ABSTRACT

HOLMES, ROBERT. Characterization of an Aflatoxin Biosynthetic Gene and Resistance in Maize Seeds to *Aspergillus flavus*. (Under the direction of Rebecca S. Boston and Gary A. Payne.)

Crop contamination with aflatoxins produced by the fungi *Aspergillus flavus* and *Aspergillus parasiticus* is a persistent problem. I addressed aflatoxin contamination from the perspectives of resistance and biosynthesis. A review of the literature on compounds that inhibit aflatoxin biosynthesis showed that many inhibitors are plant-derived and some may be amenable to pathway engineering for defense against aflatoxin contamination. Other compounds show promise as storage protectants. Inhibitors with different modes of action could be used in transcriptional and metabolomic profiling experiments to identify regulatory networks controlling aflatoxin biosynthesis.

To address host resistance in maize, I used liquid chromatography to characterize a protein fraction from kernels of the resistant maize line Tex6 that inhibits growth and aflatoxin production by *A. flavus in vitro*. Two proteins were associated with the inhibitory activity. Peptide sequencing identified them as chitinase A (ChitA) and zeamatin, members of the glycoside hydrolase 19 (GH19) and thaumatin-like protein (TLP) families, respectively. Removal of chitin-binding proteins from the fraction dramatically reduced its inhibitory effect. Adding the chitin-binding fraction back to the zeamatin-enriched fraction restored activity. We used bioinformatic, phylogenetic and gene expression analyses to investigate the GH19 and TLP gene families in maize. Phylogenetic analyses placed the maize GH19 genes into four major phylogenetic groups. The TLP gene family was larger and

was similar to the rice TLP family. Transcripts of members of each gene family were induced during *A. flavus* infection of kernels.

To address aflatoxin biosynthesis, I investigated an undescribed aflatoxin biosynthetic cluster gene, *hypE*. The predicted HypE protein contains a putative EthD domain, a domain described to date only in bacteria. A *hypE* mutant of *A. flavus* was created by gene deletion via homologous recombination with a selectable marker. The *hypE* mutant produced less aflatoxin B1 and B2 than control strains and accumulated a metabolite which we assigned as a tentative HypE substrate (HESUB). Aflatoxin biosynthesis could be restored in this *hypE* mutant by a *hypE* overexpression construct and this restoration was associated with loss of HESUB accumulation. Other Aspergilli and *Podospora anserina* have potential *hypE* orthologs. We identified additional *hypE*-like genes which were also restricted to the Pezizomycotina.

Characterization of an Aflatoxin Biosynthetic Gene and Resistance in Maize Seeds to *Aspergillus flavus*

by
Robert A. Holmes

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

Plant Biology

Raleigh, North Carolina
2008

APPROVED BY:

Co-chair of Advisory Committee Rebecca Boston	Co-chair of Advisory Committee Gary Payne
Margaret Daub	Steven Spiker

BIOGRAPHY

Robert Alexander Holmes was born February 28, 1979 in Provo, Utah to Blair Ralph Holmes and Margie Green Holmes. He attended elementary, middle and high school in Provo, graduating from Provo High School in 1997. Two excellent high school biology teachers, Jan Cooper and Merrill Webb, furthered his natural interest in biology. Botanical field work in southern Utah under the direction of Ms. Cooper led to his decision to study botany. Rob attended Brigham Young University where he completed his B.S. in Botany in 2002. From June 1998 to June 2000, Rob served in the Seoul West Mission in South Korea as a volunteer missionary for the Church of Jesus Christ of Latter-day Saints. There he had the opportunity to learn the Korean language and culture while sharing his religious beliefs.

Upon returning from South Korea, Rob married Taliatha Aldous Palmer in December 2000. They had met their freshmen year at BYU and kept in close contact while Rob was in South Korea through weekly letter writing. At BYU Rob performed research in the labs of Craig E. Coleman and Bruce N. Smith in the Department of Botany and Range Science. Prior to beginning doctoral work at North Carolina State University they were blessed with their first daughter, Jane, who accompanied them on the journey from Utah to Raleigh, North Carolina. In Raleigh they were pleased to have two more children, Paul and Mallory.

ACKNOWLEDGEMENTS

I express my gratitude to my two major professors for advising my dissertation research. They have been dedicated and accessible mentors who have helped me in my scientific and professional development. Advising the research of a graduate student who is bouncing back and forth between two disciplines required extra coordination and patience as the project took on a life of its own. I also thank Drs. Boston and Payne for pulling together the funding necessary to keep me afloat financially for the last six years.

I am particularly indebted to Dr. Jeff Gillikin and Greg OBrian for help in the lab throughout the project and critical assistance in completing the final experiments. Other members of the Boston lab (Norma Houston, Mariana Kirst, and Jian Wu) and Payne lab (Mike Price, Kim Schwartzburg, Carrie Smith, Ryan Georgianna, and Andrea Dolezal) have been fantastic friends and colleagues. My other committee members, Drs. Margaret Daub and Steven Spiker, provided excellent advice and constructive criticism.

I also acknowledge the support I received from 2003-2004 by an NIH/NCSU Biotechnology Traineeship and for most of 2006 by an IFAFS Fellowship. I was also supported by grants from the Multi-crop Aflatoxin Working Group for which I am grateful.

Finally, I thank my dear wife and best friend, Taliatha, for supporting me through the ups and downs of graduate school and making our family life one of joy and refuge.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AF	aflatoxin
AFB1	aflatoxin B1
AFB2	aflatoxin B2
AFG1	aflatoxin G1
AFG2	aflatoxin G2
BLAST	basic local alignment search tool
BLASTP	protein basic local alignment search tool
Chit	chitinase
CTL	chitinase-like
DHOMST	dihydro-*O*-methylsterigmatocystin
DMST	demethylsterigmatocystin
GH19	glycoside hydrolase 19
gpdA	glyceraldehyde phosphate dehydrogenase
HAP	hydroxyapatite
hel	*hypE*-like
HESUB	HypE substrate, putative
LC-MS/MS	liquid chromatography-mass spectrometry/mass spectrometry
MLS	minimal media low salts
OMST	*O*-methylsterigmatoycstin
PDA	potato dextrose agar
PKS	polyketide synthase
PSI-BLAST	position specific iterative basic local alignment search tool
ST	sterigmatocystin
TBLASTN	translated nucleotide basic local alignment search tool
TLC	thin layer chromatography
TLP	thaumatin-like protein
VERA	versicolorin A

PREFACE

This dissertation is an investigation of aflatoxin biosynthesis in the filamentous fungus *Aspergillus flavus* and resistance factors from maize and other sources that may be useful in controlling contamination of crops with aflatoxin. Chapter 1 is a literature review published in Applied Microbiology and Biotechnology that synthesizes a large literature on inhibitors of aflatoxin biosynthesis. In it I attempt to bring together bioactivity data for a chemically diverse group of inhibitors so that they can be compared. I also address current ideas on how aflatoxin inhibitors function and investigate promising applications of inhibitors for control of aflatoxin contamination of crops and dissection of the regulatory controls of aflatoxin.

Chapter 2 is a genetic, bioinformatic and biochemical analysis of the previously undescribed aflatoxin biosynthetic cluster gene *hypE,* which represents a previously undescribed fungal gene family. I am responsible for all data except for the LC-MS/MS analyses which were performed at the North Carolina State University Genomic Sciences Laboratory. I was assisted in designing and running the phylogenetic analyses by Dr. Ignazio Carbone (Plant Pathology, NCSU).

Chapter 3 describes a protein fraction from seeds of the resistant maize line Tex6 that is inhibitory to growth and aflatoxin production of *A. flavus*. I identified two major components of the inhibitory fraction as chitinase A and zeamatin. I then identified the respective gene families of these two proteins and explored the phylogenetic relationships of the gene families in plants. I also investigated transcription of these gene families in response to *A. flavus* infection in the field. I performed the protein purifications and bioassays in this chapter with assistance from Dr. Jeff Gillikin (Plant Biology, NCSU) in the final hydroxyapatite purification step for zeamatin and in purification of chitinase A for enzyme assays. Dr. Gillikin also performed the in-gel chitinase assay under my direction. The field studies on maize infection by *A. flavus* were designed and carried out by Andrea Dolezal (Plant Pathology). She extracted RNA for microarrays and performed the global analysis of the data. I identified the glycoside hydrolase 19 (chitinase) and thaumatin-like protein family

probes and interpreted the expression data. I also performed the RNA extractions and RT-PCR analyses of gene expression. I performed the trypsin digestion in collaboration with the Genomic Sciences Laboratory and they performed the LC-MS/MS identification of chitinase A and zeamatin peptides. I was assisted in performing phylogenetic analyses by Dr. Carbone and Dr. Isaac Winkler (Entomology, NCSU).

CHAPTER 1

Diverse inhibitors of aflatoxin biosynthesis

R. A. Holmes, R. S. Boston and G. A. Payne

This work was published in Applied Microbiology and Biotechnology **78**:559-572.

CHAPTER 2

Characterization of maize seed chitinases and thaumatin-like proteins in resistance to *Aspergillus flavus*

ABSTRACT

Contamination of maize, peanuts, and tree nuts with carcinogenic aflatoxins produced by the fungal plant pathogens *Aspergillus flavus* and *Aspergillus parasiticus* is a recurrent agronomic problem for which effective pre-harvest resistance remains elusive. Using liquid chromatography, we isolated and characterized a protein fraction from kernels of the resistant maize line Tex6 that inhibits growth and aflatoxin production by *A. flavus in vitro*. Two prominent protein bands were consistently associated with the inhibitory activity, and LC-MS analysis following trypsin digestion identified them as chitinase A (ChitA) and zeamatin, members of the glycoside hydrolase 19 (GH19) and thaumatin-like protein (TLP) plant gene families. We found that removal of chitin-binding proteins from the active fraction dramatically reduced its inhibitory effect, and adding the chitin-binding fraction back to the zeamatin-enriched fraction restored the activity. Because both of our major candidate proteins are encoded by genes belonging to gene families that have not been described in maize we used bioinformatics and analysis of gene expression to investigate other maize GH19 and TLP genes that may be involved in seed defense responses. We analyzed the phylogenetic relationship of both gene families. Maize GH19 genes grouped into four major phylogenetic groups similar to those described in rice (37). TLP genes from maize generally followed their previously described phylogeny in rice (46). Transcripts of several members of each gene family were induced during *A. flavus* infection of maize kernels.

INTRODUCTION

Maize seeds synthesize and store defense proteins during maturation (5, 15). While maize seeds are susceptible to few diseases, fungal ear rot pathogens such as *Aspergillus flavus* and *Fusarium verticillioides* that can successfully overcome seed resistance are particularly problematic because they contaminate grain with aflatoxins and fumonisins, respectively (7, 41). There have been extensive efforts to identify kernel resistance factors that are associated with resistance to *A. flavus* infection and aflatoxin production and could be introduced into maize cultivars through breeding and transgenic approaches (6, 9, 10, 16, 20, 25, 34). These efforts have resulted in the identification of chitinases, thaumatin-like

proteins (TLPs), a pathogenesis-related 10 protein, and a 14 kDa trypsin inhibitor in maize seeds that are either induced in response to infection or possess antifungal activity (11, 12, 17, 25).

Studies of antifungal proteins performed *in vitro* and in transgenic plants indicate that a combination of diverse classes of defense proteins is more effective in achieving resistance than over-expression of a single defense protein (18, 27, 32). In addition to the identification of the Tex6 chitinase (35) other work has shown that Tex6 maize seeds contain additional activities inhibitory to *A. flavus* growth and aflatoxin production. Huang et al. (20) identified an inhibitory fraction that affected primarily aflatoxin production and a second fraction that affected primarily growth. Subsequently Moore et al. (35) identified a chitinase in kernels of the resistant maize inbred Tex6 that had strong growth inhibitory activity against *A. flavus*, and Ji et al. (25) showed induction of chitinase isoforms in response to *A. flavus* infection. ChitA and ChitB belong to the glycoside hydrolase 19 (GH19) family of chitinases, and the Tex6 chitinase is closely related to these two proteins. Beyond these chitinases the role of other GH19 family chitinases in seed resistance is unclear. Zeamatin, an abundant TLP in kernels, has also been implicated in resistance in seeds and is associated with maize genotypes with ear rot resistance (8, 49). Zeamatin is thought to function as a membrane permeabilizer but its exact mode of action remains unresolved (44, 48). Another TLP, osmotin, also has membrane permeabilizing activity (1). Fungal cell wall targets have been implicated in the activity of osmotin against yeast, *Aspergillus nidulans* and *F. oxysporum* (13, 23, 24, 30, 38, 45). Like the GH19 family, the TLP family has not been characterized in maize.

We report here the characterization of an inhibitory fraction from Tex6 kernels that is enriched in ChitA and zeamatin and was inhibitory to *A. flavus* growth and aflatoxin production. Removal of chitin-binding proteins from this fraction greatly decreased its inhibitory activity. We also show that ChitA and zeamatin are members of larger GH19 and TLP gene families in maize and members of these gene families are induced in response to *A. flavus* infection.

MATERIALS AND METHODS

Maize kernels and *A. flavus* strain. All kernel proteins were extracted from the maize inbred line Tex6 grown at Central Crops Research Station in Clayton, NC. The *A. flavus* strain GAP26 (tan, *pyrG, leu-, omtA::uidA*) was used in bioassays. For field inoculation studies the maize inbred line B73 was grown at Central Crops Research Station in Clayton, NC and inoculated with *A. flavus* strain NRRL 3357.

Protein purification and SDS-PAGE. For protein extraction mature kernels were ground to a fine meal in a coffee grinder, added 1:5 (w/v) to cold 10 mM sodium phosphate, pH 6.8, and extracted by magnetic stirring at 4ºC for 4 hours. The resulting extract was filtered through cheesecloth and subjected to centrifugation at 4ºC for 20 minutes at 14,500 x g. The resulting supernatant was filtered through two layers of Miracloth (Calbiochem, La Jolla, CA). Extracts were stored at -20º C until use. Prior to chromatography extracts were again subjected to centrifugation (14,500 x g, 20', 4º C) and the supernatant was vacuum-filtered through 0.45 μm filters (Pall, Ann Arbor, MI) to yield the crude extract fraction.

For the initial identification of the activity inhibitory to aflatoxin production 10 mL Tex6 crude extract was applied to a 5 mL High S Econo-Pac strong cation exchange column (Bio-Rad, Hercules, CA) equilibrated in 10 mM sodium phosphate, pH 6.8. Bound material was eluted with a 1 M NaCl step gradient in 10 mM sodium phosphate, pH 6.8 with a flow rate of 2 mL/min.

Purification of ChitA and zeamatin for LC-MS analysis was initiated by application of 500 mL Tex6 crude extract to tandem High S Econo-Pac columns equilibrated in 10 mM sodium phosphate, pH 6.8. Bound material was eluted with a 60 mL linear gradient (0-0.2 M NaCl in 10 mM sodium phosphate, pH 6.8) with a flow rate of 2 mL/min. The first major peak was pooled (14 mL) and adjusted to 1.6 M $(NH_4)_2SO_4$ in 10 mM sodium phosphate, pH 6.8 by addition of 56 mL 2 M $(NH_4)_2SO_4$ in 10 mM sodium phosphate, pH 6.8 to a final volume of 70 mL. This pool was subjected to centrifugation (14,500 x g, 20', 4ºC) following which the supernatant was applied to a t-Butyl Econo-Pac hydrophobic interaction column (Bio-Rad, Hercules, CA) and eluted with a 30 mL linear gradient (1.6-0 M $(NH_4)_2SO_4$ in 10 mM sodium phosphate, pH 6.8) with a flow rate of 1.3 mL/min. The bound material eluted as

a single peak at ~1.2 M $(NH_4)_2SO_4$ that was collected (10 mL) and exchanged into an equal volume of 10 mM sodium phosphate, pH 6.8 using a Centricon-10 10,000 MWCO centrifugal filtration device (Millipore, Bedford, MA). One mL of the Centricon retentate was applied to an UnoS-1 cation exchange column (Bio-Rad, Hercules, CA) equilibrated in 10 mM sodium phosphate, pH 6.8 and eluted with a 39 mL linear gradient (0-0.25 M NaCl in 10 mM sodium phosphate, pH 6.8) with a flow rate of 3 mL/min.

Chitin affinity chromatography was carried out with a chitin column prepared from chitin (Sigma, St. Louis, MO) that had been ground to a fine powder in a mortar with a pestle. 0.4 g ground chitin was placed in a 2 mL polyprep chromatography column (Bio-Rad, Hercules, CA) and equilibrated with 10 mM sodium phosphate, pH 6.8. Two hundred mL Tex6 crude extract was applied to tandem High S columns and eluted with a 60 mL linear gradient (0-0.2 M NaCl in 10 mM sodium phosphate, pH 6.8) with a flow-rate of 2 mL/min. Five mL of the High S peak was applied to the chitin column by gravity flow and the flow-through was collected. The column was rinsed with 20 mL 10 mM sodium phosphate, pH 6.8 and chitin-binding proteins were eluted with 6 mL 0.5% (v/v) acetic acid. The column was washed with 20 ml 10 mM sodium phosphate, pH 6.8 and 1 M NaCl and equilibrated in 10 mM sodium phosphate, pH 6.8. The initial column flow-through was reapplied to the column to remove residual chitin-binding proteins. Chitin-binding and unbound fractions were exchanged into 10 mM sodium phosphate, pH 6.8 using a Centricon-10 device and adjusted to equal final volumes (2 mL). The unbound fraction was applied to a UnoS-1 column and eluted with a 50 mL linear gradient (0-0.25 M NaCl in 10 mM sodium phosphate, pH 6.8) with a flow rate of 3 mL/min. The major peak which contained the zeamatin fraction was exchanged into 10 mM sodium phosphate, pH 6.8 using a Centricon-10 device and the resulting fraction was used in the bioassay.

