtained.

Thanks to Dr. Barbara Sherry, Dr. Stephanie Curtis, and Dr. David Bird for being directors of the functional genomics program during my tenure as a graduate student. Thanks to Dr. Bob Kelly, director of the NC State Biotechnology Program, for heading the NCSU/NIH Biotechnology Traineeship that awarded me a two-year fellowship. Thanks to Dr. Shannon Conners at JMP Genomics, part of the SAS institute, for allowing me to participate in an internship.

Thanks to the many lab members and fellow graduate students that I have interacted with over the past several years. Your advice and fellowship have aided in the research endeavors I have pursued. Special thanks to Greg Obrian, the Payne Lab Manager, and to Dr. Norm Glassbrook for helping spur my interest in mass spectrometry and providing excellent metabolite analyses.

Finally, I thank the support I have received from my family while a graduate student. I especially acknowledge my future wife and best friend, Wesleigh Edwards, your companionship, support, and research advice has been indispensable.
TABLE OF CONTENTS

LIST OF TABLES ........................................................................................................ viii

LIST OF FIGURES ................................................................................................... ix

1 Genetic regulation of aflatoxin biosynthesis: From gene to genome ........ 1
   Abstract .................................................................................................................. 2
   Introduction ............................................................................................................ 2
   Regulation of aflatoxin biosynthesis ................................................................. 3
   Regulation within the aflatoxin biosynthetic cluster ...................................... 3
   Transcriptional regulators (AflR and AflS) ....................................................... 4
   Other possible cluster regulatory mechanisms ............................................... 5
   External control of the aflatoxin cluster and global regulation of secondary metabolism .......... 6
   Heterotrimeric G-protein signaling and PKA .................................................. 6
   cAMP signaling ................................................................................................... 7
   Ras family GTPase signaling ............................................................................. 7
   Chromosomal remodeling and silencing ......................................................... 8
   Environmental conditions regulating aflatoxin biosynthesis ...................... 9
   Light mediated regulation of aflatoxin biosynthesis ..................................... 9
   Regulation of aflatoxin by carbon source ...................................................... 10
   Lipids – key precursors and important signaling molecules ....................... 10
   Regulation of aflatoxin by nitrogen source .................................................... 11
   Regulation of aflatoxin by pH ........................................................................... 11
   Regulation of aflatoxin by metals/trace elements .......................................... 11
   Regulation of aflatoxin by temperature .......................................................... 11
   Regulation of aflatoxin by oxidative stress ..................................................... 12
   Outlook for genomics ......................................................................................... 12
   References ........................................................................................................... 12

2 The effect of elevated temperature on gene transcription and aflatoxin biosynthesis .......................................................... 15
   Abstract ................................................................................................................ 16
   Introduction .......................................................................................................... 16
   Materials and Methods ....................................................................................... 17
   Growth, media and aflatoxin analysis ............................................................... 17
   RNA isolation ...................................................................................................... 18
   QPCR .................................................................................................................... 18
   Microarrays ......................................................................................................... 18
   Results .................................................................................................................. 18
Aflatoxin production is regulated by temperature ................. 18
Transfer of cultures from 28°C to 37°C stops aflatoxin biosynthesis . 18
The transcription profile is different at 28-37°C ................. 19
Aflatoxin biosynthetic genes are more highly expressed at 28°C ... 19
Quantitative PCR is consistent with microarray data ............. 19
Discussion ............................................................ 20
Acknowledgements .................................................... 22
Literature Cited ...................................................... 22

3 Temperature dependent regulation of proteins in Aspergillus flavus:
Whole organism stable isotope labeling by amino acids .......... 24
Abstract .............................................................. 25
Introduction .......................................................... 25
Materials and Methods ............................................. 26
Culture of A. flavus for SILAC .................................... 26
Protein isolation and sample preparation ............................ 26
Nanoflow LC-MS/MS ................................................. 27
Data searching and analysis ......................................... 27
RNA isolation and microarray hybridization ....................... 27
Results and Discussion ............................................. 29
Labeling efficiency and method variation .......................... 29
Comparison of protein expression at 28 and 37°C ............... 29
Expression of specific genes and proteins involved in aflatoxin biosynthesis ........................................ 29
Conclusions ............................................................ 30
Data availability ....................................................... 30
Acknowledgements ................................................... 30
Supporting information available .................................... 30
References ............................................................. 30
Supporting information .............................................. 32

4 Beyond aflatoxin: four distinct expression patterns and functional roles
associated with Aspergillus flavus secondary metabolism gene clusters .... 33
Summary ............................................................... 34
Introduction ........................................................... 35
Results ................................................................. 38
Transcriptional analysis of 55 predicted gene clusters in A. flavus . 38
Characterization of the CPA cluster ................................ 43
Accumulation of CPA and AF ....................................... 47
Discussion ............................................................. 49
Experimental Procedures ........................................... 58
Fungal strains ........................................................... 58
Gene deletion, fungal transformation, DNA isolation ............... 59
<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Culture media for CPA and AF production</td>
</tr>
<tr>
<td>RNA isolation and RT-PCR</td>
</tr>
<tr>
<td>Microarray analysis</td>
</tr>
<tr>
<td>Detection and quantification of CPA and AF</td>
</tr>
<tr>
<td>Acknowledgements</td>
</tr>
<tr>
<td>References</td>
</tr>
</tbody>
</table>

5 Comparative genome hybridization reveals few genetic differences between *Aspergillus flavus* and *A. oryzae* | 71 |
| Abstract | 72 |
| Introduction | 73 |
| Materials and Methods | 74 |
| Fungal strains | 74 |
| Affymetrix GeneChip microarrays | 75 |
| Isolation of genomic DNA | 75 |
| Biotin labeling for Affymetrix GeneChip genomic DNA hybridization | 76 |
| Correlation analysis for strain phylogeny | 76 |
| Determining gene content using CGH | 77 |
| ANOVA for species-specific polymorphism prediction | 79 |
| Results and Discussion | 79 |
| Predicted strain phylogeny | 79 |
| Prediction of core genes—genes potentially unique to AFL or AO | 81 |
| Predicting species enriched gene polymorphism | 84 |
| Functional changes and polymorphism | 90 |
| Secondary metabolism gene clusters | 92 |
| Physical location of polymorphic and unique genes | 94 |
| Outlook for comparative genomics of AFL and AO | 97 |
| Acknowledgements | 99 |
| References | 99 |

6 *Aspergillus flavus* and the domesticated *A. oryzae*: Two species with nearly identical genomes but distinctive gene expression | 103 |
| Abstract | 104 |
| Introduction | 105 |
| Materials and Methods | 107 |
| Field inoculation of maize | 107 |
| Isolation of RNA from infected kernels | 107 |
| Preparation of wheat bran koji fungal cultures | 108 |
| Isolation of RNA from wheat bran koji fungal cultures | 109 |
| Microarray hybridization and analysis | 110 |
| Results | 111 |
| Performance of the AFL Affymetrix GeneChip for
hybridization of AO RNA samples ........................................ 112
Differential expression of genes between AFL and AO .......... 114
Differential expression of maize genes in response to
AFL and AO infection ..................................................... 116
Gene expression for core and polymorphic genes ............... 118
Expression of secondary metabolism gene clusters ............. 119
Functional differences in AFL and AO gene expression ........ 121
Physical properties of gene expression ............................. 124
Discussion ..................................................................... 127
Acknowledgements ......................................................... 130
References ..................................................................... 131

Appendix ........................................................................ 134

Appendix A: The effect of elevated temperature on the
localization and activity of AflR ........................................ 135
Summary ........................................................................ 136
Materials and Methods .................................................. 137
   GFP strains: microscopy for localization of AflR ............ 137
   Analysis of aflatoxin by TLC ......................................... 138
   GUS strains: aflatoxin pathway promoters fused to GUS .... 139
   Qualitative GUS Assay ................................................ 139
   Quantitative GUS Assay .............................................. 140
Results and Discussion .................................................. 141
   Localization of AflR .................................................... 141
   AflR functions in A. flavus at 37°C .............................. 142
Conclusions .................................................................. 143
References ..................................................................... 144
LIST OF TABLES

1 Genetic regulation of aflatoxin biosynthesis: From Gene to Genome
   Table 1: Nutritional and environmental effects on AF biosynthesis . . . . . 8

2 The effect of elevated temperature on gene transcription and aflatoxin biosynthesis
   Table 1: Top 20 differentially expressed genes between 28°C and 37°C . 20
   Table 2: Fold gene induction ................................................. 21

3 Temperature dependent regulation of proteins in Aspergillus flavus:
   Whole organism stable isotope labeling by amino acids
   Table 1: Differentially expressed proteins ......................... 27

4 Beyond aflatoxin: four distinct expression patterns and functional roles associated with Aspergillus flavus secondary metabolism gene clusters
   Table 1: Experimental conditions used for analysis of secondary metabolism gene clusters ......................... 39
   Table 2: Predicted secondary metabolism gene cluster and accession numbers for backbone enzymes predicted by SMURF ......................... 40

5 Comparative genome hybridization reveals few genetic differences between Aspergillus flavus and A. oryzae
   Table 1: Aspergillus strains used for CGH ......................... 75
   Table 2: Descriptions and EST evidence for core genes ........... 85

6 Aspergillus flavus and the domesticated A. oryzae: Two species with nearly identical genomes but different gene expression
   Table 1: Genome statistics for Aspergillus species .................. 107
   Table 2: Induction of genes in AFL and AO when grown on maize
            or wheat bran koji for fifteen biological processes .......... 123
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Figures</th>
</tr>
</thead>
</table>
| 1       | Genetic regulation of aflatoxin biosynthesis: From gene to genome                                                                      | Figure 1: Changes in the expression of genes within the AF cluster in response to conditions favorable or not favorable for AF biosynthesis. 2  
Figure 2: Signaling pathways regulating AF biosynthesis and their ties to growth and development. 4  
Figure 3: Epigenetic control of the aflatoxin cluster. 6                      |
| 2       | The effect of elevated temperature on gene transcription and aflatoxin biosynthesis                                                   | Figure 1: Microarray experimental design. 18                            |
|          |                                                                                                                                        | Figure 2: Effect of temperature on aflatoxin production in solid and liquid media. 19  
Figure 3: Temperature shift. 19                                                |
|          |                                                                                                                                        | Figure 4: Expression levels of aflatoxin pathway genes. 20              |
| 3       | Temperature dependent regulation of proteins in *Aspergillus flavus*: Whole organism stable isotope labeling by amino acids           | Figure 1: (A) Experiment design (B) SDS-PAGE. 26                        |
|          |                                                                                                                                        | Figure 2: Distribution of SILAC ratios. 28                              |
|          |                                                                                                                                        | Figure 3: Scatterplot of shared data points between 28°C:28°C and 28°C:37°C data sets. 28  
Figure 4: Scatterplot of gene and protein expression. 28                      |
|          |                                                                                                                                        | Figure 5: Extracted ion chromatogram (EIC) for selected peptides from data set of differentially expressed proteins. 29  
Supplement1: Gel slice number for common peptides. 32                         |
| 4       | Genetic and environmental conditions differentially affect expression of the secondary metabolite gene clusters in *Aspergillus flavus* | Figure 1: Hierarchical clustering of backbone enzymes. 42                |
|          |                                                                                                                                        | Figure 2: Location of cluster 55 in the genome of *A. flavus*. 43       |
|          |                                                                                                                                        | Figure 3: Overlap PCR strategy for *dnaA* deletion construct and predicted integration into the genome. 44  
Figure 4: Selected ion chromatogram for CPA. 45                                 |
|          |                                                                                                                                        | Figure 5: RT-PCR for gene products in Δ*hypG*, Δ*hypF* and parent strain. 46  
Figure 6: CPA and AF concentrations in maize kernels after 48 h, 60 h, and 96 h infection. 48  
Figure 7: Response of CPA and AF concentrations to alteration of temperature, pH, nitrogen source, carbon source, and addition of the antioxidant gallic acid. 48 |
5 Comparative genome hybridization reveals few genetic differences between *Aspergillus flavus* and *A. oryzae*

Figure 1: Strain phylogeny represented by correlation of arrays used for CGH analysis ........................................ 80
Figure 2: Core genes for each species .................................................. 83
Figure 3: Polymorphic genes ................................................................. 89
Figure 4: Biological process GO term analysis ............................... 91
Figure 5: Strain phylogeny represented by correlation of probes across only genes predicted to be part of secondary metabolism clusters ................................................................. 94
Figure 6: Location of *A. flavus* unique “core” genes and polymorphic genes along the chromosomes .......... 96
Figure 7: Location of *A. oryzae* unique “core” genes and polymorphic genes along the chromosomes .......... 96
Supplement 1: Distribution of mismatch type for probes based on NRRL3357 and RIB40 genome sequence ............. 102

6 *Aspergillus flavus* and the domesticated *A. oryzae*: Two Species with Nearly Identical Genomes but Different Gene Expression

Figure 1: Gene expression and aflatoxin accumulation in wheat bran koji ................................................................. 112
Figure 2: Gene (probe set) mismatch significance data for evaluation of probe sequence bias in *A. oryzae* .......... 114
Figure 3: Distribution analysis of log2 transformed microarray data ................................................................. 115
Figure 4: Gene expression differences between AFL and AO ........... 116
Figure 5: Infection of maize line B73 kernels with *A. flavus* NRRL3357 and *A. oryzae* RIB40 .................................................. 118
Figure 6: Expression of secondary metabolism gene clusters ............ 122
Figure 7: Biological process GO term enrichment ........................................ 123
Figure 8: Physical mapping of gene expression differences between AFL and AO during maize infection .............. 125
Figure 9: Physical mapping of gene expression differences between AFL and AO during growth in wheat bran koji .......... 126
Figure 10: Distribution of absolute differences between AFL and AO for genes found significantly higher in AFL or AO ........... 126

Appendix A: The effect of elevated temperature on the localization and activity of AflR

Figure 1: Plasmid pRG2A3 (gpdA::sGFP:AflR) .......................................... 138
Figure 2: Localization of AflR and production of aflatoxin by GFP strains measured by TLC ................................. 141
Figure 3: Qualitative GUS assay ................................................................. 143
Figure 4: Quantitative GUS assay ................................................................. 143
Chapter 1

Genetic Regulation of Aflatoxin Biosynthesis: From Gene to Genome

D. Ryan Georgianna and Gary A. Payne

Kojic acid, antibiotics such as penicillin, and cholesterol reducing oligos made by species of Aspergillus as toxins or as beneficial compounds. Beneficial secondary metabolites impact our daily lives either extensively studied due to their bioactive nature and economic importance. Secondary metabolites influence the fitness of the organism in its natural environment. Understanding and appreciation of fungal secondary metabolites remained relatively obscure until the discovery of penicillin in 1929 by Alexander Fleming. Since that time they have been studied in detail because of their impact on human and animal health, and their effects on the marketability of key grain and nut crops. Understanding aflatoxin biosynthesis is the focus of a large and diverse research community. Concerted efforts by this community have led not only to a well-characterized biosynthetic pathway, but also to the discovery of novel regulatory mechanisms. Common to secondary metabolism is the clustering of biosynthetic genes and their regulation by pathway specific as well as global regulators. Recent data show that arrangement of secondary metabolite genes in clusters may allow for an important global regulation of secondary metabolism based on physical location along the chromosome. Available genomic and proteomic tools are now allowing us to examine aflatoxin biosynthesis more broadly and to put its regulation in context with fungal development and fungal ecology. This review covers our current understanding of the biosynthesis and regulation of aflatoxin and highlights new and emerging information garnered from structural and functional genomics. The focus of this review will be on studies in Aspergillus flavus and Aspergillus parasiticus, the two agronomically important species that produce aflatoxin. Also covered will be the important contributions gained by studies on production of the aflatoxin precursor sterigmatocystin in Aspergillus nidulans.

1. Introduction

Secondary metabolism is a hallmark of filamentous fungi. The diversity and complexity of secondary metabolites is astounding, and species of Aspergillus are rich in genes for secondary metabolism (Kobayashi et al., 2007; Nierman et al., 2005; Rokas et al., 2007). In Aspergillus oryzae there are 48 predicted NRPS and PKS proteins and in Aspergillus fumigatus there are at least 26 different secondary metabolism clusters (Nierman et al., 2005). As noted by Keller et al. (2005), it is often difficult to define secondary metabolites, and even more difficult to ascribe biological roles for these compounds. A common working definition is that secondary metabolites are not required for growth of the organism in culture, but do contribute to the fitness of the organism in its natural environment.

Understanding and appreciation of fungal secondary metabolites remained relatively obscure until the discovery of penicillin in 1929 by Alexander Fleming. Since that time they have been extensively studied due to their bioactive nature and economic importance. Secondary metabolites impact our daily lives either as toxins or as beneficial compounds. Beneficial secondary metabolites made by species of Aspergillus include food additives such as kojic acid, antibiotics such as penicillin, and cholesterol reducing drugs such as lovastatin (Adrio and Demain, 2003; Endo et al., 1976).

In contrast, the repertoire of fungal secondary metabolites also includes harmful products known as mycotoxins. Aflatoxin (AF), originally discovered as the cause of Turkey-X disease, is a classic mycotoxin that has been studied for nearly 50 years (Nesbitt et al., 1962). It is a stable, small molecule that contaminates major grain and nut crops colonized by species of Aspergillus. It is both toxic and carcinogenic and exposure to aflatoxins has been associated with many veterinary toxic syndromes (Bressac et al., 1991; Hsu et al., 1991; Richard and Payne, 2003). Comprehensive studies have shown that AF is a risk factor for human hepatocellular carcinoma, especially in Asia and sub-Saharan Africa (Greenman et al., 2005). While death is an uncommon outcome of aflatoxicosis in humans, several deaths were attributed to aflatoxicosis as recently as 2004 (Nyikal et al., 2004). Because of its toxicity, over 100 countries restrict the content of AF in the food and feed supplies (van Egmond et al., 2007). A guideline of 20 parts AF per billion parts of food or feed substrate is the maximum allowable limit imposed by the US Food and Drug Administration for interstate shipment. In the United States, agricultural economic losses due to AF contamination of food and feed amount to $270 M annually (Richard and Payne, 2003).

Secondary metabolites owe their diverse structures to the many pathways involved in their biosynthesis, and the versatility of large multifunctional proteins, such as polyketide synthases (PKSs) and non ribosomal peptide synthases (NRPSs). There are several path-
ways leading to the production of secondary metabolites, with polyketides being the most abundant class (Keller et al., 2005). The mycotoxins AF and aflatoxin represent a classic polyketide and non ribosomal peptide, respectively, produced by species of Aspergillus (Duran et al., 2007; Keller et al., 2005).

AF remains one of the best-characterized fungal secondary metabolites. The genes for AF biosynthesis along with the pathway specific regulator aflR reside in a 70 kb DNA cluster near the telomere of chromosome 3 (Chang et al., 1993; Payne et al., 2006; Woloshuk et al., 1994; Yu et al., 2004). The seminal research on AF biosynthesis was conducted with Aspergillus flavus and Aspergillus parasiticus, the two species most commonly associated with AF contamination of agricultural commodities. AF is produced by several species of Aspergilli including A. flavus, A. parasiticus, Aspergillus nomius, Aspergillus pseudograminearum, and Aspergillus bombycis (Cary et al., 2005; Frisvad et al., 2005; Varga et al., 2003). Elegant research on production of the AF precursor sterigmatocystin (ST) with the genetic model Aspergillus nidulans, has contributed greatly to our understanding of the aflatoxin pathway, and has been essential for our understanding of the global regulatory mechanisms discussed later.

There are excellent reviews describing the chemistry and enzymology of AF biosynthesis as well as its genetic regulation (Bennett and Klich, 2003; Bhatnagar et al., 2003, 2008; Hicks et al., 2002; Keller et al., 2005; Yu et al., 2004). A goal of this review is to cover recent research on structural and functional genomics and put this information in context with current knowledge about aflatoxin biosynthesis.

1.1. Regulation of aflatoxin biosynthesis

The biosynthesis of AF has many layers of regulation, some of which are almost entirely specific to the pathway while others display a more global regulation of secondary metabolism. The use of genetics and tools from genomic sciences over the past decade (Bhatnagar et al., 2008) has expanded our knowledge of regulatory machinery beyond the AF pathway specific regulator AflR to show that the AF pathway is regulated by many mechanisms (Bok and Keller, 2004; Perrin et al., 2007). Many environmental factors control AF biosynthesis, including development, light (Calvo et al., 2002), carbon source, temperature, and pH (Fig. 1) (O’Brien et al., 2007; Price et al., 2005). We anticipate that many of these higher order processes regulating the biosynthesis of AF will also regulate the synthesis of other secondary metabolites as has been shown for those linked with the proteins VeA and LaeA discussed later in this review (Bayram et al., 2008; Duran et al., 2007; Keller et al., 2005).

2. Regulation within the aflatoxin biosynthetic cluster

The regulation of AF biosynthesis is complex and involves several interconnecting networks. At the risk of being somewhat arbi-
trary as to how we describe this complex regulation, we have divided the regulation into three parts, starting with regulation within the biosynthetic cluster. AF was one of the first fungal secondary metabolites shown to have all its biosynthetic genes organized within a DNA cluster. These genes, along with the pathway specific regulatory genes aflR and aflS, reside within a 70 kb DNA cluster. We now know that clustering of genes for secondary metabolism is a common feature (Keller and Hohn, 1996), although in some cases biosynthetic and pathway regulatory genes reside outside the cluster (Desjardins and Proctor, 2007). As will be discussed later, the position of genes within a DNA cluster has implications for its regulation. 

Research on A. flavus, A. parasiticus and A. nidulans has led to our current understanding of the enzymatic steps in the AF biosynthetic pathway, as well as the genetic organization of the biosynthetic cluster. A. nidulans does not produce AF but has all of the genes and enzymatic steps preceding the production of sterigmatocystin (ST). The AF and ST pathways appear to have a common biosynthetic scheme up to the formation of ST, and thus information gained from both pathways has been used to study AF regulation (Hicks et al., 2002; Yu et al., 2004). Interestingly, the order of the genes in A. nidulans differs from that of A. flavus. This shuffling of genes reflects the relative age of the cluster, based on estimates of protein similarities between A. nidulans, A. fumigatus and A. oryzae the cluster could be as old as 450 million years (Galagan et al., 2005). However, within section Flavi evidence exists suggesting that for at least 25 million years gene order, location of AFR binding sites, and intergenic distances have been conserved (Ehrlich et al., 2005).

2.1. Transcriptional regulators (AFR and AFS)

Two genes, aflR and aflS, located divergently adjacent to each other within the AF cluster are involved in the regulation of AF/ST gene expression (Fig. 1). The gene aflR encodes a sequence-specific DNA-binding bimetallic zinc cluster (Zn(II)$_2$Cys$_6$) protein, required for transcriptional activation of most, if not all, of the structural genes (Chang et al., 1993, 1995; Chang et al., 1999a,b; Ehrlich et al., 1998; Flaherty and Payne, 1997; Payne et al., 1993; Price et al., 2006; Woloshuk et al., 1994; Yu et al., 1996b). The aflR locus has been compared among strains within section Flavi, especially among isolates of AF producers such as A. parasiticus and A. flavus (Chang et al., 1995; Lee et al., 2006). These comparisons revealed differences in many promoter regulatory elements such as PacC and AreA binding sites (Ehrlich and Cotty, 2002; Lee et al., 2006). The aflR gene is also found in A. nidulans and A. fumigatus (Carbone et al., 2007). Despite clear differences in the sequence of AFR among A. nidulans and A. flavus, function is conserved: AFR from A. flavus is able to drive expression of the ST cluster in the ST cluster of the A. nidulans aflR deletion strain (Yu et al., 1996a).

To more carefully examine gene regulation by AFR, Price et al. (2006) examined gene expression in a wild-type and aflR deletion strain of A. parasiticus using a 5002 element microarray. They found some genes in the AF biosynthetic cluster that showed essentially no expression in the AflR strain, suggesting that AFR is absolutely required for their transcription. Other genes in the AF pathway showed some transcription, albeit lower than in the wild-type. The first column in the microarray expression summary of Fig. 1 shows six genes within the AF cluster whose expression appears to be more independent of AFR regulation. These results argue that the expression of these genes is either not tightly regulated, i.e., they have a moderate level of basal expression, or that their expression is modulated by regulatory factors other than AFR.

One explanation for these changes in expression for some genes but not others may be differences in AFR consensus binding sites. AFR binds to the palindromic sequence 5’-TCGN6CGA-3’ (also called AFR binding motif) in the promoter region of many of the structural genes in A. parasiticus, A. flavus, and A. nidulans (Fernandes et al., 1998). While a good predictor of AFR binding, the AFR binding motif is not sufficient to predict the regulatory efficiency, and an additional three AF genes for which AFR binding sites could not be demonstrated (Ehrlich et al., 1999b). Among these genes, aflQ, aflJ, aflP, aflM, aflK, aflR, aflG, aflS, aflR, and aflG all have predicted AFR binding in EMSA assays. All of these genes were differentially expressed between WT and the AflR mutant (Fig. 1), suggesting that AFR is required to activate their expression (Price et al., 2006). When differences in gene expression were examined across several transcriptional profiling studies these differences do not always correspond with the differences seen in the comparison between WT and the AflR mutant (Fig. 1). This suggests that there are other regulatory factors modulating AF gene cluster expression.

AF biosynthesis is also regulated by AFS (formerly aflj), a gene that resides next to aflr. The genes aflq and aflj are divergently transcribed, but have independent promoters. The intergenic region between them, however, is short and it is possible they share binding sites for transcription factors or other regulatory elements (Ehrlich and Cotty, 2002). The precise role of AFS in AF biosynthesis remains unclear. Meyers et al., (1998) initially characterized AFS through use of an aflj disruption strain and observed that mycertain levels for the aflatoxin biosynthetic pathway genes, aflQ, aflG, aflM, and aflP, remained unchanged but despite this pathway intermediates could not be converted to aflatoxin (Meyers et al., 1998). Chang (2004) showed that AFS binds AFR and that AFS is required for AFR activation in A. parasiticus (Chang, 2004). However, Du et al. (2007) were unable to show an absolute requirement for AFS in AF biosynthesis or activation of AFR. The roles of AFR and AFS were examined by studying the expression of pathway genes in transformants of A. flavus strain 649 that received the respective genes individually and in concert (Du et al., 2007). Strain 649 lacks the AF biosynthetic cluster but has the necessary upstream regulatory elements to drive the transcription of aflR (Prietó et al., 1996; Woloshuk et al., 1995). These studies showed that AFR is sufficient to initiate gene transcription of early, mid, and late genes in the pathway, and that AFS enhances the transcription of early and mid aflatoxin pathway genes. Roles for AFS have been suggested to be as diverse as aiding in transport of pathway intermediates to the interaction of AFS with AFR for altered AF pathway transcription. The observation that AFS binds to AFR argues that AFS modulates aflatoxin expression through its interaction with AFR (Chang, 2003).

Consistent with earlier data, microarray experiments have shown that aflatoxin biosynthesis is always associated with the expression of aflR. However, in experiments comparing conditions conducive and non-conducive for the biosynthesis of aflatoxin, the expression of aflq and aflj is not always significantly different (Fig. 1). This could be the result of inherently low levels of aflq and aflj transcripts, making their quantification difficult. Another possible explanation is that regulation of AF cluster gene expression is complex, and factors other than transcriptional levels of aflq and aflj are important in its regulation. AFS transcript is thought to be dependent on aflj (Du et al., 2007; Ehrlich et al., 1999b; Price et al., 2006). This presents the argument that many of the non-conducive conditions described in Fig. 1, where aflq and aflj do not change but most of the pathway is strongly down-regulated, are independent of regulation by AFR or AFS.
2.2. Other possible cluster regulatory mechanisms

There is no direct evidence for the involvement of regulatory RNAs in AF biosynthesis, however a Naturally occurring Antisense Transcript (NAT), aflRas, was identified at the same time as aflR (Woloshuk et al., 1994). Recently, a global analysis of NAT expression in A. flavus during its colonization of maize seeds incubated at 28 or 37 °C identified 32 of 352 putative cis NATs whose expression was directly or inversely correlated with expression of their respective sense gene (Smith et al., 2008). The expression of one of these NATs was inversely correlated with the expression of aflD. Thus, a higher expression of the NAT for aflD was associated with less aflatoxin production. While this does not prove a role for a NAT in the regulation of aflatoxin biosynthesis, it is tempting to speculate that they may have some role in modulating gene expression as has been shown for the NAT for frq, a gene involved in circadian rhythm in Neurospora crassa (Kramer et al., 2003). In strains that did not accumulate the NAT for frq, the internal clock was delayed and unable to be reset by light. Kramer et al. (2003) observed that maximal accumulation of the NAT for frq occurred directly opposite of the sense transcript for frq which reaches highest levels in the dark.

In addition to the binding sites for AflR, there are binding sites within the cluster for other transcriptional factors that may play important roles in transcriptional regulation of the AF cluster. A novel cAMP-response element, CRE1, site has been studied specifically in the aflD (nor-1) promoter of A. parasiticus. The CRE1 site and the regulation of aflatoxin by cAMP and glucose concentrations are discussed in further detail in Sections 3.2 and 4.2. In addition AeaA binding sites (Section 4.4) and PacC binding sites (Section 4.5) are found in promoter regions throughout the AF biosynthetic cluster and may impart a role in the regulation of AF biosynthesis by nitrogen and pH respectively.

3. External control of the aflatoxin cluster and global regulation of secondary metabolism

Although there is no known role for AF in the ecology of the fungus, AF biosynthesis is tightly regulated by environmental and development cues. The entire signaling network for processes regulating AF biosynthesis is unclear, but components of these networks have been characterized (Yu and Keller, 2005). Fig. 2 summarizes our current understanding of the regulatory networks that transmit the environmental and development signals acting on AF biosynthesis. One regulatory element shown in Fig. 2, LaeA, appears to be involved in a novel type of global regulation and will be discussed in Section 3.4.

3.1. Heterotrimetic G-protein signaling and PKA

Early research on AF biosynthesis uncovered an association between development and AF production (Kale et al., 1996). This was not surprising as a tight link between secondary metabolism and development in bacteria had been established with the bldA gene controlling aerial mycelia, conidia, and secondary metabolites (De-main, 1992). In Aspergillus, several observations linked a fluffy phenotype to loss of AF/ST production. The available well-characterized fluffy mutants in A. nidulans were instrumental in the discovery of a signal transduction pathway regulating both

![Fig. 2. Signaling pathways regulating AF biosynthesis and their ties to growth and development. Dashed lines indicate hypothesized pathways. Adapted from Calvo (2008), Seo and Yu (2006), Keller et al. (2003), Yu and Keller (2005), and Shimizu et al. (2003).]
conidiation and ST/AF biosynthesis. Proteins identified as belonging to this signaling pathway include FhaA, an RGS (Regulator of G-protein Signaling) protein, FlgG, an early acting development regulator (Dohlman and Thorner, 1997; Hicks et al., 1997), FadA, an oncogene; and a factor of a heterotrimeric Ras (Som and Kolaparthi, 1994). Examination of both FadA and PhiA inhibited conidiation and ST biosynthesis in A. nidulans, whereas FhaA suppresses FadA activation (Calvo et al., 2002). FlgG activation of gene expression is dependent on FlgG (Hicks et al., 1997). FadA is thought to activate adenylate cyclase, resulting in an upregulation of FadA which is able to inhibit AFR activity by preventing nuclear localization through phosphorylation of AflR (Hicks et al., 2001). Later experiments have also implicated the Gj (Stab) and Gy (GpGA) subunits of this heterotrimeric G-protein and PhiA, a phosphokinase-like protein that is an activator of Gj-g-mediated signaling, as positive regulators of Afl expression needed for ST biosynthesis (Fig. 2) (See and Yu, 2006).

Although the initial characterization of the heterotrimeric G-protein signaling pathway that regulates ST production and conidiation was established in A. nidulans, evidence shows that a similar signaling pathway exists in both A. parasiticus and A. flavus to regulate AFR (Hicks et al., 1997; McDonald et al., 2005; Roze et al., 2004a; Tag et al., 2000; Yu and Keller, 2005). Rather than responding to a specific external stimulus, this G-protein mediated pathway appears to play a role in regulation of AF as a function of vegetative growth and cellular development (See and Yu, 2006; Yu et al., 1996c).

Recent genomic evidence has revealed G-protein coupled receptors (GPCRs) to be alpha subunits conserved among A. oryzae, A. fumigatus, and A. nidulans, but G-proteins themselves to be more divergent (Lafon et al., 2006). A. oryzae and A. flavus are nearly identical at the genomic level, with no detected differences in G-proteins and their receptors. These receptors enable the fungus to respond to various environmental and developmental stimuli through heterotrimeric G-protein signals. The GPCR tied to the FadA-mediated signaling to control AF and ST still remains to be discovered. A total of nine GPCRs (GprA-I) in A. nidulans have been grouped into five different classes, class I and II (GprA and GprB) analogous to yeast pheromone receptors, class III (GprC-E) possibly involved in carbon source sensing, class IV (GprF and GprG) and class V (GprH and GprI) which are proposed to play a role in cAMP signaling (Han et al., 2004; Lafon et al., 2006). Six of the nine identified GPCRs have been disrupted in A. nidulans, with GprD proposed to control sexual development and PKA, but not through FadA (Han et al., 2004). No reports have been made on whether any of these receptor knockout mutants affected AF or ST production.

3.2. cAMP signaling

Cyclic-AIMP is an important signaling molecule. The synthesis, by adenylate cyclase, and degradation, by phosphodiesterase, of cAMP is regulated by various G-protein signaling cascades. The addition of cAMP to Aspergillus growth media causes an increase in AF biosynthesis (Tice and Buchanan 1982). Roze et al. (2004b) measured PKA activity as a function of exogenously applied cAMP and found total PKA activity to increase in vitro with added cAMP, however when cultures were grown in the presence of 5mM cAMP, basal and total PKA activity levels did not increase, but instead appeared to decrease (Roze et al., 2004a). It appears that high physiologically relevant cAMP levels are able to inhibit AFR through an increase in phosphorylation by cAMP dependent PKA activity. However, greatly exceeding these levels inhibits cAMP dependent PKA activity. cAMP dependent PKA activity is closely examined in A. flavus (cAMP-response element) site necessary for aflD induction by cAMP in the aflD promoter that was dependent on AFR (Roze et al., 2004b). This work examined how close the CRE1 site must be to an AFR binding site for p32 to interact with and stimulate AFR binding.

It has been shown that PKA and cAMP levels are regulated through a negative feedback loop affected by the phosphodiesterase in Cryptococcus neoformans (Hicks et al., 2005). Phosphodiesterases have also been implicated to control cAMP levels in Saccharomyces cerevisiae with this regulation postulated to be controlled by stress responses triggered by the Ras-cAMP pathway (Park et al., 2005). Recent evidence has shown that a potential phosphatidylinositol (PI) 3-kinase pathway regulates aflatoxin through control of cAMP levels and PKA activity (Lee et al., 2007). Lee et al. (2007) came to these conclusions after observing that wortmannin, a secondary metabolite commonly made by Penicillium and Fusarium species, inhibited aflatoxin biosynthesis and promoted cAMP activity (Lee et al., 2007). No effect on the expression of other genes was reported specifically but it would be interesting to know whether wortmannin affects AflR activity (Lee et al., 2007). Presumably, treatment with wortmannin would activate PKA causing phosphorylation of AFR and inhibition of aflatoxin biosynthesis. This research has led to a proposed regulatory pathway where PI 3-kinase is able to activate phosphodiesterase, causing a decrease in cAMP levels and correspondingly a decrease in PKA activity (Fig. 2).

3.3. Ras family GTPase signaling

The Ras family is a highly conserved group of GTPases composed of six sub-families: Ras, Rho, Ran, Rab, Arf, and Kir/Rem. Rad (Ehrhardt et al., 2002; Reuther and Der, 2000). This family of GTPases composes a key intracellular signaling network having the ability to transmit extracellular signals to the nucleus and control transcription in response to certain stimuli (Roze et al., 1999). GTPases are in their active signaling state when associated with GTP and inactive state when bound to GDP. The Ras-cAMP pathway mediates global responses connected to stress tolerance (temperature, osmolality, oxidative, and others) and nutrient sensing (Park et al., 2005).

In comparison to heterotrimeric G-protein signaling in Aspergil- lus, relatively little work has been reported on the possibilities of Ras-like GTPase involvement in AF biosynthesis. Examination of the A. flavus genome using HMM analysis for Ras superfamily proteins has predicted a total of 31 potential Ras family GTPases. Ras members and their regulators are known to be highly similar within and across species (Cetkovic et al., 2007; Colman and Field, 2005). Amino acid similarities between Ras family members within a species are usually between 40% and 95%, yet most if not all have their own unique set of biological functions (Ehrhardt et al., 2002; Reuther and Der, 2000).

The A. nidulans rasA gene encodes a homolog of the human ras proto-oncogene (Som and Kolaparthi, 1994). Examination of A. nidulans mutants overproducing a dominant active form of RasA...
(RasA<sup>17V</sup>) supports a signaling role for this protein in spore germination and carbon sensing (Osherov and May, 2000; Som and Kolaparthi, 1994). Further, RasA has been shown to regulate ST biosynthesis by both transcriptional and post-transcriptional control of aflR (Shimizu et al., 2003). Interestingly, RasA transcriptional control of aflR is PkaA independent, but the post-transcriptional control is at least partially regulated by PkaA. How this partial regulation works is unclear, but RasA post-transcriptional control is not affected by the ABR phosphorylation state at PKA sites. This has been shown by use of an A. nidulans phosphorylation mutant AFR<sup>p32,381,382</sup> overexpressed in a dominant active (permanent GTP bound state) RasA<sup>17V</sup> background (Shimizu et al., 2003). Despite many hypotheses, the pathway for transcriptional control of aflR by RasA is not known. Other non-PKA mediated systems of AFR phosphorylation or other post-translational modifications also remain unknown.

Regulation of GTPase activity and its relation to PKA is a complicated multi-component process and could explain the partial mediation of Ras signalling by AFR. Crosstalk between Ras and cAMP could prove to be very important in tying together many of the described signals regulating AF and has potential involvement with results such as the increased levels of AF resulting from exogenously applied cAMP in culture. Interestingly, PD98059, an inhibitor of a mitogen-activated protein kinase (MAPK), inhibits PD98059, an inhibitor of a mitogen-activated protein kinase, inhibited the stimulatory effect of PD98059 on cAMP induction, but not AF (Roze et al., 2004). Ras proteins are not limited to activation of PKA but can activate numerous other kinases often tied to stress-sensing pathways needed for survival (Thevelein and de Winde, 1999). While the specific kinase responsible is not known, it is clear that another factor must be involved in Ras control of AFR, as shown by the aforementioned deletion of PKA phosphorylation sites in ABR which cannot restore AF in dominant active Ras mutants (Shimizu et al., 2003).

### 3.4. Chromosomal remodeling and silencing

LaeA, for “loss of aflR expression”, controls the expression of genes involved in ST production in A. nidulans (Bok and Keller, 2004; Butchko et al., 1999). Deletion of laeA results in loss of aflR gene expression and ST/AF synthesis in A. nidulans and A. flavus. Furthermore, LaeA was hypothesized to be a global regulator of secondary metabolism when it was shown to regulate penicillin and lovastatin biosynthesis in A. nidulans (Bok and Keller, 2004). Using microarray data from AFR deletion and overexpression mutants, Bok et al. (2006) mapped expression patterns to genomic regions to identify new secondary metabolite gene clusters regulated by LaeA. These data supported the role of LaeA as a global regulator of secondary metabolism (Bok et al., 2006). The mechanism of this control is not currently known; because of its homology to arginine and histone methyltransferase, LaeA may function through chromatin remodeling of metabolic gene clusters (Bok and Keller, 2004; Yu and Keller, 2005; Zhang et al., 2004). In A. flavus LaeA is clearly a strong regulator of aflR and aflS and has inhibitory effects for every gene in or around the AF cluster. The same is true for AFR mutants, Bok et al. (2006) mapped expression patterns to genomic regions to identify new secondary metabolite gene clusters regulated by LaeA (Shimizu et al., 2003). Furthermore, AFR-mediated AFR activity through use of an A. parasiticus aflR deletion strain, AFS10. As noted by the authors, this conclusion may be complicated by the presence of a
A. flavus strain 649, which contains the aflatoxin. From what the author's called "silent biosynthetic pathways" (Williams et al., 2008). While many of these new compounds are known inhibitors to histone deacetylases and methyltransferases, A. flavus and repetitive DNA sequence.

The biosynthesis may be due to its lack of proximity to the telomere that the lack of an effect of the hdaA deletion on terraquinone A expression of metabolite clusters. Shwab et al. (2007) suggested that the lack of an effect of the hda deletion on terraquinone A biosynthesis may be due to its lack of proximity to the telomere and repetitive DNA sequence.

Complementary to the work on histone acetyltransferase, a histone deacetylase (HDAC) gene (Fig. 3), hdaA, was deleted in A. nidulans. Deletion mutants of hdaA showed increased accumulation of secondary metabolites and gene expression for both the ST and penicillin gene clusters (Shwab et al., 2007). The hdaA deletion was also able to restore ST as well as penicillin in a ΔaflR mutant. Additionally, the terraquinone A gene cluster was not affected by deletion of hdaA alone. Results from Roze et al. (2007a) also displayed a dependence on physical gene order for activation that results in a temporal pattern of gene expression corresponding with histone H4 acetylation, where the "early", "middle", and "late" stages of the biosynthetic pathway are expressed in an order corresponding to their stage in the biosynthetic pathway. Interestingly, the gene order of the ST cluster in A. nidulans is not identical to A. flavus or A. parasiticus and a predicted sequence for the CRE1 binding protein, p32, is not present in A. nidulans (Roze et al., 2007a). This difference between species could prove to be important in explaining regulatory processes not observed in A. nidulans, such as the regulation of aflatoxin biosynthesis by temperature, which is discussed in Section 4.7.

Not surprisingly, many of the secondary metabolites in fungi appear to be under epigenetic control. In a recent study in which A. flavus and other fungi were grown in the presence of several known inhibitors to histone deacetylases and methyltransferases, a wide variety of new compounds were observed to accumulate (Williams et al., 2008). While many of these new compounds are from what the author's called "silent biosynthetic pathways" which never appear to be expressed, it is clear from previously described work that epigenetics is capable of playing a major role in regulation of actively expressed secondary metabolites such as aflatoxin.

Chromosomal silencing/remodeling may also be involved in the silencing of aflatoxin pathway genes in diploids formed between A. flavus strain 649, which contains the afl-1 mutation, and wild-type strains (Pietro et al., 1996; Woloshuk et al., 1995). Papa (1979) described the afl-1 mutation in 649 and it remains the only known dominant mutation for AF production. Further characterization of this strain showed that it is missing 317 kb of chromosome III, including the known genes for aflatoxin biosynthesis (Pietro et al., 1996; Woloshuk et al., 1995). In addition, 939 kb of chromosome II is present as a duplication on chromosome III in the region that originally contained the aflatoxin gene cluster (Smith et al., 2007). Inhibition of AF biosynthesis in diploids formed with 649 and aflatoxin producing strains is not due to the activation of a repressor of afl or the result of currently described silencing mechanisms in fungi such as repeat-induced point mutation (RIP), quelling (RNAs), or transvection (Smith et al., 2007). The silencing appears to be restricted to the aflatoxin cluster; genes on either side of the cluster are expressed in the afl-1 diploids. Interestingly, diploids formed between a wild-type strain and a transformant of 649 containing an ectopic copy of afl confer AF. Thus the presence of an additional copy of afl not in the biosynthetic cluster prevents silencing of the genes in the cluster. This mechanism may not be related to the ΔaflRΔ mutants but it is interesting that an ectopic copy of afl restores AF both in silenced diploids and in ΔaflR mutants.

4. Environmental conditions regulating aflatoxin biosynthesis

Several environmental and cultural conditions modulate AF biosynthesis including light, temperature, pH, nitrogen, carbon source and metals (Calvo et al., 2004; Luchese and Harrigan, 1993; Price et al., 2005). Table 1 summarizes the effect of known environmental and nutritional conditions affecting AF biosynthesis. For a more detailed discussion refer to the following reviews: (Bhatnagar et al., 2006; Cotty and Jaime-Garcia, 2007; Luchese and Harrigan, 1993; Payne and Brown, 1998). An understanding of how these factors impact AF biosynthesis is critical as it will be important in determining the role of AF in the ecology of the producing organism, and it may identify target sites for control of AF formation. Unfortunately, the regulatory networks involved in sensing and transmitting environmental and nutritional stimuli are not well understood. Price et al. (2005) examined the effect of four cultural and environmental conditions on gene transcription specific to the AF pathway. They found temperature to have the most profound effect followed by pH, nitrogen source, and carbon source (Price et al., 2005). Schmidt-Heidt et al. (2008) examined temperature and water activity in relation to growth on secondary metabolism genes in several fungal species, including the AF cluster in A. parasiticus. They observed that conditions of intermediate stress to the organism were more favorable for mycotoxins (Schmidt-Heidt et al., 2008). Calvo et al. (2004) showed light to affect the transcription of several genes, including genes in the AF gene cluster and genes putatively involved in the development of sclerotia in A. flavus (Calvo et al., 2004). The effect of nutrition and environmental conditions on the transcription of the AF cluster is summarized in Fig. 1.

4.1. Light mediated regulation of aflatoxin biosynthesis

Of the many environmental conditions that modulate AF biosynthesis, the effect of light is becoming among the best-characterized. Production of cyclopiazonic acid, aflatrem, and AF by A. flavus is regulated by VeA, a gene necessary for sclerotial formation (Kato et al., 2003). Velvet A protein (VeA) has been studied in many Aspergillus species including A. flavus, A. parasiticus, and A. nidulans and has been shown by these researchers to regulate, in response to light, qfR expression as well as formation of the environmentally resilient structures, cleistothecia in A. nidulans and sclerotia in A. flavus and A. parasiticus. Recent work has begun to uncover how VeA may function at the protein level. VeA is included in Fig. 2 to show that it is needed for qfR expression and linked to development and condiation, these two are likely not independent processes. Purschwitz et al. (2008) has shown that FbA, a phytochrome in A. nidulans that responds to blue and red light, interacts with VeA. They found that red light and white light inhibited mycotoxin biosynthesis whereas blue light had a stimulatory affect.

<table>
<thead>
<tr>
<th>Conditions</th>
<th>Conducive (+)</th>
<th>Non-conducive (−)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carbon source</td>
<td>Simple sugars</td>
<td>Complex sugars</td>
</tr>
<tr>
<td>Nitrogen source</td>
<td>Reduced</td>
<td>Oxidants</td>
</tr>
<tr>
<td>Oxidative stress</td>
<td>Oxidants</td>
<td>Anti-oxidants</td>
</tr>
<tr>
<td>Temperature</td>
<td>&gt;35 °C</td>
<td>&lt;30 °C</td>
</tr>
<tr>
<td>pH</td>
<td>Acidic (pH 4.5)</td>
<td>Basic (pH 8)</td>
</tr>
</tbody>
</table>
(Purschwitz et al., 2008). For a detailed review of VeA see the recent review by Calvo (2008). Recently, Bayram et al. (2008) discovered through tandem affinity purification that VeA functions in a complex consisting of another velutin-like protein, VelB, and a global regulator of secondary metabolism, LaeA. These interactions allow for the coordinated control of both development and secondary metabolism by light (Bayram et al., 2008).

Microarray analysis of an A. flavus veA deletion strain has confirmed that VeA regulates the transcription of several genes in the fungus. Cary et al. (2007) used a 5002 element DNA microarray to assess transcriptional regulation by VeA by comparing the deletion strain to a wild-type strain. A total of 136 genes were differentially expressed between the two strains (Cary et al., 2007). Included in these differentially expressed genes were gene products that were regulated in other array experiments by O’Brian et al. (2007) and Wilkinson et al. (2007) such as the AAA family ATPase (clone NAF-BED02TV), two different condensation-specific proteins (clones NAG-DC44TV and NAFEY43TV), and a Rieske 2Fe-2S family-like protein (clone NACCL73TV). Many of the genes in the AF cluster were similarly affected when compared to prior array experiments (Fig. 1). The gene for the transcriptional regulator ABR was not found to be differentially expressed in the array study, however quantitative PCR results indicated a significant decrease in aqr expression in the VeA mutant compared to the wild-type strain; this would be expected if VeA is controlling AF biosynthesis through activation of LaeA (Bayram et al., 2008; Cary et al., 2007). An interesting experiment to further reiterate that VeA functions through LaeA would be to see if the haad deletion, described in Section 3.4, could partially relieve repression of the ST cluster in a veA deletion background as done with previously with laeA.

4.2. Regulation of aflatoxin by carbon source

While many investigators have examined the effect of carbon sources on AF biosynthesis (Abdollahi and Buchanan, 1981a, 1981b; Buchanan, 1984; Davis and Diener, 1968; Mateles and Adye, 1965; Wiseman and Buchanan, 1987), it is unclear if carbon source directly regulates AF biosynthesis or modulates its synthesis through general metabolism. Unlike the production of most secondary metabolites that are repressed by simple sugars, production of AF biosynthesis is stimulated by simple glucose containing sugars, such as sucrose, or derivatives of glucose, such as fructose (Davis and Diener, 1968). A high degree of glucose utilization has been correlated with production of AF (Applebaum and Buchman, 1979; Shih and Marsh, 1974).

In an attempt to better understand the source of carbon for AF biosynthesis, researchers have studied rates of labeled acetate incorporation into AF relative to glucose (Hsieh and Mateles, 1970). Interestingly, it was found that acetate preferentially incorporated into AF during conditions where the concentration of glucose was 4-fold higher than acetate. This result suggests that the aflatoxins are synthesized extracellularly from glucose-derived acetyl coenzyme A (Hsieh and Mateles, 1970; Shantha and Murthy, 1981). This conclusion is supported by additional studies measuring the flux of the tricarboxylic acid cycle (TCA) intermediates. TCA intermediates were found to inhibit AF production and an active functioning of the TCA cycle was linked to lack of AF production (Wilkinson et al., 2007). These results suggest that introduction of intracellular acetyl-CoA concentrations through complete oxidation of Acetyl-CoA via the TCA cycle limits AF production (Buchanan and Ayres, 1977; Gupta et al., 1977; Shantha and Murthy, 1981).

Interestingly, carbon sources non-conducive for AF biosynthesis appear to result in less repression of the cluster genes as compared to non-conducive temperature, nitrogen source, and pH (Price et al., 2005). Similar results were obtained when AF production was examined in A. flavus grown with and without sucrose in yeast extract media (Wilkinson et al., 2007). While carbon source clearly has many effects on gene/protein expression, a lack of strong gene down-regulation within the AF cluster could suggest that carbon source is not having a specific effect on AF biosynthesis. VeA has also been shown to alter hdaA and ahlR transcription levels (Price et al., 2007). One possibility is that instead of a direct regulatory effect there is an alteration in the metabolic pool of precursor products needed for unique secondary metabolites such as AF. One possible role for glucose in AF biosynthesis is its influence on cAMP signaling/catabolite repression. Glucose and cAMP levels are typically inversely related; however in Aspergillus it was shown that cAMP levels were high during exponential growth, a period of high glucose utilization, and then lowered later during stationary phase, with the majority of AF accumulating during a period where cAMP levels are shifting (Applebaum and Adye, 1965; Wiseman and Buchanan, 1987), it is unclear if carbon source is not having a specific regulatory effect on the AF pathway

4.3. Lipids-key precursors and important signaling molecules

The AF polyketide synthase uses hexanoate, formed by a fatty acid synthase encoded by two genes within the AF biosynthetic cluster, as the starter unit, unlike most PKSs which use acetyl-CoA (Yabe et al., 2003). Early studies revealed that lipid synthesis was highly correlated with AF production (Shih and Marsh, 1974). The first step in fatty acid biosynthesis is the conversion of Acetyl-CoA to Malonyl-CoA by Acetyl-CoA carboxylase. Surprisingly, evidence has shown that Malonyl-CoA is needed for biosynthesis of AF (Dutton, 1988). In corn, A. flavus is preferentially associated with the germ region of infected kernels (Fennell et al., 1973; Lillehoj et al., 1974); this lipid rich germ tissue also appears more amenable to AF production than the whole kernel (Mellon et al., 2005).

Oxylipins (also called psi factors) are signaling molecules derived from fatty acids that have been shown to regulate developmental and pathogenic processes as well as secondary metabolism in A. nidulans (Fig. 1) (Tsitsigiannis and Keller, 2006, 2007). Three genes (pp0A, pp0B, and pp0C) encoding fatty acid oxygenases were found in the A. nidulans genome and predicted to be responsible for production of these psi factors. Deletion mutants were made for each of these three genes (Tsitsigiannis et al., 2005). Tsitsigiannis et al. (2005) presented evidence that the asex- ual conidiation regulatory gene hbrA was clearly regulated by these genes and additionally showed evidence that veA was possibly displaying a different temporal pattern of gene expression in a triple deletion mutant background of pp0A, pp0B, and pp0C. VeA has also been shown to alter hbrA transcripts however it is not clear whether VeA must function through oxylipins to exhibit this regulation (Fig. 2) (Calvo, 2008). Interestingly, certain ppo deletion mutants were found to have increased expression of aqr and ctsA (aqr equivalent) while other ppo deletion strains showed a decrease in expression of these genes in the ST pathway. It was shown that this regulation of aqr linked to psi factors was not by LaeA (Tsitsigiannis and Keller, 2006). Tsitsigiannis and Keller (2006) proposed that aqr inhibition by ppoA and ppoC products may function in a PKA mediated manner since PKA has been shown to inhibit ST and increase penicillin accumulation. They also propose that this is regulated by a C-protein coupled receptor (GCPR) signaling cascade, possibly involving the heterotrimeric Go-subunit FadA (Tsitsigiannis and Keller, 2006). Clearly use of previously described ABR PKA phosphorylation mutants and FadA mutants in a delta-ppoA, ppoC
background would provide excellent evidence as to whether there is a link between oxylipins and IRA or G-Protein mediated regulation of the ST/AF cluster.

4.4 Regulation of aflatoxin by nitrogen source

Nitrogen source has been widely reviewed as important in the biosynthesis of AF (Luchese and Harrigan, 1993; Payne and Brown, 1998). Various sources of nitrogen have been tested and it has been observed that organic nitrogen sources are superior for AF biosynthesis (Davis et al., 1967) and that nitrate inhibits AF production (Kachholz and Demain, 1981). However, ST production in A. nidulans is not regulated similar to AF in response to nitrogen source, with nitrate inducing ST and ammonia causing a reduction (Feng and Leonard, 1998). Among various amino acids, proline has been reported as supporting the most AF. Asparagine, casein, and ammonium sulfate also support robust AF production (Payne and Hagler, 1983). Nitrate has been shown to result in decreased expression of genes involved in AF biosynthesis (Fig. 1) (Feng and Leonard, 1998; Price et al., 2005). Worth noting is the presence of AreA binding sites in the intergenic region between aflR and aflS. Results have shown that the AreA protein from A. parasiticus is able to bind to this region (Chang et al., 2000b). Nitrate resulted in active AreA binding to this region, and actually increased expression of aflS.

In a comparison between divergent populations of A. flavus from North America and West Africa nitrate regulation of aflatoxin biosynthesis was found to differ. North American strains were found to be less inhibited of AF biosynthesis due to the presence of additional AreA binding sites in these strains in the intergenic region between aflR and aflS when compared to West African strains (3 sites compared to 6 sites) (Ehrlich and Cotty, 2002). Expression of the regulatory gene aflR was found to be 2-6-fold higher in nitrate media for North American strains examined (Ehrlich and Cotty, 2002). Interestingly, researchers have tried to overcome nitrate repression of AF by increasing expression of aflR, but not aflS. Initially, Chang et al. (1995) found that insertion of an additional copy of aflS into A. parasiticus was able to restore expression of cluster genes when grown in nitrate; however, AF levels were not reported. Flaherty and Payne (1997) used a constitutively-expressed aflR which resulted in a similar restoration of transcripts for aflD, aflM and aflS. However, they found that AF was still not made when nitrate was the sole nitrogen source (Flaherty and Payne, 1997). These results show that while gene expression is controlled by nitrate, the actual biosynthesis of AF is inhibited by another control point of the pathway. One hypothesis was that a change in redox potential caused by nitrate increasing the activity of mannitol dehydrogenases (Niehaus et al., 2007). The intracellular cytoplasmic pH does not appear to fluctuate in fungi so regulation must come from ambient pH (Parton et al., 1997).

It is not clear whether PacC sites have a specific role in the control of aflatoxin biosynthesis by pH. PacC binding sites have been identified in the promoter of aflR so it would have been expected that alkaline pH would enhance AF cluster gene expression (Ehrlich et al., 1999a). In addition to a down-regulation of expression for AF cluster genes by pH, metals cause an increase in transcription among many genes, several of which were involved in cell growth and division. The expression of one gene encoding an alcohol dehydrogenase was measured by RT-PCR and found to increase in metal treated cultures of both F. graminearum and A. flavus (Cuero and Ouellet, 2005). Alcohol dehydrogenase, aflH (aflhA) is a gene located in the AF biosynthetic gene cluster needed for the conversion of 5-hydroxyaverantin to averufin (Chang et al., 2000a). Cuero and Ouellet (2005) measured a gene they called aflh1, but it is not clear whether this is the cluster gene or another alcohol dehydrogenase in the fungus; if it is the cluster gene then the transcription of at least two genes in the cluster aflR and aflS are stimulated by the addition of metals.

4.5. Regulation of aflatoxin by pH

Acidic conditions are more favorable for the biosynthesis of AF than alkaline conditions. Keller et al. (1997) showed that transcripts for cluster genes were down-regulated in basic conditions relative to acidic (Keller et al., 1997). The regulation of AF by pH stems from the alkaline transcription factor PacC and appears to function inversely to the role of pH regulation on penicillin production. PacC is a zinc finger transcription factor that regulates gene expression depending on ambient pH (Tilburn et al., 1995). PacC is synthesized as an inactive precursor requiring a two-step proteolytic activation for its activity under alkaline conditions. At acidic pH it adopts an inactive conformation through an intramolecular inter- action in the C-terminus of the protein (Drez et al., 2002; Penas et al., 2007). PacC has been shown to negatively regulate AflR through phosphorylation (Shimizu et al., 2003; Shimizu and Keller, 2001) and is discussed in Section 3.1.

4.6. Regulation of aflatoxin by metals/trace elements

It has been known since the 1960’s that metals are important factors needed for high levels of AF accumulation in synthetic med- ia, especially zinc (Davis et al., 1967; Mateles and Adyar, 1965; Nesbitt et al., 1962). Early on it was observed that corn steep liquor resulted in a significant increase in AF production for fungi grown in synthetic liquid media (Schoenfeld, 1966). This increase was thought to come from a nutrient, possibly a metal, in the corn li- quor that was needed for efficient synthesis of AFs. Several studies were done on important metals to analyze metal content in relation to Aspergillus infection and AF production. Leliehoj et al. (1974) found that levels of metals stayed the same or increased when corn seeds were infected with A. flavus. They demonstrated that biologically available trace elements were limited in corn seed because of phytate levels, and found that the addition of the metals zinc or lead caused significant increases in AF production in corn germ (Leliehoj et al., 1974). Metal availability appears critical; addition of EDTA to A&M media metal mix inhibits AF biosynthesis, presumably due to the chelation of available metal ions (unpublished data). Researchers also noted that copper has an inhibitory effect on the production of AF (Leliehoj et al., 1974; Marsh et al., 1975).

Cuero et al. (2001) explored the molecular significance of met- als and AF. They found that a mixture of copper, iron, and zinc caused a significant increase in total RNA, biomass, aflD (wtM) transcription, and AF (Cuero et al., 2005). A subtractive cDNA expression library was used in Fusarium graminearum to show that metals cause an increase in transcription among many genes, sev- eral of which were involved in cell growth and division. The expression of one gene encoding an alcohol dehydrogenase was measured by RT-PCR and found to increase in metal treated cul- tures of both F. graminearum and A. flavus (Cuero and Ouellet, 2005). Alcohol dehydrogenase, aflH (aflhA) is a gene located in the AF biosynthetic gene cluster needed for the conversion of 5'-hydroxyaverantin to averufin (Chang et al., 2000a). Cuero and Ouellet (2005) measured a gene they called aflh1, but it is not clear whether this is the cluster gene or another alcohol dehydrogenase within the fungus; if it is the cluster gene then the transcription of at least two genes in the cluster aflR and aflS are stimulated by the addition of metals.
Research into the effects of temperature on AF production is ongoing. Early studies showed that AF was produced maximally at 24°C, but was not produced at temperatures lower than 18°C or higher than 35°C (Schindler et al., 1967). Diener and Davis (1957) showed by Northern analysis that the AF polyketide synthase gene of *A. parasiticus* was expressed at 27°C but not at 37°C (Feng and Leonard, 1995). Another AF pathway gene, aflP, was similarly transcribed in *A. parasiticus* at 29°C but not at 37°C (Li and Chu, 1998). Temperature shift studies in which cultures are moved from a conducive temperature to a non-conductive temperature, or vice versa, appear to have no lasting effect on AF biosynthesis, indicating that the inhibition of AF is caused by temperature and not the lack of a needed precursor molecule nor presence of an inhibitory molecule (Georgianna and Payne, unpublished).

Microarray data from a 5002 element array revealed that among 144 differentially expressed genes between 28 and 37°C, a majority of the AF biosynthetic cluster genes are repressed at the higher temperature, despite transcript levels of *aflR* and *aflS* remaining constant (Feng and Leonard, 1995). Another AF pathway gene, aflD, was similarly transcribed in *A. parasiticus* at 29°C but not 37°C (Li and Chu, 1998). Temperature shift studies in which cultures are moved from a conducive temperature to a non-conductive temperature, or vice versa, appear to have no lasting effect on AF biosynthesis, indicating that the inhibition of AF is caused by temperature and not the lack of a needed precursor molecule nor presence of an inhibitory molecule (Georgianna and Payne, unpublished).

Current data argue that inhibition of AF biosynthesis at 37°C is not due to an inactivation of AF by phosphorylation. The localization of AF using a strain expressing a GFP::AFR fusion protein was recently examined and AflR was found to be localized in the nucleus at both 28 and 37°C (Georgianna and Payne, unpublished data). These data effectively rule out phosphorylation by PKA since the localization does not change and AflR is unable to enter the nucleus when phosphorylated (Shimizu et al., 2003). It seems likely that temperature regulation of AF biosynthesis is more complex than simple being inactive at 37°C; particularly the work by Kim et al. using *A. flavus* mutants has shown that protein stability/folding at higher temperatures is also not affected (Kim et al., 2005). Interestingly, differential expression for several *APK* genes has been found in the aforementioned temperature comparison arrays. One pitfall of the yeast studies is that for many of the genes identified only one strain has actually been shown to have an effect on AF production. A knockout of a superoxide dismutase gene in *A. flavus* has been shown to decrease AF production (He et al., 2007). One of the most important recent discoveries stemming from the work by Kim et al. (2005) was that the anti-oxidant response transcription factor Yap1 has an apparent role in regulation of AF biosynthesis (Reverberi et al., 2008). Reverberi and coworkers (2008) found that deletion mutants of yap1 in *A. parasiticus* had an increased accumulation of AF. It is not clear whether the increase in AF can be directly related to deletion of Yap1 or a result of increased reactive oxygen, however, it was proposed that there may be Yap1 binding sites in the *aflR* promoter. *aflR* was induced earlier in yap1 deletion strains compared to wild-type and a known binding site for the human ortholog of Yap1, AP-1, is found in the promoter region of *aflR* (Reverberi et al., 2008).

5. Outlook for genomics

Functional genomics promises to provide important information on metabolites with no known function. For example, gene expression experiments have been valuable in characterizing regulatory mechanisms controlling the AF cluster (Fig. 1). Further analyses of these types of array experiments may lead to identification of secondary metabolism gene clusters under similar regulatory controls. The newly created *A. flavus* database (Smith and Payne, unpublished) should help facilitate complex genomic mining studies to discover important functional traits and further our understanding of the complex genetic relationships described throughout this review.