Additional hydroxyapatite purification of the zeamatin fraction following UnoS-1 separation was performed on a 1 mL CHT-II Econo-Pac column (Bio-Rad, Hercules, CA). The zeamatin-fraction was eluted with a 26 mL linear gradient (10-400 mM sodium phosphate, pH 6.8 gradient) at a flow rate of 0.6 mL/min. Fractions corresponding to the

single zeamatin peak were pooled and concentrated in 10 mM sodium phosphate, pH 6.8 using a Centricon-10 device.

SDS-PAGE analysis of proteins was performed as described by Laemmli (28). Samples were diluted 3:1 in 4X-SDS loading buffer and boiled for 5 min prior to electrophoresis for 1 hr at 150 V. Gels were stained with Coomassie Brilliant Blue for 1 hr at room temperature and de-stained in 10% (v/v) acetic acid.

Protein quantification. Protein concentrations were determined using the Coomassie Plus protein assay kit with BSA as a standard (Pierce, Rockford, IL). Twenty-five µL of a protein fraction was mixed with 750 µL assay reagent and incubated at room temperature for 5 minutes before measurement of the absorbance at 595 nm.

Bioassays for aflatoxin production and fungal growth. For initial identification of the cation exchange binding activity (Fig. 1) bioassays were performed in 96-well microtiter plates by combining 150 µL of a protein fraction and 50 µL of sucrose low salts (SLS) medium supplemented with 10 µg/mL each of uracil (U) and leucine (L) as previously described (20). In subsequent experiments, 100 µL of a protein fraction was combined with 50 µL 3X SLSUL, inoculated with $3x10^6$ spores/mL of *A. flavus*, and incubated at 28º C for 4 days. Aflatoxin was extracted by vortexing the culture medium from each well with 50 µL chloroform. Following centrifugation for 5 minutes at 10,000xg the chloroform layer was removed and evaporated prior to resuspending the residue in 8 µL chloroform for spotting onto Whatman Partisil K6 Silica Gel 60 Å TLC plates (Fisher, Pittsburgh, PA). TLC plates were developed in toluene-methanol-acetic acid (80:15:5 [vol/vol/vol]) for one hour in a saturated environment and imaged over a UV transilluminator. Dry weight determinations were made after drying mycelial mats overnight in an incubator oven at 60ºC in a pre-weighed Eppendorf tube. The chitin-binding fraction and the zeamatin fraction were assayed at approximately equimolar concentrations in bioassays (5 µM in Fig. 3 and 3 µM in Fig. 4).

Chitinase enzyme assays. Glycol chitin was prepared by acetylation of glycol chitosan (Sigma, St. Louis, MO) with acetic anhydride as previously described (47). Gels for chitinase assays were prepared as follows. A 0.01% glycol chitin, 15% SDS-polyacrylamide gel was loaded with equimolar amounts of protein from the chitin-binding fraction (900 ng)

and HAP-purified zeamatin fraction (750 ng) without a reducing agent and subjected to electrophoresis for 1 hour at 150 V. The gel was incubated at 28º C for 4 hours in 10 mM sodium acetate, pH 5.0 containing 1% (v/v) Triton X-100 and stained with 0.01% Calcofluor White (Polysciences, Warring, PA), 500 mM Tris-HCl, pH 8.9 for 5 minutes at room temperature. The gel was then incubated for 1 hour at room temperature in dH_2O and imaged using a UV transilluminator. For spot chitinase assays 0.04% glycol chitin, 7.5% polyacrylamide gels buffered with 10 mM sodium acetate, pH 5.0 were polymerized by adding 113 µL of 10% (w/v) ammonium persfulate and 30 µL TEMED to 9.1 mL of the gel solution. Gels were spotted with 4 µL of a given protein fraction dispensed directly onto the surface of the gel and then incubated at 28ºC for 1 hour. Spot assay gels were stained and imaged as described above.

Microarray analysis. During the 2006 and 2007 growing seasons maize ears at the specified stage of development were either mock-inoculated or inoculated with a pinbar contaminated with a spore suspension of *A. flavus*. Each kernel was pierced once with a pin delivering approximately 12 conidia per kernel. In 2006 a stage-specific induction study was conducted in which mock-inoculated and inoculated kernels were harvested four days post-inoculation and quick frozen in liquid nitrogen. In 2007 a time-course study was conducted in which ears were inoculated as above and harvested at 12, 24, 36, 48, 60, 72 and 96 hours after inoculation.

RNA was extracted from either 4 mock-inoculated or 4 infected kernels after grinding the kernels with a pestle in a mortar containing liquid nitrogen. Ground tissue was added to 5 mL water saturated phenol, pH 6.6 (Ambion, Austin, TX) in an Oak Ridge tube and vortexed thoroughly. Five mL Tris-EDTA, pH 8.0 buffer (Acros Organics, Geel, Belgium) was then added and the solution was vortexed thoroughly and subjected to centrifugation (14,500xg, 20 min, 4ºC). The aqueous layer was transferred to another Oak Ridge tube containing 5 mL 125:24:1 acid-phenol:chloroform with isoamyl alcohol (pH 4.5; Ambion, Austin, TX). This solution was vortexed thoroughly and subjected to centrifugation (14,500xg, 20 min, 4ºC). The last step was repeated once more and the resulting aqueous phase was transferred to a new Oak Ridge tube. Two volumes ice cold 95% (v/v) ethanol were added and the sample

was incubated overnight at -20ºC. The sample was subjected to centrifugation (14,500xg, 30 min, 4ºC) and the precipitate was dried for 5 min before resuspension in 450 µL buffer RLT from a QIAGEN Plant RNeasy Kit (QIAGEN, Valencia, CA). The RNA in buffer RLT was then purified further according to the manufacturer's protocol. RNA was treated with DNase (Promega, Madison, WI) and then sent to the Purdue Core Genomics Facility (http://www.genomics.purdue.edu/~core/) where cRNA synthesis and microarray hybridization were performed. The cRNA was hybridized to a custom Affymetrix GeneChip DNA microarray (42) containing the genome of *A. flavus* and approximately 8,000 genes known to be expressed in inbred line B73 seeds. The gene expression data were analyzed and expression values were normalized using Loess normalization in JMP Genomics (SAS, Cary, NC). For measurement of inductions in the stage-specific infection a p-value<0.05 cutoff was used. For the time course an F-test with a 0.05 cutoff was used to analyze if differences in least squared means expression values between mock-inoculated and infected kernels showed a positive or negative correlation with time.

RT-PCR. RNA extraction was the same as for microarrays. One µg of RNA was used for cDNA synthesis using Stratascript reverse transcriptase and random hexamers (Stratagene, La Jolla, CA). Real-time PCR for GH19 and TLP genes was carried out in 25 µl reactions with 12.5 µl of SYBR Green Mastermix (Applied Biosystems, Foster City, CA), 3 µl of cDNA (167 ng of starting RNA), and 0.5 µM primers. Control reactions for 18S ribosomal RNA were carried out with 1 µL cDNA (56 ng of starting RNA) and 0.17 µM primers. All reactions were done in triplicate. Real-time PCR was performed in a DNA Engine Opticon2 (MJ Research, Reno, NV) with cycle parameters as follows: 95 °C for 10 min followed by 40 cycles of 95°C, 15 s and 60°C, 1 min. Normalization against 18S ribosomal RNA controls and calculation of fold-changes of transcripts between infected and mock-inoculated control followed the $2^{-\Delta\Delta CT}$ method (31). Primers for chitinase A and TLP genes were designed using gene sequences from nucleotide accessions listed in Tables 1 and 2. The primer pairs were as follows: chitA forward, 5'-GAGATCGCCCGCCTTCTTC, chitA reverse, 5'-AGTAGGCGTTGCTCTTGTTGA; zeamatin forward, 5'-CAGGTACTTCAAGGGGCAGT; zeamatin reverse, 5'-GGGCAGAAGACGACCTTGTA;

TLP2 forward, 5'-CGTGTTCAAGACGGACCACT; TLP2 reverse, 5'-GCAGAGCCCCTTGAAGAACT; TLP4 forward, 5'-CTCCTCGTCCTCGGTCCT; TLP4 reverse, 5'-TGTTGGTGATGGTGAAGGTG; TLP10 forward, 5'-GCCCCAACTCGTACAGCTAC; TLP10 reverse, 5'-GGGCAGAAGGTGATGGTGTA; 18S forward, 5'-TGGTGACGGGTGACGGAGAA; 18S reverse, 5'-GCCCTCCAATGGATCCTCGT.

Bioinformatics and phylogenetics. Using the protein sequences of ChitA (AAA33444) and zeamatin (P33679) we performed TBLASTN searches against Genbank, maize GSS and BACs sequences (http://www.plantgdb.org), and the maize EST database hosted at Iowa State (http://magi.plantgenomics.iastate.edu). Rice and Arabidopsis GH19 and TLP sequences were obtained by searches for genes annotated with the word "chitinase" or "thaumatin" and TBLASTN searches (http://rice.plantbiology.msu.edu and http://www.arabidopsis.org) using ChitA and zeamatin as queries. Gene structures were predicted using FgenesH (http://www.softberry.com), Eukaryotic Genemark (http://exon.gatech.edu/GeneMark), and GENSCAN (http://genes.mit.edu/GENSCAN.html) and compared to EST sequences by alignment with ClustalW2 (http://www.ebi.ac.uk/clustalw/index.html). Maize GH19 and TLP family members that were not in Genbank were named based on their similarity to ChitA and zeamatin, respectively.as determined by TBLASTN score. Maize, rice and Arabidopsis GH19 and TLP sequences were aligned using ClustalW. Alignments were trimmed and gaps were closed and manually adjusted using Sequencher (http://www.genecodes.com/). For GH19 sequences only the catalytic domain was included in the matrix. For TLP sequences the signal peptide and C-terminal extensions longer than zeamatin were excluded from the matrix. Phylogenetic inferences were performed using MrBayes v3.0B4 (21).

RESULTS

Identification of inhibitory activities in Tex6 kernels. We fractionated crude extracts of mature kernels of the resistant maize line Tex6 to identify protein fractions with inhibitory activity against AF production by *A. flavus*. We applied 10 mL of Tex6 crude

extract to a low pressure strong cation exchange column and assayed material eluted with a 1 M NaCl step gradient for inhibition of AF production in a microtiter plate bioassay (Fig. 1A). We found AF inhibitory activity in both the bound and unbound fractions (Fig. 1B). This is consistent with previous findings that Tex6 kernels contain multiple inhibitory activities (20, 35). Additional characterization of the activity in the unbound fraction showed that it likely was caused by a non-proteaceous compound(s) (Appendix B).

Because we were interested in identifying inhibitory proteins we pursued purification of the cation exchange binding activity. We used three liquid chromatography steps (High S cation exchange, t-butyl hydrophobic interaction and UnoS-1 cation exchange; supplemental data in Appendix A). Bioassays of fractions from the final purification step identified inhibitory activity that correlated with the presence of two candidate protein bands of ~27 kDa and ~22 kDa (Fig. 2A, B). However, the elution profile of neither major candidate band correlated directly with AF inhibition. Trypsin digestion and LC-MS analysis of the two bands indicated that they corresponded to chitinase A (ChitA, AAA33444) and zeamatin precursor (P33679). The peptide sequences for ChitA indicated that this chitinase from Tex6 corresponds to the major seed chitinase A purified by Huynh et al. (22). This is likely not the Tex6 chitinase purified previously as the peptide sequences are different (35). Both ChitA and zeamatin are abundant seed defense proteins (22, 44) and the observation of AF inhibitory activity in fractions where both proteins were present suggested that the AF inhibitory activity may result from a combination of activities rather than a protein corresponding to a single band.

Chitin affinity separation of chitin-binding and zeamatin fractions. We reasoned that if chitinase activity was important for the inhibitory activity its removal would reduce the activity of the fraction. To test this hypothesis we applied an inhibitory fraction containing ChitA and zeamatin to a chitin affinity column and separated the material into unbound and chitin-binding fractions (Fig. 3A, lanes 2 and 3). We subjected the unbound fraction to an additional cation exchange chromatography step to give the zeamatin fraction (Fig. 3A, lane 4). We then assayed equimolar concentrations of the chitin-binding fraction and zeamatin fraction alone and in combination for inhibition of aflatoxin production (Fig.

3B). Little if any effect on aflatoxin production was seen with either fraction alone (compare panels 3 and 4 to panel C). However, together they were able to reconstitute the initial activity (compare panels T6 and 1 to panel 3+4). We also found that inhibition of aflatoxin production correlated positively with reductions in fungal growth, as measured by dry weight of the fungal cultures, and in mycelial mat formation, with the exception of T6 (Fig. 3C, D). The fact that the crude extract was able to inhibit production of aflatoxin without influencing growth was consistent with our earlier observation that the Tex6 crude extract contains multiple inhibitory activities including activities that do not interact with the cation exchange column (Fig. 1, Appendix B). We used a hydroxyapatite column (HAP) to gain additional purification of the zeamatin fraction (Fig. 3E, compare to Fig. 3A lane 4). This additional purification step removed all but one contaminating band which did not co-migrate with the ChitA band. This fraction was then assayed alone and in combination with the same chitin-binding fraction from the previous experiment (Fig. 3F). As observed before, the chitin-binding fraction and the highly enriched zeamatin fraction inhibited aflatoxin production more efficiently in combination than alone.

Verification of chitinase enzyme activity. We performed in-gel chitinase assays to verify that the chitin-binding fraction contained chitinase enzyme activity (Fig. 4A). The chitin-binding fraction showed activity as four bands. We observed three bands in the zeamatin fraction at positions similar to the top three bands, although at much lower intensity. The detection of chitinase activity bands in the zeamatin fraction was unexpected since a Coomasse gel of the zeamatin fraction showed no staining at the position of the ChitA band (Fig. 3E). Because the bands appeared to match the chitinase activity bands observed in the chitin-binding fraction we suspected the activity was residual chitinase activity corresponding to the major activity separated out by chitin affinity purification. We performed a serial dilution spot test of the chitin-binding fraction and the zeamatin fraction to assess their relative chitinase activities (Fig. 4B). Chitinase activity in the chitin-binding fraction at 4 ng was approximately equivalent to activity in the zeamatin fraction from 50-150 ng. Chitinase activity was detected with as little as 160 pg of protein in the chitin-binding fraction (Fig. 4C). These results indicated that the chitin-binding fraction was at least

12-fold more active than the zeamatin fraction and that low levels of chitinase activity could be detected in protein fractions at concentrations below the detection limit of Coomassie staining. These data suggested that the ChitA-enriched fraction may contain additional isoforms of chitinase and that the inhibitory activity may result from interaction between multiple chitinases and TLPs such as zeamatin.

Maize glycoside hydrolase 19 and thaumatin-like protein gene families. Because ChitA and zeamatin both belong to large plant gene families we were interested to see if other maize chitinases might also be involved in resistance to *A. flavus* in seeds. The GH19 (ChitA) and TLP (zeamatin) gene families had not been described in maize so we searched GSS, BAC and EST sequence databases for sequences related to ChitA and zeamatin and identified 14 and 33 related sequences, respectively, that appeared to encode functional proteins with significant similarity (e-values < e-9 based on BLASTP alignment of predicted proteins; Tables 1 and 2). We aligned predicted protein sequences to form a matrix and excluded apparent pseudo-genes with retro-transposon insertions, a TLP sequence with a premature stop codon and two members of the TLP family with kinase domains and predicted coding regions that were much larger than all other members of the family. The final matrix generated for phylogenetic analysis of each family contained 15 and 36 members of the GH19 and TLP families, respectively.

Phylogenetic analysis of maize, rice and Arabidopsis GH19 families. Bayesian analysis of 16 maize, 16 rice and 14 Arabidopsis GH19 protein sequences yielded a 50% strict consensus tree with four major phylogenetic groups (Fig. 5). The GH19 family contains seven members that lack chitin-binding domains. With the exception of *Chit21*, which is homologous to GH19 members involved in cell wall formation in Arabidopsis and cotton (50, 51), all members fit into structural class *Chia1*, *Chia2* or *Chia4* based on the presence or absence of conserved gaps and 50 percent similarity of the catalytic domain to tobacco *Chia1* chitinase (29). While members of structural classes *Chia2* and *Chia4* clustered together, members of structural class *Chia1* fell into two distinct clades (B and C-1). These results indicate that the current grouping of GH19 chitinases by structural class does not reflect correctly the evolutionary history of the gene family. *Chit21* grouped with members of the

chitinase-like (*CTL*) group. We chose to root our tree with the *CTL* sequences based on their distinction from other GH19 members and on functional data showing that this class of GH19 family members are involved in cell wall development rather than defense (50, 51). Groups B and C may in fact reflect four distinct lineages, but the nodal support for such a division is tentative. Maize GH19 proteins without chitin-binding domains are found in distinct lineages within groups B and C. All maize group B and C-2 members have lost the domain while in group C-1 only *Chit18* has lost the domain. Nodal support for the separation of C-1 and C-2 is weak so the loss of the chitin-binding domain in *Chit18* may or may not be distinct from the loss in group C-2. At any rate, it is intriguing that the chitin-binding domain has been lost at least twice in the evolution of maize and that those lineages have been maintained.

Phylogenetic analysis of maize, rice and Arabidopsis TLP families. We performed Bayesian analysis on alignment of 32 maize, 34 rice and 21 Arabidopsis TLP protein sequences (Fig. 6). We also included 2 sequences from insects and 3 from *Caenorhabditis elegans* that belong to the TLP family (46). We found that the rice and maize TLP genes were generally represented equally in most clades. Many clades were enriched in either eudicot (Arabidopsis) or monocot (rice and maize) sequences which may indicate expansion of these families since the divergence of eudicots and monocots. Zeamatin and PR5 from maize grouped with the well described tobacco TLP osmotin.

Measurement of GH19 and TLP transcripts during *A. flavus* infection. We investigated transcript accumulation of maize GH19 and TLP genes in developing seeds of the susceptible line B73 infected with *A. flavus* in the field. Transcription was profiled using a multigenome *A. flavus*/maize custom Affymetrix GeneChip microarray and real-time RT-PCR analysis of mock-inoculated and infected kernels. Comparisons were made at two stages in kernel development: early (blister) and middle (dough). We also used microarrays to monitor the response of GH19 and TLP genes over a four day infection time course at 22 days after pollination (late milk/early dough). The analysis of the GH19 and TLP genes is a subset of a global analysis of the *A. flavus*/maize seed interaction (Andrea L. Dolezal, unpublished data). The responses of the seven GH19 genes and eleven TLP genes

represented on the array differed (Table 3). In general, transcript increases in the blister stage were stronger than in the dough stage for genes that responded to infection. Genes that were induced at either stage also tended to show increased transcript accumulation in correlation with fungal infection over the four day time course. RT-PCR was performed on dough stage kernels. The fold change values obtained by RT-PCR are reported in the dough fold-change column (Table 3). RT-PCR results for *ChitA* and *TLP10* showed higher induction values than for the microarrays. *TLP2* showed the greatest induction of any gene measured (56-fold). Taken as a whole, the most highly responsive GH19 genes were *ChitA*, *Chit6* and *Chit13*. The most highly responsive TLP genes to infection were *TLP2*, *TLP4*, and *TLP10*. Transcripts for *TLP18*, *TLP22*, and *TLP25* were also responsive to infection. These data suggest that defense genes from both families are more strongly induced at earlier stages of kernel development.

DISCUSSION

We observed that an inhibitory protein fraction from maize kernels required the cooperative activities of chitin-binding proteins and a zeamatin-enriched fraction to inhibit aflatoxin production and growth of the fungal pathogen *A. flavus*. Although these two activities co-purify it is most likely not through physical association as they can be separated by chitin affinity chromatography. The fact that the two major components of the inhibitory fraction, ChitA and zeamatin, co-purify during ion exchange and hydrophobic interaction chromatography probably results from their similar theoretical pIs (8.34 and 8.16, respectively).

The two major seed chitinases, chitinases A (ChitA) and chitinase B (Chit B) were first purified by Huynh et al. (22) and shown to have antifungal activity against *Trichoderma reesei*, *Alternaria solani* and *Fusarium oxysporum*. Although ChitA is the most abundant chitin-binding protein in maize seeds (22), the maize genome contains at least 22 additional genes predicted to encode secretory proteins with chitin-binding domains. Of these 22 genes, six, including ChitA, are predicted to have molecular weights between 26-30 kDa and a theoretical pI between 7.8 and 8.9 following processing of the signal peptide. Three of these

are GH19 chitinases (C*hitB*, C*hit5, and Chit15*) and two are homologous to ZmPRm3 (accession # NP_001105541), an acidic *Chib1* chitinase of the PR-8 family of chitinases. Thus, although ChitA is the predominant protein in the chitin-binding fraction it is possible that the inhibitory activity involves a more complex mixture of chitinases in combination with the zeamatin-enriched fraction.

Although we detected chitinase activity in the zeamatin-enriched fraction (Fig. 4) this does not explain the cooperative effect between the two fractions since the level of chitinase activity was reduced by at least 12-fold and appeared to be the same isoform(s) present in the highly active chitin-binding fraction. The four bands observed in the in-gel chitinase assay (Fig. 4A) may indicate the presence of additional chitinase isoforms in the chitin-binding fraction. Alternatively, electrophoresis of a chitinase with multiple conserved disulfide-bonds (4) in the absence of a reducing agent may alter its electrophoretic mobility versus that observed in reducing conditions. Although we cannot rule out contributions of proteins present at low concentrations, our data are consistent with a model suggesting interaction between chitinase activities and zeamatin.

The interaction of antifungal proteins to achieve additive or synergistic inhibitory activity has been demonstrated *in vitro* and in field studies with transgenic lines transformed with pairs of genes encoding antifungal proteins (2, 27, 32, 43). Lorito et al. (32) observed synergistic activity against *Botrytis cinerea* spore germination by combining the tobacco PR-5 protein osmotin with various fungal cell wall degrading enzymes including chitinases. Thus, the combinatorial activity we observed is not unprecedented. Zeamatin has membrane permeabilizing activity although the precise mechanism for this activity is still unknown (44, 48). It is easy to envision a model whereby degradation of the fungal cell wall by the chitin-binding fraction could facilitate access to the plasma membrane by zeamatin. It has even been proposed that membrane disruption by permeabilizing agents may feed back to enhance cell wall degradation through perturbation of chitin synthase complexes (32). Indeed, tomato and rice plants transformed with a combination of a TLP and a GH19 chitinase showed enhanced resistance against fungal pathogens (33, 40).

Previous work has implicated maize seed chitinases and zeamatin in resistance to *A. flavus*. Another chitinase from Tex6 kernels has been reported (34) but it is unlikely that this protein corresponds to ChitA because the Tex6 chitinase had strong growth inhibitory activity against *A. flavus* at concentrations where the ChitA-enriched chitin-binding fraction did not inhibit growth. At a concentration of 133 µg/mL the ChitA-enriched chitin-binding did not inhibit growth while the Tex6 chitinase at a concentration of 100 µg/mL inhibited growth to ~20% of control (Fig. 3C, 34). Also, the first peptide sequence we obtained contains amino acid differences with a similar peptide sequence from the Tex6 chitinase. Ji et al. (25) observed induction of multiple chitinase isoforms in milk stage kernels seven days after inoculation with *A. flavus*. They also found that mature kernels contaminated with *A. flavus* as determined by detection of bright-green-yellow fluorescence had accumulated more chitinase than control kernels. Spore germination by *A. flavus* was reduced after a 12 hour incubation with zeamatin (17). We observed little effect of zeamatin-enriched fractions against *A. flavus* growth, but this difference is probably attributable to the greater length of our bioassay (4 days) which was necessary to observe aflatoxin production.

Our observation that inhibition of aflatoxin correlates with reduced fungal growth (Fig. 3B, C) is consistent with effects reported in other studies of aflatoxin inhibitory compounds. Of the described inhibitors of aflatoxin production, many show growth inhibitory activity at high concentrations while still inhibiting aflatoxin production at concentrations that do not reduce growth (19). This may result from a lag in the growth curve or in alterations in development (for example hyphal tip branching) or some combination of both.

We identified the maize GH19 and TLP gene families using bioinformatics tools (Tables 1 and 2). GH19 fell into four major phylogenetic groups. Phylogenetic analyses revealed that the current classification of the GH19 genes based on structural features is incongruent with the evolutionary history of the gene family. We noted that loss of the chitin-binding domain has occurred independently and been maintained in separate lineages. How chitinases that lack chitin-binding domains might function is an intriguing question although it seems unlikely that they are involved in defense. Roles outside of defense have been

established for GH19 chitinases. For example, Arabidopsis and cotton chitinase-like genes are involved in cell wall formation and the active site within the catalytic domain in the cotton homologs differs from that of other GH19 chitinases (14, 26, 50). Little is known about TLP genes in maize beyond zeamatin and *PR5*, which are induced in maize leaves following treatment with salicylic acid analogs and inoculation with common rust (36). Less is known about functions of maize TLP proteins that are unrelated to defense.

RNA profiling of infected kernels in the field showed that transcripts for a subset of the GH19 and TLP families are induced during infection. While the genes that were not induced may function outside of defense, it is also possible that their transcription is developmentally regulated in coordination with seed maturation. Maize ribosome-inactivating protein-1 (*RIP1*) is an antifungal protein that is under endosperm-specific regulation by the transcriptional activator Opaque-2 and accumulates to high levels in the seed (3, 39). Defense proteins that are abundant in the seed and are under developmental regulation may in fact be important components of pre-formed resistance. Of the genes that were induced during infection, C*hit6* and C*hit13* may be promising candidates for new proteins involved in defense against *A. flavus*. The TLP family also contains several candidates that may be important in defense. *TLP2*, *TLP4*, *TLP10*, *TLP24*, and *TLP30* were all highly responsive to infection.

Knowledge of the GH19 and TLP gene families would be improved by recombinant production of promising candidates for screening *in vitro*. Unfortunately, both families have highly conserved disulfide bonds and can prove difficult to express in *E. coli*. If proteins cannot be produced recombinantly and screened *in vitro* in fungal bioassays, co-expression of candidate genes in transgenic plants for resistance studies is an alternative. This approach is not without limitations either since interactions between endogenous and transgenic proteins would also need to be evaluated.

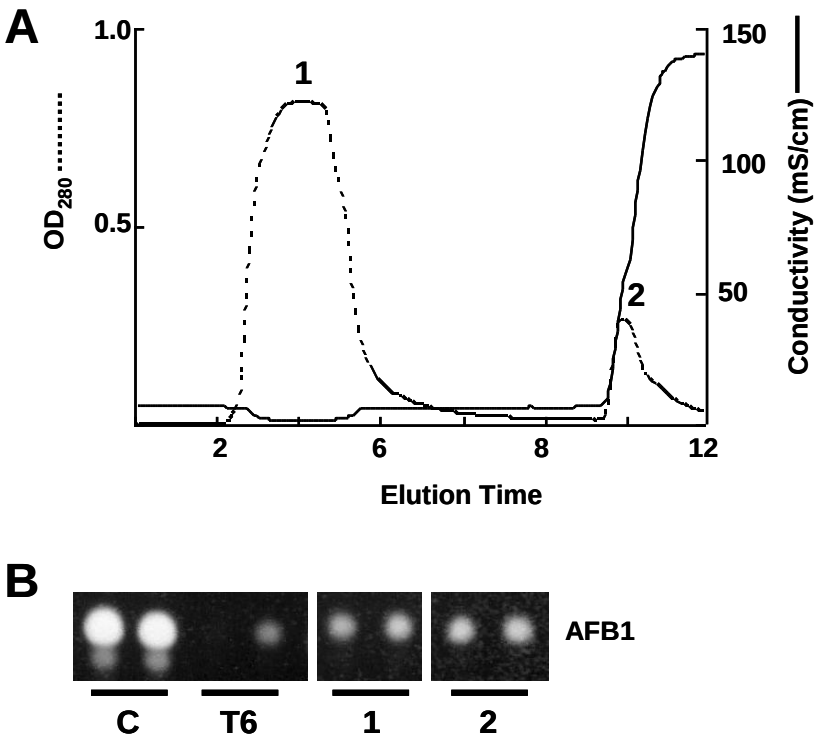
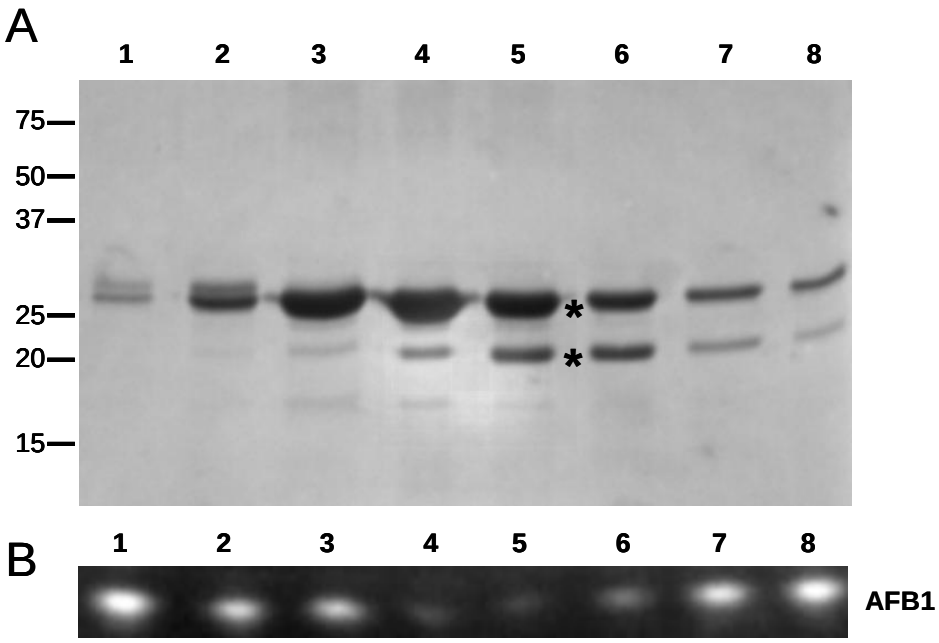


Figure 1. Cation exchange separation of major activities inhibitory to aflatoxin production. (A) Chromatograph of low pressure strong cation High S separation of Tex6 crude extract into an unbound fraction (peak 1) and bound fraction that was released by 1 M NaCl step elution (peak 2). (B) TLC plate showing aflatoxin extracted from *A. flavus* cultured in the presence the buffer control (C), the Tex6 crude extract (T6), unbound (1) or bound (2) fraction.

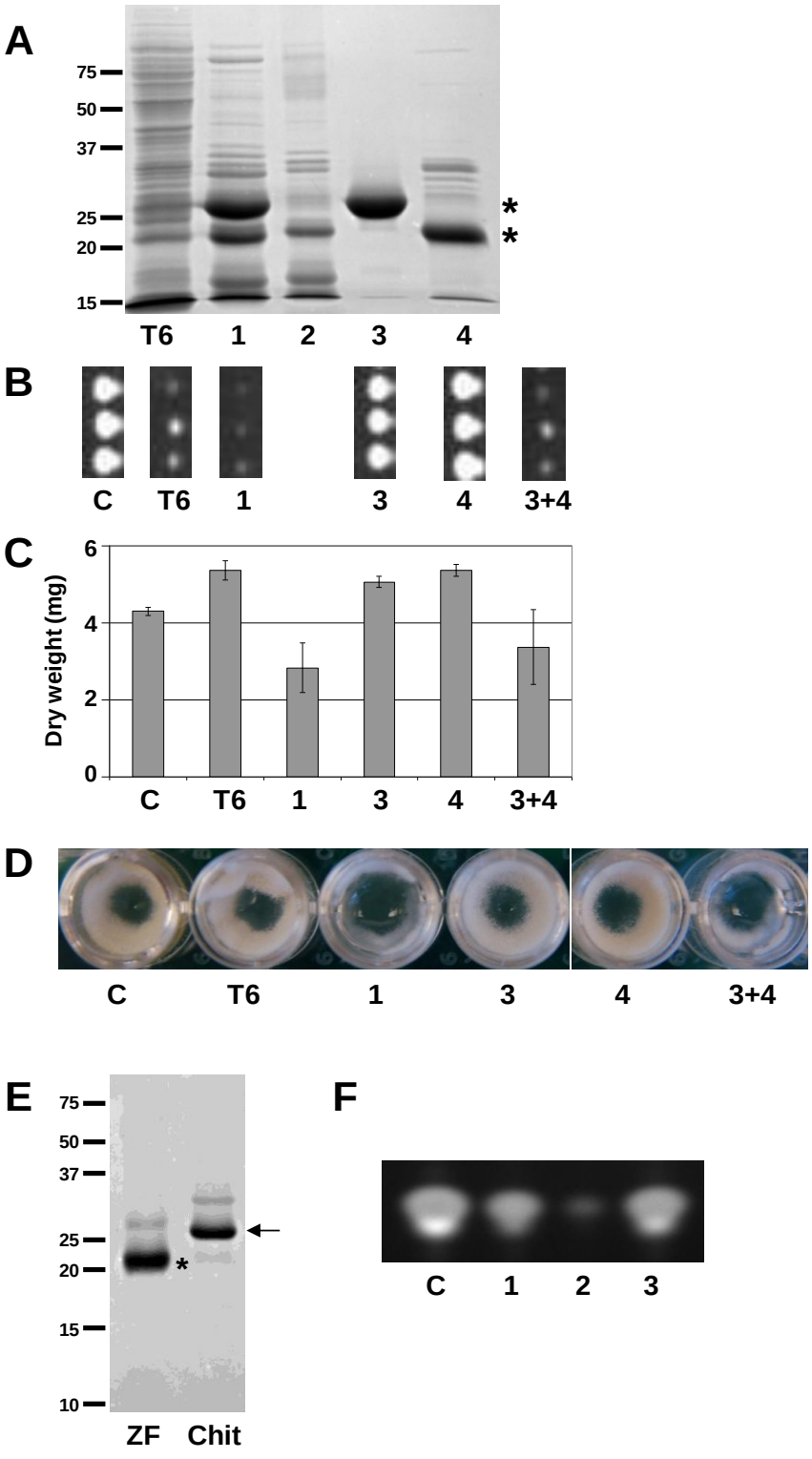


C

Band	Peptide Sequences	Accession #
~26 kD band	SAFLSAVNAYPGFAHGGTEVEGK	AAA33444 Chitinase A
	VAQDAVIAFK	
	DIGFNGLADPNR	
	GPLQISWNYNYGPAGR	
~22 kD band	APNTLAEYALK	P33679 Zeamatin precursor
	ITAPAGTTAAR	

Figure 2. Identification of proteins associated with inhibition of aflatoxin production in the cation exchange binding fraction. (A) 12% SDS polyacrylamide gel of proteins in inhibitory fractions obtained by final high pressure cation exchange chromatography step. Bands analyzed by LC-MS are marked with asterisks. (B) TLC plate showing aflatoxin from *A. flavus* cultured in the presence of the cation exchange fractions. Lanes 1-8 correspond to lanes 1-8 in panel A. (C) Peptide sequences obtained following trypsin digestion and LC-MS analysis of candidate bands.