Perhaps the most immediate impact of functional genomics will be a better understanding of the regulatory networks controlled by global regulators like VeA and LaeA. These two genes, and perhaps additional genes, appear to be important in the regulation of several pathways of secondary metabolism. Genome sequences and whole genome DNA microarrays will undoubtedly provide new insight into how these genes modulate development and metabolism. It is logical to speculate that other global regulators that control secondary metabolism may also be discovered. The striking control of AF biosynthesis, but not condensation, by temperatures within the cardinal ranges for growth provides strong evidence that the fungus uses temperature as a cue to regulate secondary metabolites in addition to AF, and in the case of AF overdosing ABIs. Environmentally relevant volatiles have emerged as a promising area of research for regulation of AF biosynthesis with ethylene and cresyl alcohol both exhibiting some inhibition of accumulation of AF cluster gene transcripts, potentially independent of AflR (Roze et al., 2007b, 2004b).

The genomes of several *Aspergillus* species have recently been sequenced; comparative genomics between these species are revealing many new insights about the genetic diversity within the genus (Rokas et al., 2007). The *Aspergillus* community is just starting to reap the benefit of many new functional genomics tools.

---

4.8. Regulation of aflatoxin by oxidative stress

Research over the past decade has made clear that oxidative stress stimulates AF production (Jayashree and Subramanyam, 2000; Reverberi et al., 2008); and anti-oxidants, such as gallic acid in walnut, have an inhibitory effect on AF production (Mahoney and Molyneux, 2004). It has been reported that gallic acid causes an inhibition of several AF pathway genes, including aflM (ver-1) and aflD (nor-1) (Mahoney and Molyneux, 2004). More recently, Kim et al. (2008) analyzed the effects of the anti-oxidant caffeic acid on gene expression using microarrays. They found that most genes in the AF biosynthetic pathway appeared to be down-regulated by caffeic acid treatment (Fig. 1). Reverberi et al. (2005) also reported that the induction of anti-oxidant enzymes by β-glucans from *Lentinula edodes* inhibited AF biosynthesis and delayed transcription of *aflR* as well as AF cluster genes. *A. flavus* has also been grown under anaerobic conditions and still shown to make AF, albeit at much lesser quantities (Clevstrom et al., 1983).

Kim et al. (2005) identified several genes needed to cope with oxidative stress using *Saccharomyces cerevisiae* as a model for fungal stress. Identified genes included many oxidative stress genes (glutaredoxins, catalases, superoxide dismutases) and others such as a MAPK gene like the yeast hog1 (Kim et al., 2005). Interestingly, differential expression for several *MAPK* genes has been found in the aforementioned temperature comparison arrays. One pitfall of the yeast studies is that for many of the genes identified only one strain has actually been shown to have an effect on AF production. A knockout of a superoxide dismutase gene in *A. flavus* has been shown to decrease AF production (He et al., 2007). One of the most important recent discoveries stemming from the work by Kim et al. (2005) was that the anti-oxidant response transcription factor Yap1 has an apparent role in regulation of AF biosynthesis (Reverberi et al., 2008). Reverberi and coworkers (2008) found that deletion mutants of yap1 in *A. parasiticus* had an increased accumulation of AF. It is not clear whether the increase in AF can be directly related to deletion of Yap1 or a result of increased reactive oxygen, however, it was proposed that there may be Yap1 binding sites in the *aflR* promoter. *aflR* was induced earlier in yap1 deletion strains compared to wild-type and a known binding site for the human ortholog of Yap1, AP-1, is found in the promoter region of *aflR* (Reverberi et al., 2008).

Functional genomics promises to provide important information on metabolites with no known function. For example, gene expression experiments have been valuable in characterizing regulatory mechanisms controlling the AF cluster (Fig. 1). Further analyses of these types of array experiments may lead to identification of secondary metabolism gene clusters under similar regulatory controls. The newly created *A. flavus* database (Smith and Payne, unpublished) should help facilitate complex genomic mining studies to discover important functional traits and further our understanding of the complex genetic relationships described throughout this review.

Perhaps the most immediate impact of functional genomics will be a better understanding of the regulatory networks controlled by global regulators like VeA and LaeA. These two genes, and perhaps additional genes, appear to be important in the regulation of several pathways of secondary metabolism. Genome sequences and whole genome DNA microarrays will undoubtedly provide new insight into how these genes modulate development and metabolism. It is logical to speculate that other global regulators that control secondary metabolism also will be uncovered. The striking control of AF biosynthesis, but not condensation, by temperatures within the cardinal ranges for growth provides strong evidence that the fungus uses temperature as a cue to regulate secondary metabolites in addition to AF, and in the case of AF overdosing ABIs. Environmentally relevant volatiles have emerged as a promising area of research for regulation of AF biosynthesis with ethylene and cresyl alcohol both exhibiting some inhibition of accumulation of AF cluster gene transcripts, potentially independent of AflR (Roze et al., 2007b, 2004b).

The genomes of several *Aspergillus* species have recently been sequenced; comparative genomics between these species are revealing many new insights about the genetic diversity within the genus (Rokas et al., 2007). The *Aspergillus* community is just starting to reap the benefit of many new functional genomics tools.
In the A. flavus community, several different array platforms (cDNA and Affymetrix) are currently available. Additionally, high throughput quantitative proteomics methods have been applied to A. flavus (Georgianna et al., 2008). With the push for more information on the complex processes being studied in the fungal community, it will be important for consistent methodology to be applied between different platforms and to make data easily accessible. To this end, in addition to the newly created A. flavus database, the A. flavus community has designed the AflaGenomics Laboratory as the center for all AflaMetrix array hybridization and scanning. Systems databases, for incorporation of genomic, transcriptomic, metabolomic, proteomic, and phenomic data will prove to be a very powerful tool for functional analysis of the complex processes being studied in A. flavus.

References


Chapter 2

The effect of elevated temperature on gene transcription and aflatoxin biosynthesis

O'Brian GR, Georgianna DR, Wilkinson JR, Yu J, Abbas HK, Bhatnagar D, Cleveland TE, Nierman W, Payne GA.

The effect of elevated temperature on gene transcription and aflatoxin biosynthesis

G.R. O'Brian
Department of Plant Pathology, Box 7567, North Carolina State University, Raleigh, North Carolina 27695-7567

D.R. Georgianna
Department of Plant Pathology, Box 7567, North Carolina State University, Raleigh, North Carolina 27695-7567, and Functional Genomics Graduate Program, Box 7567, North Carolina State University, Raleigh, North Carolina 27695-7567

J.R. Wilkinson
Department of Biochemistry and Molecular Biology, Box 9650, Mississippi State University, Mississippi State, Mississippi 39762

J. Yu
USDA/ARS, Southern Regional Research Center, New Orleans, Louisiana 70124

H.K. Abbas
USDA/ARS, Crop Genetics & Production Research Unit, Stoneville, Mississippi 38776

D. Bhatnagar
T.E. Cleveland
USDA/ARS, Southern Regional Research Center, New Orleans, Louisiana 70124

W. Nierman
The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850

G.A. Payne¹
Department of Plant Pathology, Box 7567, North Carolina State University, Raleigh, North Carolina 27695-7567

Abstract: The molecular regulation of aflatoxin biosynthesis is complex and influenced by several environmental conditions; one of these is temperature. Aflatoxins are produced optimally at 28–30°C, and production decreases as temperatures approach 37°C, the optimum temperature for fungal growth. To better characterize the influence of temperature on aflatoxin biosynthesis, we monitored the accumulation of aflatoxin and the expression of more than 5000 genes in Aspergillus flavus at 28°C and 37°C. A total of 144 genes were expressed differentially (P < 0.001) between the two temperatures. Among the 103 genes more highly expressed at 28°C, approximately 25% were involved in secondary metabolism and about 30% were classified as hypothetical. Genes encoding a catalase and superoxide dismutase were among those more highly expressed at 37°C. As anticipated we also found that all the aflatoxin biosynthetic genes were much more highly expressed at 28°C relative to 37°C. To our surprise expression of the pathway regulatory genes aflR and aflS as well as aflR antisense, did not differ between the two temperatures. These data indicate that the failure of A. flavus to produce aflatoxin at 37°C is not due to lack of transcription of aflR or aflS. One explanation is that AFLR is nonfunctional at high temperatures. Regardless, the factor(s) sensing the elevated temperatures must be acute. When aflatoxin-producing cultures are transferred to 37°C they immediately stop producing aflatoxin.

Key words: AFLR, biosynthesis, micro-arrays

INTRODUCTION

Aflatoxin is a highly carcinogenic polyketide secondary metabolite produced by several species of Aspergillus, including A. flavus, A. parasiticus, A. nomius, A. pseudotamarii, A. bombycis, (Samson 2001, Varga et al. 2003) and others (Cary et al. 2005). Aspergillus flavus along with A. parasiticus is a well known pathogen of many economically important commodities including corn, peanuts, cotton and tree nuts. Infection of seeds and contamination with aflatoxin is not limited to the field but also can occur postharvest if seeds are stored improperly. Exposure to aflatoxins has been associated with liver cancer and many veterinary toxic syndromes (Bressac et al. 1991, Hsu et al. 1991, Wang et al. 2001, Lewis et al. 2005). The United States, along with many other developed nations, have imposed regulatory limits on aflatoxin in food and feed. In the US agricultural economic losses due to aflatoxin contamination of food and feed are estimated to be $270 million annually (Richard and Payne 2003). The regulation of aflatoxin biosynthesis is influenced by several environmental and cultural conditions such as temperature, pH, adenylate concentration and energy charge, and nitrogen and carbon source (Luchese and Harrigan 1993, Payne and Brown 1998, Price et al. 2005). The influence of temperature on aflatoxin biosynthesis has intrigued researchers because only low concentrations of aflatoxin are produced at the temperatures that are optimal for fungal growth (37°C).
Early studies by Schindler et al. (Schindler et al. 1967) showed that aflatoxin was produced maximally at 24°C and not at all at temperatures lower than 18°C or higher than 35°C. However, Diener and Davis reported aflatoxin production in peanuts at 40°C by A. flavus (Diener and Davis 1967). Mayne et al. (1967) suggested that the effect of temperature is more dependent on substrate than on strain. It also has been reported that the production of aflatoxin or its pathway intermediates is regulated differently among some aspergilli species. Feng and Leonard compared A. parasiticus to A. nidulans under varying culture conditions. They detected aflatoxin at 27°C and lesser amounts at 33°C but were unable to detect aflatoxin in A. parasiticus at 37°C. In contrast, they found that A. nidulans produced sterigmatocystin at similar levels at all three temperatures (Feng and Leonard 1998).

While the cardinal range for aflatoxin production differs among strains and culture conditions, most research shows temperatures between 24–30°C favor aflatoxin biosynthesis.

The mechanism underlying the temperature-dependent regulation of aflatoxin production is unclear. Both transcriptional and posttranscriptional regulation mechanisms control aflatoxin gene transcription (for reviews see Payne and Brown 1998; Bennett and Klich 2000; Chang 2003). The AFLR protein binds to aflatoxin polyketide synthase gene of A. parasiticus (Yu et al. 2004c) provides the opportunity to better understand its role in aflatoxin biosynthesis (OBrian et al. 2003; Price et al. 2005, 2006). The AFLS protein binds to AFLR and modulates its expression. More recently microarrays have been developed to evaluate gene transcription during aflatoxin biosynthesis (OBrian et al. 2003; Price et al. 2005, 2006). The availability of DNA micro-arrays of A. flavus containing more than 5000 elements from an EST library (Yu et al. 2004c) provides the opportunity to better examine the effect of temperature on the pathway regulatory genes as well as nearly half of the genes in the A. flavus genome (www.aspergillusflavus.org).

**MATERIAL AND METHODS**

**Growth, media and aflatoxin analysis.**—Aspergillus flavus strain NRRL 3357 (ATCC 200028; SRR 167), a wild type A. flavus strain widely used in laboratory and field studies, as well as the strain of choice for the whole genome sequencing project, was used for gene transcription analysis and growth studies. For gene transcription and temperature shift experiments, a mother culture supplemented with 0.4% agar was seeded with 1 × 10⁶ spores/mL. Mother cultures were grown in A&M media at 37°C and 200 rpm for 16 h, and a 20 mL aliquot was used to inoculate 200 mL of fresh A&M media. Daughter cultures were grown at 28°C, 33°C or 37°C, and shaken at 200 rpm. Aflatoxin concentrations were determined as indicated either by HPLC or LC/MS at the Proteomics and Metabolomics Center (N.C. State University). To determine the effect of temperature on aflatoxin production in liquid grown cultures, 1 × 10⁶ spores/mL were seeded into 100 mL A&M media and incubated at selected temperatures. For similar studies performed on solid media, fresh spores were generated by plating 50 μL of 10⁶ spores onto Difco potato-dextrose agar (PDA) (American Scientific Products, Charlotte, North Carolina) and incubated at 30°C for 5 d. The spores were collected from 5-d cultures with sterile 0.05% Triton X-100. Samples were grown and collected according to Abbas et al. (2004) with minor modifications by plating 100 μL of 107 spores of each isolate on PDA enriched with 0.3% β-CD (CD-PDA) (Carasol®WTM, Wacker-Chemie GmbH, Burghausen, Germany). Duplicate cultures of each isolate were incubated for 24 h in total darkness at 28°C, 29°C, 30°C, 31°C, 32°C, 33°C, 34°C, 35°C, 36°C and 37°C. Fungal biomass and agar were removed from each sample with an inverted 1 mL pipette tip placed in glass scintillation vials (20 mL) and fresh weights were recorded (typically 0.5–1.0 g). A 10:1 volume of methanol-water (70:30, v/v) was added to the samples and the vials were shaken for 1 h at low speed on a reciprocal shaker. A 1 mL aliquot of extract was removed and centrifuged at 12000 g for 10 min and the supernatant was cleaned with an Altech 1.5 mL Extract-Clean reservoir containing 200 μg of aluminum oxide. The extract was eluted by gravity and a total of 20 μL was examined by HPLC (Sobolev and Dorner 2002). Quantitation of aflatoxins was determined by the external standard method where the standard curve was 0.5–20 ng mL⁻¹ (AFB1, AFG1) and 0.2–6 ng mL⁻¹ (AFB2, AFG2) (Sigma).

**RNA isolation.**—For micro-array and QPCR experiments, RNA was isolated from lyophilized cultures with Trizol (Life Technologies, Rockville, Maryland) according to the manufacturer’s instructions. Isolated RNA was purified further by precipitation on ice overnight in 2 M LiCl. The RNA was pelleted, washed with 70% ethanol and air-dried about 10 min. The RNA pellet was resuspended in 50 μL DEPC-dH₂O with 40 units RNase inhibitor (Promega Corporation, Madison, Wisconsin) and quantified by spectrophotometry.
QPCR.—RNA (1 µg) isolated from 28 C, 35 C and 37 C cultures grown 24 h was used in a reverse transcription reaction (Stratascript) to synthesize the cDNA template. QPCR reactions were performed in triplicate with a DNA Engine Opticon 2 System (MJ Research) and data were collected with Opticon Monitor Software version 2.02 (MJ Research). SYBR-green master mix (Applied Biosystems) was used to monitor expression with a 96well format. Expression levels were measured in triplicate and calculated by a variation of comparative C(t) method (Livak and Schmittgen 2001) with 18s rRNA as the endogenous reference for sample normalization. For each set of temperatures, the mean of the normalized C(t) values for a given gene was used to measure fold increase relative to that gene across conditions tested. To provide a conservative estimate of the mean, the maximum delta C(t) score possible given the number of cycles run was assigned for samples where no expression was detected.

Micro-arrays.—Micro-arrays used in this study were printed at The Institute for Genome Research (TIGR) with amplicons (approx. 500 bp) from EST clones (Yu et al 2004c). A total of 5002 genes were arrayed at least three times each for a total of 17991 spots. Total RNA from each treatment studied was converted to cDNA and labeled as described by Price et al (2005). Each treatment was labeled with each dye, removing effects on measurements caused by the individual dyes. The hybridized slides were scanned with a Perkin Elmer ScanArray Express Lite scanner (Perkin Elmer Life and Analytical Sciences Inc., Boston, Massachusetts). Spot intensity data were extracted from the images with UCSF-Spot (Jain et al 2002). The resulting spot-intensity data were analyzed with the mixed procedure in SAS (SAS v8, SAS Institute, Cary, North Carolina) as described by Price et al (2005). Briefly, least squares estimates of gene-specific treatment effects were obtained for each gene under each treatment. Differences between treatment effects (least squares estimates) for pairs of treatments can be considered as log2-transformed fold changes (Wolflinger et al 2001). Comparisons were made between cultures grown at 28 C and 37 C. The experimental design is provided (Fig. 1). The experiment was performed in three phases. In the first phase a dye-flip experiment was performed to compare expression levels at 28–37 C after 24 h. Next, a time course experiment was performed with cultures grown at 37 C for 8 h, 16 h and 24 h. Finally, the same time course was performed with cultures grown at 28 C. Data from these arrays were analyzed together using temperature as the treatment effect.

RESULTS

Aflatoxin production is regulated by temperature.—Time course experiments were performed to evaluate aflatoxin production in liquid cultures of A. flavus grown at various temperatures (Fig. 2A). In these experiments, particularly after 36 h of growth, aflatoxin production was the highest at 28 C. Decreasing amounts of aflatoxin were produced as temperature increased from 34 C to 37 C. Minimal amounts were produced at 37 C. This trend in aflatoxin production also occurred when cultures were grown on solid media. The results show that aflatoxin production peaked at 30 C and then decreased as Aspergilli were exposed to increasing temperatures (Fig. 2B).

Transfer of cultures from 28 C to 37 C stops aflatoxin biosynthesis.—A series of A. flavus 3557 cultures were grown in A&M media mother cultures at 37 C for 16 h and transferred to 28 C daughter cultures for varied amounts of time before being moved back to 37 C. The total postmother culture incubation time was 24 h regardless of temperature. Aflatoxin concentrations were determined by HPLC from two replicates that were pooled after 24 h incubation. The results (Fig. 3) indicated that aflatoxin biosynthesis begins at 28 C after 12 h. If cultures were allowed to grow at 28 C for the entire 24 h, 465 ng/mL aflatoxin was made. However transfer of cultures to 37 C resulted in no further significant accumulation of aflatoxin. Results were equivalent when repeating the experiments with A. parasiticus strain SU-1 (data not shown). In a separate experiment daughter cultures were grown at 28 C for 24 h and...
then placed at 37°C. Aflatoxin concentrations were determined at 0 h, 0.5 h, 1 h, 2 h, 4 h, 8 h and 24 h after transfer. The results show that no additional aflatoxin was made after transfer to 37°C (data not shown).

The transcription profile is different at 28–37°C.— Cultures were grown in A&M media for 8 h, 16 h and 24 h at 28°C and 37°C. Micro-array experiments were performed according to the design shown (Fig. 1). Only cultures grown for 16 h and 24 h at 28°C produced aflatoxin. A two-stage ANOVA approach was used to analyze the data (Wolflinger et al. 2001). In this analysis temperature was the only treatment considered. Differences in least squares estimates of transcription levels were ranked according to p-value. A listing of the top 20 differentially expressed genes between the two temperatures are provided (Table I). Included in this list are three genes from the aflatoxin biosynthetic cluster, which were spotted as controls (known to be expressed during aflatoxin production). Three additional genes were also from the aflatoxin cluster. These were norsolorinic acid reductase, O-methyltransferase B and the aflatoxin polyketide synthase (also spotted as a control).

A total of 144 genes were differentially expressed (P < 0.001) between the two temperatures. Of these 103 were more highly expressed at 28°C. Among the genes more highly expressed at 28°C, approximately 25% were involved in secondary metabolism and about 30% were classified as hypothetical. Genes encoding a catalase and superoxide dismutase were among those more highly expressed at 37°C. These two genes also were found to be induced at 37°C in Cryptococcus neoformans (Kraus et al. 2004).

Aflatoxin biosynthetic genes are more highly expressed at 28°C.— The least square means estimates for the aflatoxin biosynthetic cluster genes are illustrated (Fig. 4). Most aflatoxin genes were more highly expressed at 28°C relative to 37°C. However aflR and aflS did not follow this pattern. AflR and aflS showed about equal expression at both temperatures.

Quantitative PCR is consistent with micro-array data.— We used quantitative PCR and determined expression levels from cultures grown at 28°C, 35°C and 37°C for aflR, aflR antisense, aflS and aflP. As shown (Table II) levels of aflS, aflR and aflR antisense were relatively constant across each temperature tested. However the aflatoxin biosynthetic gene aflP was significantly more highly expressed at 28°C. There was also some expression of aflP at 35°C but no detectable expression at 37°C.

![Figure 2](image1.png)  
**Fig. 2.** Effect of temperature on aflatoxin production in liquid and solid media. A. 1 x 10^6 spores/mL were seeded into 100 mL A&M medium and incubated at selected temperatures. The medium was sampled at 48 h and aflatoxin concentrations were determined by LC/MS. B. A total of 1 x 10^6 spores were plated on PDA medium and incubated at different temperatures. The medium was sampled at 24 h and aflatoxin concentrations were determined by HPLC.

![Figure 3](image2.png)  
**Fig. 3.** Temperature shift. Cultures of *A. flavus* were grown in A&M medium at 37°C and moved to 28°C for indicated times before being returned to 37°C.
DISCUSSION

The ecological significance of aflatoxin biosynthesis to Aspergillus flavus is unknown. Production of this secondary metabolite presumably contributes to the competitiveness of the fungus (Wilkinson et al. 2004) and a large body of information shows that aflatoxin production can be induced by several substrates and environmental conditions. A better understanding of environmental regulation of this toxin may provide information regarding its role in the ecology of the fungus and lead to new approaches to reduce aflatoxin contamination of food.

In this study we focused on temperature as a modulator of aflatoxin production because it has one of the most striking effects of any environmental factor yet examined (Price et al. 2005). While Aspergillus flavus grows over a wide range of temperatures in culture, its optimum temperature for growth is 37°C. To our surprise essentially no aflatoxin is produced at this temperature. Because temperature can have broad effects on fungi, several factors could account for reduced aflatoxin production at higher temperatures including changes in the metabolite partitioning, energy status of the cell or a direct effect of temperature on transcriptional regulatory circuits. We took advantage of a 5002

TABLE I. Top 20 differentially expressed genes between 28°C and 37°C

<table>
<thead>
<tr>
<th>Aflatoxin genes</th>
<th>Higher at 28°C</th>
<th>Higher at 37°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>aflK (vbs)</td>
<td>NAFD24TV</td>
<td>NAGAX81TV</td>
</tr>
<tr>
<td>aflT (hypA)</td>
<td>NAGA89TV</td>
<td>NAFD73TV</td>
</tr>
<tr>
<td>aflC (phaA)</td>
<td>NAGA194TV</td>
<td>NAFD40TV</td>
</tr>
<tr>
<td>NAFFER13TV</td>
<td>NAFCH99TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFL32TV</td>
<td>NAFDO13TV</td>
<td>NAFCH39TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFDT13TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFL32TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
<tr>
<td>NAFFI76TV</td>
<td>NAFCH39TV</td>
<td>NAGDF01TV</td>
</tr>
</tbody>
</table>

*Clone IDs were obtained from the TIGR website (http://www.tigr.org/tigr-scripts/Tgi/T_report.cgi?species=a_flavus).

FIG. 4. Expression levels of aflatoxin pathway genes. Least square means estimates of expression levels obtained from the mixed-model analysis of the micro-array data were plotted for genes within the aflatoxin biosynthetic cluster. The analysis compared expression levels from cultures grown at 28°C and 37°C.
element DNA micro-array to better characterize the effect of temperature on gene transcription. We observed the differential expression of 144 genes when temperature was increased from 28 C to 37 C. Of interest, most of these genes (103) were expressed more highly at 28 C. Approximately 25% of these genes are involved in secondary metabolism including aflatoxin biosynthesis (Fig. 4). Our study shows that temperature, directly or indirectly, affects transcription of genes for secondary metabolism. These data are consistent with the transcription profiles reported for aflP (Liu and Chu 1998) and aflRC (phkA) (Feng and Leonard 1995).

Even though little to no aflatoxin is produced at 37 C, we observed a low level of gene transcription for some of the pathway genes. This is probably due to basal transcription levels for these genes because aflatoxin pathway gene expression at 37 C followed a similar profile to that previously observed in an aflR deletion mutant (Price et al. 2006). We also were interested in learning if naturally occurring antisense of aflR discovered several years ago (Woloshuk et al 1994) played any role in the temperature response. Expression of this antisense along with other pathway genes is shown (Table II). The results of the quantitative PCR showed that levels of aflS, aflR, and aflR antisense were relatively constant across each temperature tested. In contrast, the aflatoxin biosynthetic gene aflP was significantly more highly expressed at 28 C, with some expression at 35 C but no detectable expression at 37 C. These data obtained from quantitative PCR were consistent with those obtained from micro-array studies.

Our data argue that the failure of A. flavus to produce aflatoxin at 37 C is not due to the effect of temperature on the transcription of the pathway regulatory genes because transcript levels of aflR and aflS did not change significantly between 28 C and 37 C. One explanation for the temperature effect might be that less AFLR is produced at 37 C. Liu and Chu (1998) reported a lower concentration of AFLR at 37 C compared to 29 C. Another possibility is that AFLR is nonfunctional at higher temperatures. It is known that phosphorylation of AFLR interferes with the regulatory protein’s activity because it may prevent the movement of AFLR into the nucleus (Shimizu et al 2005). Another possibility is that at elevated temperature, AFLS and AFLR are unable to interact; Chang (2005) has shown that AFLR and AFLS interact and together regulate transcription of the aflatoxin biosynthetic pathway. Additional studies are needed to determine the effect of temperature on AFLR and possibly on AFLS.

It is possible that other factors in addition to the nonfunctionality of AFLR affect aflatoxin production at elevated temperatures. For example the temperature response could be due to a modification of one or more of the pathway enzymes. This seems likely because aflatoxin production was greatly diminished in an aflatoxin producing culture after transferring from 28 C to 37 C (Fig. 3). At 28 C all of the necessary aflatoxin biosynthetic enzymes already had been made and were functioning to produce aflatoxin. After the shift to 37 C, production virtually ceased.

Our observations cannot rule out a direct effect of temperature on metabolic pathways that support aflatoxin biosynthesis. However it seems unlikely that an effect on these pathways would lead to such a rapid cessation of aflatoxin biosynthesis in cultures moved from 28 to 37 C. We found that the transfer to 37 C of aflatoxin producing cultures resulted in the almost immediate inhibition of aflatoxin synthesis. Others also have observed decreases in aflatoxin production at these temperatures (Schindler et al 1967, Schroeder and Hein 1967) or when cultures were exposed to elevated temperatures for short periods (Schroeder and Hein 1968).

Another interesting observation from this study is that the expression levels of aflR and aflS are relatively constant at both temperatures. This argues that the two genes may be transcriptionally coregulated. Perhaps they are both regulated by LAEA as has been proposed by Bok and Keller (2004). It also has been reported that AFLR does not transcriptionally regulate aflS or vice versa (Chang 2003).

In summary we have shown in this study that temperature affects aflatoxin production and the transcriptional profile of A. flavus. Transfer of an aflatoxin producing culture from 28 C to 37 C quickly turns off aflatoxin biosynthesis. The speed by which this occurs suggests that one or more of the pathway enzymes are posttranslationally regulated and are nonfunctional at 37 C. There is also a transcriptional component to temperature regulation. A larger number of genes are more highly expressed at 28 C relative to 37 C. We focused on the aflatoxin cluster genes and demonstrated a significant reduction in transcription at 37 C compared to 28 C. Although transcripts (and presumably protein) for the transcriptional regulator, aflR, and aflS are present at

<table>
<thead>
<tr>
<th>Temperature</th>
<th>aflS</th>
<th>aflR</th>
<th>aflRas</th>
<th>aflP</th>
</tr>
</thead>
<tbody>
<tr>
<td>28 C</td>
<td>1.06</td>
<td>1.08</td>
<td>0.83</td>
<td>64.4</td>
</tr>
<tr>
<td>35 C</td>
<td>1.03</td>
<td>0.97</td>
<td>1.12</td>
<td>2.79</td>
</tr>
<tr>
<td>37 C</td>
<td>0.92</td>
<td>0.94</td>
<td>1.07</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Fold induction measured relative to mean expression across a gene.
37°C, the function of AFLR is inhibited at this temperature. We propose that one or both of these proteins may be nonfunctional at elevated temperatures in *A. flavus*.

**ACKNOWLEDGMENTS**

The authors thank James Burroughs III for his critique of this manuscript. This research was supported by Grant 2002-35201-12562 from the USDA/NRI Competitive Grants Program.

**LITERATURE CITED**


Chapter 3

Temperature-Dependent Regulation of Proteins in *Aspergillus flavus*: Whole Organism Stable Isotope Labeling by Amino Acids

D. Ryan Georgianna, Adam M. Hawkridge, David C. Muddiman, and Gary A. Payne

Temperature-Dependent Regulation of Proteins in Aspergillus flavus: Whole Organism Stable Isotope Labeling by Amino Acids

D. Ryan Georgianna,1 Adam M. Hawkridge,1 David C. Muddiman,1 and Gary A. Payne*†‡

Department of Plant Pathology, Center for Integrated Fungal Research, and W.M. Keck FT-ICR-MS Laboratory, Department of Chemistry, North Carolina State University, Raleigh, North Carolina 27695

Received February 7, 2008

Stable isotope labeling by amino acids in cell culture (SILAC) has been used in many different organisms including yeast, mammalian cells, and Arabidopsis cell culture. We present an adaptation of this method to quickly quantify protein changes in response to environmental stimuli regulating biosynthesis of the carcinogen aflatoxin in the fungus Aspergillus flavus. Changes in relative protein concentrations in response to temperature were quantified and compared to changes in aflatoxin biosynthesis and the transcription of the aflatoxin biosynthetic genes. In a comparison between conducive (28 °C) and nonconducive (37 °C) temperatures for aflatoxin biosynthesis, 31 proteins were found to be more abundant at 37 °C and 18 more abundant at 28 °C. The change in expression of the aflatoxin pathway enzymes closely followed the strong repression of both aflatoxin biosynthesis and transcription of the aflatoxin pathway genes observed at 37 °C. Transcripts corresponding to the 379 proteins quantified by SILAC were analyzed using microarrays, but their expression did not always correlate well with transcript levels of encoding genes. This is the first reported labeling of a multicellular free-living prototroph using the SILAC procedure to compare 13C6-arginine-labeled samples to 12C6-arginine-labeled samples for quantitative proteomics. The data presented shows the utility of this procedure in quantifying changes in protein expression in response to environmental stimuli.

Keywords: SILAC • mass spectrometry • proteomics • Aspergillus flavus • Aflatoxin

Introduction

The genus Aspergillus represents a large group of filamentous fungi well-studied for their production of a diverse array of economically important compounds, both beneficial and harmful. Aspergillus flavus is recognized globally for its production of the mycotoxin aflatoxin in several commodities including maize, peanuts, cotton, and various important tree nuts. Contamination with aflatoxin creates a serious health risk as aflatoxins are potent carcinogens and liver toxins, affecting many domestic and wild animals as well as humans. In addition to the significant health risks, aflatoxin contamination also creates a large economic burden. Because of the carcinogenicity of aflatoxins, the U.S. Food and Drug Administration has placed an action level for aflatoxins in food at 20 ppb and in milk at 0.5 ppb. Furthermore, A. flavus is an important and unique opportunistic pathogen in that it is also virulent in animals across many phyla, causing the fungal disease aspergillosis in many species including humans.