Figure 3. Chitin affinity fractionation of activity inhibitory to aflatoxin production. (A) 12% SDS polyacrylamide gel of protein fractions assayed for inhibition of AF production. T6, Tex6 crude extract; 1, initial strong cation exchange fraction; 2, chitin flow-through (unbound fraction); 3, chitin-binding fraction eluted by acetic acid wash; 4, unbound fraction following additional UnoS cation exchange purification step. Asterisks denote bands corresponding to ChitA (upper) and zeamatin (lower). (B) TLC plate showing aflatoxin production from triplicate *A. flavus* cultures. Images have been cropped and rotated to align with protein gel. Labels correspond to protein gel except for C, buffer control. The chitin-binding fraction and the zeamatin-enriched fraction were assayed at roughly equimolar concentrations (~5 μM). Bioassays were not carried out on the chitin flow-through (lane 2) because it was used for the UnoS-1 column (lane 4). Final protein concentrations (μg/mL) in bioassay were as follows: T6, 430; 1, 193; 3, 133; 4, 120; 3+4, 127. (C) Mean dry weights (± SD) of *A. flavus* cultures. (D) Photograph of *A. flavus* cultures following four days growth in bioassay. (E) 15% SDS polyacrylamide gel showing hydroxyapatite purification of unbound, zeamatin fraction (ZF). ChitA is shown for reference (Chit, arrow). (F) TLC plate showing aflatoxin production from *A. flavus* cultures treated as follows: C, buffer control; 1, chitin-binding fraction as shown in 3A, lane 3; 2, chitin-binding fraction (CB) and HAP-purified zeamatin fraction (ZF) combined; 3, HAP purified ZF alone. CB and ZF were assayed at ~3 μM concentration. Final protein concentrations (μg/mL) in the bioassay were as follows: 1, 80; 2, 68; 3, 56.


A
75
50
37
25
20
15
*
*
T6
1
2
3
4
B
C
T6
1
3
4
3+4
C
Dry weight (mg)
6
4
2
0
C
T6
1
3
4
3+4
D
C
T6
1
3
4
3+4
E
75
50
37
25
20
15
10
*
ZF
Chit
F
C
1
2
3

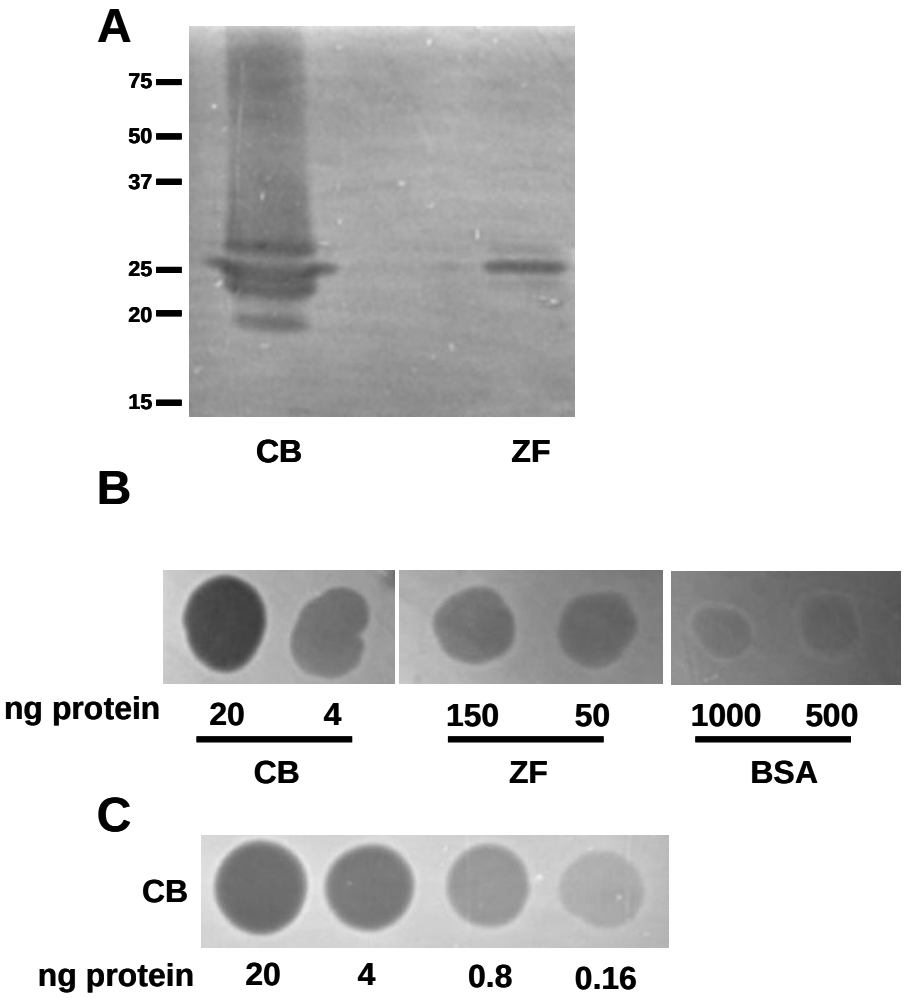


Figure 4. Comparison of chitinase activity in chitin-binding and zeamatin fractions. (A) 15% SDS polyacrylamide/0.01% (w/v) glycol chitin gel of chitin-binding fraction (CB) and zeamatin-enriched fraction (ZF) loaded at equimolar concentrations. Bands were visualized by the lack of fluorescence following chitin staining with Calcofluor White. (B) 7.5% SDS-polyacrylamide/0.04% (w/v) glycol chitin gel spotted with protein fractions. Labels as in 4A. BSA, bovine serum albumin control. Gel was visualized as in 4A. (C) Serial dilution of chitin-binding fraction (CB) to sub-nanogram amounts of protein of gel as in 4B.

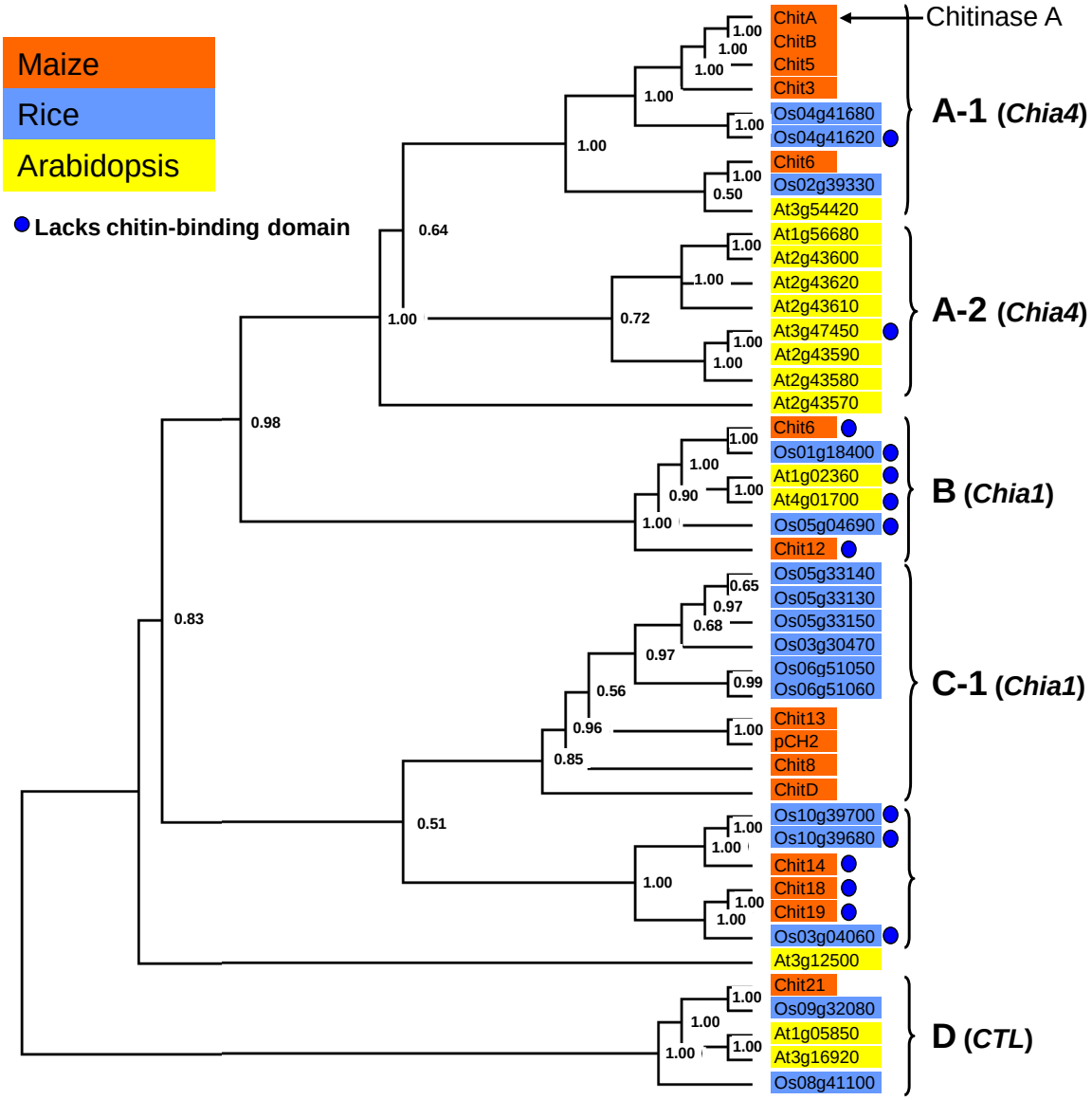


Figure 5. Plant GH19 phylogeny (50% strict consensus Bayesian tree). Plant GH19 family members fell into four distinct phylogenetic groups (A, B, C and D). Group A and Group C were divided into subgroups rather than separate groups due to low support at those nodes. The tree was rooted with the chitinase-like (CTL) group of GH19 sequences based on their distinct amino acid sequence and divergent function. Sequences were color-coded based on species of origin (maize, red; rice, blue; Arabidopsis, yellow). Blue dots indicate protein sequences that lack chitin-binding domains.

Figure 6. Bayesian tree of the TLP family. The phylogenetic tree was generated using protein sequences of maize, rice, and Arabidopsis. Sequences were color-coded based on species of origin (maize, red; rice, blue; Arabidopsis, yellow). *Caenorhabitans elegans* and two insect sequences are not colored.

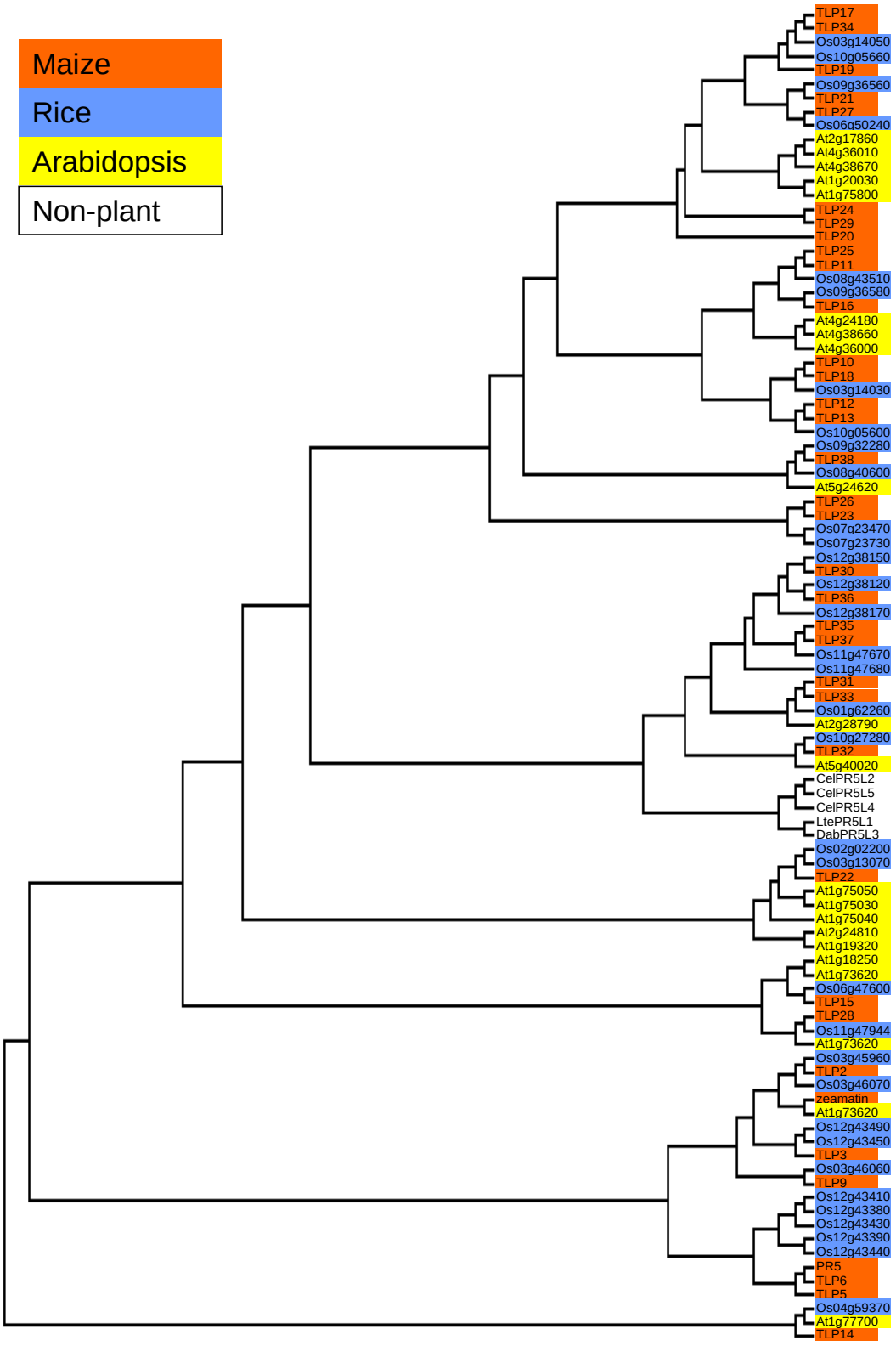

Maize
Rice
Arabidopsis
Non-plant
TLP17
TLP34
Os03g14050
Os10g05660
TLP19
Os09g36560
TLP21
TLP27
Os06g50240
At2g17860
At4g36010
At4g38670
At1g20030
At1g75800
TLP24
TLP29
TLP20
TLP25
TLP11
Os08g43510
Os09g36580
TLP16
At4g24180
At4g38660
At4g36000
TLP10
TLP18
Os03g14030
TLP12
TLP13
Os10g05600
Os09g32280
TLP38
Os08g40600
At5g24620
TLP26
TLP23
Os07g23470
Os07g23730
Os12g38150
TLP30
Os12g38120
TLP36
Os12g38170
TLP35
TLP37
Os11g47670
Os11g47680
TLP31
TLP33
Os01g62260
At2g28790
Os10g27280
TLP32
At5g40020
CelPR5L2
CelPR5L5
CelPR5L4
LtePR5L1
DabPR5L3
Os02g02200
Os03g13070
TLP22
At1g75050
At1g75030
At1g75040
At2g24810
At1g19320
At1g18250
At1g73620
Os06g47600
TLP15
TLP28
Os11g47944
At1g73620
Os03g45960
TLP2
Os03g46070
zeamatin
At1g73620
Os12g43490
Os12g43450
TLP3
Os03g46060
TLP9
Os12g43410
Os12g43380
Os12g43430
Os12g43390
Os12g43440
PR5
TLP6
TLP5
Os04g59370
At1g77700
TLP14

TABLE 1. Properties of maize GH19 family

Gene	Structural Class	Phylogenetic Group	Chitin Binding Domain	No. of Exons	Predicted Protein Length	Chromo-some	Nucleotide Accession	Protein Accession	UA 5' cDNA / Assembled EST[a]
ChitA	*Chia4*	A-1	+	2	280	2	M84164	P29022	MEC_19222_P95-Mar06
ChitB	*Chia4*	A-1	+	2	281	10	M84165	AAA33445	DV536632
Chit3	*Chia4*	A-1	+	2	278	2	AC186025[b]		none
Chit5	*Chia4*	A-1	+	2	278	8	27265.1[c]		none
Chit6	*Chia4*	A-1	+	2	271	5	AC197114[b]		DR819933
Chit8	*Chia1*	C-1	+	1	344	6	AC208338[b]		MEC_14583_P95-Mar06
ChitD	*Chia1*	nd	+	1	336	5	L16798[d]	AAA62420[e]	none
Chit11	*Chia1*	B	-	1	285	8	AC208042[b]		none
Chit12	*Chia1*	B	-	1	311	8	AC205328[b]		none
Chit13	*Chia1*	C-1	+	1	322	6	AC208338[b]	AAT40012[f]	MEC_30393_P95-Mar06
Chit14	*Chia2*	C-2	-	4	253	1	AC212113[b]		none
pCH2	*Chia1*	C-1	+	na	318	?	L00973	AAA62421	L00973
Chit18	*Chia2*	C-2	-	3	259	1	AC195193[b]		EE045723
Chit19	*Chia2*	C-2	-	3	261	1	AC195193[b]		DV025500
Chit21	*CTL*[g]	D	+	3	328	7	AC203444[b]		MEC_63395_P95-Mar06

[a] MEC accessions available at http://magi.plantgenomics.iastate.edu/. [b] Maize BAC. [c] GSS ID: ZmGSStuc11-12-04 (http://www.plantgdb.org/). [d] mRNA sequence not full length. Complete nucleotide sequence resides on BAC AC211441. [e] Protein accession not full length. [f] Protein accession from *Zea mays* ssp. *parviglumis*. [g] Chitinase-like (homology with *GhCTL-1*, *GhCTL-2* and *AtCTL1*).