An available annotated genome sequence for this organism, as well as the characterized aflatoxin biosynthetic pathway, provides an ideal system for the identification of proteins that may regulate the biosynthesis of this compound. While the proteins involved in pathway specific regulation have been identified,1-3 more global regulatory elements that affect biosynthesis remain to be characterized. We anticipate that some of the regulatory proteins may be constitutively expressed at a low level and may be activated by protein localization or post-translational modification. To identify these proteins, procedures are needed to quantify changes in protein content during conditions conducive and nonconducive for aflatoxin biosynthesis. Aflatoxin biosynthesis is influenced by several environmental conditions such as temperature, pH, adenylate concentration and energy charge, nitrogen and carbon source.3-5

The SILAC method has been applied to a variety of different organisms for relative quantification of protein levels. It has the benefit of being a relatively simple and inexpensive way to label proteins without applying chemical modifications to isolated protein or peptide samples. The SILAC method was first developed by Ong et al. in mouse C2C12 cells using Leu-3.6 Subsequently, Saccharomyces cerevisiae cells,7 the human HeLa cell line,8 and most recently Arabidopsis thaliana cell culture9 have all been studied using the SILAC method. The labeling of Arabidopsis cell cultures represented the first time that a prototroph was labeled, with all others being auxotrophs labeled by an essential amino acid. Labels have progressed from
research articles

A. Experimental Design

Figure 1. (A) Experiment design. Comparison 1, 28 °C (\(^{13}\text{C}_6\)-arginine) vs 27 °C (\(^{12}\text{C}_6\)-arginine) and comparison 2, 28 °C (\(^{13}\text{C}_6\)-arginine) vs 28 °C (\(^{12}\text{C}_6\)-arginine) were separated by SDS-PAGE. (B) SDS-PAGE. Lanes 1 and 2 show SDS-PAGE separations of respective sample comparisons. For each comparison, 50 μg of each protein sample was combined. After electrophoresis, the gel was sliced into 40 bands and the proteins in each band were digested with trypsin. After digestion, all 80 samples were sampled uninterrupted on nanoflow reverse-phase LC directly coupled with the Thermo LTQ-FT Ultra.

The initial use of Lex-1 to a variety of different amino acids using \(^{13}\text{C}\) and \(^{15}\text{N}\) isotopes. The two most commonly used amino acids are arginine\(^{10}\) and lysine.\(^{11}\) Arginine and lysine both benefit from being the target site in proteins for trypsin cleavage, meaning that some resulting peptides can have one C-terminal label, making analysis much simpler. Only arginine labeling was used to avoid any differential arginine and lysine labeling efficiencies as a result of A. flavus being a prototroph. Increased labels within an organism also increases sample complexity; an arginine or lysine label within a tryptic digest will double the number of peptide masses containing isotopes of these amino acids.

We report here the successful labeling of amino acids in A. flavus, a free-living multicellular prototroph, by growing the strain on medium containing \(^{15}\text{C}_6\)-arginine. Further, we show that the SILAC procedure can be used to reliably quantify changes in protein concentrations in the fungus in response to changes in temperature.

Materials and Methods

Culture of A. flavus for SILAC. Stock cultures of A. flavus strain NRRL 3357 were stored at ~80 °C in 35% glycerol.

Inoculum was prepared by streaking an aliquot of the stock onto potato dextrose agar (PDA) (DIFCO, Sparks, MD) supplemented with either (i) 340 μg/mL \(^{13}\text{C}_6\)-arginine (SIGMA, St. Louis, MO) or (ii) 340 μg/mL \(^{15}\text{C}_6\)-arginine (Cambridge Isotopes, Andover, MA). After incubation of the culture at 37 °C for 5 days, conidia were dislodged with a glass rod and suspended in 0.05% Triton X-100. The concentration of conidia in the suspension was quantified using a hemacytometer. Flasks containing 100 mL of A& medium\(^{12}\) supplemented with 340 μg/mL of either \(^{13}\text{C}_6\)-arginine or \(^{15}\text{C}_6\)-arginine were inoculated with the conidial suspension to a final concentration of 10° conidia/mL. These liquid cultures were grown for 24 h at 28 or 37 °C (Figure 1A) with shaking at 200 rpm. Fungal mycelium was collected for protein and mRNA analysis and the filtrate was assayed for the presence of aflatoxin.

Protein Isolation and Sample Preparation. Fungal mycelium was separated from the medium by filtration through miracloth (Calbiochem, San Diego, CA). The retained fungal mat was washed with cold PBS (Sigma, St. Louis, MO), removed from the filter, immediately frozen in liquid nitrogen, lyophilized, and stored at ~80 °C until extracted for proteins. A 0.2 g sample of lyophilized tissue was ground with a pestle in a mortar containing liquid nitrogen and 0.1 g of 120–200 μm glass beads (Sigma, St. Louis, MO). The resultant fine powder was placed into 1.5 mL microcentrifuge tubes, resuspended with 1 mL of cold PBS, and ground an additional 5 min at 4 °C in a Vortex Disruptor Genie (Scientific Industries, Inc. Bohemia, NY). The homogenate was spun for 5 min at 8000 g to pellet large cellular debris. The supernatant was transferred to a new tube and spun at 100 000 g for 45 min to pellet membranes and other debris. Relative protein concentrations of the samples were measured at an absorbance of 280 nm using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE).

On the basis of the relative protein concentrations of the samples, approximately 50 μg of protein from the 37 °C \(^{15}\text{C}_6\)-arginine sample and 50 μg of protein from the 28 °C \(^{15}\text{C}_6\)-arginine sample were combined and added to the same well in a 12.5% SDS-PAGE 1D gel (Bio-Rad Laboratories, Inc., Hercules, CA). Similarly, 50 μg each of a 28 °C \(^{15}\text{C}_6\)-arginine protein sample and identically treated 28 °C \(^{13}\text{C}_6\)-arginine sample were applied to an adjacent well (Figure 1B). Samples were separated on the 12.5% SDS-PAGE gel for 1 h at 120 V. The SDS-PAGE gel was stained according to the manufacturer’s suggestions using the Bio-Rad Bio-Safe coomassie stain (Bio-Rad, Hercules, CA). Forty bands from each lane were sliced from the gel using a gel cutter (The Gel Company, San Francisco, CA); placed in individual 1.5 mL tubes, washed with 50 mM NH\(_4\)HCO\(_3\), pH 7.8, dehydrated with MeCN and dried in a vacuum centrifuge. Samples were reduced with 10 mM DTT for 45 min at 56 °C until extracted for proteins. A 0.2 g sample of lyophilized tissue was ground with a pestle in a mortar containing liquid nitrogen and 0.1 g of 120–200 μm glass beads (Sigma, St. Louis, MO). The resultant fine powder was placed into 1.5 mL microcentrifuge tubes, resuspended with 1 mL of cold PBS, and ground an additional 5 min at 4 °C in a Vortex Disruptor Genie (Scientific Industries, Inc. Bohemia, NY). The homogenate was spun for 5 min at 8000 g to pellet large cellular debris. The supernatant was transferred to a new tube and spun at 100 000 g for 45 min to pellet membranes and other debris. Relative protein concentrations of the samples were measured at an absorbance of 280 nm using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE).

The basis of the relative protein concentrations of the samples, approximately 50 μg of protein from the 37 °C \(^{15}\text{C}_6\)-arginine sample and 50 μg of protein from the 28 °C \(^{15}\text{C}_6\)-arginine sample were combined and added to the same well in a 12.5% SDS-PAGE 1D gel (Bio-Rad Laboratories, Inc., Hercules, CA). Similarly, 50 μg each of a 28 °C \(^{15}\text{C}_6\)-arginine protein sample and identically treated 28 °C \(^{13}\text{C}_6\)-arginine sample were applied to an adjacent well (Figure 1B). Samples were separated on the 12.5% SDS-PAGE gel for 1 h at 120 V. The SDS-PAGE gel was stained according to the manufacturer’s suggestions using the Bio-Rad Bio-Safe coomassie stain (Bio-Rad, Hercules, CA). Forty bands from each lane were sliced from the gel using a gel cutter (The Gel Company, San Francisco, CA); placed in individual 1.5 mL tubes, washed with 50 mM NH\(_4\)HCO\(_3\), pH 7.8, dehydrated with MeCN and dried in a vacuum centrifuge. Samples were reduced with 10 mM DTT for 45 min at 56 °C following reduction, alkylation was performed using an equivalent volume of 55 mM iodoacetamide in 50 mM NH\(_4\)HCO\(_3\), pH 7.8, at room temperature for 30 min. Reduced and alkylated samples were dehydrated in excess MeCN and dried using a vacuum centrifuge. The resulting dried gel pieces were hydrated on ice for 30 min with 20 μg/μl trypsin (Trypsin Gold Mass Spectrometry Grade, Promega, Madison, WI) in 50 mM NH\(_4\)HCO\(_3\), pH 7.8, and then digested for 18 h at 37 °C before peptide extraction.

For peptide extraction, the digested samples were rinsed with 50 μM NH\(_4\)HCO\(_3\), pH 7.8 and the rinse solution was dehydrated twice in 50% MeCN with 5% formic acid in 50 μM NH\(_4\)HCO\(_3\), pH 7.8. After each rinse, the supernatant was
hydrated in a minimal volume of 50 mM NH₄HCO₃, pH 7.8. The two rinse samples were pooled and evaporated to near dryness using a vacuum centrifuge. The dried samples were resuspended in water and used directly for LC-MS/MS analysis.

**Table 1. Differentially Expressed Proteins**

<table>
<thead>
<tr>
<th>Description Affymetrix ID</th>
<th>SILAC Log(2°C:37°C)</th>
<th>peptides quantified</th>
<th>std. dev. microarray Log(2°C:37°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AflH</td>
<td>2911.009990</td>
<td>2.06</td>
<td>1</td>
</tr>
<tr>
<td>AflO</td>
<td>2911.008990</td>
<td>1.86</td>
<td>6</td>
</tr>
<tr>
<td>AflD</td>
<td>2911.009805</td>
<td>2.08</td>
<td>5</td>
</tr>
<tr>
<td>AflF</td>
<td>2911.008998</td>
<td>1.60</td>
<td>6</td>
</tr>
<tr>
<td>UBA/TS-N domain containing protein</td>
<td>2689.00103</td>
<td>1.19</td>
<td>1</td>
</tr>
<tr>
<td>lysine-rich protein, putative</td>
<td>2091.00402</td>
<td>1.03</td>
<td>1</td>
</tr>
<tr>
<td>conserved hypothetical protein</td>
<td>2541.00489</td>
<td>0.95</td>
<td>1</td>
</tr>
<tr>
<td>aldehyde dehydrogenase, putative</td>
<td>1866.00243</td>
<td>0.84</td>
<td>1</td>
</tr>
<tr>
<td>Aim</td>
<td>2911.008997</td>
<td>0.91</td>
<td>40</td>
</tr>
<tr>
<td>AflG</td>
<td>2911.008999</td>
<td>0.70</td>
<td>1</td>
</tr>
<tr>
<td>casein kinase 1, putative</td>
<td>2911.003550</td>
<td>0.59</td>
<td>1</td>
</tr>
<tr>
<td>Hsp20/alpha Crystallin family protein</td>
<td>1918.00016</td>
<td>0.57</td>
<td>1</td>
</tr>
<tr>
<td>acetylglutamate kinase, putative</td>
<td>1866.00400</td>
<td>0.56</td>
<td>1</td>
</tr>
<tr>
<td>AIC</td>
<td>2911.009908</td>
<td>0.55</td>
<td>1</td>
</tr>
<tr>
<td>tyrosyl-tRNA synthetase, putative</td>
<td>2541.00673</td>
<td>0.42</td>
<td>1</td>
</tr>
<tr>
<td>proteasome regulatory particle subunit Rp4</td>
<td>1918.00050</td>
<td>0.39</td>
<td>1</td>
</tr>
<tr>
<td>Sec23/Sec24 trunk domain containing protein</td>
<td>2504.01235</td>
<td>0.39</td>
<td>1</td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>2258.00170</td>
<td>0.38</td>
<td>1</td>
</tr>
<tr>
<td>UDP-Glc/Gal ER nucleotide sugar transporter</td>
<td>2091.00358</td>
<td>-0.80</td>
<td>2</td>
</tr>
<tr>
<td>adenosine phosphoribosyltransferase 1, putative</td>
<td>2368.00082</td>
<td>-0.80</td>
<td>1</td>
</tr>
<tr>
<td>esterase, putative</td>
<td>2504.00286</td>
<td>-0.80</td>
<td>3</td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>2689.00704</td>
<td>-0.80</td>
<td>2</td>
</tr>
<tr>
<td>conserved hypothetical protein</td>
<td>2541.00373</td>
<td>-0.82</td>
<td>1</td>
</tr>
<tr>
<td>To bile acid ABC transport protein, putative</td>
<td>2634.00189</td>
<td>-0.81</td>
<td>3</td>
</tr>
<tr>
<td>aspartate transaminase, putative</td>
<td>2043.010176</td>
<td>-0.91</td>
<td>3</td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>2091.006114</td>
<td>0.89</td>
<td>1</td>
</tr>
<tr>
<td>Eukaryotic aspartyl protease family protein</td>
<td>2504.012113</td>
<td>-0.92</td>
<td>1</td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>1866.00380</td>
<td>-0.98</td>
<td>2</td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>2043.010355</td>
<td>-0.96</td>
<td>1</td>
</tr>
<tr>
<td>Calreticulin family protein</td>
<td>2091.00281</td>
<td>-0.96</td>
<td>1</td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>2258.00680</td>
<td>-0.96</td>
<td>2</td>
</tr>
<tr>
<td>60S ribosomal protein L27, putative</td>
<td>2634.00642</td>
<td>-0.96</td>
<td>1</td>
</tr>
<tr>
<td>sugar-phosphate isomerase, RpiB/LacA/LacB family</td>
<td>2911.00386</td>
<td>-0.95</td>
<td>6</td>
</tr>
<tr>
<td>TIM-barrel enzyme family protein, putative</td>
<td>2368.001190</td>
<td>-1.05</td>
<td>1</td>
</tr>
<tr>
<td>conserved hypothetical protein</td>
<td>2911.00270</td>
<td>-1.05</td>
<td>2</td>
</tr>
<tr>
<td>O-methyltransferase family protein</td>
<td>2504.00514</td>
<td>-1.07</td>
<td>2</td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>2043.010127</td>
<td>-1.10</td>
<td>1</td>
</tr>
<tr>
<td>28S proteasome regulatory subunit mts4, putative</td>
<td>2541.003389</td>
<td>-1.10</td>
<td>1</td>
</tr>
<tr>
<td>To nuclear protein SA-1</td>
<td>1866.00182</td>
<td>-1.15</td>
<td>1</td>
</tr>
<tr>
<td>Copper/zinc superoxide dismutase family protein</td>
<td>2043.01368</td>
<td>-1.13</td>
<td>3</td>
</tr>
<tr>
<td>Fungal specific transcription factor domain protein</td>
<td>2258.00807</td>
<td>-1.22</td>
<td>1</td>
</tr>
<tr>
<td>MAP kinase (Osm1), putative</td>
<td>2091.00189</td>
<td>-1.30</td>
<td>1</td>
</tr>
<tr>
<td>expressed protein-related</td>
<td>2368.00812</td>
<td>-1.30</td>
<td>3</td>
</tr>
<tr>
<td>PQD domain containing protein</td>
<td>2504.01409</td>
<td>-1.30</td>
<td>1</td>
</tr>
<tr>
<td>hydroxymethylglutaryl-CoA synthase, putative</td>
<td>1918.003565</td>
<td>-1.40</td>
<td>1</td>
</tr>
<tr>
<td>aspartyl aminopeptidase, putative</td>
<td>2504.01399</td>
<td>-1.60</td>
<td>2</td>
</tr>
<tr>
<td>heat shock protein HSP30, putative</td>
<td>2541.00334</td>
<td>-1.51</td>
<td>20</td>
</tr>
<tr>
<td>aspartyl-tRNA synthetase</td>
<td>2911.00342</td>
<td>-1.70</td>
<td>1</td>
</tr>
<tr>
<td>Variant SE3 domain containing protein</td>
<td>2642.00662</td>
<td>-1.70</td>
<td>1</td>
</tr>
</tbody>
</table>
performed in the ICR cell and 8 data-dependent tandem mass spectra obtained in the ion-trap.

Data Searching and Analysis. All mass spectrometry data were analyzed using the SEQUEST search algorithm\textsuperscript{13} using a protein database generated from the \textit{A. flavus} genome\textsuperscript{14} within the Bioworks v3.1 software suite (Thermo Fisher Scientific, Somerset, NJ). Data were searched at 3 ppm mass measurement accuracy with variable modifications for methylation, oxidation, deamidation or $^{13}$C$_6$-arginine and fixed modification for carboxamidomethylation of cysteine. Searches allowed for 1 missed cleavage and up to 3 modifications. Peptides were initially filtered using the requirement that XCorr values are greater than one and peptide probability values are less than 0.05. SILAC ratios were determined using the PepQuan feature (mass tolerance, 0.02; minimum threshold, 50 000; smoothing 0.05. SILAC ratios were determined using the PepQuan feature greater than one and peptide probability values are less than 0.05. SILAC ratios were determined using the PepQuan feature.

Figure 2. Distribution of SILAC ratios. Ninety-five percent of the values for the log ratio of the 28 vs 28 °C distribution fall between −0.7774 (2.5% quantile) and 0.3276 (97.5% quantile); these cutoffs fall within the bins marked by dashed lines (bin size 0.1). These values were used as cutoffs for differential protein expression in 28 vs 37 °C samples which as expected has a wider distribution of ratios. A total of 31 proteins were below the −0.7774 cutoff and determined to be higher at 37 °C and 18 proteins were above the 0.3276 cutoff, indicating higher levels at 28 °C. Labeling efficiency determined by the 28 vs 28 °C data was 78.0 ± 3.4% (95% CI).

Figure 3. Scatterplot of shared data points between 28 °C:28 °C and 28 °C:37 °C data sets. A total of 210 proteins were shared between both data sets. Dashed lines represent cutoff assigned based on the 28 °C:28 °C data. The 28 °C:37 °C has a much wider distribution. Lack of data points in corner quadrants indicate no shared significant proteins; if these were present, it would indicate one of the masses for the peptides used in the quantification may have had an interfering species making quantification difficult.

Figure 4. Scatterplot of gene and protein expression. The X-axis corresponds to the transcript ratios for 28 °C:37 °C for each of 379 proteins/genes (DNA probes for genes encoding two identified proteins were not represented on the Affymetrix GeneChip microarray). Pairwise correlation for all 379 proteins/genes is 0.4334. Dashed lines represent assigned cutoff for SILAC data based on 28 vs 28 °C data and a fold change of 2 for transcript data. Black triangle markers (\textbullet) represent proteins/genes for aflatoxin biosynthesis; these are the most differentially expressed genes/proteins of all examined using either measure and appear to respond similar at a protein or gene level. Lack of points in the upper left and lower right quadrants reveals there are no gene/proteins that are in clear conflict, only cases where significance was seen using one measure or the other. The lower left and upper right indicate protein levels that share regulation with the transcripts.

In addition to the automated SILAC ratio determination by PepQuan, we chose to manually examine peptides with ratios in the upper and lower 2.5% of the total distribution for our relative quantification data in both the 28 °C vs 28 °C data and 28 °C vs 37 °C data. We defined the points in the outer 2.5% of the ratio distribution as outliers and used Xcalibur to manually verify peptides and their SILAC ratios. We continued to look at the outer 2.5% of the SILAC ratio distribution for peptides to check manually until all peptides within these outer ranges appeared to be correct.

Manually calculated values were different from those automatically obtained from PepQuan because some peaks were not amenable to accurate ICIS automated peak detection (baseline window = 200, area noise factor = 20, peak noise factor = 50, minimum peak width = 3, multiplet resolution = 10, area tail extension = 10, area scan window = 0). The outer ranges of each distribution were the only proteins examined since for this experiment concerns were greater for preventing false positives (claiming something is different when it is not) than false negatives (claiming something is not different when it really is).

RNA Isolation and Microarray Hybridization. RNA was isolated from lyophilized cultures as previously described using a Trizol (Invitrogen, Carlsbad, CA) procedure according to the manufacturer’s suggestions\textsuperscript{15} Lyophilized fungal tissue was ground with a pestle in a mortar containing liquid nitrogen. The ground tissue was suspended in Trizol reagent and extracted with chloroform, and the protein was precipitated.
using a high salt isopropanol precipitation followed by a LiCl precipitation. A NanoDrop spectrophotometer was used to quantify RNA. RNA was sent to the Purdue University Genomics Core Facility (West Lafayette, IN) for further quality analysis of the RNA and hybridization and chip scanning according to standard procedures used for Affymetrix GeneChips (Affymetrix, Santa Clara, CA). Ratios between 28 and 37 °C were determined using the gene set intensity values from the output in the Affymetrix GeneChip operating software (GCOS).

Results and Discussion

Labeling Efficiency and Method Variation. Labeling efficiency was determined on samples grown identically except for the different isotopes of arginine supplemented in the defined medium. Arginine was chosen over lysine for the label for the different isotopes of arginine supplemented in the defined medium. Arginine was chosen over lysine for the label because it occurs at the relatively high frequency of about 6% in the proteome, and enzymes such as trypsin will cleave C-terminal to arginine. A total of 12 847 proteins from the recently sequenced A. flavus genome were included in our SEQUEST search database. In Figure 2, the distribution of recently sequenced C-terminal to arginine. A total of 12 847 proteins from the proteome, and enzymes such as trypsin will cleave

$C$-terminal to arginine. A total of 12 847 proteins from the recently sequenced A. flavus genome were included in our SEQUEST search database. In Figure 2, the distribution of recently sequenced C-terminal to arginine. A total of 12 847 proteins from the proteome, and enzymes such as trypsin will cleave C-terminal to arginine. A total of 12 847 proteins from the recently sequenced A. flavus genome were included in our SEQUEST search database. In Figure 2, the distribution of recently sequenced C-terminal to arginine. A total of 12 847 proteins from the proteome, and enzymes such as trypsin will cleave C-terminal to arginine. A total of 12 847 proteins from the recently sequenced A. flavus genome were included in our SEQUEST search database. In Figure 2, the distribution of recently sequenced  

Figure 5. Extracted ion chromatogram (EIC) for selected peptides from data set of differentially expressed proteins. The top row of shows EIC from peptides (12C6-Arg) at 37 °C and the bottom row shows the EIC from peptides (13C6-Arg) at 28 °C. (A and C) The MAP kinase (Osm1) protein was more abundant at 37 °C than 28 °C. (B and D) The Casein kinase I protein was more abundant at 28 °C than 37 °C; neither of the above proteins showed differences in transcript levels in response to temperature. To further our comparison of these two data sets (28 vs 28 °C), we found 210 proteins to be in common between experiments. When quantification values were compared for the same proteins, it was apparent that the distribution of the 28 vs 37 °C data set was wider than the 28 vs 28 °C data set (Figure 3). As expected, temperature difference did not result in a significant change in the relative abundance of many proteins. Peptides within each data set also correlated well with the predicted intact protein size based on the gel slice they came from and gel slice number was highly linear ($R^2 = 0.866$) between the two data sets (Supplementary Figure 1).

Expression of Specific Genes and Proteins Involved in Aflatoxin Biosynthesis. The same fungal culture conditions used for deciphering proteome changes with SILAC were used...
research articles
to assay the effect of temperature on gene transcription. Gene transcription was measured on an Affymetrix GeneChip
cs array containing elements representing all of the predicted genes in A. flavus. Ratios for gene expression at the two
temperatures were compared to ratios of protein expression from the same genes (Figure 4). The customary 2-fold change
in gene expression was selected as a biologically relevant value for assigning expression differences in the transcript data,
whereas the aforementioned direct comparison between the 28 vs 28 °C data set was used to determine acceptable SILAC
values. Interestingly, expression levels for all the detected enzymes in the aflatoxin pathway except for one met the
assigned cutoff for both data types (black triangles, ▲, Figure 4). There was a moderate correlation between transcript and
protein levels for the same gene (pairwise correlation = 0.4334). A strong correlation between transcript level and protein
concentration would not be expected as translation rates and protein turnover influence protein levels and alter the final
observed ratio. Although the majority of values were similar with regards to whether a gene or protein was differentially
expressed, there were several cases where protein expression was different and the encoding transcripts showed no dif-
ferential expression. Figure 5 shows two examples of extracted ion chromatograms for two identified proteins that do not show
differential gene expression. Indeed, previous studies have shown that for many genes, transcript and protein concentra-
tions differ in their levels.16 These observed differences rein-
force the importance of being able to monitor both transcript and protein accumulation in studies aimed at understanding
complex biological processes. The modified SILAC procedure described here will allow quantification of protein changes in
complex prototrophs, and will allow investigators to identify responses that would not have been observed if only one
method was applied. While there were some interesting proteins that did not show a change in transcripts (see Table 1),
there were no examples in which transcript accumulation and protein accumulation for the same gene were in direct
conflict with one another.

Conclusions
We report the use of a modified SILAC procedure for a complex prototroph that allows the quantification of protein
expression. The procedure was powerful enough to easily detect an increase in the concentration of aflatoxin biosynthetic
enzymes at 28 °C, a temperature that is conducive for the transcription of the pathway genes and the accumulation of aflatoxin.
The measurement of changes in aflatoxin enzymes validates that the SILAC method is an ideal approach to quantitatively measure biological changes in response to environmental stimuli. Additionally, these results reveal that regulation of the aflatoxin cluster is at the transcript level and that proteins do not accumulate from any lowly expressed
genomes in the aflatoxin cluster at 37 °C. Current studies are underway to specifically characterize the mechanism control-
ing gene expression in the aflatoxin cluster in response to temperature. Many of the proteins identified in this study may
prove to play a role in this regulation. Interestingly, many proteins whose concentration changed in response to temper-
ature were encoded by corresponding RNA transcripts whose expression did not appear to change. Some of these proteins
(Table 1) are involved in signaling and may not be regulated at a transcriptional level. They potentially represent interesting
proteins for further study because their concentration may be

regulated by factors other than gene transcription. Given that
the analyzed samples were prepared to include mainly cytosolic proteins (nondenaturing gel banding and 100 000 g spin),
it is possible some of the changes in concentration were due
to cellular localization or the interaction of these proteins with
hydrophobic membrane associated proteins.

Whereas SILAC has been reported to take several days in
other cell culture systems,17 our labeling strategy for A. flavus
was very rapid. Further, we were able to achieve a high level of
15N-arginine incorporation into proteins of this prototrophic
fungus. We anticipate that even greater labeling could be
achieved by increased subculturing of the fungus on labeled
media. A more efficient strategy to prevent dilution of the
labeled arginine by endogenously produced arginine would be
the development of an arginine auxotroph that could not
synthesize arginine. Arginine auxotrophs are available and can
be utilized in cases where appropriate; however, we chose to
label a wild-type strain as auxotrophic strains are often not as
pathogenic and we would like to use this procedure to study
changes in proteins levels during fungal–host interactions. We
were surprised at how quickly the fungus accumulated the
labeled arginine. Apparently, the fungus quickly sensed the
presence of arginine in the medium and shut down its own
biosynthesis. It is also apparent from our high labeling ef-

iciency that the starting inoculum did not contain enough
15N-arginine to significantly dilute the added labeled arginine.
The procedure should have broad application for studying
protein changes in fungi. Most fungi can be cultured from
microscopic spores that contain low concentrations of stored
amino acids. Likewise, most fungi share with A. flavus the
ability to efficiently sense nutrients in their environment and
shut down endogenous production of the amino acid. These
two traits allow the efficient labeling of proteins through the
addition of targeted amino acids to the culture medium. These
biological traits along with the increasing number of fungi with
complete genome sequences make this an ideal procedure to
increase our understanding of the biology of these organisms.

Data Availability. The data associated with this manuscript may
be downloaded from Tranche (http://tranche.proteomecommons.
org) using the following hash: 9B9UVG+OQG+eG5+yTYSo
B2p3yF1wHwFp3bkkZa33C6ytv9MomOcn1H6boZDzfl2aQxsnc
CH6b2ICnOjowoAAAAAAAdGc=. The Tranche hash can also
be used to verify that files have not changed since publication.

Acknowledgment. The authors thank Timothy S.
Collier for his assistance in revising the manuscript. We
acknowledge the W.M. Keck Foundation and the NIH/NCSU
Molecular Biotechnology Training Program for their
generous support of this research.

Supporting Information Available: Comparison of the
values for the gel slice number for peptides in common
between the 28 vs 37 °C and 28 vs 28 °C data set. This material
is available free of charge via the Internet at http://pubs.acs.org.

References
Cleveland, T. E.; Payne, G. A. Appl Environ Microbiol. 1994, 60,
2408–2414.
362.

43
Temperature-Dependent Regulation of Proteins in A. flavus

SUPPORTING INFORMATION

**Supplement 1.** Values for the gel slice number for peptides in common between the 28 °C vs. 37 °C and 28 °C vs. 28 °C dataset were compared. Values for gel slice correlated well between data sets as denoted by the strong linear relationship ($R^2 = 0.866$). Gel slice also correlated well with molecular weight in both datasets (28 °C vs. 37 °C dataset $R^2 = 0.75$ and 28 °C vs. 28 °C dataset $R^2 = 0.69$).
Chapter 4

Beyond aflatoxin: four distinct expression patterns and functional roles associated with *Aspergillus flavus* secondary metabolism gene clusters

D. Ryan Georgianna¹, Natalie D. Fedorova², James L. Burroughs¹, Andrea L. Dolezal¹, Bok, J.³, Horowitz-Brown, S.³, Charles P. Woloshuk⁴, Jiujiang Yu⁵, Nancy P. Keller³, and Gary A. Payne¹*

¹Department of Plant Pathology, Center for Integrated Fungal Research, North Carolina State University, Raleigh, NC 27695, USA

²J. Craig Venter Institute, Rockville, MD 20850, USA

³Department of Medical Microbiology and Immunology, Department of Plant Pathology, University of Wisconsin, Madison, WI 53706, USA

⁴Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907, USA

⁵Southern Regional Research Center, USDA/ARS, New Orleans, LA 70124, USA

Keywords: Secondary Metabolism, Cyclopiazonic Acid, Aflatoxin

*To whom correspondence should be addressed: Campus Box 7244 Raleigh, NC 27695, (919)515-6994, gary_payne@ncsu.edu
SUMMARY

Species of *Aspergillus* produce a diverse array of secondary metabolites, and recent genomic analysis predicts that these species have the capacity to synthesize many more compounds. It has been possible to infer the presence of 55 gene clusters associated with secondary metabolism in *A. flavus*, however, only three metabolic pathways - aflatoxin, cyclopiazonic acid (CPA), and aflatrem - have been assigned to these clusters. To gain insight into the regulation of, and infer ecological significance for the 55 secondary metabolite gene clusters predicted in *A. flavus*, we examined their expression over 28 diverse conditions. Variables included culture media and temperature, fungal development, colonization of developing maize seeds, and misexpression of laeA, the global regulator of secondary metabolism. Hierarchical clustering analysis of expression profiles allowed us to categorize the gene clusters into four distinct clades. Gene clusters for the production of aflatoxins, CPA, and seven other unknown compound(s) were identified as belonging to one clade. To further explore the relationships found by gene expression analysis, aflatoxin and CPA production were quantified under five different cell culture environments known to be conducive or non-conducive for aflatoxin biosynthesis and during colonization of developing maize seeds. Results from these studies showed that secondary metabolism gene clusters have distinctive gene expression profiles. Aflatoxin and CPA were found to have unique regulation but are similar enough that they would be expected to co-occur in commodities colonized with *A. flavus*.
INTRODUCTION

Aspergillus flavus first came to notoriety for its production of the highly carcinogenic secondary metabolite aflatoxin. Aflatoxins represent a family of mycotoxins that occur worldwide in several important food and feed crops, including maize, peanuts, and seeds of nut trees, colonized by Aspergillus flavus or A. parasiticus. Aflatoxin B1 (AF) is the most carcinogenic member of the family and the one most carefully studied. The presence of aflatoxins in food and feed is regulated by the Food and Drug Administration in the U.S., and by other agencies in most countries. Because of health concerns and the worldwide distribution of AF contamination, environmental and genetic regulation of the AF biosynthetic pathway has been studied extensively since its discovery in the early 1960s (Bhatnagar et al. 2006; Georgianna and Payne 2008; Yu and Keller 2005).

Aflatoxins are not the only mycotoxins produced by A. flavus. Over fourteen described mycotoxins are known to be produced by A. flavus (see www.aspergillus.org.uk), although ten of these are derived from the AF biosynthetic pathway. All characterized mycotoxins in A. flavus are secondary metabolites produced by genes organized in physical clusters throughout the genome. Genomic analysis of A. flavus has led to the prediction of 55 secondary metabolism gene clusters by the Secondary Metabolite Unknown Regions Finder (SMURF) (Khaldi et al. unpublished).
Cluster prediction was using the SMURF program which searches for genes encoding multifunctional enzymes (encoded by backbone genes) associated with four classes of secondary metabolites. These included nonribosomal peptide synthases (NRPSs) for nonribosomal peptides, polyketide synthases (PKSs) for polyketides, hybrid NRPS-PKS enzymes for hybrids, and prenyltransferases (PTRs) for indole alkaloids (Hoffmeister and Keller 2007; Keller et al. 2005). Once these putative multifunctional enzymes are identified, SMURF explores neighboring genes for domains commonly found in enzymes associated with secondary metabolism. This process defines the extent of the physical genomic regions for each identified cluster.

The metabolites produced by most of 55 predicted secondary metabolism clusters in *A. flavus* are unknown. The fact that one secondary metabolism cluster can produce numerous products suggests that *A. flavus* has the potential to produce a very diverse repertoire of secondary metabolites. Although secondary metabolites are by definition not essential they can have important functions aside from being dangerous mycotoxins. As an example they can be important in host-pathogen relationships with some providing protection against environmental stresses as well as fungivory (Rohlfs et al. 2007). Many secondary metabolites also include beneficial products for human health such as the antibiotic penicillin and the cholesterol lowering drug lovastatin (Endo et al. 1976; Hoffmeister and Keller 2007; Keller et al. 2005). Thus, understanding the conditions under which the genes of these clusters are expressed may allow us to predict when these compounds may be found in food and feed as well as give potential clues for ecological roles.
We recently began a series of studies that examine gene expression in *A. flavus* with a whole genome Affymetrix GeneChip. Expression data sets have been collected from a variety of experimental conditions. These conditions include field-inoculation of maize kernels, inoculation of kernels in different stages of development, inoculation of autoclaved mature kernels and kernel tissues, incubation of infected kernels at temperatures conducive and non-conducive for AF biosynthesis, and field-inoculation of maize with the domesticated species *A. oryzae*. Other experiments have examined the effect of AF conducive and non-conducive culture media and temperature, and the effect of *laeA*, the global secondary metabolism regulator of gene expression (Kale et al. 2008). LaeA misexpression was also examined during conditions favorable for sclerotia development (Horowitz-Brown et al. unpublished). The overall goal of these studies has been to better understand the factors governing morphological development, secondary metabolism biosynthesis and pathogenicity of *A. flavus*.

The goal of our study was to establish expression patterns of the putative secondary metabolism clusters under a variety of experimental conditions and to associate gene clusters with known functions to others having no currently accepted roles. In addition, we used data obtained from our microarray studies to examine the expression of genes in all 55 gene clusters, predicted by SMURF, during growth under conditions known to influence AF biosynthesis. We hypothesized that the variety of conditions examined would result in the identification of groups of secondary metabolites that are likely to be produced under similar conditions or environments. We show that the
expression of genes in the cluster responsible for the secondary metabolite cyclopiazonic acid (CPA) share a similar transcription profile with genes in the AF cluster.

RESULTS

Transcriptional analysis of 55 predicted gene clusters in *A. flavus*

Expression data obtained from 28 experiments (Table 1) were used to identify patterns among the genes in the 55 different secondary metabolism clusters predicted by SMURF (Table 2). A hierarchical clustering analysis was used to determine which clusters showed the most similar patterns of transcription over the 28 conditions. Attempts to use all of the predicted genes contained within each of the 55 secondary metabolism clusters yielded results that were not easily interpreted. Therefore, the genes encoding the “backbone” enzymes predicted with SMURF for each cluster were chosen. These enzymes are likely to be essential for metabolite biosynthesis, and our assumption is that expression of the corresponding gene should be representative of gene cluster biosynthetic activity. In 13 of the 55 clusters, more than one backbone enzyme was found. Another advantage of the method we used is that it is not affected by errors in prediction cluster boundaries and it is not biased by the expression of non-structural genes within the cluster such as transcription factors, which may follow a different expression pattern from other genes within a secondary metabolism cluster (O'Brian *et al.* 2007).
Distance relationships derived from the hierarchical clustering analysis revealed four discernable expression patterns, designated clades A-D (Figure 1). Notably, a few backbone enzymes from the same cluster failed to group together. For example, the three backbone enzymes of cluster 26 exhibited expression patterns associated with clades A,

<p>| Table 1: Experimental conditions used for analysis of secondary metabolism gene clusters |
|-----------------|-----------------|
| Number | Name | Description |
| 1 | Blister | Inbred maize line B73 kernels identified at the blister, milk, dough, and dent stages of development were used for each treatment. Pin-bar needles were dipped into a 1E6 conidia/mL conidial suspension of NRRL3357. The needles were then inserted into a row of kernels, introducing approximately 11-13 conidia into each kernel. Five rows were inoculated per ear and harvested after 4 d. |
| 2 | Milk | |
| 3 | Dough | |
| 4 | Dent | |
| 5 | Autoclaved Endosperm | Kernels from maize line B73 were autoclaved and dissected to include the endosperm, germ, or whole kernel. Tissue sections were placed in vials, inoculated with 1E5 conidia NRRL3357, and incubated at 29°C for 5 d. |
| 6 | Autoclaved Germ | |
| 7 | Autoclaved Whole Kernel | |
| 8 | B73 48 h | At the late milk-early dough stage of development, maize line B73 kernels were inoculated as described above for the Blister, Milk, Dough, and Dent experiments. Harvesting took place 48 h, 60 h, 72 h, and 96 h after inoculation. |
| 9 | B73 60 h | |
| 10 | B73 72 h | |
| 11 | B73 96 h | |
| 12 | 28C maize | Kernels from maize line B73 were placed in vials, inoculated with 1E6 conidia NRRL3357, and incubated at 28°C or 37°C for 5 d. |
| 13 | 37°C maize | |
| 14 | 28C A&amp;M | 100 ml of modified A&amp;M medium in a 500 ml flask were inoculated with conidia from NRRL3357 at 1E6 conidia/mL. Shake cultures (200 rpm) were grown for 24 h at either 28 or 37°C. |
| 15 | 37C A&amp;M | |
| 16 | RIB40 maize | Inbred maize line B73 ears in the milk stage were inoculated with a pin-bar dipped in 1E6 conidia/ml of strain NRRL3357 or RIB40 and harvested after 4 d. |
| 17 | NRRL3357 maize | |
| 18 | RIB40 wheat bran | A mix of 5 g wheat bran and 5 mL dH2O was autoclaved and mixed with 1E8 conidia of either strain NRRL3357 or RIB40 and analyzed after 2 d growth. |
| 19 | NRRL3357 wheat bran | |
| 20 | WT 6 h | 50 ml liquid YEP medium (6% peptone, 2% yeast extract) was inoculated with NRRL3357 or RIB40 and incubated for 24 h. After 24 hours, the mycelium was collected and incubated in the aflatoxin-stimulating YES medium for 6 and 24 hr (220 rpm, 29°C). |
| 21 | WT 24 h | |
| 22 | laeA deletion 6 h | |
| 23 | laeA deletion 24 h | |
| 24 | laeA OE 6 h | |
| 25 | laeA OE 24 h | |
| 26 | WT sclerotia | 10E5 conidia were inoculated into liquid GMM + 2% sorbitol in a 60 x 15 mm Petri plate. Plates were incubated at 29°C under continuous dark for 6 d. |
| 27 | laeA deletion sclerotia | |
| 28 | laeA OE sclerotia | |</p>
<table>
<thead>
<tr>
<th>Cluster</th>
<th>Gene Name</th>
<th>Accession</th>
<th>Clade</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>polyketide synthase, putative</td>
<td>AFLA_002000</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>dimethylallyl tryptophan synthase, putative</td>
<td>AFLA_004300</td>
<td>C</td>
</tr>
<tr>
<td>3</td>
<td>nonribosomal peptide synthase, putative</td>
<td>AFLA_005440</td>
<td>B</td>
</tr>
<tr>
<td>4</td>
<td>polyketide synthase PksP</td>
<td>AFLA_008170</td>
<td>B</td>
</tr>
<tr>
<td>5</td>
<td>nonribosomal peptide synthase, putative</td>
<td>AFLA_009120</td>
<td>C</td>
</tr>
<tr>
<td>6</td>
<td>PKS-like enzyme, putative</td>
<td>AFLA_009140</td>
<td>C</td>
</tr>
<tr>
<td>7</td>
<td>polyketide synthase, putative</td>
<td>AFLA_010000</td>
<td>C</td>
</tr>
<tr>
<td>8</td>
<td>nonribosomal peptide synthase, putative</td>
<td>AFLA_010010</td>
<td>D</td>
</tr>
<tr>
<td>9</td>
<td>nonribosomal peptide synthase, putative</td>
<td>AFLA_010020</td>
<td>D</td>
</tr>
<tr>
<td>10</td>
<td>nonribosomal peptide synthase, putative</td>
<td>AFLA_010580</td>
<td>C</td>
</tr>
<tr>
<td>11</td>
<td>NRPS-like enzyme, putative</td>
<td>AFLA_036600</td>
<td>C</td>
</tr>
<tr>
<td>12</td>
<td>enterobactin esterase IroE-like, putative</td>
<td>AFLA_041050</td>
<td>D</td>
</tr>
<tr>
<td>13</td>
<td>dimethylallyl tryptophan synthase, putative</td>
<td>AFLA_047190</td>
<td>A</td>
</tr>
<tr>
<td>14</td>
<td>L-ornithine N5-oxigenase SidA</td>
<td>AFLA_047200</td>
<td>A</td>
</tr>
<tr>
<td>15</td>
<td>PKS-like enzyme, putative</td>
<td>AFLA_053770</td>
<td>C</td>
</tr>
<tr>
<td>16</td>
<td>NRPS-like enzyme, putative</td>
<td>AFLA_053780</td>
<td>C</td>
</tr>
<tr>
<td>17</td>
<td>polyketide synthase, putative</td>
<td>AFLA_053870</td>
<td>D</td>
</tr>
<tr>
<td>18</td>
<td>polyketide synthase, putative</td>
<td>AFLA_060680</td>
<td>A</td>
</tr>
<tr>
<td>19</td>
<td>polyketide synthase, putative</td>
<td>AFLA_062860</td>
<td>B</td>
</tr>
<tr>
<td>20</td>
<td>polyketide synthase, putative</td>
<td>AFLA_064560</td>
<td>C</td>
</tr>
<tr>
<td>21</td>
<td>polyketide synthase, putative</td>
<td>AFLA_066720</td>
<td>D</td>
</tr>
<tr>
<td>22</td>
<td>hybrid NRPS/PKS enzyme, putative</td>
<td>AFLA_066980</td>
<td>A</td>
</tr>
<tr>
<td>23</td>
<td>polyketide synthase, putative</td>
<td>AFLA_069330</td>
<td>A</td>
</tr>
<tr>
<td>24</td>
<td>nonribosomal peptide synthase Pes1</td>
<td>AFLA_070800</td>
<td>D</td>
</tr>
<tr>
<td>25</td>
<td>ACV synthetase PcbAB</td>
<td>AFLA_079290</td>
<td>C</td>
</tr>
<tr>
<td>26</td>
<td>PKS-like enzyme, putative</td>
<td>AFLA_079380</td>
<td>D</td>
</tr>
<tr>
<td>27</td>
<td>polyketide synthase, putative</td>
<td>AFLA_096700</td>
<td>C</td>
</tr>
<tr>
<td>28</td>
<td>polyketide synthase, putative</td>
<td>AFLA_096710</td>
<td>C</td>
</tr>
<tr>
<td>29</td>
<td>polyketide synthase, putative</td>
<td>AFLA_096770</td>
<td>C</td>
</tr>
<tr>
<td>30</td>
<td>polyketide synthase, putative</td>
<td>AFLA_101700</td>
<td>D</td>
</tr>
<tr>
<td>31</td>
<td>polyketide synthase, putative</td>
<td>AFLA_104210</td>
<td>C</td>
</tr>
<tr>
<td>32</td>
<td>polyketide synthase, putative</td>
<td>AFLA_104240</td>
<td>C</td>
</tr>
<tr>
<td>33</td>
<td>polyketide synthase, putative</td>
<td>AFLA_104250</td>
<td>A</td>
</tr>
<tr>
<td>34</td>
<td>polyketide synthase, putative</td>
<td>AFLA_105190</td>
<td>B</td>
</tr>
<tr>
<td>35</td>
<td>polyketide synthase, putative</td>
<td>AFLA_105450</td>
<td>D</td>
</tr>
<tr>
<td>36</td>
<td>polyketide synthase, putative</td>
<td>AFLA_105850</td>
<td>D</td>
</tr>
<tr>
<td>37</td>
<td>polyketide synthase, putative</td>
<td>AFLA_114820</td>
<td>D</td>
</tr>
<tr>
<td>38</td>
<td>polyketide synthase, putative</td>
<td>AFLA_116220</td>
<td>C</td>
</tr>
<tr>
<td>39</td>
<td>polyketide synthase, putative</td>
<td>AFLA_116900</td>
<td>C</td>
</tr>
<tr>
<td>40</td>
<td>polyketide synthase, putative</td>
<td>AFLA_118440</td>
<td>C</td>
</tr>
<tr>
<td>41</td>
<td>polyketide synthase, putative</td>
<td>AFLA_118940</td>
<td>D</td>
</tr>
<tr>
<td>42</td>
<td>polyketide synthase, putative</td>
<td>AFLA_119960</td>
<td>C</td>
</tr>
<tr>
<td>43</td>
<td>polyketide synthase, putative</td>
<td>AFLA_121520</td>
<td>B</td>
</tr>
<tr>
<td>44</td>
<td>polyketide synthase, putative</td>
<td>AFLA_125630</td>
<td>C</td>
</tr>
<tr>
<td>45</td>
<td>polyketide synthase, putative</td>
<td>AFLA_126500</td>
<td>C</td>
</tr>
<tr>
<td>46</td>
<td>polyketide synthase, putative</td>
<td>AFLA_126710</td>
<td>B</td>
</tr>
<tr>
<td>47</td>
<td>polyketide synthase, putative</td>
<td>AFLA_127990</td>
<td>C</td>
</tr>
<tr>
<td>48</td>
<td>polyketide synthase, putative</td>
<td>AFLA_128060</td>
<td>D</td>
</tr>
<tr>
<td>49</td>
<td>polyketide synthase, putative</td>
<td>AFLA_135490</td>
<td>C</td>
</tr>
<tr>
<td>50</td>
<td>polyketide synthase, putative</td>
<td>AFLA_139410</td>
<td>A</td>
</tr>
<tr>
<td>51</td>
<td>hybrid PKS/NRPS enzyme, putative</td>
<td>AFLA_139490</td>
<td>A</td>
</tr>
</tbody>
</table>

Table 2: Predicted secondary metabolism gene cluster and accession numbers for backbone enzymes predicted by SMURF
C and D, respectively. It is possible that cluster 26 is more than one cluster merged together or that it contains non-functional backbone genes.

Clade A contains the aflatoxin cluster (54) and eight other predicted gene clusters (1, 16, 19, 23, 24, 26, 36, and 55). Characteristic to this clade is moderate to high levels of gene expression in most of the 28 experimental conditions examined. Based on the predicted multifunctional synthases (Table 2) in these clusters, we inferred their products to include diverse classes of secondary metabolites including polyketides, non-ribosomal peptides, pigments, and siderophores.

Clade B contains eleven gene clusters (4, 5, 9, 10, 11, 20, 23, 31, 37, 48, and 50). Characteristic of this group is low gene expression in liquid shake culture media (the A&M media (Mateles and Adye 1965) experiments and laeA experiments). The expression values in other experiments also were not as high as the ones associated with genes in clade A.

Clade C contains genes from 25 of the 55 predicted clusters. Gene expression in these clusters appeared low across almost all experimental conditions. These expression results suggest that products from these clusters are not being actively synthesized under the conditions tested. For example, these metabolites may be involved in competition with other fungi, bacteria and other soil inhabitants and only synthesized during co-cultivation with these microorganisms.

Clade D contains 18 different gene clusters. Most of the backbone enzymes monitored for these clusters appeared only slightly more expressed than those in clade C. A group of eight clusters within clade 2 were highly expressed during only two
Figure 1. Hierarchical clustering of backbone enzymes. Cluster numbers of each backbone enzyme are on the right side. Experiments 1-28 (Table 1) are listed on the bottom edge. Colors represent intensity of gene expression with increasingly darker colors indicating higher expression. Letters A-D represent the 4 clades discussed in more detail in the text.
The biosynthesis of aflatoxin is thought to require two clusters located on separate chromosomes (Zhang et al. 2004), its other cluster, number 32, does not have a backbone enzyme for representation in Figure 1.

**Figure 2.** Location of cluster 55 in the genome of *A. flavus*. Cluster 55 is located immediately distal to the aflatoxin cluster (cluster 54) near the telomere of chromosome 3. Cluster 55 spans approximately 20 kb and contains 4 genes, a major facilitator superfamily protein (*mfs1*), an FAD oxidoreductase (*maoA*), a dimethylallyl tryptophan synthase (*dmaA*), and a hybrid polyketide/non-ribosomal peptide synthase. AFLA_139450 (*hypF*) and CO1467294 (*hypG*) have also been characterized in our results. This figure was created using the genome browser (Stein et al. 2002) for AFL available at www.aspergillusflavus.org.

**Characterization of the CPA cluster**

Among the nine clusters in clade A, cluster 55 is located nearest the AF cluster in the genome. Cluster 55 spans a 20 kb region of DNA near the telomere of Chromosome 3, immediately distal to the AF cluster (Figure 2). Until recently, the genes in cluster 55 had remained uncharacterized. Tokuoka et al. (2008) reported that the PKS-NRPS encoded within this cluster is necessary for cyclopiazonic acid (CPA) production in *A. oryzae*. Based on its proximity to the polyketide synthase, we disrupted a gene (designated *dmaA*) encoding a putative dimethylallyl tryptophan synthase and studied its
function in CPA biosynthesis. By homologous recombination, we obtained a transformant, number 39, with a deleted *dmaA*. Integration of the disruption vector at the *dmaA* locus was verified by amplification of a 5 kb DNA fragment from the mutant with PCR primers 7 and 8 (Figure 3A and 3B). To determine the impact of *dmaA* deletion on CPA biosynthesis, transformant 39 and the wild-type (WT) strain were grown in potato dextrose broth medium. After 24 hr, the medium was analyzed by LC-MS/MS for the presence of CPA. As can be seen in Figure 4, a distinct peak with mass 337.2 Da at 5.45

![Figure 3. Overlap PCR strategy for *dmaA* deletion construct and predicted integration into the genome. A) Using overlap PCR with primers numbered 1 to 6, a DNA deletion construct was prepared with the upstream and downstream flanking regions of the dimethyl allyl tryptophan synthase gene flank the *argD* marker. After transformation, the presence of this construct at the *dmaA* locus was determined by screening DNA from the transformants with external PCR primers 7 and 8. B) Gel showing DNA bands amplified from wild-type (WT) and transformant 39 (39) when screened with primers 7 and 8. The PCR band of 5 kb is predicted for replacement of the WT gene, *dmaA*, with the deletion construct.](image)
min retention time, representing CPA, was observed in the analysis of culture medium from the WT strain, but not in culture medium from transformant 39. Two additional transformants with predicted gene deletions for *dmaA* were also selected, and these transformants also tested negative for CPA (data not shown).

We also examined the potential role of two putative genes, *hypF* and *hypG*, located between the AF cluster and the CPA cluster in AF and CPA production. *HypF* encodes the hypothetical protein AFLA_139450, and *hypG* (CO146294.10) ([Figure 2](#)) was identified in an EST library. The *hypG* gene was not predicted during annotation of

![Figure 4](#). Selected ion chromatogram for CPA. Full scan LC-MS chromatogram displayed for detection of mass 337 da corresponding to CPA.
the *A. flavus* genome. The predicted gene product from *hypG* was found to have weak homology to transcription factors in *A. niger* and *P. chrysogenum*. We hypothesized that they may be necessary for CPA cluster gene transcription. Both genes were deleted in strain CA14Δku70ΔpyrG. The deletion strains ΔhypF and ΔhypG were grown in potato dextrose broth medium for 48 hr. RT-PCR analysis (Figure 5) showed that neither deletion of *hypF* nor *hypG* had an effect on transcription of genes within the AF (*aflK*) or CPA (*dmaA*) clusters. Additionally, mycotoxin analysis showed no effect on accumulation of CPA or AF in either gene deletion mutant (data not shown). The *hypG* gene was highly expressed whereas *hypF* was not (Figure 5). Neither of the predicted genes was expressed in their respective deletion strain, however it appears that *hypF* may either be needed for expression of *hypG* or the promoter for *hypG* may be contained within *hypF* as expression of *hypG* was not detected in ΔhypF. Regardless, the data indicate that these two genes do not play an essential role in AF or CPA biosynthesis.

*Figure 5.* RT-PCR for gene products in ΔhypG, ΔhypF and parent strain CA14Δku70ΔpyrG. Each strain was assayed for expression of AF cluster gene *aflK*, CPA cluster gene *dmaA*, *hypG*, and *hypF*. The *gpdA* gene was used as a control.
Accumulation of CPA and AF

Studies have shown many strains of *A. flavus* to commonly produce both AF and CPA (Chang *et al.* 2005; Gallagher *et al.* 1978; Martins and Martins 1999; Widiastuti *et al.* 1988). We further examined the production of these mycotoxins more closely by testing the effects of specific culture conditions that are known to affect AF biosynthesis.

Concentrations of both AF and CPA increased over time (48 h, 60 h, and 96 h) in maize kernels infected by strain NRRL3357, indicating that this substrate is favorable for the production of the two mycotoxins (Figure 6). Growth of the fungus on defined media, however, showed AF and CPA biosynthesis to respond differently to the carbon and nitrogen source of the medium and to culture temperature.

The medium most conducive for AF production contained sucrose as the carbon source, ammonium salt as the nitrogen source, and had a pH of 4.5 (shown as the 28°C treatment in Figure 7). While CPA was produced on all media examined, the most favorable carbon and nitrogen sources for CPA production differed from that for aflatoxin production. Lactose, for example supported the greatest amount of CPA production. Increasing the temperature from 28°C to 37°C or the pH from pH 4.0 to pH 8.0 supported more CPA production whereas it inhibited AF production. Addition of the antioxidant gallic acid to the medium reduced AF production but had no measurable affect on CPA production.
Figure 6. CPA and AF concentrations in maize kernels after 48 h, 60 h, and 96 h infection.

Figure 7. Response of CPA and AF concentrations to alteration of temperature, pH, nitrogen source, carbon source, and addition of the antioxidant gallic acid. Each medium used for growth of *A. flavus* was identical to the most aflatoxin conducive condition, listed as 28C, except for modification of pH, nitrogen source, carbon source, temperature, or addition of antioxidant.
DISCUSSION

The results presented in Figure 1 allow further insights into the regulation and ecological roles of secondary metabolites in *A. flavus*. Functions for most secondary metabolites, including CPA and AF, are not known. They are assumed to be important for niche adaptation and thus confer a fitness advantage in particular environments. From a functional perspective, multiple secondary metabolites may have overlapping or synergistic effects to optimize fitness (Challis and Hopwood 2003). Characterization of fungi lacking the ability to produce multiple mycotoxins may provide further insight into the potential ecological roles for these compounds.

The four different gene clusters (1, 16, 26, 36) represented in the upper grouping of clade A of Figure 1 appear to be constitutively expressed at moderate to high levels. It is plausible that their products may have important roles in basal functions such as growth and development of the fungus. As an example, secondary metabolites can be important effectors of differentiation, often acting as signaling molecules known to influence sporulation and germination (Demain and Fang 2000). Siderophores are known to influence germination through their ability to solubilize iron (Horowitz *et al.* 1976). Consistent with our hypothesis of these secondary metabolites being involved in basal functioning, the *sidA* gene, an ortholog of the *A. flavus* siderophore biosynthesis gene from cluster 16, has been shown to be essential for growth of *A. nidulans* (Eisendle *et al.* 2003).
A possible ecological role for the products from the three better characterized gene clusters within the remainder of clade A (24, 54, and 55) could be tolerance to oxidative stress. CPA, produced by cluster 55, was found to prevent lipid peroxidation caused by patulin (Riley and Showker 1991). Mutants for pes1, the NRPS in cluster 24, show increased sensitivity to oxidative stress (Reeves et al. 2006). Finally, AF, produced by cluster 54, has been proposed to be involved in quenching reactive oxygen (Campbell 2005). Researchers have shown that aflatoxin is induced by oxidative stress (Jayashree and Subramanyam 2000). In A. parasticus, deletion of the antioxidant enzyme gene yapA resulted in altered timing of AF biosynthesis in culture media and maize seeds (Reverberi et al. 2007; Reverberi et al. 2008). A possible functional relatedness of the products from three of the gene clusters in clade A as antioxidants may be one possible explanation for their close grouping. We predict that the other secondary metabolite gene clusters in this clade may also produce metabolites that are related to oxidative stress responses as well as basal functioning.

In plants, exposure to a fungal pathogen results in many defense responses including production of reactive oxygen species (ROS) such as superoxide, hydrogen peroxide, and hydroxyl free radicals (Ferreira et al. 2006). The potential shared functions of the three members of clade A may explain the strong expression of each of these gene clusters in living maize kernels (Figure 1). Deletion of laeA in Aspergillus species has been shown to result in reduced pathogenicity (Bok et al. 2005; Rohlfs et al. 2007; Sugui et al. 2007), including A. flavus (Amaike and Keller 2009; Kale et al. 2008). While regulation of gene clusters by LaeA extends beyond just those in clade A, it is
conceivable that lack of one of the metabolites or an antagonistic affect from lack of all of these metabolites in the previous LaeA studies could have resulted in the decreased potential for pathogenicity/survival of A. flavus.

We observed that the metabolites in clade B of Figure 1 were more highly expressed on solid media compared to liquid media. This led us to hypothesize that some of these clusters are involved in conidiogenesis. Within clade B, the gene pksP (alb1) from cluster 5 has been implicated in A. fumigatus for production of dihydroxynaphthalene (DHN) – like melanin (Tsai et al. 1998). Consistent with our hypothesis, conidia from A. fumigatus strains lacking the pksP gene appear white instead of the typical bluish-green and smooth rather than echinulate. Strains lacking pksP were also less virulent than wild-type strains (Langfelder et al. 1998; Tsai et al. 1998). The DHN cluster contains six different genes in a 19 kb region in A. fumigatus (Tsai et al. 1999). This cluster does not appear to be intact in A. flavus; arp1, another gene in the A. fumigatus DHN cluster is found in A. flavus cluster 10 (clade A). The gene arp2 does not appear to be present in A. flavus and abr1, abr2, and ayg1 have genes sharing significant homology located elsewhere in the A. flavus genome. Among the clusters in clade B, cluster 31 appeared to be the most highly coordinated with the expression of cluster 5. The only conditions that appeared to contrast were expression in A. oryzae and expression in the Δl aeA strain under sclerotia favoring conditions. Interestingly, Δl aeA strains have been found to show slightly reduced conidiation (Kale et al. 2008).

For clades A and B we observed that expression often appeared greater in A. flavus strain NRRL3357 compared to A. oryzae strain RIB40. The different expression
patterns of these identified clusters between the two species could provide clues for their function. Expression in *A. flavus* and not *A. oryzae*, a domesticated fungus important for food fermentations, may indicate a role for these compounds unique to the ecology of *A. flavus*. Alternatively, but not exclusive of the role postulated above, is that these compounds are toxic. During the domestication of *A. oryzae*, strains would have been selected for the lack of toxic compounds.

Many of the biosynthetic pathways present in *A. flavus* may be silent. Our results suggest that several of the 25 different clusters in clade C may be silenced at the epigenetic level. Researchers have observed enhanced chemical diversity when treating *A. flavus* and other fungal species with DNA methyltransferase inhibitors and histone deacetylase inhibitors (Shwab *et al.* 2007; Williams *et al.* 2008). A genomic analysis of gene expression changes in response to epigenetic modifiers could prove to be very revealing. It is also conceivable that secondary metabolism pathways represented in clade C could produce their respective metabolites in very low quantities, or simply be expressed in response to conditions not encountered within these experiments.

Clade D appears to contain some metabolites that are only expressed in the dark during conditions favorable for sclerotia production. While aflatrem has no known role, deletion of *veA*, a gene necessary for sclerotia formation, was found to inhibit biosynthesis of aflatrem (Duran *et al.* 2007). Interestingly several antiinsectan products thought to prevent fungivory have been found in sclerotia (Whyte *et al.* 1996; Wicklow *et al.* 1996). We hypothesize that the six clusters within clade D (12, 15, 27, 35, 39, 46) from the subgroup that shows expression under conditions 26 and 28 will be involved in
development or survival of sclerotia. These six clusters were not in the ΔlaeA strain from condition 27 which was grown under the same conditions as 26 and 28. Fittingly, ΔlaeA strains do not produce sclerotia (Kale et al. 2008).

We more closely characterized cluster 55, which has now been shown to be necessary for CPA production (Chang et al. 2009; Tokuoka et al. 2008). While the function of the MFS predicted to be part of the CPA cluster has not yet been verified, Chang et al. (2009) recently characterized dmaA, which they named dmaT. They also described an essential role in CPA biosynthesis for the predicted FAD oxidoreductase in cluster 55 and named its gene maoA. The shared regulation we have observed for the four genes from cluster 55 shown in Figure 2 provides strong evidence that they are part of a gene cluster necessary for CPA biosynthesis. Whether there are additional structural genes required for the biosynthesis of CPA located elsewhere in the genome is not yet known.

Several transcriptional profiles have been performed on a 5002 element oligo array including an experiment analyzing a veA deletion strain for affects on secondary metabolism and fungal development (Duran et al. 2007) as well as an experiment analyzing a deletion strain for aflR, the AF pathway specific transcription factor (Price et al. 2006). These 5002 element microarrays contained only a limited number of probes for secondary metabolite genes aside from those for AF biosynthesis. Only 90 of the 443 secondary metabolism genes from the 55 gene clusters predicted by SMURF had probes on the microarrays; however, three of the four genes predicted for the CPA cluster were included on these microarrays.
Interestingly, no pathway-specific transcription factor has been described for CPA biosynthesis and no gene with predicted sequence to a transcription factor resides in cluster 55. Chang et al. (2009) deleted a predicted zinc-finger transcription factor found distal to the PKS-NRPS but saw no effect on CPA. This gene is on the A. flavus Affymetrix GeneChip and was not significantly transcribed under any of the 25 experimental conditions examined in this study. Additionally we have shown that the two predicted genes located between the CPA and AF clusters do not appear to be essential for biosynthesis of either metabolite.

Since the AF and CPA gene clusters are located next to each other we investigated whether the AF pathway specific transcription factor AflR could control genes in the CPA cluster. Price et al. (2006) studied the effect of AflR on gene expression by deleting the AF pathway transcription factor aflR. A closer look at the microarray data revealed no significant differences for any of the three CPA biosynthetic genes contained on the 5002 element arrays used for this study. There is one caveat to the aflR deletion experiment regarding CPA. These studies were performed using A. parasiticus, which is not known to produce CPA (Dorner et al. 1984); however, a closer examination of the data shows a strong expression signal in both the WT and deletion mutant for 2 of the 3 genes needed for CPA production present on the 5002 element array. This result indicates that the lack of CPA production in A. parasiticus is likely not a result of a deficiency in gene expression from these two genes.