TABLE 2. Properties of maize TLP family

Gene	No. of Exons	Predicted Protein Length	Chromosome	Nucleotide Accession	Protein Accession	UA 5' cDNA/ Assembled EST[a]	Predicted Localization[b]
zeamatin	1	227	7	NM_001111886	P33679	EC891318	S
TLP2	1	230	1	AC195987[c]		EE046233	S
TLP3	1	220	3	AC187558[c]		EE020462	S
PR5	1	172	1	AC204824[c]	AAC25630	EC882756	S
TLP5	1	177	1	AC189052[c]		EE153676	S
TLP6	1	171	?	80434[d]		EC890131	S
TLP9	1	220	5	AC206943[c]		EE046233	S
TLP10	3	335	?	36432[d]		EC889345	M
TLP11	3	319	1	AC205545[c]		none	S
TLP12	2	511	1	AC210844[c]		DT653546	S
TLP13	2	255	5	AC185304[c]		MEC_25250_P95-Mar06	S
TLP14	2	252	7	6661[d]		MEC_13495_P95-Mar06	S
TLP15	2	249	5	AC209874[c]		EE039818	S
TLP16	3	245	?	AC183933[c]		EC891083	S
TLP17	3	324	9	AC209690[c]		DV167199	S
TLP18	3	333	9	AC206261[c]		EC895806	S
TLP19	2	326	1	AC204301[c]		DT647358	M
TLP20	2	276	6	AC209843[c]		none	C
TLP21	3	321	?	AC197544[c]		none	S
TLP22	1	249	9	AC22535[c]		EE155502	C
TLP23	3	259	10	AC193403[c]		EE013176	O
TLP24	2	256	4	AC211531[c]		DT647358	S
TLP25	2	319	3	AC214485[c]		DR795501	S
TLP26	2	250	2	AC195146[c]		EE013176	S
TLP27	2	269	5	AC190922[c]		DR815004	S
TLP28	2	256	4	AC211224[c]		EC895806	S
TLP29	3	318	6	AC194900[c]		none	S
TLP30	2	205	3	AC210168[c]		MEC_25135_P95-Mar06	O
TLP32	1	288	1	AY664413[c]	Q5GAV4	DV516626	M
TLP34	-	244[e]	?	37511[d]		DV167199	S
TLP35	1	240	1	AC214267[c]		EE188955	S
TLP36	1	219	1	AC206682[c]		MEC_24088_P95-Mar06	S
TLP37	1	254	?	AC204621[c]		EE160761	S
TLP38	4	298	2	AC190600[c]		none	C

[a] MEC accessions available at http://magi.plantgenomics.iastate.edu/. [b] Localization predicted with TargetP 1.1 (http://www.cbs.dtu.dk/services/TargetP/) set to winner-takes-all cutoff: C, chloroplast; M, mitochondria; O, other; S, secretory. [c] Maize BAC. [d] GSS ID: ZmGSStuc11-12-04 (http://www.plantgdb.org/). [e] No full length sequence available. Note: Accession numbers DQ245263, DQ245963, DQ245143 are likely misannotated in Genbank as maize TLPs since they not match any available maize sequences and BLAST searches showed multiple hits for DQ245263 and DQ245963 in soybean and DQ245143 in *Brassica* species.

TABLE 3. Expression of maize GH19 and TLP genes in response to infection by *A. flavus*

Gene	Affymetrix ID	Blister Infected/Mock	p-value	Dough Infected/Mock	p-value	Time Course[a]	p-value
GH19 Family							
ChitA	TC292341_ZM_s_at	6.00	0.0015	1.22	0.6088	+	<.0001
				8.5 (5.9-12.2)[b]			
ChitB	NP1132056_ZM_s_at	3.15	0.0086	1.05	0.8000	+	0.0003
Chit6	TC292372_ZM_s_at	4.32	0.0019	3.31	0.0058	+	0.0001
Chit8	TC282901_ZM_at	1.30	0.2366	1.01	0.9653	ns	0.6020
	TC282901_ZM_s_at	2.88	0.0002	1.22	0.2418	ns	0.0589
ChitD	TC290057_ZM_at	0.52	0.0031	0.800	0.1863	ns	0.0689
Chit13	TC293152_ZM_s_at	25.70	0.0002	6.84	0.0059	+	0.0011
Chit21	TC305511_ZM_at	0.36	0.0052	0.555	0.0586	ns	0.7494
	TC305512_ZM_s_at	0.58	0.0041	0.776	0.1009	ns	0.2924
TLP Family							
zeamatin				3.8 (2.8-5.3) [b]			
TLP2				56.5 (45.5-70.1) [b]			
PR5	TC312694_ZM_at	4.02	0.0011	1.77	0.0742	+	0.0218
	TC312694_ZM_x_at	27.30	0.0013	1.35	0.1914	ns	0.2110
				12.4 (10.1-15.2) [b]			
TLP10	TC295916_ZM_at	1.46	0.0379	1.09	0.6004	ns	0.4158
				3 (2.1-4.3) [b]			
TLP14	TC287963_ZM_at?	0.83	0.3194	0.846	0.3799	+	0.0021
TLP15	CO457118_ZM_at	0.73	0.1533	0.696	0.1102	-	<.0001
TLP16	TC312695_ZM_at	1.85	0.0018	1.19	0.2408	ns	0.2070
TLP18	TC311640_ZM_at	3.89	0.0014	1.87	0.0575	+	0.0005
	TC311640_ZM_x_at	3.62	0.0012	1.75	0.0655	+	0.0002
TLP22	TC315842_ZM_at	2.47	0.0013	1.41	0.1055	+	0.0003
	TC315842_ZM_x_at	2.56	0.0012	1.54	0.0541	+	<.0001
TLP24	CO532586_ZM_at	1.09	0.3138	0.818	0.0385	ns	0.4489
	CO532586_ZM_x_at	1.58	0.0260	1.14	0.4484	ns	0.4186
	TC295231_ZM_at	2.04	0.0020	1.36	0.0868	+	0.0051
TLP25	TC295219_ZM_at	3.52	0.0013	1.56	0.1225	+	0.0016
	TC295219_ZM_x_at	2.64	0.0013	1.36	0.1618	+	0.0003

[a] Transcript accumulation was rated as positively (+) or negatively (-) correlated with time based on a p-value < 0.05 for the linear regression model. Array elements with no significant change designated as ns. [b] Fold-change as determined by RT-PCR (numbers in parentheses represent the lower and upper standard deviation).

REFERENCES

1. **Abad, L. R., M. P. D'Urzo, L. Dong, M. L. Narasimhan, M. Reuveni, K. Z. Jian, N. Xiaomu, N. K. Singh, P. M. Hasegawa, and R. A. Bressan.** 1996. Antifungal activity of tobacco osmotin has specificity and involves plasma membrane permeabilization. Plant Sci **118:**11-23.

2. **Anand, A., T. Zhou, H. N. Trick, B. S. Gill, W. W. Bockus, and S. Muthukrishnan.** 2003. Greenhouse and field testing of transgenic wheat plants stably expressing genes for thaumatin-like protein, chitinase and glucanase against *Fusarium graminearum*. J Exper Bot **54:**1101-1111.

3. **Bass, H. W., C. Webster, G. R. OBrian, J. K. Roberts, and R. S. Boston.** 1992. A maize ribosome-inactivating protein is controlled by the transcriptional activator Opaque-2. Plant Cell **4:**225-234.

4. **Beintema, J. J.** 1994. Structural features of plant chitinases and chitin-binding proteins. FEBS Lett **350:**159-163.

5. **Bowles, D. J.** 1990. Defense-related proteins in higher plants. Ann Rev Biochem **59:**873-907.

6. **Brown, R. L., Z. Y. Chen, A. Menkir, and T. E. Cleveland.** 2003. Using biotechnology to enhance host resistance to aflatoxin contamination of corn. Afric J Biotechnol **2:**557-562.

7. **Bullerman, L. B.** 1996. Occurrence of *Fusarium* and fumonisins on food grains and in foods. Adv Exp Med Biol **392:**27-38.

8. **Chen, Z. Y., R. L. Brown, T. E. Cleveland, K. E. Damann, and J. S. Russin.** 2001. Comparison of constitutive and inducible maize kernel proteins of genotypes resistant or susceptible to aflatoxin production. J Food Protect **64:**1785-1792.

9. **Chen, Z. Y., R. L. Brown, K. E. Damann, and T. E. Cleveland.** 2007. Identification of maize kernel endosperm proteins associated with resistance to aflatoxin contamination by *Aspergillus flavus*. Phytopathology **97:**1094-1103.

10. **Chen, Z. Y., R. L. Brown, K. E. Damann, and T. E. Cleveland.** 2002. Identification of unique or elevated levels of kernel proteins in aflatoxin-resistant maize genotypes through proteome analysis. Phytopathology **92:**1084-1094.

11. **Chen, Z. Y., R. L. Brown, A. R. Lax, T. E. Cleveland, and J. S. Russin.** 1999. Inhibition of plant-pathogenic fungi by a corn trypsin inhibitor overexpressed in *Escherichia coli*. Appl Environ Microbiol **65:**1320-1324.

12. **Chen, Z. Y., R. L. Brown, K. Rajasekaran, K. E. Damann, and T. E. Cleveland.** 2006. Identification of a maize kernel pathogenesis-related protein and evidence for its involvement in resistance to *Aspergillus flavus* infection and aflatoxin production. Phytopathology **96:**87-95.

13. **Coca, M. A., B. Damsz, D. J. Yun, P. M. Hasegawa, R. A. Bressan, and M. L. Narasimhan.** 2000. Heterotrimeric G-proteins of a filamentous fungus regulate cell wall composition and susceptibility to a plant PR-5 protein. Plant J **22:**61-69.

14. **De Jong, A. J., J. Cordewener, F. L. Schiavo, M. Terzi, J. Vandekerckhove, A. Van Kammen, and S. C. De Vries.** 1992. A carrot somatic embryo mutant is rescued by chitinase. Plant Cell **4:**425-433.

15. **Di Fonzo, N., H. Hartings, M. Brembilla, M. Motto, C. Soave, E. Navarro, J. Palau, W. Rhode, and F. Salamini.** 1988. The b-32 protein from maize endosperm, an albumin regulated by the O2 locus: nucleic acid (cDNA) and amino acid sequences. Mol Gen Genet **212:**481-487.

16. **Guo, B. Z., R. L. Brown, A. R. Lax, T. E. Cleveland, J. S. Russin, and N. W. Widstrom.** 1998. Protein profiles and antifungal activities of kernel extracts from corn genotypes resistant and susceptible to *Aspergillus flavus*. J Food Protect **61:**98-102.

17. **Guo, B. Z., Z. Y. Chen, R. L. Brown, A. R. Lax, T. E. Cleveland, J. S. Russin, A. D. Mehta, C. P. Selitrennikoff, and N. W. Widstrom.** 1997. Germination induces accumulation of specific proteins and antifungal activities in corn kernels. Phytopathology **87:**1174-1178.

18. **Gurr, S. J., and P. J. Rushton.** 2005. Engineering plants with increased disease resistance: what are we going to express? Trends Biotechnol **23:**275-282.

19. **Holmes, R. A., R. S. Boston, and G. A. Payne.** 2008. Diverse inhibitors of aflatoxin biosynthesis. Appl Microbiol Biotechnol **78:**559-572.

20. **Huang, Z., D. G. White, and G. A. Payne.** 1997. Corn seed proteins inhibitory to *Aspergillus flavus* and aflatoxin biosynthesis. Phytopathology **87:**622-627.

21. **Huelsenbeck, J. P., and F. Ronquist.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:** 754-755.

22. **Huynh, Q. K., C. M. Hironaka, E. B. Levine, C. E. Smith, J. R. Borgmeyer, and D. M. Shah.** 1992. Antifungal proteins from plants: purification, molecular cloning, and antifungal properties of chitinases from maize seed. J Biol Chem **267:**6635-6640.

23. **Ibeas, J. I., H. Lee, B. Damsz, D. T. Prasad, J. M. Pardo, P. M. Hasegawa, R. A. Bressan, and M. L. Narasimhan.** 2000. Fungal cell wall phosphomannans facilitate the toxic activity of a plant PR-5 protein. Plant J **23:**375-383.

24. **Ibeas, J. I., D. J. Yun, B. Damsz, M. L. Narasimhan, Y. Uesono, J. C. Ribas, H. Lee, P. M. Hasegawa, R. A. Bressan, and J. M. Pardo.** 2001. Resistance to the plant PR-5 protein osmotin in the model fungus *Saccharomyces cerevisiae* is mediated by the regulatory effects of SSD1 on cell wall composition. Plant J **25:**271-280.

25. **Ji, C., R. A. Norton, D. T. Wicklow, and P. F. Dowd.** 2000. Isoform patterns of chitinase and beta-1,3-glucanase in maturing corn kernels (*Zea mays* L.) associated with *Aspergillus flavus* milk stage infection. J Agric Food Chem **48:**507-511.

26. **Kasprzewska, A.** 2003. Plant chitinases: regulation and function. Cell Mol Biol Lett **8:**809-824.

27. **Kim, J. K., I. C. Jang, R. Wu, W. N. Zu, R. S. Boston, Y. H. Lee, P. Alm, and B. H. Nahm.** 2003. Co-expression of a modified maize ribosome-inactivating protein and a rice basic chitinase gene in transgenic rice plants confers enhanced resistance to sheath blight. Transgen Res **12:**475-484.

28. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227:**680-685.

29. **Levorson, J., and C. A. Chlan.** 1997. Plant chitinase consensus sequences. Plant Mol Biol Rep **15:**122-133.

30. **Liu, D., D. Rhodes, M. P. DUrzo, Y. Xu, M. L. Narasimhan, P. M. Hasegawa, R. A. Bressan, and L. Abad.** 1996. In vivo and in vitro activity of truncated osmotin that is secreted into the extracellular matrix. Plant Sci **121:**123-131.

31. **Livak, K. J., and T. D. Schmittgen.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. Methods **25:**402-408.

32. **Lorito, M., S. L. Woo, M. DAmbrosio, G. E. Harman, C. K. Hayes, C. P. Kubicek, and F. Scala.** 1996. Synergistic interaction between cell wall degrading enzymes and membrane affecting compounds. Mol Plant-Microbe Interact **9:**206-213.

33. **Maruthasalam, S., K. Kalpana, K. K. Kumar, M. Loganathan, K. Poovannan, J. A. J. Raja, E. Kokiladevi, R. Samiyappan, D. Sudhakar, and P. Balasubramanian.** 2007. Pyramiding transgenic resistance in elite indica rice cultivars against the sheath blight and bacterial blight. Plant Cell Rep **26:**791-804.

34. **Menkir, A., R. L. Brown, R. Bandyopadhyay, Z. Chen, and T. E. Cleveland.** 2006. A USA–Africa collaborative strategy for identifying, characterizing, and developing maize germplasm with resistance to aflatoxin contamination. Mycopathologia **162:**225-232.

35. **Moore, K. G., M. S. Price, R. S. Boston, A. K. Weissinger, and G. A. Payne.** 2004. A chitinase from Tex6 maize kernels inhibits growth of *Aspergillus flavus*. Phytopathology **94:**82-87.

36. **Morris, S. W., B. Vernooij, S. Titatarn, M. Starrett, S. Thomas, C. C. Wiltse, R. A. Frederiksen, A. Bhandhufalck, S. Hulbert, and S. Uknes.** 1998. Induced resistance responses in maize. Mol Plant-Microbe Interact **11:**643-658.

37. **Nakazaki, T., T. Tsukiyama, Y. Okumoto, D. Kageyama, K. Naito, K. Inouye, and T. Tanisaka.** 2006. Distribution, structure, organ-specific expression, and phylogenic analysis of the pathogenesis-related protein-3 chitinase gene family in rice (*Oryza sativa* L.). Genome **49:**619-630.

38. **Narasimhan, M. L., H. Lee, B. Damsz, N. K. Singh, J. I. Ibeas, T. K. Matsumoto, C. P. Woloshuk, and R. A. Bressan.** 2003. Overexpression of a cell wall glycoprotein in *Fusarium oxysporum* increases virulence and resistance to a plant PR-5 protein. Plant J **36:**390-400.

39. **Nielsen, K., G. A. Payne, and R. S. Boston.** 2001. Maize ribosome-inactivating protein inhibits normal development of *Aspergillus nidulans* and *Aspergillus flavus*. Mol Plant-Microbe Interact **14:**164-172.

40. **Ouyang, B., Y. H. Chen, H. X. Li, C. J. Qian, S. L. Huang, and Z. B. Ye.** 2005. Transformation of tomatoes with osmotin and chitinase genes and their resistance to *Fusarium* wilt. J Hort Sci Biotechnol **80:**517-522.

41. **Payne, G. A.** 1992. Aflatoxin in maize. Crit Rev Plant Sci **10:**423-440.

42. **Payne, G. A., J. J. Yu, W. C. Nierman, M. Machida, D. Bhatnagar, T. E. Cleveland, and R. A. Dean.** 2007. A first glance into the genome sequence of Aspergillus flavus, p. 15-24. *In* S. A. Osmani and G. H. Goldman (ed.), The Aspergilli: Genomics, medical aspects, biotechnology, and research methods, vol. 26. CRC Press, Boca Raton.

43. **Punja, Z. K.** 2001. Genetic engineering of plants to enhance resistance to fungal pathogens: a review of progress and future prospects. Can J Plant Pathol **23:**216-235.

44. **Roberts, W. K., and C. P. Selitrennikoff.** 1990. Zeamatin, an antifungal protein from maize with membrane-permeabilizing activity. J Gen Microbiol **136:**1771-1778.

45. **Salzman, R. A., H. Koiwa, J. I. Ibeas, J. M. Pardo, P. M. Hasegawa, and R. A. Bressan.** 2004. Inorganic cations mediate plant PR5 protein antifungal activity through fungal Mnn1- and Mnn4-regulated cell surface glycans. Mol Plant-Microbe Interact **17:**780-788.

46. **Shatters, R. G., L. M. Boykin, S. L. Lapointe, W. B. Hunter, and A. A. Weathersbee.** 2006. Phylogenetic and structural relationships of the PR5 gene family reveal an ancient multigene family conserved in plants and select animal taxa. J Mol Evol **63:**12-29.

47. **Trudel, J., and A. Asselin.** 1989. Detection of chitinase activity after polyacrylamide gel electrophoresis. Anal Biochem **178:**362-366.

48. **van Loon, L. C., M. Rep, and C. M. J. Pieterse.** 2006. Significance of inducible defense-related proteins in infected plants. Ann Rev Phytopathol **44:**135-162.