With our newfound knowledge of the location of the CPA cluster we were able to further explore some previous results examining A. flavus strain 649. The afl-1 mutation
in *A. flavus* strain 649 is the only known dominant mutation inhibiting AF biosynthesis, and thus AF biosynthesis can not be restored in parasexual diploids resulting from pairing *afl-1* mutants with WT AF producing strains (Papa 1979). This dominant mutation is thought to inhibit AF biosynthesis through an unknown silencing mechanism that can be overcome through addition of ectopic copies of the pathway regulatory gene *aflR* (Smith *et al.* 2007). Strain 649 does not contain the AF cluster or the CPA cluster as a result of a large chromosomal loss. The silencing mechanism resulting from the *afl-1* mutation is likely very specific to AF biosynthesis or possibly AflR. Unknown to Smith *et al.* (2007) when they assayed for affects of *afl-1* on genes outside the aflatoxin cluster, one of the genes measured was the dimethylallyl tryptophan synthase (*dmaA*) of the CPA cluster (denoted 16TV in the Smith et al (2007) study). This gene was found to be expressed at WT levels in the diploid between the *afl-1* mutant strain 649 and AF producing strain 86. We examined the same 86x649 diploid and found it to produce CPA (data not shown), indicating that all of the genes necessary for the biosynthesis of CPA in the WT strain are expressed in the diploid. Based on these results, we hypothesize that if there is a regulatory gene specific to the CPA cluster, it is at a locus outside the deleted region on chromosome 3 of the *afl-1* mutation.

The physical grouping of genes for secondary metabolism means that entire clusters may be controlled through epigenetic regulation. LaeA is a predicted histone methyltransferase that has been described as a global regulator of secondary metabolism (Bok and Keller 2004). Regulation of secondary metabolism by LaeA is thought to be through chromosomal remodeling. Another well studied protein, VeA, is thought to affect
LaeA activity through signaling of the heterotrimeric G-protein α-subunit, FadA, or by direct interaction with LaeA (Bayram et al. 2008; Calvo 2008). Deletion of veA was found to inhibit accumulation of AF, aflatoxin, and CPA (Duran et al. 2007). As expected, when we examined array data from Duran et al. (2007), the veA mutant displayed lowered transcription for some of the genes for the AF, aflatoxin, and CPA clusters. Since VeA is believed to interact and control secondary metabolism through involvement with LaeA (Bayram et al. 2008), the phenotype of strains that lack functional LaeA and VeA should appear similar. As a result of the physical proximity of the AF and CPA cluster we predict that when conditions control both CPA and AF it may function through epigenetics.

Until now, very little was known about the regulation of CPA in *A. flavus*. Previous research by Gqaleni et al. (1997) suggested an interaction between water activity and temperature, with conditions of higher water activity showing less inhibition of CPA and AF biosynthesis by temperature (Gqaleni et al. 1997). However, substantial evidence suggests AF is strongly inhibited by temperature regardless of medium (Georgianna et al. 2008; Georgianna and Payne 2008; O'Brian et al. 2007), and based on gene expression, unlike AF, CPA biosynthesis is not reduced in liquid culture (A&M) or maize kernels when grown at 37°C (O'Brian et al. 2007; Smith et al. 2008). Interestingly, *A. nidulans* produces the AF precursor sterigmatocystin more abundantly at 37°C (Feng and Leonard 1998).

CPA is also regulated similar to sterigmatocystin biosynthesis in *A. nidulans* with regard to nitrogen source. Nitrate, like temperature, affects the biosynthesis of
sterigmatocystin differently from that of AF, with more sterigmatocystin being produced on nitrate rather than the more AF conducive nitrogen source, ammonium (Feng and Leonard 1998). Nitrogen source, pH, and carbon source are thought to directly effect the expression of genes through cis-regulatory elements related to environmental sensing of nitrogen, pH and carbon. The proteins that interact with these regulatory elements include global transcription factors such as AreA for nitrogen (Caddick et al. 1994), PacC for pH (Penalva et al. 2008), and the various regulatory DNA binding proteins associated with carbon catabolite repression (Ebbole 1998). Many aflatoxin genes possess these regulatory elements but appear to respond differently than sterigmatocystin and CPA; however proteins like PacC can act as both activators and repressors of gene expression (Cary et al. 2006; Georgianna and Payne 2008; Penalva et al. 2008). From these data it appears that while CPA does not respond identically to AF for the culture conditions tested, it responds very similar to how we would expect sterigmatocystin to accumulate in *A. nidulans*. Despite the differential accumulation of CPA and AF in response to culture media environment, **Figure 6** showed us that these two metabolites accumulate almost identically in a natural environment such as maize.

The goal of our study was to establish expression patterns and associate gene clusters with known functions to others having no currently accepted roles. With this information we were able to associate gene clusters with known functions to others having no currently accepted roles. Additionally, we have been able to determine the identity of the metabolite produced by predicted gene cluster 55 as CPA. CPA has the potential to become an overlooked mycotoxin that could have significant health risks
(Nishie et al. 1985). Our data show that under conditions favorable for the production of AF, CPA also will likely be produced, assuming that the strain has the capacity to produce CPA. As not all strains of *A. flavus* produce aflatoxin, some do not produce CPA due to mutations or loss of the gene cluster (Chang et al. 2005; Geiser et al. 2000). Since both metabolites can be detected simultaneously (see experimental procedures) it is not difficult to include CPA testing when checking for AF. Additionally, CPA may be more commonly encountered than AF due to its biosynthesis by both *Aspergillus* and *Penicillium* species (Frisvad and Thrane 2000; Hermansen et al. 1984; Le Bars 1979; Pitt et al. 1986) and its production over a wide range of conditions.

**EXPERIMENTAL PROCEDURES**

**Fungal strains.** *A. flavus* strains NRRL3357, AFC-1, and CA14Δku70ΔpyrG were used in these studies. Strain 3357-5, a uracil auxotroph of NRRL3357 (He et al. 2007), was used to create a new strain with both an arginine and uracil requirement, called AFC-1. This strain was created through use of *pyr4*-blaster, a plasmid containing complementary kanamycin cassette sequences flanking the *pyr4* gene from *Neurospora crassa*. The construct is designed such that the marker gene can be forced to loop out under selection on 5-FOA (5-FOA is metabolized to the toxic compound fluorodeoxyuridine in strains containing an active 5’-orotidine decarboxylase gene like *pyr4*) similar to the method used by d’Enfert (1996) with *pyrG* in *A. fumigatus* (d'Enfert
The loss of this marker upon growth on medium containing 5-FOA and uracil restored the requirement for uracil in the fungus.

We used pyr4-blaster to delete the argD gene in 3357-5 and subsequently mutagenized with 5-NQO as described by He et al. (2007) and selected on 5-FOA (100 μg/ml) (Zymo Research, Orange, CA) containing MLS medium with nucleic acid supplements uracil (7.5 mM) and uridine (7.5 mM) plus the amino acid arginine (1.5 mM) at pH 3 to create the AFC-1 double mutant (-pyrG, -argD) (note: agar was autoclaved separately and added to a filter sterilized pH 3 solution of the 5-FOA containing media). The 5-NQO step was added because the background in strains simply selected on 5-FOA was not sufficient for genetic transformations. Strain AFC-1 shows no background when selecting against either marker plus produces AF at similar levels to the parent strains 3357 and 3357-5.

The A. flavus strain CA14Δku70ΔpyrG has a ku70 deletion background which results in high rates of homologous recombination for more efficient gene targeting. CA14Δku70ΔpyrG was kindly provided from Perng-Kuang Chang (Southern Regional Research Center, ARS, USDA, New Orleans, LA).

Gene deletion, fungal transformation, DNA isolation. An overlap PCR method (Davidson et al. 2002) was used to create gene deletion constructs with argD as the selectable marker (Figure 5A). Primers were designed from DNA sequence downloaded from the genome browser (Stein et al. 2002) for A. flavus (www.aspergillusflavus.org). Genomic DNA from strain NRRL3357 was used as a PCR template for both the 5’ and 3’
homologous flanking regions of the genes targeted for deletion in this study. Amplification of \( \text{argD} \) included sufficient upstream sequence to include the native promoter. For fungal transformation we used the methods suggested by He et al. (2007) and currently available at www.aspergillusflavus.org. The transformation selection media, MLS, was supplemented with 1.12 g/L uracil in order to use the \( \text{argD} \) containing deletion construct with strain AFC-1. DNA was isolated from transformants for PCR based screening using the CTAB method described by He et al. (2007). Screening primers were designed to amplify outside of the \( \text{dmaA} \) deletion construct to screen for its integration in the genome (Figure 5A, primers 7 and 8). Other genes were deleted in strain CA14\( \Delta \text{ku70}\Delta \text{pyrG} \) using the same method but with \( \text{pyr4} \) as the selectable marker with the appropriate ~1kb flanking regions and MLS as the selective media.

**Culture media for CPA and AF production.** A total of five different cell culture conditions known to influence AF biosynthesis (Mahoney and Molyneux 2004; Price et al. 2005) were tested for their effect on the accumulation of CPA and AF. NRRL3357 was grown in either A&M medium (Mateles and Adye 1965) or A&M medium modified for carbon source (50 g/L), nitrogen source (3 g/L), pH or the presence of antioxidants. All media were buffered at pH 4.5 with 100 mM citric acid except for the pH 8.0 treatment, which was buffered with 100 mM Tris. The following five conditions were compared: 1) the AF conducive carbon source, sucrose, was compared to lactose, 2) the AF conducive nitrogen source, ammonium sulfate, was compared to sodium nitrate, 3) the AF conducive temperature, 28°C, was compared to 37°C, 4) the AF conducive pH 4.5
was compared to pH 8.0, and 5) A&M medium containing 2.5 mM gallic acid (n-propyl gallate, MP Biomedicals, Inc. Solon, OH), an antioxidant that inhibits AF (Mahoney and Molyneux 2004), was compared to standard A&M medium. A 2-ml aliquot of the medium was placed in 24-well plates (Corning Life Sciences Inc. Lowell, MA) and inoculated with A. flavus to a final concentration of 1x10^6 conidia/ml. The plates were wrapped with parafilm and incubated at 28°C, except for the 37°C treatment. Each treatment was replicated three times.

**RNA isolation and RT-PCR.** Mycelia were filtered from 48 h PDB cultures grown at 28°C with shaking at 200 rpm using miracloth filters along with vacuum flasks. This tissue was lyophilized and used for RNA isolation with the Qiagen Plant RNeasy kit (Qiagen Inc. Valencia, CA). After RNA isolation the concentration was determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE).

A total of 3 μg of RNA was using for DNase treatment with Promega RQ1 DNase according to the manufacturer’s suggestion (Promega, Madison, WI). After DNase treatment 1.5 μg of the RNA was used for reverse transcription (RT) using the first-strand cDNA sythesis protocol for SuperScript™ II RT (Invitrogen Corp. Carlsbad, CA). PCR was performed on the resultant cDNA with the appropriate gene specific primers with Takara ExTaq (Takara Bio, Otsu, Shiga, Japan) following the manufacturer’s suggestions with a total of 30 one minute extension cycles. Primers were purchased from Sigma-Aldrich. PCR products were analyzed on a 0.8% TAE agarose gel.
Microarray analysis. Data from experiments using the *A. flavus* Affymetrix GeneChip microarrays were imported into JMP genomics (SAS, Cary, NC). All array data was deposited in the NCBI GEO database as experiment GSE15435. The arrays included data from the following experiments: 1) a 28°C vs. 37°C comparison in A&M culture media from Georgianna et al. (2008), 2) a 28°C vs. 37°C experiment on detached field-grown corn kernels from Smith et al. (2008), 3) a time course study in which developing seeds were inoculated in the field with *A. flavus* and seeds were assayed every 24 hr from 48-96 hr, 4) a study in which developing seeds were inoculated at four different corn development stages and assayed 4-5 days later (Dolezal et al unpublished), 5) *A. flavus* grown on either the endosperm or embryo of autoclaved corn kernels (Woloshuk et al. unpublished), 6) a comparison of a wild-type (WT) *A. flavus* to an laeA deletion strain and laeA overexpression strain at 6h and 24h (Bok and Keller, unpublished), and 7) a comparison of *A. flavus* (NRRL3357) to *A. oryzae* (RIB40) grown on field inoculated corn and wheat bran koji (Georgianna et al. unpublished) and 8) a comparison of a wild-type (WT) *A. flavus* to an laeA deletion strain and laeA overexpression strain under sclerotia forming conditions (Sigal S.B., Yu, J. et al. unpublished). These experiments are described in further detail in Table 1. Data from each experiment were corrected for background, normalized, and summarized in JMP genomics using the RMA normalization procedure (Irizarry et al. 2003a; Irizarry et al. 2003b). The mean value for all replicates of each individual experiment was taken. We analyzed the subset of values for all backbones enzymes associated with secondary metabolism clusters predicted in SMURF. Hierarchical clustering using the Fast Ward
procedure was performed in JMP genomics for the backbone enzyme expression values across all microarrays experiments.

**Detection and Quantification of CPA and AF.** All experiments were performed with the Thermo LTQ ion trap instrument (Thermo Scientific, Bremen, Germany) at the NCSU Genome Sciences Laboratory (GSL; http://gsl.cals.ncsu.edu/). For quantification of CPA and AF B₁, standards were purchased (Sigma-Aldrich, St. Louis, MO). The solvent system used consisted of 25 mM Morpholine and Methanol at 50 μl/min on a Thermo Hypersil Gold C18 column (3 μm particle size, 150 x 1 mm). A 15 min linear gradient was used starting at 20% MeOH increasing to 80% MeOH followed by a hold at 80% MeOH for 9 min and column equilibration at 20% MeOH for 12 min. The mass spectrometer was set to run in ESI positive mode with six different scans, one Full Scan from 200-2000 m/z and then five selected ion MS/MS mode scans for 331 Da (AFG₂), 329 Da (AFG₁), 315 Da (AFB₂), 313 Da (AFB₁), and 337 Da (CPA). All samples used for quantification in A&M media were diluted by 1/10th with 1:4 MeOH:25 mM Morpholine. Dilution of the sample and the use of the highly basic (> pH 10) morpholine limited signal suppressing matrix effects that were observed for CPA using undiluted sample or other carrier ions. A standard curve was created in a similar dilution of A&M media and for each different condition examined a single standard addition was used to verify the accuracy and comparability of quantitative measurements. Values for accumulation of CPA and AF in A&M media were adjusted for dry weights by determining the mass of lyophilized tissue for each replicate. Maize samples were
analyzed similar to the A&M experiments except concentrations were not normalized for biomass.

Qualitative analysis used for determining absence or presence of mycotoxins was performed using a similar method with a direct injection of the sample media onto the column and no reference standards. This method was used for detection of CPA in the deletion mutants created in this study where the deletion was hypothesized to contain no detectable CPA.

ACKNOWLEDGEMENTS

Thanks to USDA CSRS grant 2006-35604-16666 and the NIH/NCSU Molecular Biotechnology Training Program for their generous support of this research.
REFERENCES


Campbell, B. C., 2005 Panel Discussion: Fungal Genomics Workshop, pp. in Aflatoxin Elimination Workshop, Raleigh, NC.


78
Davidson, R. C., J. R. Blankenship, P. R. Kraus, M. de Jesus Berrios, C. M. Hull et al., 2002 A PCR-based strategy to generate integrative targeting alleles with large regions of homology. Microbiology. 148: 2607-2615.


Chapter 5

Comparative Genome Hybridization Reveals Few Genetic Differences between *Aspergillus flavus* and *A. oryzae*

D. Ryan Georgianna\(^1\), Natalie Fedorova\(^2\), Juijang Yu\(^3\), Doug Brown\(^1\), Ralph Dean\(^1\), Charles P. Woloshuk\(^4\), William Nierman\(^2\) and Gary A. Payne\(^1\)

\(^1\)Department of Plant Pathology, Center for Integrated Fungal Research, North Carolina State University, Raleigh, NC 27695
\(^2\)The J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, MD
\(^3\)Southern Regional Research Center, USDA/ARS, New Orleans, LA 70124, USA
\(^4\)Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907
ABSTRACT

*Aspergillus flavus* and *A. oryzae* are two closely related species predicted to have spent centuries in vastly different environments. *A. flavus* is an important opportunistic plant pathogen known for contaminating crops with the carcinogenic mycotoxin, aflatoxin, and *A. oryzae* is a domesticated species used in food fermentations. Remarkably, the genomes of these two species are nearly identical. We used the recently sequenced genomes of *A. oryzae* RIB40 and *A. flavus* NRRL3357 along with array based comparative genome hybridization (CGH) to compare genomes across several strains of these two species. A comparison of three strains from each species by CGH revealed only 42 and 129 genes unique to *A. flavus* and *A. oryzae*, respectively. Further, only 709 genes were identified as being polymorphic between the species. Despite the high degree of similarity between these two species we were able to use data from the CGH arrays to reveal that there are enough species derived genomic differences to separate each of the six strains into their respective species. Analysis of the CGH data for genomic differences in the subset of genes contained in secondary metabolism clusters revealed a different separation of strains inconsistent with the classification of *A. flavus* and *A. oryzae* as separate species. Our findings suggest that while these two species are nearly identical they can be distinguished as unique species in a whole genome analysis with multiple strains. Our data further suggest that genes for secondary metabolism may be evolving differently from those for the genome as a whole.
INTRODUCTION

The filamentous fungus *Aspergillus flavus* (AFL) is widely distributed throughout the world, but more commonly found in tropical to subtropical regions (Klich 1992). It is recognized as a health concern because it produces aflatoxin, a carcinogen that can contaminate economically important crops such as corn, peanuts, cotton, and several varieties of tree nuts that support growth of AFL. AFL is both a post- and pre-harvest problem and aflatoxin is currently the only mycotoxin regulated by the U.S. Food and Drug Administration. Acceptable levels in the U.S. are less than 20 parts per billion (ppb) in food or feed and less than 2 ppb in milk, resulting in losses in the U.S. alone of $270 million annually (Richard and Payne 2003). AFL is also an important opportunistic pathogen of several organisms; in addition to being a plant pathogen it is a pathogen of animals and the second leading cause of invasive aspergillosis in humans (Hedayati et al. 2007). The genome of AFL strain NRRL3357 was recently sequenced, this stain is commonly used in the lab for studying aflatoxin biosynthesis and pathogenicity.

*Aspergillus oryzae* (AO) is a koji-mold used in fermentation for a variety of food products. Its cultivation on rice grains, soybeans, and wheat bran dates back at least 2-3 thousand years (Machida et al. 2008). *Aspergillus oryzae* is thought to be a domesticated species or ecotype derived from AFL (Payne et al. 2008) and is not usually prevalent in the environment. Traditional fermentation foods requiring the secreted enzymes made by AO include miso, soy sauce, and sake. The genome of AO strain RIB40 has been sequenced (Machida et al. 2005).
For several decades it has been known that the genomes of AFL and AO are similar (Kurtzman et al 1986). Comparison of AO and AFL at the genomic level has revealed strong conservation of gene content and order between these species (Payne et al. 2006). This led us to examine the question of whether hallmarks of AO's domestication can be detected in its genome. This analysis focuses on characterizing a set of “core” genes that potentially may be unique between the species and a set of genes that may be uniquely polymorphic between the two species.

**MATERIALS AND METHODS**

*Fungal strains*

Six different strains were used in this study, three characterized as AFL and three characterized as AO. These strains were selected to include diverse members of each species. Details for each strain are described in Table 1. All three AO strains were provided by the Japanese National Research Institute for Brewing. NRRL3357, the sequenced strain of AFL, and NRRL1957, the type species strain, were obtained from the USDA-ARS culture collections. Strain IC277 was provided by Ignazio Carbone (NC State University, Dept. of Plant Pathology) and is now in the USDA-ARS culture collections as NRRL29506.
**Affymetrix GeneChip microarrays**

Custom AFL Affymetrix GeneChip microarrays (Santa Clara, CA) were used for CGH analysis; these arrays were originally designed to examine gene expression. The arrays contain 13,051 different probesets designed for hybridization to genes from the AFL NRRL3357 genome. Each probeset contains eleven 25 bp probes to represent one gene. The arrays also include probes tiled across the intergenic regions of the aflatoxin gene cluster. Of particular interest for studies in this manuscript is the presence of probes for 426 different AO RIB40 genes on these arrays. These probes are for genes predicted to be missing from the AFL NRRL3357 genome by *in silico* comparison of NRRL3357 and RIB40 genome sequences. The AFL GeneChip also includes probes designed towards genes from *Zea mays* and *Mus musculus* that were not used in this study.

**Isolation of genomic DNA**

Genomic DNA was extracted from lyophilized mycelial tissue from the six *Aspergillus* strains using the Qiagen Plant Maxi Kit (Qiagen, Valencia, CA). The Qiagen protocol was modified to include an initial step of grinding tissue with a pestle in a
mortar containing liquid nitrogen and 150-200 micron glass beads. After the final elution from the Qiagen columns, DNA samples were further cleaned by a phenol:chloroform extraction, followed by EtOH precipitation.

Biotin labeling for Affymetrix GeneChip genomic DNA hybridization

DNA from each strain was labeled using previously described methods (Lieu et al. 2005). A total of 300 ng of genomic DNA for each sample was labeled using Klenow polymerase and random primers to incorporate a dNTP mix containing biotin conjugated dCTP (Stratagene BioPrime Kit, Agilent Technologies, Cedar Creek, TX). The labeling reaction was performed at 25C for 16 h to create short fragments of labeled DNA. The reactions were cleaned using a phenol chloroform extraction followed by DNA precipitation. DNA was re-suspended in 100 μl of water and 5 μl was analyzed on a 2% agarose gel to check for the presence of fragments of approximately 30 bp. The remaining DNA was sent to the Purdue Genomics Laboratory for hybridization to the AFL Affymetrix GeneChip using standard laboratory protocols. Data from all genome hybridization arrays are available from the NCBI Gene expression omnibus (www.ncbi.nlm.nih.gov/geo) as data series GSE1560. This analysis was done for six different strains of Aspergillus, three AO strains and three AFL strains (Table 1).

Correlation analysis for strain phylogeny

All Aspergillus CGH array data were used to perform a correlation analysis between arrays. Using data imported from Affymetrix .CEL files, a phylogeny was
created by a Pearson correlation with the correlation and principal components function in JMP Genomics v6 (SAS Institute Inc, Cary, NC). Correlations from intensities for each of the six *Aspergillus* strains from every *Aspergillus* probe on the Affymetrix GeneChip were converted into distances that were used for clustering to construct the dendrogram representing a phylogenetic tree.

**Determining gene content using CGH**

Genomes were analyzed *in silico* using the GMAP module of the PASA package (J. Craig Venter Institute, Rockville, MD) to identify unique genes by mapping coding sequences to the other species genomes (AO RIB40 and AFL NRRL3357).

All CGH data were Loess Normalized using JMP genomics v6. To develop a threshold for the Affymetrix GCOS Present/Absent (P/A) algorithm calls in CGH data from NRRL3357 and RIB40 we limited our dataset to the predicted coding sequences that were analyzed by GMAP and also spotted on the microarray. A set of 12,313 genes predicted to be present in both strains RIB40 and NRRL3357 was used to determine False Negatives and a set of 307 genes predicted to be absent in RIB40 (unique to 3357) and 240 genes predicted to be absent in NRRL3357 (unique to RIB40) to control for false positives. Sixteen of the 24,616 genes (the number of genes predicted to be present in both species based on sequence analyses) in the *in silico* present set were called Absent (A) or Marginal (M) using the Affymetrix GCOS software. This represents a 0.065% occurrence of false negatives based on GCOS P/A calls. The unique NRRL3357 dataset contained 26/306 genes (306 being the number of genes predicted to be missing in
RIB40) called as P/M for RIB40 and the unique RIB40 dataset contained 28/241 (241 being the number of genes predicted to be missing in NRRL3357) genes called as P/M for NRRL3357. Combined this gave a result of 54 out of 547 RIB40 or NRRL3357 specific genes potentially misidentified for a false positive rate of 9.87% based on GCOS P/A calls. Since we had in silico genome data and experimental CGH data for strains RIB40 and NRRL3357 we decided to improve upon GCOS P/A calls by creating an adjustment threshold for GCOS predictions based on probe intensities.

The Loess normalized mean intensity for the 24,600 data points representing genes predicted to be present in both NRRL3357 and RIB40 by in silico analysis was 2797.006 with standard deviation of 874.3796. We wanted to see if we could improve on the GCOS P/A calls and lower the false positive rate. A probe intensity limit of two standard deviations less than the average intensity (2797.006) was used to assign a threshold for changing genes marked as A/M to P. The upper threshold value for these data set was 1048.247, meaning that genes predicted as absent by GCOS with CGH intensity values above this limit were counted as present.

The average Loess normalized intensity for the 493/547 experimentally absent genes in the unique genes dataset had an average Loess normalized intensity of -208.574 with standard deviation 139.397. Again, to control for false positives we assigned a threshold for changing genes marked as P/M to A based on a limit of two standard deviations greater than the average intensity (-208.574) for absent genes. This lower threshold values was 56.213, meaning that genes assigned as present by GCOS below this value were counted as absent.
ANOVA for species-specific polymorphism prediction

CGH data from the six different strains were log2 transformed and Loess normalized. Each probe within the probe set was assigned a unique accession by combining the Affymetrix probe set ID and the individual probe number. A one-way ANOVA model examining species effects was run on this dataset for each individual probe with the false discovery rate (FDR) multiple testing procedureset to 0.05 (Benjamini and Hochberg 1995). This analysis was performed in JMP Genomics v6.

RESULTS AND DISCUSSION

Predicted strain phylogeny

Before proceeding with the analysis to determine core and polymorphic genes, the correlation between microarray probes across all strains was compared and used to create a predicted phylogeny. Six different strains were used in this study, three characterized as \textit{A. flavus} (AFL) and three characterized as \textit{A. oryzae} (AO) (Table 1). Rather than use only the sequenced strains of AFL and AO, NRRL3357 and RIB40 respectively, we chose to use multiple strains in order to prevent bias due to a small sample size. AFL and AO are both polyphyletic with strains of each species falling into distinct clades containing both AO and AFL (Chang \textit{et al.} 2006; Geiser \textit{et al.} 2000). By including multiple strains we increased the possibility of differences being the result of species differentiation rather than differences unique to an individual strain. It was clear from
this analysis that overall the genomes of the three strains of AFL were more similar to each other than to the genomes of AO strains, and the genomes of the AO strains were more similar to each other than those of the AFL strains (Figure 1). These results showed that our genomic comparisons by CGH were able to distinguish described species of AFL from those described for AO.

Previous studies using amplified fragment length polymorphism analysis were able to identify species specific markers between several Aspergillus species yet were unsuccessful in finding markers to consistently distinguish between AFL and AO (Montiel et al. 2003). Since it is difficult to differentiate between AFL and AO, and AO is thought to be domesticated from AFL, many researchers have suggested that there is

---

**Figure 1.** Strain phylogeny represented by correlation of arrays used for CGH analysis.
no basis to classify AFL and AO as separate species (Geiser et al. 1998). Because it is important to understand the relationship between the two species from a regulatory perspective, and new genomic tools are available for comparing fungal species, we reexamined the genetic relationship between AFL and AO. Both gene content and gene polymorphism among three strains of AFL and three strains of AO (Table 1) were examined using CGH. CGH methods provide a much greater resolution of the genome than traditional marker based approaches. While our method contains fewer strains than some previous studies, the Affymetrix GeneChip provides greater coverage of the genome through the examination of 150,865 DNA probes representing 3,771,625 bp of genome sequence (more than 10% of the genome). Our analysis provides the most inclusive evaluation of DNA sequence performed so far between multiple strains of AFL and AO.

Prediction of core genes—genes potentially unique to AFL or AO

We were interested in determining core genes for AFL and AO because we hypothesized that many of these gene losses/gains could be unique adaptations as a result of the diverse environments encountered by both species and possibly be a signature of domestication. A previous examination of Streptococcus species for unique genes has indicated that many of the identified species-specific genes are probably adaptive by increasing fitness of the organism in its ecological niche (Marri et al. 2006).

Core genes for AFL and AO were determined by the adjusted Affymetrix GCOS P/A calls as described in the materials and methods. The adjustment took into account the
deviation from the mean expression values for *in silico* predicted genes from each sequenced strain (AFL NRRL3357 and AO RIB40). This was important because it allowed us to determine gene presence/absence and control for false positives based on known differences between the NRRL3357 and RIB40 genomes. When these limits were applied to the absent and present datasets predicted by GCOS, the false negative rates changed from 0.065% to 0.118%. This was due to the change of several present calls to absent based on our adjustments defined by the *in silico* data. This increase in predicted false negatives may indicate that certain probe sets on the array were not amenable to CGH, meaning that they hybridized poorly or that sequence divergence resulted in lower intensities through less efficient probe hybridization. The adjustment determined by the *in silico* analysis decreased the false positive rate from 9.87% to 2.38%. Barth *et al.* (2000) explored the false positive rate in experiments using CGH to determine copy number on oligo arrays of human lymphoma tissue or blood specimens. Depending on the measure they used to define changes in copy number, false positive rates varied from 3.7% to 19.3%. A fixed threshold procedure, a similar concept to our adjustments based on *in silico* data, resulted in the lowest false positive rate of 3.7% (Barth *et al.* 2000). Thus, our observation of a 2.38% false positive rate appears to be very low. It is possible that some of the remaining false positive gene detections may have had DNA sequence present in the other species but were either pseudogenes or unannotated genes resulting in a higher false positive than negative rate. The intensity thresholds from *in silico* analysis were applied to the GCOS P/A calls for the entire Loess normalized dataset for the six strains examined.
Our analysis showed a total of 1142 genes to be absent in at least one of the six strains examined. Genes that were present in all strains of AO and absent from all strains of AFL were called AO core genes. Similarly, genes that were present in all strains of AFL and absent in all strains of AO were called AFL core genes. There were 129 core genes predicted to be unique to AO and 42 core genes predicted to be unique to AFL (Figure 2). As can be seen in Table 2, the majority of the genes unique to AO or AFL encode predicted proteins with no known function. Only 11 and 12 unique genes from AO and AFL, respectively, encoded proteins with predicted functions. There is EST evidence (Table 2) for 7 of the 42 genes unique to these AFL strains and 14 of the 129 genes unique to these AO strains, thus these genes appear to be expressed.

Fedorova et al. (2008) suggested that in three closely related Aspergillus species, A. fumigatus, A. clavatus, and Neosartorya fischeri, much of the gene loss/gain occurred in genomic islands that represented “gene dumps” or “gene factories” serving to diversify
the genus. This led us to believe that the majority of the gene loss/gain would not occur at random in the genome.

*Predicting species enriched gene polymorphism*

Our previous analysis showed little difference between the two species in core genes. To further compare these strains gene polymorphism was examined among the three strains from each species. It was reasoned that these polymorphisms may be useful for distinguishing AFL from AO, and that they may reveal hallmarks of the domestication of AO that could be important in describing functional differences between the AFL and AO genomes.
Table 2: Descriptions and EST evidence for core genes

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Descriptions</th>
<th>ESTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1569.m00165</td>
<td>conserved hypothetical protein</td>
<td>TC9754 CO134550 CO151757</td>
</tr>
<tr>
<td>1569.m00166</td>
<td>conserved hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00167</td>
<td>Protein kinase domain containing protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00180</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00181</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00183</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00184</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00185</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00186</td>
<td>hypothetical protein</td>
<td>NAFCJ86TV CO137648</td>
</tr>
<tr>
<td>1569.m00187</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00188</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00189</td>
<td>ATPase, AAA family protein</td>
<td>CA747483</td>
</tr>
<tr>
<td>1569.m00190</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00191</td>
<td>hypothetical protein</td>
<td>TC8882 NAFET18TV CO135842 CO142745</td>
</tr>
<tr>
<td>1569.m00192</td>
<td>hypothetical protein</td>
<td>CO144710 CO147467 CO152207</td>
</tr>
<tr>
<td>1569.m00193</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00194</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1569.m00194</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1739.m00052</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1918.m00088</td>
<td>eIF-3A, putative</td>
<td>CA747982</td>
</tr>
<tr>
<td>1918.m01157</td>
<td>hypothetical protein</td>
<td>TC10467 CO134943 CO140525 CO143343</td>
</tr>
<tr>
<td>1918.m01197</td>
<td>hypothetical protein</td>
<td>CO148013 CO148163 CO152019 CO152146</td>
</tr>
<tr>
<td>2043.m00239</td>
<td>conserved hypothetical protein</td>
<td>NAFDD02TV CO139225</td>
</tr>
<tr>
<td>2043.m00396</td>
<td>DUF821 domain protein</td>
<td></td>
</tr>
<tr>
<td>2043.m00398</td>
<td>AMP-dependent synthetase and ligase, putative</td>
<td></td>
</tr>
<tr>
<td>2541.m00106</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2541.m00159</td>
<td>FAD binding domain containing protein</td>
<td></td>
</tr>
<tr>
<td>2689.m00829</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2689.m00917</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2842.m00541</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2856.m00001</td>
<td>pseudogene</td>
<td></td>
</tr>
<tr>
<td>2856.m00002</td>
<td>conserved hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2856.m00003</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2856.m00004</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2856.m00005</td>
<td>Glycosyl hydrolase family 3 N terminal</td>
<td></td>
</tr>
<tr>
<td>2856.m00006</td>
<td>domain containing protein</td>
<td></td>
</tr>
<tr>
<td>2856.m00007</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2856.m00009</td>
<td>hydrolase, CocE NonD family-related nicotinamide mononucleotide permease,</td>
<td></td>
</tr>
<tr>
<td>2856.m00010</td>
<td>,putative</td>
<td></td>
</tr>
<tr>
<td>2856.m00011</td>
<td>Adenosine AMP deaminase family protein</td>
<td></td>
</tr>
<tr>
<td>2856.m00012</td>
<td>NCS1 nucleoside transporter family protein</td>
<td></td>
</tr>
<tr>
<td>2856.m00013</td>
<td>hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>2911.m00742</td>
<td>hypothetical protein</td>
<td></td>
</tr>
</tbody>
</table>
Table 2 continued

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Description</th>
<th>ESTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AO0702270000020</td>
<td>Amino acid transporters</td>
<td></td>
</tr>
<tr>
<td>AO070177000001</td>
<td>Ankyrin repeat</td>
<td></td>
</tr>
<tr>
<td>AO070316000011</td>
<td>Cytochrome P450</td>
<td></td>
</tr>
<tr>
<td>AO070316000107</td>
<td>Cytochrome P450 CYP4 CYP19 CYP26 subfamilies</td>
<td></td>
</tr>
<tr>
<td>AO09013800000151</td>
<td>Nucleoside phosphorylase</td>
<td></td>
</tr>
<tr>
<td>AO070227000003</td>
<td>Predicted NTPase (NACHT family)</td>
<td></td>
</tr>
<tr>
<td>AO09001000000300</td>
<td>Reverse transcriptase</td>
<td></td>
</tr>
<tr>
<td>AO0901380000125</td>
<td>Reverse transcriptase</td>
<td></td>
</tr>
<tr>
<td>AO07022700000118</td>
<td>Serine threonine protein kinase</td>
<td></td>
</tr>
<tr>
<td>AO0702970000018</td>
<td>Synaptic vesicle transporter SVOP and related transporters (major facilitator superfamily)</td>
<td></td>
</tr>
<tr>
<td>AO07032800000151</td>
<td>Translation initiation factor 5A (eIF-5A)</td>
<td>ta0552f</td>
</tr>
<tr>
<td>AO07033600000220</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0703360000219</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0703360000218</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0702920000050</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0701940000001</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0702350000003</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0703190000036</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0703190000036</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0703190000035</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO07033500000302</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0703160000105</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO07031600000106</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO07031600000108</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0703160000112</td>
<td>Predicted protein</td>
<td>ns0528f</td>
</tr>
<tr>
<td>AO0702840000024</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0702840000025</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0702840000026</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0702750000044</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO09002300000941</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070217000011</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0702970000080</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0702970000081</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000418</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0701770000003</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0701770000002</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000422</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000423</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000424</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000425</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000426</td>
<td>Predicted protein</td>
<td>AC9104F</td>
</tr>
<tr>
<td>AO090102000427</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0702060000011</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000429</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000430</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000431</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0701980000008</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AC7630f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AC7824f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN0451f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN0504f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN0629f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AN1581f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JZ1927f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JZ2161f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JZ4212f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JZ4216f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JZ4646f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH0548f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SH1935f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ns0543f</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ns0921f</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 2 continued

**A. oryzae core genes**

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Description</th>
<th>ESTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AO090102000433</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000434</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000435</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000436</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070198000004</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000438</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000439</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000440</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090102000441</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070266000053</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070238000011</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070333000029</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070333000030</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070325000013</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO07021900010</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070219000005</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070219000004</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070219000003</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070219000001</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070313000037</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070313000038</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090120000199</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070314000064</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070221000005</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070221000004</td>
<td>Predicted protein</td>
<td>SH2203F</td>
</tr>
<tr>
<td>AO070221000003</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070221000002</td>
<td>Predicted protein</td>
<td>JZ5304F</td>
</tr>
<tr>
<td>AO070221000001</td>
<td>Predicted protein</td>
<td>JZ2513F</td>
</tr>
<tr>
<td>AO070192000005</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070192000006</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070192000006</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090138000140</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070206000008</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090138000150</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090138000152</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090138000155</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090138000156</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000019</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000016</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000016</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000015</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000013</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000001</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000002</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000003</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000004</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000005</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000019</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000016</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000016</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000015</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070257000013</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000001</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000002</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000003</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000004</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000005</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070313000037</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070313000038</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO090120000199</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070314000064</td>
<td>Predicted protein</td>
<td></td>
</tr>
</tbody>
</table>
### Table 2 continued

**A. oryzae core genes**

<table>
<thead>
<tr>
<th>Identifier</th>
<th>Description</th>
<th>ESTs</th>
</tr>
</thead>
<tbody>
<tr>
<td>AO070211000006</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000007</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000008</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000009</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000010</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000011</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000012</td>
<td>Predicted protein</td>
<td>ns0704f</td>
</tr>
<tr>
<td>AO070211000013</td>
<td>Predicted protein</td>
<td>ns0704f</td>
</tr>
<tr>
<td>AO070211000014</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070211000014</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000031</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000030</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000029</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000027</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000026</td>
<td>Predicted protein</td>
<td>AC1067F</td>
</tr>
<tr>
<td>AO070227000025</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000023</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000022</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000021</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000019</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000018</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000018</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000017</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000016</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000015</td>
<td>Predicted protein</td>
<td>AC9278F AN1606F AN1621F JZ1930F</td>
</tr>
<tr>
<td>AO070227000014</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000011</td>
<td>Predicted protein</td>
<td>HG1177F</td>
</tr>
<tr>
<td>AO070227000010</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000007</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000006</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000005</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000004</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000002</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO070227000001</td>
<td>Predicted protein</td>
<td></td>
</tr>
<tr>
<td>AO0900100000300</td>
<td>Reverse transcriptase</td>
<td></td>
</tr>
<tr>
<td>AO090138000125</td>
<td>Reverse transcriptase</td>
<td></td>
</tr>
<tr>
<td>AO070257000018</td>
<td>Serine threonine protein kinase Synaptic vesicle transporter SVOP and related transporters (major facilitator superfamily)</td>
<td></td>
</tr>
<tr>
<td>AO070316000109</td>
<td>Serine threonine protein kinase Synaptic vesicle transporter SVOP and related transporters (major facilitator superfamily)</td>
<td></td>
</tr>
<tr>
<td>AO070328000151</td>
<td>Translation initiation factor 5A (eIF-5A)</td>
<td>ta0552f</td>
</tr>
</tbody>
</table>
Figure 3. Polymorphic Genes. Identification of polymorphism shared between several members of AO and AFL using CGH analysis. Yellow dots represent probes from absent genes. Green dots represent the polymorphic probes, as expected these are on the AFL side of the plot, given that the probes were designed to be perfect matches to the AFL NRRL3357 genome. Blue dots represent false positives for identifying polymorphic genes based on finding perfect matches to the NRRL3357 and RIB40 Genome sequences. Black dots represent the probes with no significant differences observed. There are 2845 significant probes (-log10 p-value > 3.03), 1708 probes are from absent genes (1512 of which are in the core set for either species), 1052 predicted polymorphic probes and 85 false positive probes.

The approach described in the methods predicted up to 2842 significantly polymorphic regions (probes) between the AO and AFL strains (Figure 3). A total of 1708 probes were removed from this data set as they were from genes predicted to be absent in any one of the individual strains. These probes were removed because we wanted to identify polymorphic regions identical within each species and it is conceivable that a significant result for identifying a polymorphism could occur when the polymorphic region is identical between two strains of a species and then absent in the
other strain. After purging our data of absent genes only 1137 probes remained, representing 745 different probesets (709 different genes). To increase our confidence in the prediction of polymorphisms, each probe sequence was searched against the AO RIB40 and AFL NRRL3357 genomes to identify false positives. A total of 80 probes (7.0% false positive rate from 1137 remaining probes) showed perfect matches to both genomes. These probes were likely mismatches in the non-sequenced strains of AFL or AO, or had high levels of cross-hybridization with other similar DNA sequences. Out of the 1057 verified polymorphic probes (1137 minus 80 false positives) most tend to be single mismatches. The distribution for bit scores from blastn for all predicted polymorphic probes can be seen in supplementary figure S1. A bit score of 42 was most common and these are typically single nucleotide polymorphisms. As expected, given the majority of the probes on the microarray being designed as perfect matches to AFL, 1010 of the 1057 verified polymorphic probes were significantly higher by the ANOVA model for AFL with 41 of the probes higher in AO being from the probes on the microarray designed for the AO RIB40 genome.

Functional changes and polymorphism

To determine possible functional changes that may have resulted from gene polymorphism we examined the occurrence of fifteen different gene ontology (GO)
Figure 4. Biological process Gene Ontology (GO) term analysis. (Top) Percent occurrence of 15 GO terms in the AFL genome. (Bottom) Percent occurrence of 15 GO terms in those described as polymorphic by CGH. Transport shows enrichment for polymorphisms, with an increase of close to 10% overall in the polymorphic subset compared to the AFL genome. Similarly, there was a higher percentage of polymorphism in genes for secondary metabolism. Genes for translation, amino acid phosphorylation, protein catabolism, and protein folding show a lack of polymorphism relative to the AFL genome.
terms that were highly prevalent in the annotated AFL genome. The results of these analyses are summarized in Figure 4. The two groups containing the most polymorphic genes were transport and secondary metabolism. Nearly 10% of the overall identified polymorphisms were in genes with predicted functions in transport. Genes for secondary metabolic processes also were enriched for polymorphic sites. Secondary metabolic processes accounted for only 4.7% of the annotated genes, but accounted for 7.4% of the polymorphic genes. In general, genes for translation, amino acid phosphorylation, protein catabolism, and protein folding showed a lack of polymorphism. This might be expected as these genes likely code for products with essential functions, and little polymorphism can be tolerated for these genes to be functional. In contrast, diversity gained by a higher degree of polymorphism in transport and secondary metabolism genes may be advantageous, especially for adaptation to new niches (Price-Whelan et al. 2006).

Additionally, an examination of the aflatoxin cluster for members of Aspergillus section Flavi with intact clusters (includes AFL and AO) revealed more positive Ka/Ks values than for other Aspergillus species with partial clusters (Carbone et al. 2007b). This indicates evidence of adaptation through the observation of more positive selection.

**Secondary metabolism gene clusters**

We further examined polymorphism in genes for secondary metabolism because of their importance to food safety and possible role in ecological niche adaptation (McDonagh et al. 2008). A total of 55 different secondary metabolism gene clusters have been predicted for AFL (Khaldi et al unpublished). Out of these, 18 different secondary
metabolism gene clusters contained polymorphic genes, including the gene clusters for the mycotoxins, aflatoxin, aflatrem and cyclopiazonic acid (CPA). It is unclear if these polymorphisms contribute to reduced secondary metabolism in AO. However, mutations in the aflatoxin pathway genes have been described in some strains of AO (Lee et al. 2006a; Lee et al. 2006b; Tominaga et al. 2006). This may be true for many of the other secondary metabolism gene clusters as well, although we know that CPA is made in certain strains of AO (Benkhemmar et al. 1985; Tokuoka et al. 2008). Only one secondary metabolism gene, AFLA_005430, a hypothetical protein was predicted to be unique to AFL.

Based on the frequent polymorphism in genes for secondary metabolism observed in this study, it is tempting to speculate that production of secondary metabolites is under strong selection in Aspergillus. This would be consistent with the findings of Carbone et al. (2007a), who has proposed that distinct chemotype lineages underlie divergence and speciation in Aspergillus.

A subset of all genes predicted to be involved in secondary metabolism (Khaldi et al. unpublished) was used to investigate whether CGH could ascertain that secondary metabolism genes may have unique evolutionary history. Interestingly, if we analyze the arrays for correlation based only on intensity among probes in predicted secondary metabolism gene clusters, AFL strain IC277 appears to be more similar to the AO strains (Figure 5). This again suggests that unique selection processes may have occurred on the secondary metabolism genes compared with the rest of the genome. Previous work examining the trichothecene cluster of Fusarium graminearum and the aflatoxin cluster
has shown that the evolution of secondary metabolite clusters can be misleading for phylogenetic analysis (Carbone et al. 2007a; Ward et al. 2002).

**Figure 5.** Strain phylogeny represented by correlation of probes across only genes predicted to be part of secondary metabolism clusters.

*Physical location of polymorphic and unique genes*

As mentioned earlier we observed a physical grouping for core genes from AO and AFL. Experiments examining gene content between *A. fumigatus* and two closely related non-pathogenic species, *Neosartorya fischeri*, and *Aspergillus clavatus*, have revealed genomic islands of species-specific genes located near the telomere (Fedorova et al. 2008). In these experiments they mentioned horizontal gene transfer (HGT) as a possibility for explaining species-specific genes but based on phylogenetic relationships with other species came to the conclusion that gene duplication and divergence was the
more likely scenario for acquiring these genes. This may not be the case for AFL and AO which are much more closely related than *A. fumigatus*, *Neosartorya fischeri*, and *A. clavatus*.

Interestingly, Khaldi and Wolfe (2008) observed that AO has up to 20% more genes than *Aspergillus nidulans* and *A. fumigatus* possibly as the result of HGT, the same would also be true for AFL. They have treated AFL and AO as separate species and suggested that HGT took place before the split of AFL and AO (Khaldi and Wolfe 2008).

In *Lactobacillus* species researchers have found that HGT did not typically affect the synteny of transferred genes. Additionally, for those genes shared between the species high levels of HGT were observed (Nicolas *et al.* 2007). Genes from HGT may be expected to be enriched in polymorphism. For example, many of the genes proposed to be transferred from *Wolbachia*, an endosymbiont, to the nuclear genome of the beetle *Callosobruchus chinensis* have decayed to the point that they are mostly pseudogenes (Nikoh *et al.* 2008). Another possibility for enrichment of polymorphism is the location of recombination hotspots. In humans, it has been observed that polymorphisms may accumulate around areas of intense recombination (Spencer 2006). A recently identified sexual stage in *A. parasiticus* as well as AFL (and likely AO too) make the occurrence of recombination no longer just a proposed ancestral event for these species (Horn *et al.* 2009; Moore and Carbone 2009).

We examined the location of polymorphic and unique genes along the AFL and AO chromosomes to determine if we could identify hot spots for increased
polymorphism. These results are presented in Figures 6 and 7. Most of the AO and AFL genomes have shared synteny. The polymorphic genes reveal this synteny as they are

**Figure 6.** Location of *A. flavus* unique “core” genes and polymorphic genes along the chromosomes. Chromosome arms are the 16 largest scaffolds from JCVI afl1 v2.0 assembly. Secondary metabolism gene clusters are marked with horizontal bars below each chromosome.

**Figure 7.** Location of *A. oryzae* unique “core” genes and polymorphic genes along *A. oryzae* chromosomes.
mainly in the same location for either species. The unique genes could only be mapped to one species or the other since they are not present in both species. We have also marked the location of the 55 different secondary metabolism gene clusters predicted for AFL in Figure 6.

Polymorphic genes can be found throughout the genome and sometimes appear highly concentrated in regions near the telomere. Interestingly, these concentrated regions of polymorphism also tend to overlap with secondary metabolism clusters. The overlap of polymorphic regions with secondary metabolism clusters is consistent with the logic that a fungal species may aim to diversify its arsenal of secondary metabolites. It is thought that this diversity in secondary metabolism could result in a competitive advantage if the environment changes. The unique or core genes can also be found throughout the genome however there appears to be one large cluster of AFL core genes at the telomere of chromosome two and clusters of AO core genes near the telomeres of chromosomes four and six. Maintenance of telomeres is important for genome stability (Kolodner et al. 2002). We might suspect that these genes could have arisen or been lost during a strong selective event in the evolutionary history of AFL and/or AO, one possibility is domestication.

**Outlook for comparative genomics of AFL and AO**

AFL and AO are two extremely important fungal species and a better understanding of their genomes could greatly impact the safety and quality of food. In
this study we have identified key genomic features, both shared and unique to each species, and have gained insight into the dynamics of the AFL and AO genomes through an analysis of multiple strains for each species.

Comparisons of the two sequenced genomes (both \textit{in silico} and by CGH) coupled with CGH data from four other genomes, have identified genomic changes differing between these two species. Now that we have described many of the genomic sequence changes that may define what make AO and AFL two different species, it is clear that there are few. In fact, one may argue that these differences are not sufficient to justify splitting AFL and AO into separate species; however, our data do show sufficient differences between the strains we examined to distinguish between these fungi. DNA sequence variation in secondary metabolism clusters, including those for mycotoxins, do not appear to be able to accurately distinguish between strains described as AFL from AO. It is not known whether metabolomic analysis would able to distinguish between these species. A metabolomic analysis between AO and AFL strains could reveal interesting in determining whether the synthesis of different metabolites is linked to their apparent unique evolutionary lineages.

The core genes and polymorphisms described in this study provide important information for future studies on these two species. It is likely that many key advances can be made by focusing on identified core genes and polymorphisms, some of which we propose could play important roles in pathogenicity and/or mycotoxin biosynthesis. Microarray analysis of gene expression between these two species will be an important first step in determining which unique and polymorphic genes are utilized in either
species. Additionally, we hypothesize that expression analysis will reveal much greater diversity between AFL and AO as compared to the genome sequences.

ACKNOWLEDGEMENTS

Thanks to Chris Smith of the NCSU Bioinformatics Research Center for his assistance in creating Figures 6 & 7.

REFERENCES


Nicolas, P., P. Bessieres, S. D. Ehrlich, E. Maguin and M. van de Guchte, 2007 Extensive horizontal transfer of core genome genes between two Lactobacillus species found in the gastrointestinal tract. BMC Evol Biol. 7: 141.


Supplementary Figure S1. Distribution of mismatch type for probes based on NRRL3357 and RIB40 genome sequence. The red and blue lines represent bit scores from blastn for the 1052 polymorphic probes, the black line represents bit scores for all probes matched against the RIB40 genome. Bit scores of 50 represent a perfect match for the standard 25 bp probes on the Affymetrix GeneChip. Most polymorphic probes were at a bit score of 42, these tend to be a single mutation. A bit score of 48 represents a single mismatch at the end of the sequence, 46 represents two mismatches at the ends of the probe sequence, 44 represents three mismatches at the ends of the probe sequence and 42 represents as mentioned either a SNP or 4 mismatches at the ends of the probe sequence. The small peak at 34 tend to represent two mutations within a probe sequence.
Chapter 6

_Aspergillus flavus_ and the domesticated _A. oryzae_: Two species with nearly identical genomes but distinctive gene expression

D. Ryan Georgianna¹, Natalie Fedorova², Juijiang Yu³, Doug Brown¹, Ralph Dean¹, Charles P. Woloshuk⁴, William Nierman² and Gary A. Payne¹

¹Department of Plant Pathology, Center for Integrated Fungal Research, North Carolina State University, Raleigh, NC 27695
²The J. Craig Venter Institute, 9704 Medical Center Drive, Rockville, MD
³Southern Regional Research Center, USDA/ARS, New Orleans, LA 70124, USA
⁴Department of Botany and Plant Pathology, Purdue University, West Lafayette, IN 47907
ABSTRACT

*Aspergillus oryzae* and *Aspergillus flavus* are two nearly identical members of the genus *Aspergillus*. *A. oryzae* is thought to have been domesticated from the species *A. flavus* many centuries ago. *A. flavus* is most infamous for the contamination of crops with mycotoxins such as the highly carcinogenic aflatoxin. Surprisingly, despite the strong similarity to *A. flavus*, *A. oryzae* is generally regarded as safe (GRAS); this defining characteristic of *A. oryzae* is crucial for its use in industrial food fermentations. We monitored gene expression differences between the sequenced strains of *A. flavus* and *A. oryzae* using whole genome Affymetrix GeneChip microarrays in environmental niches commonly encountered by either species. Gene expression was compared in the two species during growth on wheat bran koji solid state cultures, prepared similar to how *A. oryzae* is propagated prior to being introduced into fermentations, and in field-inoculated B73 maize kernels, a living substrate that is commonly infected by *A. flavus*. Both experimental substrates were conducive to aflatoxin production. Gene expression in the species was compared at a time-point in which the aflatoxin biosynthetic genes were highly expressed in *A. flavus*. Our results showed that despite the very high-level of genome similarity between these two fungi, there was a relatively large set of differentially expressed genes (15-20% of the predicted genes) between the species on either substrate. One of the most striking observations of this study was the highly enriched expression of secondary metabolism genes in *A. flavus* compared to *A. oryzae*. 
INTRODUCTION

Genome sequencing projects for members of the genus *Aspergillus* have lead to several important comparative genomics studies (Galagan *et al.* 2005; Kobayashi *et al.* 2007; Machida *et al.* 2005; Nierman *et al.* 2005b; Payne *et al.* 2006; Rokas *et al.* 2007). Results from these studies have helped define species boundaries, provided insight into the evolution of the diverse species within the genus, and have allowed us to speculate about the biological significance of speciation. Comparative genomics also has increased our understanding of pathogenicity and ecological fitness of species within the genus (Fedorova *et al.* 2005; Nierman *et al.* 2005a; Rokas *et al.* 2007).

The goal of this study was to better characterize the relatedness of two economically important species of *Aspergillus*, one a pathogen and toxin producer, and the other a workhorse for the food industry. *Aspergillus oryzae* (AO) is a koji-mold that has been used in fermentations through solid state cultivation on rice grains, soybeans, and wheat bran dating back at least 2-3 thousand years (Machida *et al.* 2008). Traditional fermentation foods requiring the secreted hydrolases from *A. oryzae* include miso, soy sauce, and sake. Improvement of industrial strains for use in fermentation as well as heterologous protein expression has led to a number of experiments monitoring important classes of genes for fermentation such as the hydrolases and catabolic enzymes (Kimura *et al.* 2008; Maeda *et al.* 2004).

*Aspergillus flavus* (AFL) is the most agronomically important plant pathogen of the *Aspergillus* genus, causing millions of dollars in crop loss every year (Richard and
Payne 2003). AFL is both a pre- and post-harvest pathogen of many crops including maize, cotton, tree nuts, and peanuts. Crop loss is not directly attributed to Aspergillus rot but rather the contamination of these crops with aflatoxin, the only mycotoxin currently regulated by the United States Food and Drug Administration. AFL is also an opportunistic pathogen of animals, including humans, and causes the disease aspergillosis.

AO and AFL are the two closest members of the Aspergilli to have sequenced genomes. Strain RIB40 of AO has a completed genome sequence and recently the genome sequence of strain NRRL3357 from AFL has been released. A comparison between these genomes has provided insight into the genetic diversity of these fungi, showing that at the genomic level the two species are almost identical (Payne et al. 2006). The genomes of the two sequenced strains contain a similar number of predicted genes and share similar genome sizes (Table 1). Further, a comparison of three strains from each species by comparative genome hybridization (CGH) revealed only 43 and 129 genes unique to three strains each of AFL and AO, respectively (referred throughout as “core genes”). CGH analysis also identified only 709 genes as being distinctly polymorphic between the same three strains of each species (Georgianna et al. unpublished). The small differences uncovered by genomic examinations so far cannot solely explain differences between the species based on morphology, chemotype, and ecotype. In this study we examined global gene expression in the two species to determine if they respond differently when cultured on two ecologically relevant substrates.
MATERIALS AND METHODS

Field inoculation of maize

Maize line B73 was grown in Clayton, NC. Kernels at the dough stage of development (R4) were inoculated by pricking the top of the kernel with a stainless steel pin that had been dipped into a solution containing 1 x 10^6 conidia/ml of AFL strain NRRL3357 or AO strain RIB40. Approximately 30 kernels were inoculated per ear and three ears were inoculated in each of three replicates. Kernels were removed from the ears 4 days after inoculation, placed into specimen cups that were then frozen in liquid nitrogen, and stored at -80°C before isolating RNA.

Isolation of RNA from infected kernels

RNA isolated from frozen infected maize kernels contained both fungal and maize RNA. Five kernels for each replicate were placed in a mortar along with 250-300 μm glass beads and ground using a pestle under liquid nitrogen. The ground tissue was placed into 50 ml tubes containing 10 ml phenol pH 6.2 and 10 ml 1.0 M Tris-EDTA pH 8.0. The sample was mixed by vortexing and spun at 12000 x g for 10 min to pellet the

<table>
<thead>
<tr>
<th>Species</th>
<th>Chrom.</th>
<th>Size Mb</th>
<th>Gene No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. fumigatus</td>
<td>8</td>
<td>29.4</td>
<td>9,926</td>
</tr>
<tr>
<td>A. nidulans</td>
<td>8</td>
<td>30.1</td>
<td>9,541</td>
</tr>
<tr>
<td>A. oryzae</td>
<td>8</td>
<td>36.7</td>
<td>12,735</td>
</tr>
<tr>
<td>A. flavus</td>
<td>8</td>
<td>36.8</td>
<td>12,497</td>
</tr>
</tbody>
</table>

Table 1: Genome statistics for *Aspergillus* species
debris. The resulting aqueous layer was removed, added to 5 ml of phenol pH 6.2 plus 5 ml 24:1 chloroform:isoamyl alcohol and mixed by vortexing. The sample was spun again at 12000 x g for 10 min and the aqueous layer removed to a new tube where it was mixed for a second time with 10 ml 24:1 chloroform:isoamyl alcohol and spun at 12000 x g for 10 min. The aqueous layer was removed and added to 20 ml of 95% EtOH and allowed to precipitate overnight at -20 C. After precipitation the sample was spun for 30 min at 12000 x g. The supernatant was removed and the pellet washed with 20 ml of 70% EtOH followed by spinning at 12000 x g for 30 min. After washing with 70% EtOH, the supernatant was removed and the pellet was dried for approximately 15 min. The dried pellet was further purified using the Qiagen RNeasy Plant Kit RNA cleanup protocol according to the manufacturer’s instructions (Qiagen, Valencia, CA), including the optional on-column DNase treatment. The final step the samples were eluted twice from the spin column with 30 μl of RNase free water. Sample RNA concentrations were determined using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE).

*Preparation of wheat bran koji fungal cultures*

Wheat bran Koji was prepared by mixing 5 g of Red Winter Wheat Bran flakes (ShopNatural Tucson, AZ) and 5 ml of dH2O in a 125 ml flask. A total of 18 flasks of this medium were prepared and autoclaved. After cooling, nine flasks were inoculated with a total of 1 x 10⁸ conidia of either strain NRRL3357 or RIB40. After inoculation, the cultures were thoroughly mixed by first stirring the medium with a sterile 1 ml
serological pipet and then vortexing the flask for 1 min. The cultures were incubated at 28C and medium containing the fungus was sampled daily from 1-3 days. Three replicates were sampled at each time point, placed into 50 ml conical tubes, frozen in liquid nitrogen and stored at -80C until extracted for RNA.

**Isolation of RNA from wheat bran koji fungal cultures**

RNA was isolated from wheat bran Koji cultures using a Trizol reagent RNA prep followed by RNA cleanup using the Qiagen Plant RNeasy kit. Frozen tissue was ground in a mortar with glass beads using a pestle. The ground material was added to a 1.7 ml tube containing 1 ml of Trizol, shaken in a disruptor genie for 5 min, and centrifuged at 12000 x g for 5 min to remove the debris. The supernatant was removed and mixed with 5 ml of chloroform. After mixing, the sample was spun at 12000 x g for 5 min and the aqueous layer removed. One additional chloroform extraction was performed before precipitating the sample at -20 C for 1h by addition of 2 volumes 95% EtOH and 1/10 volume of 3 M Sodium Acetate, pH 5.5. After precipitation the sample was spun for 10 min at 12000 x g and the resulting pellet was washed with 1 ml of 70% EtOH followed by spinning 5 min at 12000 x g. The resulting pellet was then used in the Quiagen Plant RNeasy kit following the manufacturer’s instructions and including the optional on-column DNase treatment. The final sample was eluted twice with 30 μl of RNase free water. Sample concentrations were determined using the NanoDrop spectrophotometer.
Microarray hybridization and analysis

RNA was sent to the Purdue University Genomics Core Facility (West Lafayette, IN) for further quality analysis of the RNA quality and for hybridization and chip scanning according to standard procedures used for Affymetrix GeneChip microarrays (Affymetrix, Santa Clara, CA).

Data for each experiment were imported into JMP genomics software. The subset of fungal genes or alternatively the subset of maize genes was normalized according to the RMA normalization procedure (Irizarry et al. 2003a; Irizarry et al. 2003b). Background was corrected using RMA background correction, values were log2 transformed, normalized using quantile normalization, and probe values summarized using a median polish procedure (Holder et al. 2001). Data from the microarrays are available from the NCBI GEO database in experiment GSE15435.

The normalized data were analyzed using a one-way ANOVA specific to each experiment to examine the differences between the species AFL (NRRL3357) and AO (RIB40). An FDR= 0.05 multiple testing procedure was used to assess significant gene expression differences using a –log_{10} p-value cut-off. Empirical Bayes analysis was applied to the model to estimate the variance component and give a slight increase in power/sensitivity (Feng 2006).
RESULTS

Comparative analysis of the DNA sequence of the two sequenced strains of AFL and AO, and comparative genomic analysis (CGH) of three strains each of AFL and AO showed the genomes of AFL and AO to be very similar (Georgianna, unpublished). CGH analysis also revealed little gene polymorphism between the two species. To further characterize these two species we examined their gene expression on two media. We chose wheat bran because AO is commonly cultured on this medium in the food industry, and developing maize seeds, because this is a plant organ commonly infected by AFL. We predicted that we may see a different profile of gene expression on the two media, and that AFL and AO would respond differently from each other.

We were particularly interested in identifying any differences in expression of genes for secondary metabolites, as these compounds are thought to be important for niche adaptation (McDonagh et al. 2008). To ensure that we measured transcription at a time we would expect the production of secondary metabolites, we sampled the cultures when transcripts for the aflatoxin biosynthetic genes were present and aflatoxin was being produced. In the maize kernel experiments sample time was based on previous data (Georgianna et al. unpublished). To determine the appropriate time to harvest the wheat bran cultures, the medium was assayed daily from 1-3 days after inoculation by RT-PCR for expression of the aflatoxin biosynthetic gene aflK and by TLC for the accumulation of aflatoxin (Figure 1). The two day time-point was chosen for microarray analysis as this was the only time in which aflK was expressed and aflatoxin had accumulated.
We chose to examine the two sequenced strains because more genomic data are available for these strains, and these two strains have similar morphology and growth rates. The greatest observable difference between the two strains was a slightly lighter color of the conidia produced by AO. Thus, we have no reason to think that morphology or growth rate influenced the results we obtained in these comparisons. The media were chosen to provide one medium known to be conducive for growth of each species. *Aspergillus oryzae* is commonly grown on wheat bran and AFL is commonly found on developing maize kernels.

![Gene expression and aflatoxin accumulation in wheat bran koji. Gel of RT-PCR (TOP) shows gene expression for 1, 2, and 3 days in wheat bran koji for *gpdA* and the aflatoxin biosynthesis gene *aflK* in *A. flavus* NRRL3357. TLC plate (BOTTOM) shows aflatoxin accumulation after 1, 2, and 3 days in wheat bran koji for both NRRL3357 and *A. oryzae* strain RIB40.](image)

**Figure 1.** Gene expression and aflatoxin accumulation in wheat bran koji. Gel of RT-PCR (TOP) shows gene expression for 1, 2, and 3 days in wheat bran koji for *gpdA* and the aflatoxin biosynthesis gene *aflK* in *A. flavus* NRRL3357. TLC plate (BOTTOM) shows aflatoxin accumulation after 1, 2, and 3 days in wheat bran koji for both NRRL3357 and *A. oryzae* strain RIB40.