49. **Vigers, A. J., W. K. Roberts, and C. P. Selitrennikoff.** 1991. A new family of plant antifungal proteins. Mol Plant-Microbe Interact **4:**315-323.

50. **Zhang, D., M. Hrmova, C. H. Wan, C. Wu, J. Balzen, W. Cai, J. Wang, L. D. Densmore, G. B. Fincher, H. Zhang, and C. H. Haigler.** 2004. Members of a new group of chitinase-like genes are expressed preferentially in cotton cells with secondary walls. Plant Mol Biol **54:**353-372.

51. **Zhong, R., S. J. Kays, B. P. Schroeder, and Z. H. Ye.** 2002. Mutation of a chitinase-like gene causes ectopic deposition of lignin, aberrant cell shapes, and overproduction of ethylene. Plant Cell **14:**165-179.

CHAPTER 3

An *Aspergillus flavus* aflatoxin biosynthetic cluster gene encoding a predicted EthD domain protein is involved in aflatoxin biosynthesis

ABSTRACT

Aflatoxins are carcinogenic, polyketide-derived mycotoxins produced by members of the genus *Aspergillus*. Their biosynthesis is performed by the coordinated action of more than 25 genes contained within a ~70 kb aflatoxin gene cluster. We identified a small open reading frame, *hypE*, situated between *aflM(ver-1)* and *aflN(verA)* that is predicted to encode a 127 amino acid protein. The predicted HypE protein contains a putative EthD domain, a domain described to date only in bacteria. A *hypE* deletion mutant of *A. flavus* produced less aflatoxin B1 and B2 than control strains and accumulated an unknown metabolite which we assigned as a tentative HypE substrate (HESUB). Aflatoxin biosynthesis could be restored in this *hypE* mutant by a *hypE* overexpression construct. Restored aflatoxin biosynthesis was associated with loss of HESUB accumulation. These data suggest that the *hypE* gene product is involved in formation of aflatoxins, although its precise biochemical function remains unclear. Bioinformatic searches revealed the presence of potential *hypE* orthologs within aflatoxin and sterigmatocystin producing Aspergilli and, *Podospora anserina*. In addition to *hypE* orthologs, we identified other *hypE*-like genes within the Pezizomycotina. Among eukaryotes this gene family appears to be restricted to this subphylum of the Ascomycota.

INTRODUCTION

Aflatoxins are carcinogenic, polyketide-derived mycotoxins produced by filamentous fungi of the genus *Aspergillus*. Aflatoxin biosynthesis is controlled by enzymes and regulatory proteins encoded within a ~70 kb gene cluster that has been the subject of intensive study. The *A. flavus* aflatoxin gene cluster contains at least 30 genes, most of which have been assigned a function in aflatoxin biosynthesis with the exception of *aflE*, *aflF*, *aflT*, (25, 27) and four small predicted genes: *hypB*, *hypC*, *hypD*, and *hypE*. *hypB* was identified by transcriptional profiling of the aflatoxin biosynthetic cluster (18). A subsequent microarray study showed that *hypC*, *hypD* and *hypE* are transcribed and responsive to temperature regulation like other genes within the aflatoxin cluster (16). It is unclear how these small hypothetical proteins might function in aflatoxin biosynthesis, but the conservation of *hypB*, *hypD* and *hypE* sequences in the sterigmatocystin cluster of *A. nidulans* is consistent with a

role in the aflatoxin pathways. Sterigmatocystin is a late pathway intermediate in aflatoxin biosynthesis, but also accumulates in many fungal species that do not produce aflatoxin (5). The biosynthetic pathway for sterigmatocystin is similar to that for aflatoxin except the final genes necessary for aflatoxin biosynthesis are missing. The two pathways also appear to have similar genetic controls. Although *hypC* is not conserved in the *A. nidulans* cluster it may still function in aflatoxin biosynthesis but potentially after the production of sterigmatocystin.

While the biosynthetic enzymes and transcriptional regulators of aflatoxin biosynthesis are becoming better understood, the evolutionary processes driving the assembly and conservation of the aflatoxin gene cluster, and fungal secondary metabolism gene clusters generally, are still the subject of investigation and speculation (2, 17). Horizontal gene transfer among fungi or, alternatively, vertical transmission, duplication and gene loss are two proposed mechanisms for the dispersal of secondary metabolism gene clusters among diverse fungal lineages. Kroken et al. (14) have shown that horizontal transfer among fungi is not necessary to explain the distribution of polyketide synthase (PKS) genes among fungal taxa. However, among fungi, and eukaryotes generally, PKS genes are restricted to the Pezizomycotina. Thus, although horizontal transfer among filamentous fungi may not explain the distribution of secondary metabolism clusters, it remains unclear why PKS genes and other secondary metabolism gene families would be restricted to one subphylum within kingdom fungi.

We investigated the function of *hypE* using genetic, bioinformatic and phylogenetic tools. We show here that *hypE* is involved in aflatoxin biosynthesis. We also place HypE within a previously undescribed group of fungal enzymes restricted to the Pezizomycotina with significant structural conservation with bacterial EthD-domain proteins.

MATERIALS AND METHODS

Fungal strains and culture conditions. *A. flavus* strain AFC-1 (*pyrG*, *argD*), derived by mutagenesis of NRRL 3357 was used for generation of Δ*hypE* strains. AFC-1 and NRRL 3357-5 (*pyrG*, also derived by mutagenesis of NRRL 3357) were used for generation of *gpdA*::*hypE* overexpression strains, and AFC-1 was used as a control in the metabolic

profiling studies. Cultures were maintained at -80º C as glycerol stocks until needed. Inoculum was prepared as needed by culturing the fungus on PDA or PDAU (PDA+10µg uracil/L) plates at room temperature. Cultures for aflatoxin analysis and metabolic profiling were grown at 28º C in shake culture.

Construction of *hypE* deletion and overexpression vectors. An overlap PCR-based strategy (2) was used to make the *hypE* gene deletion construct (Fig. 1). The following primers were used to amplify 2.3 kb and 2.0 kb flanking regions at 5' and 3' ends, respectively, of the *hypE* coding region (nucleotide accession# AY510451) and a 2.7 kb fragment containing the *argD* marker gene: *hypE* 5' flank forward primer (p1), 5'-CCTCCACAGCACCGAATTAT; *hypE* 5' flank reverse primer (p3), 5'-*TGTGATTTGCTCCGCAATTA*ATCGAATACAAACTTAGCGA; *hypE* 3' flank forward primer (p4), 5'-*GAATCCCTGCATCAGAGGAA*TTCGCGCCGGTACCTTGGAG; *hypE* 3' flank reverse primer (p6) , 5'-GCTAGAAACAGTGCCCAAGC; *argD* forward primer (p2), 5'- TCGCTAAGTTTGTATTCGAT*TAATTGCGGAGCAAATCACA*; *argD* reverse primer (p5), 5'-*CTCCAAGGTACCGGCGCGAA*TTCCTCTGATGCAGGGATTC. The primers were designed to create regions of overlap between the resulting flanking regions and selectable marker PCR products as described (6). Nucleotides corresponding to *argD* are italicized. The three resulting PCR fragments were then combined to produce a template for an overlap PCR reaction to produce the 6.9 kb gene deletion construct. To generate the *hypE* overexpression construct a 0.4 kb fragment containing the *hypE* open reading frame (ORF) was amplified with primers containing endonuclease restriction sites: *hypE* 5' forward primer, 5'-CTGGTA<u>CCATGG</u>GCATGTCGACCGACGGCTTCAC (*Nco*I); *hypE* 3' reverse primer, 5'-ATGGATCCAAATTCGCAGTTCATTGTG (*Bam*HI). The pNOM102 plasmid (20) was cleaved with the restriction enzymes *Nco*I and *Bam*HI to remove the GUS reporter gene and the linear plasmid fragment was gel-purified and dephosphorylated with shrimp alkaline phosphatase. The *hypE* PCR product was also cleaved with *Nco*I and *Bam*HI and the cut product was ligated with the digested, linearized pNOM102 vector to form plasmid pHYPEoe. Insertion of *hypE* into the plasmid was verified by PCR with primers internal to

hypE: hypE internal forward, 5'- TGTTTTTGTGACTCGCAAGC; hypE internal reverse (sp2), 5'- GTTTTTCTCACGCCACCAAG.

Fungal transformation. Protoplasts of AFC-1, 3357-5, Δ*hypE*-A7, and Δ*hypE* -A16 were transformed as previously described (9). To generate gene deletion mutants, protoplasts were transformed with 1.5 μg of the linear *hypE* deletion construct. For co-transformations to generate Δ*hypE*+*gpdA*::*hypE* and *gpdA*::*hypE* overexpression strains, 1.5 μg of the overexpression plasmid (pHYPEoe) and 0.5 μg of the selectable marker construct (pAMA-*pyr4* for transformation of 3357-5 or TOPO-*argD* for transformation of AFC-1) were used.

Genomic DNA extraction and PCR screening of fungal strains. DNA was extracted from fungal cultures grown on PDA or PDAU plates using the spore CTAB method as previously described (9). Transformants were screened following selection on minimal medium low-salts (MLS; Czapeck-Dox broth plus 0.4 M $(NH_4)_2SO_4$ and 1% agar) or MLSU (MLS+10 μg uracil/L) using the following primers: hypE 5' screening primer forward (sp1), 5'-CACTTCACCACCGATAATTCG; hypE internal screening primer reverse (sp3), 5'-GTTTTTCTCACGCCACCAAG; argD internal screening primer reverse (sp4), 5'-GGTCGGGATTGGGTAGAGAT. The hypE internal screening primer reverse (sp2) was also used in combination with hypE internal screening primer forward (sp3) for detection of *hypE* in the leaky A16 *hypE* mutant.

TLC analysis of secondary metabolites. To test for production of aflatoxin and related metabolites of the selected strains, we inoculated 25 mL or 100 mL PDB or PDBU medium with 1 x 10^6 spores/mL of *A. flavus* and incubated the cultures at 28° C with shaking at 200 rpm. For the time course comparing AFC-1, A16, and A16+*gpdA::hypE* triplicate cultures of each strain were grown in 100 mL PDU supplemented with 10 μg/mL each of uridine and arginine. For TLC analysis 1 mL of culture medium was extracted with 0.5 mL chloroform by vortexing for two minutes followed by centrifugation at 10,000xg for five minutes. The chloroform layer was removed and evaporated and the residue was resuspended in 8 μL chloroform for spotting onto Whatman Partisil K6 Silica Gel 60 Å TLC plates (Fisher, Pittsburgh, PA). TLC plates were developed in toluene-methanol-acetic acid (80:15:5 [vol/vol/vol]) for one hour and imaged over a UV transilluminator.

Metabolic profiling. Identification of VERA accumulation in strain A20 was made after TLC separation of chloroform extracted culture medium and LC-MS/MS analysis of the colored compound isolated from the TLC separation. Target compound analysis of aflatoxins and identification of 328X and 328Y in AFC-1, A16 and A16+*gpdA::hypE* culture medium was performed at the North Carolina State University Genomic Sciences Laboratory (Raleigh, NC).

RNA isolation and cDNA synthesis. RNA was extracted with the RNeasy plant mini kit (QIAGEN, Valencia, CA) and genomic DNA was removed by DNase treatment (Promega, Madison, WI). One μg of RNA was used for cDNA synthesis using Stratascript reverse transcriptase and random hexamers (Stratagene, La Jolla, CA).

Bioinformatic and phylogenetic analysis of genes related to *hypE*. Fungal sequences similar to *hypE* were obtained by TBLASTN and BLASTP searches of Genbank and other publicly available fungal genomes (http://genome.jgi-psf.org/, http://www.broad.mit.edu/annotation/fgi/, http://fungal.genome.duke.edu/) using the predicted AfHypE protein as query. Bacterial sequences related to *hypE* were obtained by five iterative PSI-BLAST searches against the Genbank non-redundant database using default parameters. Phylogenetic analyses were performed on *hypE*-like predicted proteins with an e-value cutoff of 2e-10 or less based on alignment with AfHypE by BLASTP. Predicted protein alignments were generated with ClustalW, trimmed in Sequencher, and manually adjusted using the SNAP Workbench bioinformatics toolkit (17). SNAP Workbench was also used to generate 500 bootstrap datasets for Maximum Likelihood analysis. Maximum Likelihood and Bayesian analysis were performed using PAML 3.14beta and MrBayes v3.0B4 (11, 26), respectively. The distance tree of fungal and bacterial proteins was generated from the NCBI Genbank PSI-BLAST search tool. EthD domains in fungal *hypE*-like predicted proteins were identified using the conserved domain search function (http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) against the CDD v2.14-24291 PSSM database.

RESULTS

Identification of the *hypE* open reading frame. The identification of *hypB*, a small transcribed ORF located between *aflI* and *aflL* and transcriptionally regulated by AflR (18), prompted us to explore other intergenic regions within the aflatoxin biosynthetic cluster for additional predicted genes. This search revealed the presence of three additional small ORFs, *hypC*, *hypD*, and *hypE*, between *aflC* and *aflD*, between *aflN* and *aflG*, and between *aflM* and *aflN*, respectively, each with EST support. The *hypC* and *hypE* ORF were predicted by gene annotation software but the *hypD* ORF was not. The *hypE* ORF, which encodes a putative protein of 127 amino acids with no targeting domains for subcellular compartments (TargetP 1.1), was selected as a candidate for functional analysis.

Mutation, complementation and overexpression of *hypE*. Overlap PCR (6) was used to generate a linear *hypE* deletion construct with the *A. flavus argD* gene centered between 5' and 3' *hypE* flanking regions of ~2 kb each. This construct was then used to transform protoplasts of *A. flavus* strain AFC-1 (*pyrG, argD*; Fig. 1). Transformants selected for their ability to grow on medium without arginine were screened by PCR for the lack of *hypE* (Fig. 2A) and the presence of the *hypE* disruption construct at the *hypE* locus (Fig. 2B). Four putative Δ*hypE* mutants (A7, A16, A20 and A22) and a transformant (A9) in which the *hypE* deletion construct integrated ectopically were tested for aflatoxin production (Fig. 2C). Strains A7, A20 and A22 produced no detectable aflatoxin and accumulated a red-orange intermediate. Strain A16 produced aflatoxin at lower levels than the recipient strain AFC-1 or A9 and did not accumulate the colored intermediate. The colored intermediate was removed from the TLC plate and analyzed by LC-MS/MS which indicated that it had a molecular weight of 338 (Fig. 2D).

Two aflatoxin pathway intermediates, versicolorin A (VERA) and *O*-methyl sterigmatocystin (OMST), have a molecular weight of 338. The observation that the unknown compound had a visible color is consistent with an assignment as VERA since OMST is visible only under UV illumination (C. Woloshuk, personal communication). Also, the resulting MS/MS spectrum was more consistent with an expected fragmentation pattern of VERA than OMST (Fig. 2E; N. Glassbrook personal communication). The accumulation

of VERA in *hypE* mutants was unexpected considering that four other aflatoxin pathway genes (*aflM*, *aflN*, *aflX*, and *aflY*) have already been shown to be involved in the conversion of VERA to demethylsterigmatocystin (DMST) based on gene disruption or mutant complementation (3, 7, 12, 13, 21). These four enzymes theoretically are sufficient to catalyze the anthraquinone to xanthone conversion (VERA→OMST; 10). One possible explanation for the accumulation of VERA is disruption of either of the flanking genes *aflM* and *aflN* during deletion of *hypE*. To rule out this possibility we attempted to complement a mutant with each phenotype [A7 (VERA accumulation, no aflatoxin) and A16 (no VERA accumulation, low aflatoxin)] by co-transformation with a plasmid containing the selectable marker *pyr4* (complements the *pyrG* mutation) combined with a construct containing *hypE* driven by the constitutive glyceraldehyde phosphate dehydrogenase promoter (*gpdA*::*hypE*). The *gpdA::hypE* construct failed to restore aflatoxin production in strain A7 (Fig. 3A). These data suggested that the VERA phenotype was not a result of mutation of *hypE*. However, aflatoxin production was restored to near control levels in the low aflatoxin producer A16 (Fig. 3B). The fact that the low aflatoxin phenotype could be complemented by the *gpdA::hypE* construct in A16 suggested that *hypE* is important for aflatoxin biosynthesis. TLC analysis also revealed the accumulation of a metabolite (HESUB, for HypE substrate) in the low aflatoxin producing mutant (A16) that was not visible in the complemented strain (Fig. 3B). If HESUB is an aflatoxin precursor compound, its accumulation would support a function for HypE as a pathway enzyme. Overexpression of *hypE* by transformation of a wild-type strain with the *gpdA::hypE* construct did not noticeably increase aflatoxin production based on TLC analysis (Fig. 3C). Thus, if HypE is in fact a pathway enzyme it is probably not a limiting step for aflatoxin biosynthesis.

Analysis of *hypE*, *aflM* and *aflN* transcripts. Because deletion of *hypE* did not appear to be responsible for VERA accumulation we asked if *aflM* and *aflN* were functional in the VERA-accumulating strain. We performed PCR on cDNA from a toxigenic control strain (A9) and the *hypE* mutant A22. Transcripts for *aflM* and *aflN* were detected in both strains, although slightly less in A22 than in A9, while no transcript for *hypE* was detected in A22 (Fig. 4A). These observations suggested that if *aflM* or *aflN* functions were disrupted in

the VERA-accumulating strains it was not through loss of transcription of *aflM* or *aflN*. We also investigated transcription of *aflM* and *hypE* in the low aflatoxin mutant A16. We detected transcription of *hypE* at reduced levels relative to AFC-1 and tentatively assigned A16 as a leaky *hypE* mutant (Fig. 4B). Overexpression of *hypE* in A16 led to a dramatic increase in *hypE* transcript. Accumulation of the *aflM* transcript was relatively uniform in all backgrounds. No PCR products were detected when RNA that had not been incubated with reverse transcriptase was used as a template with primers for *gpdA* (Fig. 4C). Because we detected low *hypE* transcript in A16 (Fig. 4B) we performed PCR with primers internal to *hypE* (sp2 and sp3) on genomic DNA of AFC-1, A9 and A16 to test for the presence of *hypE*. We observed a faint band in A16 that was the same size as the *hypE* product in AFC-1 and A9 (Fig. 4D). The presence of this faint band and the detection of *hypE* transcript indicated that at least one copy of *hypE* was present and transcribed in strain A16 although *hypE* transcript accumulated at lower levels than in AFC-1 and in the overexpression strain (A16+*gpdA::hypE*).

Targeted analysis of aflatoxin production of the *hypE* mutant. We measured AFB1 and AFB2 production in AFC-1, A16 and A16 + *gpdA::hypE* by LC-MS/MS to obtain a quantitative comparison among these strains as opposed to the qualitative results presented in Fig. 3B. AFB1 production in A16 and A16 + *gpdA::hypE* was 26% and 80% of AFC-1, respectively (Fig. 5A). AFB2 production was reduced even further in A16 (8% of control) and complementation restored it to only 36% of control (Fig. 5B). LC-MS/MS analysis for aflatoxins also identified two compounds with interesting accumulation patterns and chemical properties. Because both compounds were MW 328 we named them 328X and 328Y. Accumulation of 328X in A16 and A16 + *gpdA::hypE* paralleled the accumulation of AFB1 and AFB2 (Fig. 5C), but it was not abundant. In AFC-1 the peak area of 328X corresponded to 0.19% of the peak areas of AFB1, and it was below the limit of detection in A16. 328Y was also not abundant in AFC-1 (0.35% peak area vs. AFB1), but its accumulation pattern was opposite of AFB1 and AFB2 and it accumulated dramatically in A16 (Fig. 5D). This pattern is similar to that observed for HESUB. Whether 328Y is in fact HESUB remains to be determined by additional LC-MS/MS analysis.

328Y is particularly interesting because its retention time on the C_{18} column during LC-MS/MS analysis (4.10 min) was similar to that for B and G aflatoxins (2.6-4.8 min, Fig. 6A). Also, its MS/MS fragmentation pattern was similar, but distinct, in comparison to that of AFG1 (Fig. 6B). Fragmentation of both compounds yielded two daughter ions with m/z values of 283 and 301. The similar retention time, MS/MS fragmentation pattern and molecular weight are all consistent with 328Y being a furanocoumarin with fused ring structures similar to those found in the aflatoxins.

Identification of fungal genes related to *hypE*. To learn more about the possible function of *hypE* we performed bioinformatic searches to identify related gene sequences. BLAST (Basic Local Alignment Search Tool) searches with the predicted HypE protein revealed the presence of three related predicted genes within *A. flavus*. BLAST searches of Genbank (http://www.ncbi.nlm.nih.gov/) and sequenced fungal genomes (http://fungal.genome.duke.edu/, http://genome.jgi-psf.org/) revealed the presence of an additional 25 predicted genes with sequence similarity to *hypE* (e-value <1e-04) all belonging to organisms within the Pezizomycotina subphylum of the Ascomycota (Table 1). Related sequences have been given the abbreviation *hel* for *hypE-like* except for three which likely represent orthologous genes as they are found in the aflatoxin/ST clusters of *A. parasiticus*, *A. nomius*, and *A. nidulans*. The *hypE* gene is transcribed divergently from *aflN* and this relationship is conserved in the *A. nidulans* sterigmatocystin cluster. All of the predicted proteins range in size from 119 to 167 amino acids with the exception of AniHelC which forms a domain within a larger protein. We investigated the genomic location of the two most closely related *hypE*-like genes from *P. anserina* and *Fusarium solani* to see if they were located within polyketide metabolic clusters. We found that the *P. anserina* gene sequence is located in a previously undescribed gene cluster highly similar to the sterigmatocystin cluster in *A. nidulans* and its position relative to other sterigmatocystin genes is conserved (RAH, unpublished data). We did not identify any predicted PKS genes within 40 kb upstream or downstream of *FshelA*.

Phylogenetic relationship of *hypE*-like sequences. To understand the phylogenetic relationships within this undescribed gene family we performed Maximum Likelihood and

Bayesian analyses of aligned protein sequences following manual trimming and gap closure. Maximum Likelihood analysis showed the presence of a well-supported *hypE* orthologous clade, as expected based on conserved location within aflatoxin or sterigmatocystin clusters (Fig. 7). An additional clade contained only representatives from the Aspergilli and may indicate expansion of the *hel* gene family within this genus. Alternatively, the formation of an Aspergillus-specific clade may also be due to sampling bias as Aspergilli genomes are overrepresented within available fungal genomic sequence data. Bayesian analysis yielded a concordant tree with the same *hypE* orthologous clade and also the Aspergilli-specific clade (data not shown). The separation of *hel* genes from different classes into distinct clades is consistent with duplication within the gene family prior to divergence of the major class lineages within Pezizomycotina. Gene loss or modification of function and subsequent loss of conservation may explain the absence of the *hel* family members in other sequenced Ascomycete genomes.

Relationship between fungal HEL-proteins and bacterial proteins. Conserved domain search indicated that *hypE,* predicted *hypE* orthologs, and two other related sequences (*FshelA* and *CghelA*) are predicted to encode proteins with EthD domains (Table 1). Alignment with the EthD consensus domain in the Pfam database (http://pfam.sanger.ac.uk/) and the *Rhodococcus ruber* EthD protein showed the conservation of several residues in the *hel* protein family including a histidine residue in the first half of the domain (Fig. 8). We also used mGenthreader (http://bioinf.cs.ucl.ac.uk/psipred/psiform.html) to compare the HypE protein sequence to solved protein structures and found that it had significant structural homology to the X-ray crystal structure of 2FTR from *Bacillus halodurans* which also contains an EthD domain (data not shown).

DISCUSSION

We identified the *hypE* ORF between *aflM* and *aflN* in the aflatoxin biosynthetic cluster of *A. flavus*. Transcriptional profiling indicated that *hypE* is regulated like other aflatoxin pathway genes in response to temperature (16). Bioinformatic searches revealed

that *hypE* is conserved in the aflatoxin clusters of other aflatoxigenic fungi (*A. parasiticus* and *A. nomius*) and in the sterigmatocystin biosynthetic cluster of *A. nidulans*. In addition, we found a closely related homolog in the Sordariomycete *P. anserina* which has a highly conserved sterigmatocystin cluster in its genome. We do not know if this cluster is functional as we did not find any report of *P. anserina* producing sterigmatocystin. Both Maximum Likelihood and Bayesian methods of phylogenetic inference placed the homologous *hypE* sequences from fungi with aflatoxin or sterigmatocystin clusters in an orthologous clade.

Our analysis of *hypE* gene deletion mutants showed that *hypE* is involved in biosynthesis of the B family of aflatoxins. The profile of metabolites produced in the *hypE* mutants suggested that *hypE* acts either in conversion of VERA to DMST or at a later step in the biosynthetic pathway. Several lines of evidence indicate that the latter hypothesis is more likely to be correct.

First, we found that transformation of VERA-accumulating *hypE* mutants with a *hypE* overexpression construct could not restore aflatoxin biosynthesis but transformation of a low aflatoxin, leaky *hypE* mutant with the same construct restored aflatoxin biosynthesis to a level near that in the parent strain (AFC-1). Because the low aflatoxin phenotype was present in a leaky *hypE* background it is not possible to conclude at present if *hypE* is absolutely required for aflatoxin biosynthesis. Second, the absence of a *hypE* ortholog from the genome of the sterigmatocystin producer *Chaetomium globosum* suggests that *hypE* may act after production of sterigmatocystin. Alternatively, *hypE* is involved in sterigmatocystin biosynthesis but is not absolutely essential for sterigmatocystin production in *C. globosum*. If *hypE* functions after sterigmatocystin then its conservation in the *A. nidulans* and *P. anserina* sterigmatocystin clusters is intriguing considering that other genes that function in conversion of sterigmatocystin to aflatoxin (*aflP* and *aflQ*) are not present (27). Because the residues critical for function of EthD domain proteins have not been identified it is difficult to predict if the differences between HypE protein sequences in aflatoxin and sterigmatocystin clusters would yield a less active or inactive form in the sterigmatocystin producers. There is evidence that some aflatoxin biosynthetic enzymes function in multi-protein complexes (23). If HypE functions in a multi-protein complex with other enzymes

required for synthesis of sterigmatocystin, one could envision an explanation for its retention in sterigmatocystin clusters even if its enzymatic function was after sterigmatocystin synthesis.

The accumulation of potential HypE substrates also points to a role late in aflatoxin biosynthesis. Like AFB1 and AFB2, HESUB was fluorescent under UV light with no visible pigmentation. Also, LC-MS/MS quantification of aflatoxin production identified 328Y, an unknown compound with similarities to aflatoxins that increased dramatically in the leaky *hypE* mutant versus the parent and complemented strains. Three important similarities suggest that this compound is related to aflatoxins. First its molecular weight, which is the same as AFG1, is between that of AFB1 (312) and OMST (338). Second, its retention time on the C_{18} column was similar to the aflatoxins. Third, its MS/MS fragmentation pattern revealed two daughter ions that are also produced upon fragmentation of AFG1. Whether 328Y corresponds to HESUB remains to be determined. Because the LC-MS/MS analysis targeted only aflatoxins we cannot assign 328Y as HESUB without a more detailed metabolic profiling comparison between the leaky *hypE* mutant and the parent and complemented strains. Precursors of aflatoxin with molecular weights of 328 have not been previously reported. This may be due to their low abundance relative to the B aflatoxins. Also, if they represent compounds influenced specifically by HypE function other aflatoxin pathway gene mutants may have either lacked these compounds or not had noticeable changes in their concentrations.

If HypE functions in the steps from sterigmatocystin to AFB1 it is difficult to predict where it would be necessary based on our current understanding of these steps. Sterigmatocystin is methylated to form OMST and this step was shown to require an *O*-methyltransferase encoded by *aflP* (15, 24). The conversion of OMST to AFB1 likely involves several intermediates, but Prieto and Woloshuk (19) showed that *aflQ* from *A. flavus* is sufficient to convert OMST to AFB1 when used to transform yeast along with the pathway specific regulator *aflR*. Whether AflQ is sufficient to convert DHOMST to AFB2 was not tested. Later, Udwary et al. (22) showed that the *A. parasiticus* AflQ enzyme is sufficient to make both conversions. While these studies show unequivocally that AflQ can

convert OMST to AFB1 in a heterologous system, aflatoxin biosynthesis may be more complex in aflatoxigenic Aspergilli.

The relationship between synthesis of B and G aflatoxins in *A. parasiticus* introduces more complexity. Unlike *A. parasiticus*, most *A. flavus* strains do not synthesize G aflatoxins and the sequences of *aflU* and *nadA*, two genes shown to be involved in synthesis of AFG1 and AFG2 in *A. parasiticus*, each contain a loss of function mutation (1, 4, 27). Sequence analyses comparing the *hypE* region between *aflM* and *aflN* among *A. parasiticus* populations that vary in the ratio of B aflatoxins to G aflatoxins indicate an association between polymorphisms in this region and the observed chemotype variation. Furthermore, the *A. flavus* copy of *hypE* in our genetic background (NRRL 3357) groups with the high G aflatoxin group (I. Carbone, personal communication). If HypE somehow affects biosynthesis of G aflatoxins it would be necessary to test this hypothesis in *A. parasiticus* by gene disruption and overexpression. Ideally, this would be performed in two chemotype backgrounds: a high B to G ratio strain and a high G to B ratio strain.

In addition to *hypE*, the functions of five other pathway genes (*hypB*, *hypC*, *aflE*(*norA*), *aflF(norB*) and *aflT*) remain unknown. *hypC*, *norB* and *aflT* are absent in sterigmatocystin clusters and, thus, may also function in the conversion of sterigmatocystin to B and G aflatoxins. It will be challenging to sort out the precise function of *hypE* without identifying HESUB and 328Y. The identity of these potential aflatoxin precursors will shed light on the precise role of *hypE* in this complex pathway.

Establishing the function of *hypE* in aflatoxin biosynthesis will also help provide insight into the function of the *hypE*-like gene family and EthD domains in general. Bacterial EthD domain proteins are found in degradation pathways and belong to the dimeric α/β superfamily. In the bacterium *Rhodococcus ruber* the *ethD* gene belongs to a cluster of four genes required for ethyl tert-butyl ether degradation. Although *ethD* is necessary for this metabolic activity its specific function is unknown. The *ethD* cluster also contains a cytochrome P450 monooxygenase, a ferredoxin oxidoreductase and ferredoxin. The dimeric α/β superfamily also contains a group of monooxygenases (ActVA-Orf6, TcmH, ElmH and Jad-Orf7) that do not require co-factors and are involved in polyketide biosynthesis in

bacteria (8). Thus, AfHypE shares deep structural conservation with bacterial EthD domain proteins and belongs to a superfamily that contains members with diverse enzymatic activity including monooxygenases involved in polyketide biosynthesis. It therefore seems likely that fungal HEL proteins may also be enzymes. The restriction of the *hel* gene family to the Pezizomycotina suggests that they are not involved in core, eukaryotic functions and raises the intriguing question of how fungal secondary metabolism genes were originally acquired. The distribution of PKS genes follows a similar pattern and it is tempting to speculate that gene families such as the *hel* family and PKS family may have been obtained through an ancient horizontal transfer from bacteria following the divergence of the Pezizomycotina from other subphyla within the Ascomycota, but prior to evolution of the major classes within the Pezizomycotina. It would be interesting to investigate if other fungal secondary metabolism genes that belong to gene families represented in bacteria and eukaryotes (e.g. cytochrome P450 monooxygenases) are more closely related to the bacterial or eukaryotic sequences. Regardless, the finding that *hypE* is involved in aflatoxin biosynthesis provides exciting insight into this previously undescribed gene family.

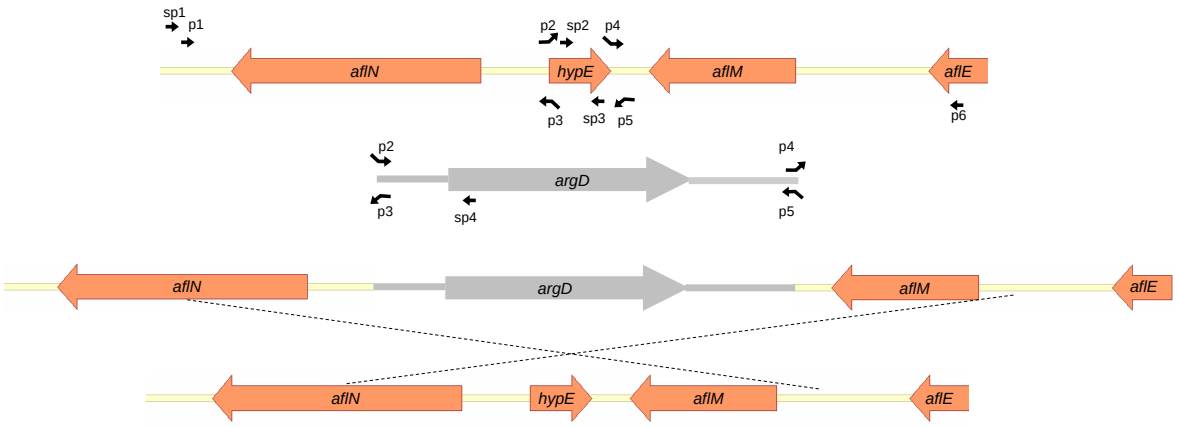
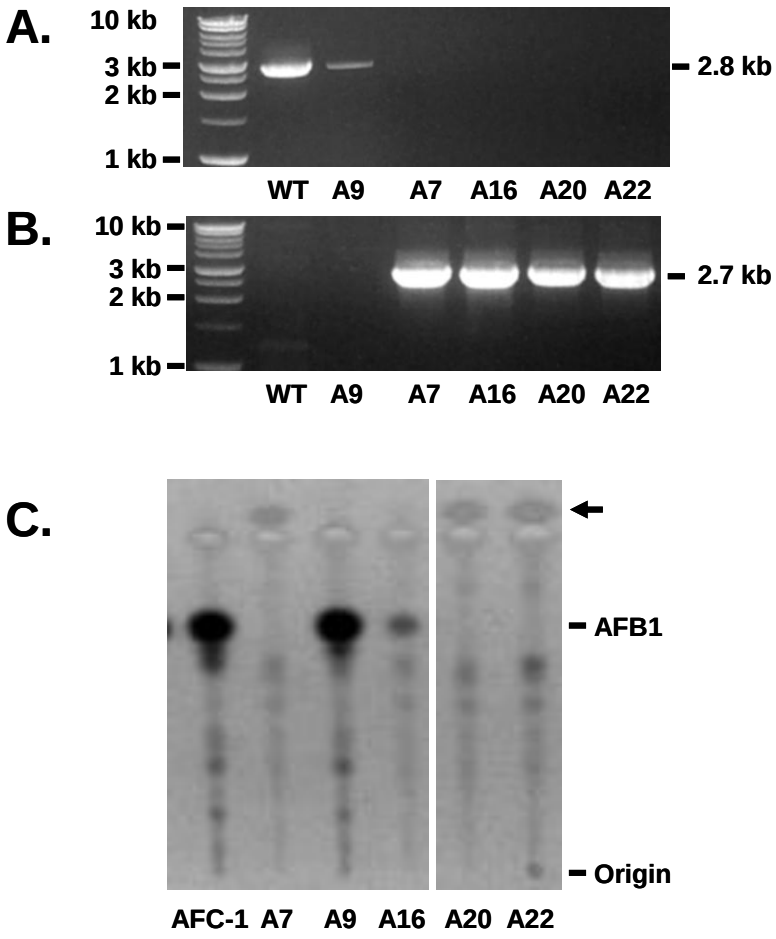
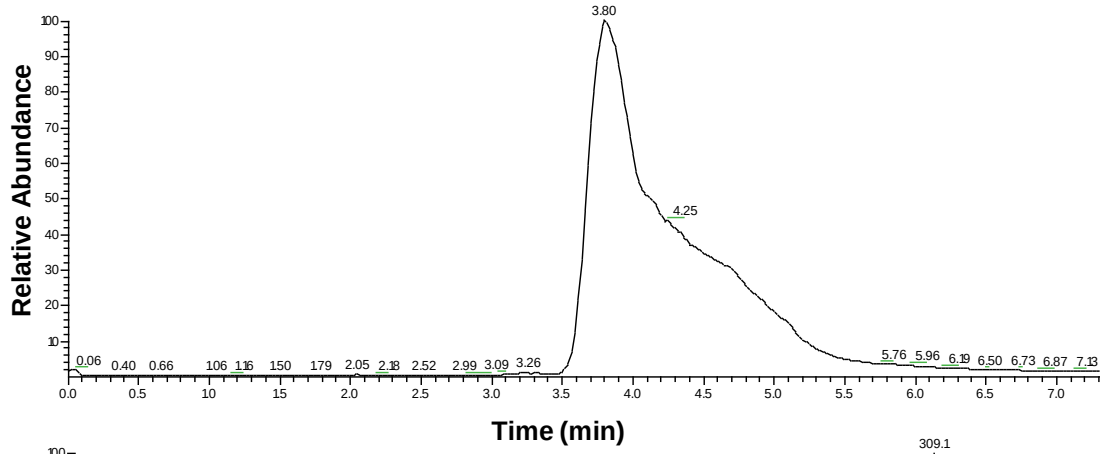


Figure 1. Gene deletion and PCR screening strategy. Initial overlap PCR template fragments were amplified using primer combinations p1+p3, p2+p5, and p4+p6 with either NRRL 3357 genomic DNA (p1+p3, p4+p6) or TOPO-argD plasmid (p2+p5). These products were then used as template with primers p1 and p6 alone to create the gene deletion construct. Transformants were screened initially for gene deletion with sp1+sp3 and sp1+sp4 (Fig. 2A, B) and later with sp2+sp3 (Fig. 4D).