**Performance of the AFL Affymetrix GeneChip for hybridization of AO RNA samples**

The AFL Affymetrix GeneChip was used to monitor gene expression in both AFL and AO grown on maize kernels and wheat bran koji. The array contains probe sets representing a total of 13,480 *Aspergillus* genes, 13,051 of which were designed with...
respect to the AFL NRRL3357 genome and an additional 429 designed with respect to the AO RIB40 genome. We used an ANOVA to determine significant differences in gene expression between the species on both media types.

Because the majority of the probes on the GeneChip were not designed with the AO RIB40 genome in mind we were interested in whether there was any bias for calling a gene significantly different between AFL and AO. Most of the probes on the AFL Affymetrix GeneChip are perfect matches to the AFL NRRL3357 genome and while the AO RIB40 genome is very similar there are still many sequence variations between the two genomes. To assess whether this variability had any effect on determining whether a gene was significantly different we examined the probe sequence for every probe on the microarray. There are a total of eleven probes that make up the probe set that represents a particular gene. A bit score from blastn was created for every probe, where a perfect match between the typical 25 bp length probe to the AO genome would result in a score of 50. In Figure 2, the bit score sum was calculated for each probe set and plotted against the number of genes found to be significantly higher in AFL or AO for either medium. This analysis showed that genes found to be higher in AFL or AO followed a similar distribution with regard to bit score with the majority of the abundant genes being found in probe sets that were perfect or near perfect matches. This result was not entirely unexpected as the median polish step used to summarize probes as described in the materials and methods should be robust enough to avoid being significantly skewed by outliers within a probe set. Additionally, as seen in Figure 3, the distribution of raw log2-transformed data from each probe across every individual array was very similar.
This indicated that the overall hybridization between AO and the AFL Affymetrix GeneChip was not globally different than hybridization of AFL RNA.

**Differential expression of genes between AFL and AO**

Our results are focused on differences in transcription between AFL and AO. These differences will serve as the basis for results describing functional differences between these species. As described in the methods we tested for significant differences in gene expression with respect to species (AFL vs. AO) using an ANOVA model.

**Figure 2.** Gene (probe set) mismatch significance data for evaluation of probe sequence bias in *A. oryzae*. Number of genes that are expressed significantly higher in AO RIB40 or AFL NRRL3357 and those not significantly different in expression plotted against the sum of the bit score for *A. flavus* NRRL3357 affymetrix microarray probes (25bp probes, 11 probes per probe set) search in blastn against the *A. oryzae* RIB40 genome.
Figure 3. Distribution analysis of log2 transformed microarray data. Distribution of gene expression intensity across all rows of data for each microarray before normalization.

A comparison of gene expression from maize kernels infected with AFL or AO showed the expression of 1908 genes to differ significantly between the species (-log10 p-value > 2.16) (Figure 4A). Of this total, 1117 genes were more highly expressed in AO and 791 genes were more highly expressed in AFL. Interestingly the differences were even greater when the fungi were grown on wheat bran. A comparison of gene expression in AFL and AO grown on wheat bran koji revealed 3459 differentially expressed genes (-log10 p-value > 1.9) (Figure 4B). A total of 2169 genes were more highly expressed in AO compared to AFL, and 1290 were more highly expressed in AFL compared to AO.

There were 871 fungal genes differentially expressed during growth in either maize or wheat bran. A total of 398 of these genes were higher in AO on both maize and wheat bran koji; and 424 genes were higher in AFL in both maize and wheat bran koji.
There were 49 genes that contrasted in significance for one species or the other with respect to media. A total of 43 of these genes were higher in AO on maize but higher in AFL on wheat bran koji and 6 genes higher in AFL on maize but higher in AO on wheat bran koji. These differences on either substrate indicate that much of the diversity we see between AFL and AO is in the ability to use their nearly identical genomes differently. There are also clear differences between each media type indicating that we have assayed two unique substrates. We will begin to explore the implications of these differences further along in the results.

**Figure 4.** Gene expression differences between AFL and AO in maize (A) wheat bran koji (B) and differences between maize gene expression when infected with either species (C). Black dots were significant points above the p-value cut-off. For A & B the blue points were significant in both experiments in the both species, red were significant in contrasting species.

*Differential expression of maize genes in response to AFL and AO infection*

In addition to the 13,480 probes for *Aspergillus* genes contained on the AFL Affymetrix GeneChip, there are probes for 8,374 genes from *Zea mays* (maize). These probes were selected from seed cDNA libraries and thus are enriched for genes expressed...
in seeds. We were curious to know if the transcriptional response of maize seeds differs when it is inoculated with AO, considered a non-pathogen, or AFL, which is a common opportunistic pathogen. Both fungi colonized the outside of the maize kernel at the site of inoculation, but AO grew poorly if at all inside the maize seed (Figure 5). AFL on the other hand, grew extensively in the maize seed, colonizing the endosperm and embryo tissue. To our surprise, the transcriptional response of maize seeds to the two fungi was similar. Only three maize genes responded with significantly different expression to the two species interactions, and all were more highly expressed in response to AFL. The predicted proteins for these genes are: a longin/synaptobrevin domain protein (TC298591); a phospholipid-hydroperoxide glutathione peroxidase (TC280418); and a hypothetical protein (TC313500 (Figure 4C). Longin/synaptobrevin proteins are involved in membrane vesicle transport and may have a regulatory function. Phospholipid-hydroperoxide glutathione peroxidase has been suggested to be involved in defense response in plants (Agrawal et al. 2002). One other data point was very close to being significant (see point near top of cut-off line in Figure 4C). This gene encoded for a monodehydroascorbate reductase (TC298323). Monodehydroascorbate reductase is an antioxidative enzyme that may be involved in defense response (Sharma and Dubey 2007).
Figure 5. Infection of maize line B73 kernels with *A. flavus* NRRL3357 and *A. oryzae* RIB40. Cross sections of representative kernels are shown. NRRL3357 is able to more thoroughly colonize and conidi ate in the internal wound of the kernel whereas RIB40 conidiates on the surface with no conidiation in the kernel.

**Gene expression for core and polymorphic genes**

The results shown above clearly show species-specific gene expression on the two substrates examined. To further define differences between the two species we examined the expression of genes shown to potentially be either unique to the species or to show species-specific polymorphism. AO is predicted to contain 129 core genes that are unique to the species and AFL has been predicted to have 43 core genes. Additionally, there are 709 genes predicted to contain polymorphisms uniquely different between AFL and AO (Georgianna *et al.* unpublished). In an examination of the 129 core AO genes we found 38 to be more highly expressed during growth on maize and 67 to be more highly expressed in wheat bran koji cultures. Among the 43 AFL core genes we observed 14 to be more highly expressed during growth on maize and 22 to have higher expression on koji cultures. It was not possible to accurately determine the expression of one AFL core gene, possibly due to non-specific binding/background. An interesting finding from the
study is that many of the core genes from each species are expressed. It remains unclear what function these core genes have in the ecology of the fungus, as the majority of the core genes have no predicted function.

Over 75% of the 709 polymorphic genes were expressed during growth on either maize or wheat bran koji. Of these 709 polymorphic genes, 194 were differently expressed during growth on maize and 305 were differently expressed in wheat bran koji cultures. Interestingly, more polymorphic genes were expressed in AO (106) than AFL (88) during colonization of maize. Similarly, of the 305 genes differently expressed in wheat bran koji, 165 were higher in AO and 140 were higher in AFL. In humans, polymorphisms affecting transcription and mRNA processing are thought account for the majority of genetic factors in human phenotypic variability (Wang and Sadee 2006). Given the close relationship of AFL and AO it is feasible to hypothesize that a similar observation would be made.

Expression of secondary metabolism gene clusters

A total of 55 different secondary metabolism (SM) gene clusters have been predicted to encode for secondary metabolites in AFL (Khaldi et al. unpublished). We chose to examine the expression of these predicted gene clusters for the maize and wheat bran koji comparisons. The 55 SM clusters contain 483 different genes, including only 26 genes that appear to be unique to AFL. We predicted that AO would express less genes for secondary metabolism than AFL. The main reason for our prediction was that no AO strains have ever been reported to make the toxic secondary metabolite aflatoxin.
This is exactly what we observed regardless of the growth medium for the fungi.

To better visualize the differential expression of SM clusters by the two fungi the data are presented in bubble plots shown in Figure 6A for growth of AFL and AO in wheat bran or maize. Each bubble represents a SM gene cluster. On the x-axis we have plotted the median of the lsmeans difference between AFL and AO for all genes within a cluster and on the y-axis we have plotted the median of the –log10 p-value for all genes within a SM gene cluster. Each bubble is sized according to the percentage of significantly different genes contained within each respective secondary metabolite cluster. From this analysis it is very clear that AFL is a much more capable producer of secondary metabolites compared to AO, at least under the conditions examined. Note that aflatoxin (cluster 54) and cyclopiazonic acid (cluster 55) are both noticeably more highly expressed in AFL. Both of these gene clusters are thought to share similar genomic regulation (Georgianna et al. unpublished).

More predicted genes for secondary metabolism were upregulated for both fungi when grown on wheat bran than on maize kernels. Growth on wheat bran resulted in the unique expression of at least one gene in 39 separate SM clusters for a total of 142 differentially expressed genes. However, only 22 of these 142 genes were expressed higher in AO. We were surprised to find fewer genes and clusters differentially expressed during growth of the fungi on developing maize kernels than on autoclaved wheat bran as secondary metabolites are thought to be necessary for seed colonization in Aspergilli (Tsitsigiannis and Keller 2006). Growth on maize resulted in differential expression of
104 different SM genes from 29 SM clusters. Only 18 of the 104 genes differentially expressed during growth in maize were higher in AO.

In Figure 6B we have used the median expression value for each SM gene cluster in each medium for each species to describe the hierarchical relationships of gene expression in these SM cluster. This plot allows us to not only look at differential gene expression but observe the intensity of gene expression summarized for each SM cluster for each species. While there are some media differences there is a group of seven different SM clusters (54, 31, 52, 23, 11, 48, and 4) where gene expression appears high in AFL and very low in AO. Cluster 54 from this group is the aflatoxin cluster. We hypothesize that a pattern where expression is absent in AO, an important fungus for food fermentations, could indicate a strong possibility that these other secondary metabolism gene cluster products may be necessary for survival in environments not encountered by AO. Additionally, these clusters could be mycotoxins since aflatoxin is in this group and mycotoxin production would not be a desirable trait in AO. Interestingly, predicted SM cluster 23 was found to associate with the aflatoxin cluster in previous studies that included not only the data presented in this manuscript but 21 other unique experimental conditions examining expression of AFL (Georgianna et al. unpublished).

Functional differences in AFL and AO gene expression

Figure 7 shows the results of analyzing the occurrence of 15 different Biological Process Gene Ontology (GO) terms within sets of genes showing significant differences between AFL and AO on wheat bran koji or maize. Table 2 includes the actual number of induced genes representing each GO term for AFL and AO on each medium. As
already mentioned, AFL appears to be a more active producer of secondary metabolites.

From the data presented in Figure 7 AO appears to be generally much more active in translation, protein folding, and protein transport. Fungal genes for cellular metabolic processes appear to be more active in AO when grown in maize and in AFL when grown in wheat bran koji. Fungal genes for amino acid phosphorylation in AFL appear similar to AO during colonization of maize however these genes are higher for AO when grown in wheat bran koji. The enrichment of processes related to protein synthesis and transport in AO seem to fit well with the need for AO to efficiently produce enzymes that break down fermentation substrates. Interestingly, from our previous CGH polymorphism
analysis, genes for translation and protein folding appeared to be highly conserved whereas transport proteins appeared to be overrepresented in the subset of genes found polymorphic between AFL and AO strains (Georgianna et al. unpublished).

**Figure 7.** Biological process GO term enrichment. Fifteen of the most abundant biological process GO terms in *A. flavus* were chosen for analysis. Values displayed represent the percentage of genes for each select biological process that were more highly expressed in AFL or AO. For instance 87% of the significantly different secondary metabolic processes were more highly expressed in AFL than AO. AO appears to generally be much more active in translation, protein folding, and protein transport. Cellular metabolic processes appear to be more active in AO grown in corn and in AFL grown in wheat bran koji. Amino acid phosphorylation appears similar for both species when grown in maize however is higher for AO grown in wheat bran koji.

**Table 2:** Induction of genes for fifteen biological processes in AFL and AO when grown on maize or wheat bran koji

<table>
<thead>
<tr>
<th>Biological Process</th>
<th>AFL Genome</th>
<th>↑AFL Maize</th>
<th>↑AO Maize</th>
<th>↑AFL Koji</th>
<th>↑AO Koji</th>
</tr>
</thead>
<tbody>
<tr>
<td>electron transport</td>
<td>560</td>
<td>45</td>
<td>46</td>
<td>97</td>
<td>54</td>
</tr>
<tr>
<td>transport</td>
<td>486</td>
<td>23</td>
<td>52</td>
<td>66</td>
<td>64</td>
</tr>
<tr>
<td>transcription</td>
<td>246</td>
<td>0</td>
<td>19</td>
<td>0</td>
<td>61</td>
</tr>
<tr>
<td>carbohydrate metabolic process</td>
<td>227</td>
<td>24</td>
<td>8</td>
<td>45</td>
<td>20</td>
</tr>
<tr>
<td>proteolysis</td>
<td>171</td>
<td>18</td>
<td>21</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>protein amino acid phosphorylation</td>
<td>169</td>
<td>4</td>
<td>5</td>
<td>10</td>
<td>40</td>
</tr>
<tr>
<td>secondary metabolic process</td>
<td>129</td>
<td>33</td>
<td>5</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>ubiquitin-dependent protein catabolic process</td>
<td>107</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>28</td>
</tr>
<tr>
<td>protein folding</td>
<td>73</td>
<td>0</td>
<td>15</td>
<td>1</td>
<td>26</td>
</tr>
<tr>
<td>cellular metabolic process</td>
<td>68</td>
<td>4</td>
<td>7</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>signal transduction</td>
<td>68</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>11</td>
</tr>
<tr>
<td>DNA repair</td>
<td>67</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>cell wall organization and biogenesis</td>
<td>62</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>intracellular protein transport</td>
<td>62</td>
<td>0</td>
<td>14</td>
<td>0</td>
<td>12</td>
</tr>
</tbody>
</table>
Physical properties of gene expression

We were interested in discerning difference in the localization of differential gene expression. Researchers examining gene expression experiments aimed at discovering changes in *A. fumigatus* during murine infection found a significant proportion of differently regulated genes located close to the telomere (McDonagh *et al.* 2008). McDonagh *et al.* (2008) also found that many induced genes were found in contiguous clusters, many of which corresponded with secondary metabolites.

The chromosomes of AO and AFL have a high level of shared synteny. We have mapped the gene expression differences between AFL and AO on maize in wheat bran along the physical genomic coordinates for the predicted chromosomes of AFL in **Figures 8 and 9**. These images also include markings indicating the location of all of the 55 different SM gene clusters that have been predicted for AFL. An interesting observation from both the maize and wheat bran experiments is that there appear to be several large groupings of genes that are much more highly expressed in AFL compared to AO. Many of these regions also appear to coincide with the predicted SM genes clusters.

Because the differences between AFL and AO seemed to be largest for those genes that were most highly expressed in AFL we examined the distribution of gene expression differences more closely. In **Figure 10** we can see that the distribution of absolute differences between AFL and AO are significantly different (shifted towards larger values) for genes that were found significantly higher in AFL compared to AO (Wilcoxon signed rank p < 0.0001 for AFL vs AO, maize or wheat bran). Part of this
higher absolute difference for genes more highly expressed in AFL is a result of the SM genes clusters because these appear to often be very highly expressed genes. However it seems that in general when there are differences they tend to be more extreme for genes more highly expressed in AFL, it is not clear to us why this is but it may be another important difference in how AFL and AO use their genomes.

Figure 8. Physical mapping of gene expression differences between AFL and AO during maize infection.
**Figure 9.** Physical mapping of gene expression differences between AFL and AO during growth in wheat bran koji.

**Figure 10.** Distribution of absolute differences between AFL and AO for genes found significantly higher in AFL or AO.
DISCUSSION

We successfully used the *A. flavus* Affymetrix GeneChip to assay for differences in gene expression between AFL and AO. One of the most important observations was higher expression of genes from more secondary metabolite clusters in AFL. Previous research has shown differences in gene content and polymorphism between the species, each of these genomic changes could be responsible for differences in SM gene expression. Gene content does not seem likely because each species contains relatively few unique genes. We did observe a larger number of polymorphisms, some of which were in predicted genes for SM. Because SM appears so extensively different between AFL and AO we think the most likely explanation for these differences is the result of changes in regulatory genes that affect SM. LaeA, a predicted histone methyltransferase, has been shown to be a near global regulator of SM through controlling gene expression and chromosomal remodeling (Bok and Keller 2004). Epigenetic regulation though DNA methylation and histone acetylation has been shown to play important roles in silencing of SM clusters in AFL and would be an enticing possibility to explain some differences in SM (Roze *et al.* 2007; Shwab *et al.* 2007; Williams *et al.* 2008). It clearly would be interesting to know what SM pathways can be expressed in AO after epigenetic remodeling. We predict that for many of these compounds, such as aflatoxin, the expression of genes in the biosynthetic pathway may be prevented by several redundant regulatory mechanisms in AO.
Analysis of the predicted core genes between AFL and AO was particularly revealing. About half of the core genes for either species were differentially expressed suggesting that they are functional. Some of these genes exhibited differential expression on only one substrate or the other, indicating that there are regulatory mechanisms in place for expression of these genes. Unfortunately, most of these core genes are annotated as hypothetical proteins so it is difficult to understand why they might be expressed on one substrate or the other. Many of the polymorphic genes were differentially expressed but not they did appear more likely to be differently expressed than any other gene, therefore we do not think polymorphism within a gene is the only reason why many of these genes may be differentially expressed between the species. However functional changes in polymorphic gene products remain a viable hypothesis for explaining differences between AFL and AO.

In addition to the clear functional disparity in SM it appears that there are other functional differences in AFL and AO. We had initially hypothesized that AO would have a higher level of carbohydrate hydrolysis (carbohydrate metabolic process) due to its use in fermentation however our analysis shows that these genes appeared to be enriched for higher expression in AFL. We now believe that while both species are probably expressing a wide variety of hydrolytic enzymes that AO may be able to use its enzymes more efficiently and thus not need to burden itself with the expression of other unnecessary enzymes. The general higher level of protein folding, protein transport, and translation activities in AO could also indicate that AO is able to accumulate higher levels of necessary proteins without the need for a high level of gene expression. The
enrichment of protein synthesis processes in AO may have possibly resulted in a
decreased need for high levels of gene expression resulting in more efficient use of the
available transcripts. Interestingly, Kato et al. (2006) reported a negative correlation
between protein accumulation and mRNA stability. They found this correlation not only
with variation in codon sequence affecting translation but found that when chloroplasts
were treated with chloramphenicol, an inhibitor of translation, both a GUS reporter
transcript and actual chloroplast genes showed increased levels and stability of transcripts
with inhibition of translation (Kato et al. 2006).

Another interesting result was that cellular metabolic processes appeared to be
more abundantly expressed in AO on maize and in AFL on wheat bran koji. This was the
only functional difference observed between the species that appeared to show an
interaction with the environmental niche (wheat bran koji or maize). We hypothesize that
AO and AFL may experience a higher metabolic load in their unfamiliar growth
environment compared to the native species. This could be the result of less efficient
adaptations to utilize the available nutrients. The repression of cellular metabolism and
protein metabolism has been observed in experiments aimed at revealing changes in A.
fumigatus during infection (McDonagh et al. 2008). Since these processes are induced in
AO in maize kernels these observations may be consistent with AFL being the better
pathogen.

In addition to the expression differences between AFL and AO we were also able
to measure the differential response of maize to infection by both species. There was an
extreme lack of significant gene expression differences in maize kernels after infection
with AFL and AO. Our lab has evidence that maize kernels have a significant gene expression response when infected with AFL compared to a mock inoculation control (Dolezal et al. unpublished data). However, the lack of significant differences in maize gene expression in response to infection by either species suggests that the response is not specific to one species. It appears most-likely that invasion of maize by either species elicits a general response to invariant pathogen associated molecular patterns (PAMP) (Ferreira et al. 2006). Despite the lack of a gene expression difference in the maize kernel AO appeared to show less colonization of the interior of the maize kernel compared to AFL (Figure 5). We suspect that this difference might either be an inability of AO to fight past the defense mechanisms put in place by the maize kernels, an inability of AO to grow as well as AFL on the substrates present inside a maize kernel, or simply a preference for growth on nutrients accumulated on the outside of the kernel after being pierced with a pin bar. While it seems unlikely that so few differentially expressed genes would result in changes in infection it may be possible since at least two of the genes that appear different in response to AFL or AO are involved in defense response. AO is not highly prevalent in maize fields; this may be simply due to AO not being ubiquitous in the environment. However, the possible lower degree of substrate utilization in a maize field by AO may explain why infectious AO material is not as commonly found.

ACKNOWLEDGEMENTS

Thanks to Chris Smith of the NCSU Bioinformatics Research Center for his assistance in creating Figures 8 & 9.
REFERENCES


Appendix A

The effect of elevated temperature on the localization and activity of AflR

D. Ryan Georgianna and Gary A. Payne

Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695
The following studies were initiated to follow leads from earlier research showing an effect of temperature on the transcription and translation of aflatoxin biosynthetic genes (Georgianna et al. 2008; O'Brian et al. 2007). Temperatures near 37°C and above strongly inhibit aflatoxin biosynthesis in A. flavus and A. parasiticus. In contrast 37°C is the optimum temperature for the production of sterigmatocystin, an aflatoxin precursor produced by a similar pathway in A. nidulans (Feng and Leonard 1998). Temperature instability of the A. flavus AflR protein at 37°C did not seem likely as the AflR from A. flavus is able to drive expression of the sterigmatocystin cluster in A. nidulans at 37°C (Yu et al. 1996).

We hypothesized that the inhibition of aflatoxin biosynthesis in A. flavus at 37°C is due to inactivity of AflR, possibly through a mechanism not present in A. nidulans. As a transcription factor, AflR must enter the nucleus to be active and thus any effect of temperature on the movement of AflR to the nucleus could affect its activity. Previous research has shown nuclear localization to be altered when AflR is phosphorylated (Shimizu et al. 2003). To address whether or not temperature affected the localization of AflR, we created a strain of A. flavus containing an AflR-GFP fusion reporter construct. We further examined the activity of AflR at 37°C by its ability to activate transcription from aflatoxin pathway gene promoters fused to the GUS reporter. Results from these studies showed that AflR is localized to the nucleus at 28°C and 37°C. Additionally, strains with aflatoxin pathway gene promoters fused to the GUS reporter exhibited GUS
activity at both 28°C and 37°C. This results indicate that AflR is neither altered in localization nor inhibited in activity at 37°C.

**MATERIALS AND METHODS**

*GFP strains: microscopy for localization of AflR*

A plasmid containing the GPDA promoter, pNOM102 (Roberts et al. 1989), was used to clone the sGFP gene from pMTSGFP (Toews et al. 2004). Both plasmids were obtained from the Fungal Genetics Stock Center, www.fgsc.net. sGFP was amplified from pMT-sGFP with primers Forward 5’-ACGTCCATGGAATTCTATGGTGAGCAAGG-3’ and Reverse 5’-ACGTGGATCCTGGGTACCCAATTTGTACAG-3’. This product was cloned into the Nco1 and BamH1 sites of pNOM102, called pRG2 (gpdA::sGFP). AflR was amplified from *A. flavus* genomic DNA with primers Forward 5’-ACGTGGTACCACGATGGTTGACCATATCTCC-3’ and Reverse 5’-ACGTGGATCCTGGGGCTTTTCTTCATTCTCC-3’ and cloned into the Acc65I and BamH1 sites of pRG2 to make pRG2A3 (gpdA::sGFP:AflR) (**Figure 1**). pRG2 and pRG2A3 were transformed into uracil auxotroph *A. flavus* 3357-5 as part of a co-transformation with plasmid pbsk-pyr4, containing the pyr4 gene from *Neurospora crassa* in pBluescript (He et al. 2007). Transformants were identified that constitutively expressed sGFP and sGFP fused to AflR. The strain containing sGFP was named RG2 and the strain with the sGFP::AflR fusion was named RG2A3-5. GFP strains were
visualized using a FITC filter and DAPI filter. The strains were grown in PDB at 28°C or 37°C for 24 h with shaking at 200 rpm then stained with 1 μg/ml DAPI nuclear stain in 4% Formaldehyde plus 0.2% Triton X-100 solution for about 5 minutes prior to visualization under the microscope. Slides were mounted in 90% glycerol with 0.1% n-propyl gallate. The 100X objective was used with a FITC fluorescent filter for GFP images; RG2 was exposed for 3s and RG2A3-5 was exposed for 15s.

**Figure 1:** Plasmid pRG2A3

Analysis of aflatoxin by TLC

For TLC analysis 1 ml of medium was sampled after 24 h, extracted with 0.5 ml chloroform, and the extract spotted on the TLC plate. The plate was placed in a solvent
chamber for one hour. Developing solvent consisted of 80% toluene, 15% MeOH, and 5% acetic acid. After drying, plates were visualized under UV light.

**GUS strains: Aflatoxin pathway promoters fused to GUS**

The GUS strains used in these studies were constructed for earlier studies conducted in our lab and contained reporter constructs in which promoters from an aflatoxin pathway gene were fused to the GUS reporter. The fusion in GAP27 is *aflS(aflJ):GUS* (Du *et al.* 2007), GAP12-19 is *aflE(norA):GUS* (Flaherty *et al.* unpublished), and GAP26 is *aflP(omtA):GUS* (Brown *et al.* 2003).

**Qualitative GUS assay**

A qualitative GUS assay was performed as described by Chiou *et al.* (2002) using strains GAP27, GAP26, and GAP12-19. Conidia from each strain were transferred with a sterile toothpick to an autoclaved circular hybridization transfer membrane (NEN Research Products) that had been placed on the surface of potato dextrose agar. This was repeated to create two identical plates. One was placed at 28°C, a temperature conducive for aflatoxin biosynthesis, and the other plate was placed at 37°C, a non-conducive temperature for aflatoxin biosynthesis. After about 48 h the membranes with attached fungal colonies were removed from the agar plates and submerged in liquid nitrogen for 30 s, thawed, and immersed again for 30 s. After thawing, membranes were incubated in a GUS substrate solution (60 mM Na₂HPO₄, 40 mM NaH₂PO₄, 10 mM KCl, 0.27% β-mercaptoethanol, 0.04% X-Gluc [Research Products International Corp.]) for 2 h at 37°C.
The presence of blue color, indicating cleavage of the X-Gluc substrate, was considered evidence of GUS activity.

**Quantitative GUS assay**

Strains containing **aflS(aflJ):GUS**, **aflE(norA):GUS**, and **aflP(omtA):GUS** were grown in 100 ml PDB medium at either 28°C or 37°C for 48 h. Tissue was separated from the medium using a miracloth filter. Protein was isolated from frozen tissue by grinding 0.2 g tissue with a pestle in a mortar containing liquid nitrogen. and the ground material was resuspended in 500 μl GUS lysis buffer (50 mM NaH₂PO₄ [pH 7.0], 10 mM EDTA, 0.1% Triton X-100, 0.1% SDS, and 10 mM β-mercaptoethanol, 25 μg PMSF per ml. The ground tissue suspension was spun at 14,000 x g for 10 min to pellet cellular debris and protein concentration was determined using the BioRad Protein Assay reagent according to the manufacturer’s protocol using a microplate reader (BioTek). GUS activity was determined by incubating 50 μl of each protein sample with 400 μl of 2 mM 4-methylumbelliferyl-β-D-glucoronide (MUG) at 37°C for 10 min. The reaction was stopped by the addition of 400 μl of sodium bicarbonate. Fluorescence was detected using a fluorescent plate reader (BioTek) with excitation 360/40 and emission 460/40. A standard consisting of a range between 0.1 nM to 700 nM methylumbelliferone was prepared to determine GUS concentration. To measure aflatoxin in these tissue samples the medium was diluted 1:2 with chloroform.
RESULTS AND DISCUSSION

Localization of AflR

Because we know that aflR is expressed at 37°C (O'Brian et al. 2007) and it has been shown to be sufficient for expression of enzymes in the aflatoxin cluster (Smith et al. 2007), we used a GFP tagged AflR to investigate whether protein abundance and localization could be associated with lack of aflatoxin pathway transcription (Figure 2). The GFP:AflR fusion protein expressed in strain RG2A3-5 showed clear nuclear localization of AflR at both 28°C and 37°C.

Figure 2. (A) Localization of GFP in mycelium of A. flavus strain RG2A3-5 (top row) and visualization of nuclei by DAPI staining (bottom row) at 28°C (left) and 37°C (right). (B) Production of aflatoxin by the GFP:AflR fusion strain RG2A3-5 and GFP control strain RG2 as determined by TLC.
Results from the TLC analysis showed that aflatoxin B1 accumulated to a higher level in the *aflR* overexpression strain, RG2A3-5, compared to the strain expressing only sGFP, labeled RG2. This indicated that our fusion protein was functional. Additionally, no aflatoxin was observed at 37°C. This experiment showed that AflR was localized to the nucleus and yet no aflatoxin accumulated. These data strongly suggest that the state of AflR phosphorylation is not the probable mechanism of AflR inactivation since prior data on inactivation of AflR after phosphorylation by PkaA has shown phosphorylation to prevent AflR from entering the nucleus (Shimizu *et al.* 2003).

*AflR functions in *A. flavus* at 37°C*

To examine whether AflR could be active in *A. flavus* at 37°C we used several strains transformed with an ectopic copy of the GUS reporter fused to the promoters of genes in the aflatoxin cluster. Figures 3 & 4 revealed that GUS activity was seen at both 28°C and 37°C, implicating that AflR is present and functional at 37°C outside the aflatoxin cluster. Because some enzymes in the aflatoxin cluster and the regulatory genes *aflR* and *aflS* were still expressed at 37°C we did not think that lack of transcription was a result of chromosomal remodeling near the location of the aflatoxin cluster at 37°C, as might be expected when LaeA is inactivated (Bok and Keller 2004). The higher GUS activity would be more consistent with the production of sterigmatocystin in *A. nidulans*, where production is increased at higher temperatures (Feng and Leonard 1998).
CONCLUSIONS

Overexpression of GFP:aflR did not affect aflatoxin production at 37°C even though AflR was clearly localized to the nucleus at both temperatures examined. It does not appear likely that AflR is inactivated at 37°C but rather there is a more global factor
causing regulation of the aflatoxin cluster at 37°C. This regulation is not present in the sterigmatocystin cluster of \textit{A. nidulans}, even when AflR from \textit{A. flavus} has replaced the native AflR (Yu \textit{et al.} 1996). While LaeA seems like an ideal candidate since it is involved in chromosomal remodeling it is not likely the main player in inhibition of aflatoxin at 37°C because LaeA inhibition results in loss of \textit{aflR} expression and our previous microarray data showed no difference in expression of \textit{aflR} at 28°C and 37°C (O'Brian \textit{et al.} 2007). The CPA cluster, located adjacent to the aflatoxin cluster is also regulated by LaeA and CPA is made at 37°C in \textit{A. flavus} (Georgianna et al unpublished). These results suggest that there is another control specifically inhibiting expression of the aflatoxin cluster at 37°C. Interestingly, evidence in the literature suggests that there is a currently unknown repressor of AflR (Chang \textit{et al.} 1999; O'Brian \textit{et al.} 2007). Since our data indicate that AflR is still active at 37°C we hypothesize that the temperature dependent regulation of aflatoxin may be functioning through novel epigenetic or silencing mechanisms that are inactive in \textit{A. nidulans} at 37°C.

REFERENCES