Figure 2. PCR, TLC and LC-MS/MS analysis of *hypE* transformants. (A) PCR analysis of WT and five transformants that received the *hypE* gene deletion construct with screening primers sp1 and sp3. Among the transformants the 2.8 kb PCR fragment predicted for the WT gene was detected only in strain A9 and not in strains A7, A16, A20 or A22. (B) PCR analysis of WT *A. flavus* and five transformants with screening primers sp1 and sp4. The predicted 2.7 kb PCR fragment was detected in A7, A16, A20 and A22 but not in WT or A9. This showed that the WT gene was replaced with *argD*. (C) TLC chromatogram of filtrate from recipient strain and 4 putative deletion mutants. The intense spot for AFC-1 and A9 was confirmed to be aflatoxin by LC-MS/MS. All deletion strains except A16 accumulated an unknown, orange-red compound at the top of the TLC (denoted by arrow). A16 made a small amount of AFB1. The TLC image was inverted to improve visibility of the unknown compound. (D) LC-MS profile of the unknown compound from lane A20 in the TLC shows a compound with a retention time of 3.80 minutes that was determined to have an m/z of 337 and formula weight of 338. Retention time is shown above the peak. (E) MS/MS fragmentation of MW 338 compound. The fragmentation pattern is consistent with an assignment as versicolorin A (VERA).


A.
10 kb
3 kb
2 kb
1 kb
2.8 kb
WT A9 A7 A16 A20 A22
B.
10 kb
3 kb
2 kb
1 kb
2.7 kb
WT A9 A7 A16 A20 A22
C.
AFB1
Origin
AFC-1 A7 A9 A16 A20 A22

D.


Relative Abundance
Time (min)
3.80
4.25
0.06
0.40
0.66
1.06
1.16
1.50
1.79
2.05
2.18
2.52
2.99
3.09
3.26
5.76
5.96
6.19
6.50
6.73
6.87
7.13

E.

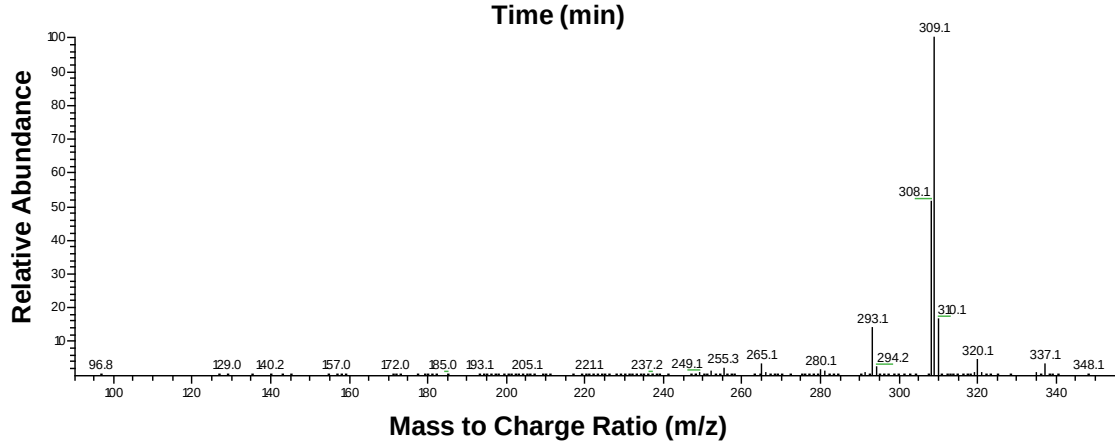
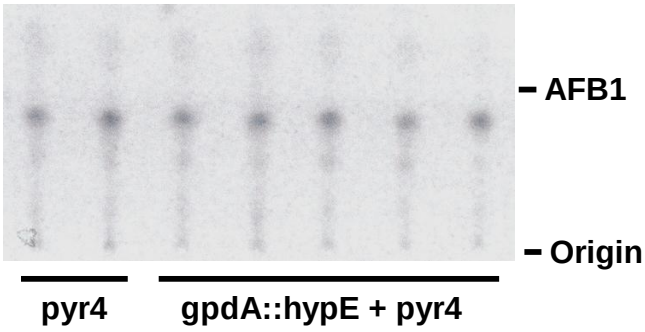
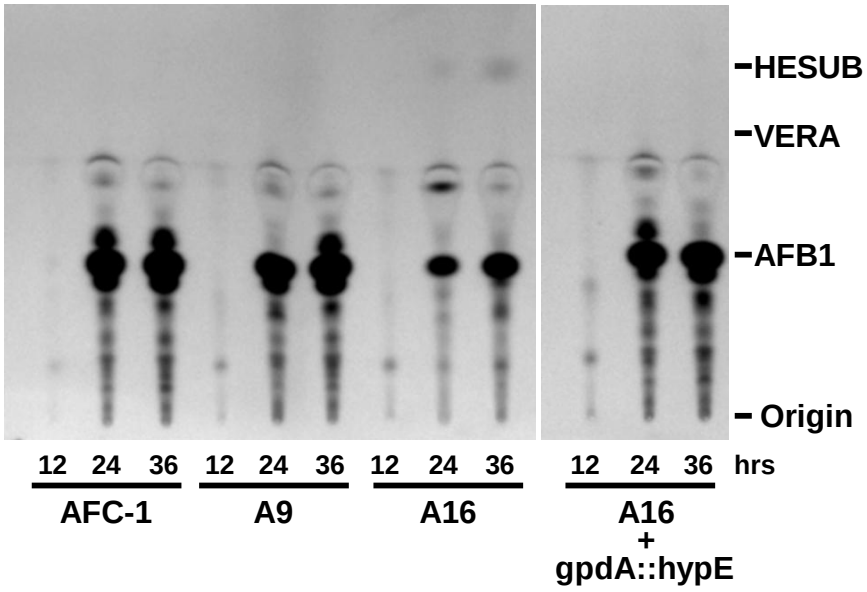

Relative Abundance
Mass to Charge Ratio (m/z)
309.1
308.1
310.1
293.1
294.2
320.1
337.1
348.1
96.8
129.0
140.2
157.0
172.0
185.0
193.1
205.1
221.1
237.2
249.1
255.3
265.1
280.1

Figure 3. TLC analysis of *hypE* mutants, recipient strain and strain 3357-5 transformed with the *gpdA::hypE* overexpression construct. (A) TLC image of filtrate from seven *hypE* A7 mutant transformants receiving the selectable marker (*pyr4*) alone (lanes 1-2) or with *gpdA::hypE* (lanes 3-7). (B) TLC image of metabolites extracted from media at 12, 24 and 36 hours of culture from AFC-1, non-disrupted strain A9, A16 mutant, and A16 mutant transformed with *gpdA::hypE*. HESUB accumulates in A16 at 24 and 36 hours. Predicted migration of VERA which did not accumulate is also marked. (C) TLC image of aflatoxin production by 3357-5 transformed with *pyr4* alone (lane 1) or with *gpdA::hypE* (lane 2). All TLC images were inverted.

A.


AFB1
Origin
pyr4
gpdA::hypE + pyr4

B.


HESUB
VERA
AFB1
Origin
12 24 36 12 24 36 12 24 36 12 24 36 hrs
AFC-1
A9
A16
A16 + gpdA::hypE

C.

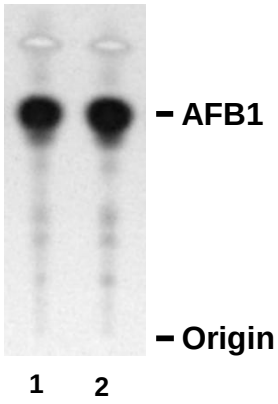
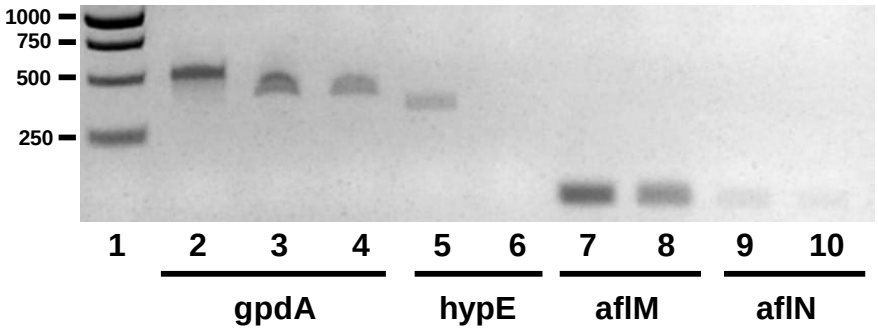
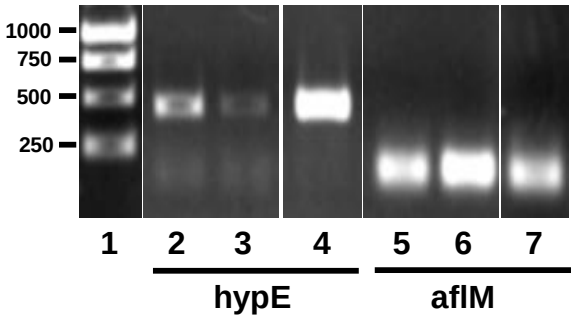

AFB1
Origin
1 2

Figure 4. RT-PCR analysis of gene expression. (A) Analysis of transcript accumulation of *gpdA*, *hypE*¸ *aflM*, and *aflN* in strain A9 (lanes 3, 5, 7 and 9) and VERA-accumulating strain A22 (lanes 4, 6, 8 and 10) after 36 hr culture. Genomic DNA control for *gpdA* is shown in lane 2. A PCR product was detected for each gene with the exception of *hypE* in A22 (lane 6). The gel image was inverted to improve visibility of faint *aflN* bands. (B) Analysis of transcripts of *hypE* and *aflM* in AFC-1 (lanes 2 and 5), A16 (lanes 3 and 6) and A16 + *gpdA::hypE* (lanes 4 and 7). Transcript for *hypE* was detected in A16. (C) PCR on RNA not treated with reverse transcriptase in AFC-1 (lane 2), A16 (lane 3) and A16 + *gpdA::hypE* (lane 4). Genomic DNA control is shown in lane 5. (D) PCR with internal *hypE* primers (sp2 and sp3) detected the presence of *hypE* in AFC-1 (lane 1), A9 (lane 2) and A16 (lane 3).

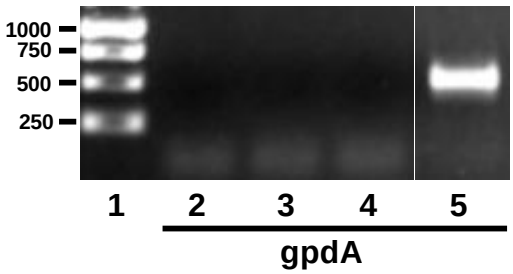
A.


1000
750
500
250
1 2 3 4 5 6 7 8 9 10
gpdA
hypE
aflM
aflN

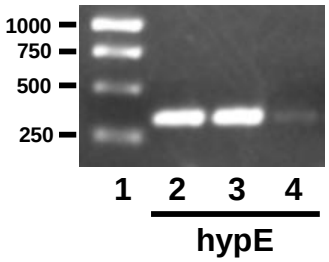
B.


1000
750
500
250
1 2 3 4 5 6 7
hypE
aflM

C.


1000
750
500
250
1 2 3 4 5
gpdA

D.


1000
750
500
250
1 2 3 4
hypE

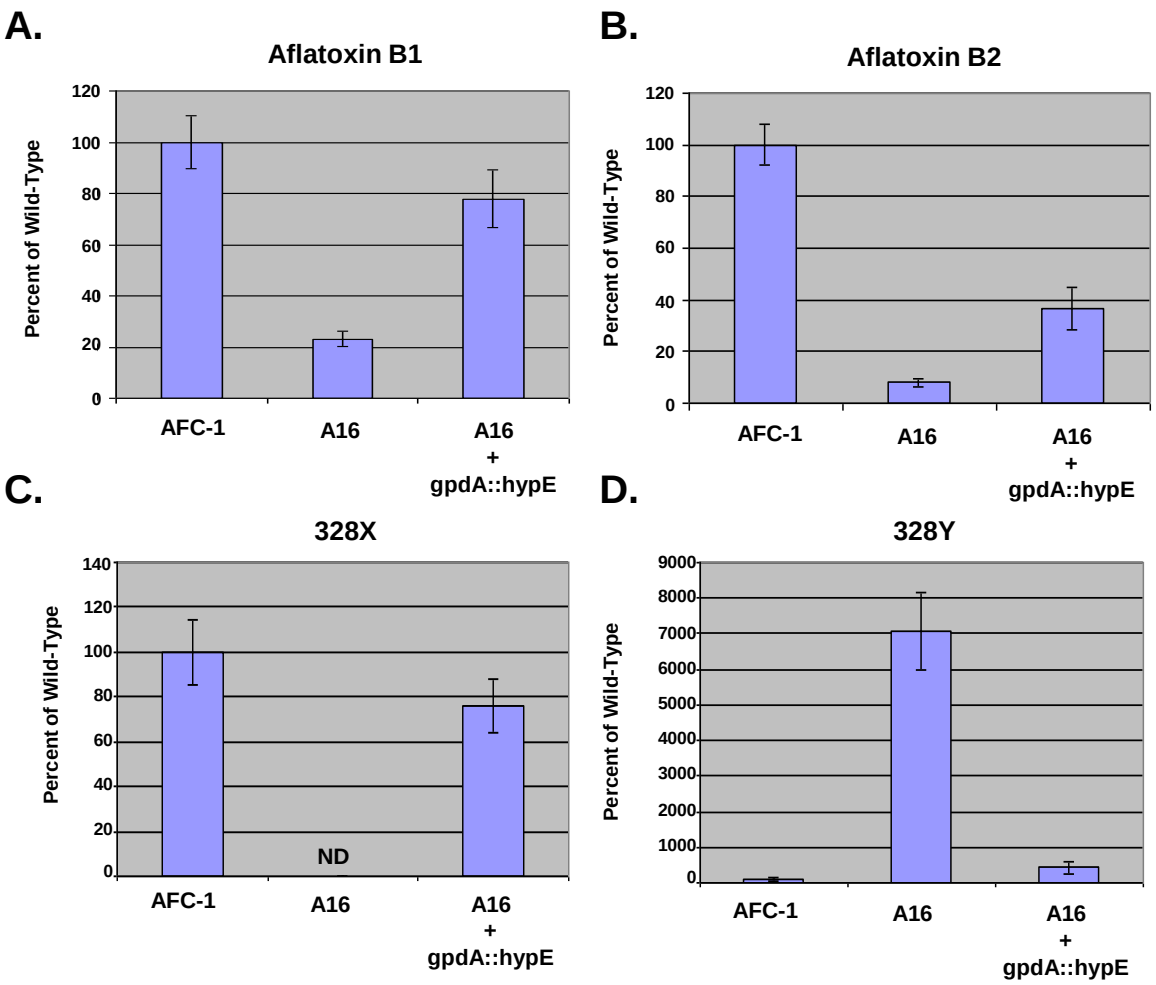


Figure 5. Quantification of AFB1, AFB2 and the unidentified compounds 328X and 328Y in AFC-1, A16 and A16 + *gpdA::hypE* by LC-MS. (A) AFB1 production in A16 and A16 + *gpdA::hypE* is 23% and 78% of AFC-1 control, respectively. (B) AFB2 production in A16 and A16 + *gpdA::hypE* is 8% and 36% of AFC-1 control, respectively. (C) 328X production in A16 and A16 + *gpdA::hypE* is 0% (not detected, ND) and 76% of AFC-1 control, respectively. (D) 328Y production in A16 and A16 + *gpdA::hypE* is 7071% and 425% of AFC-1 control, respectively. Relative abundance for each compound was obtained by integrating peak area in the reverse-phase C_{18} chromatograph.

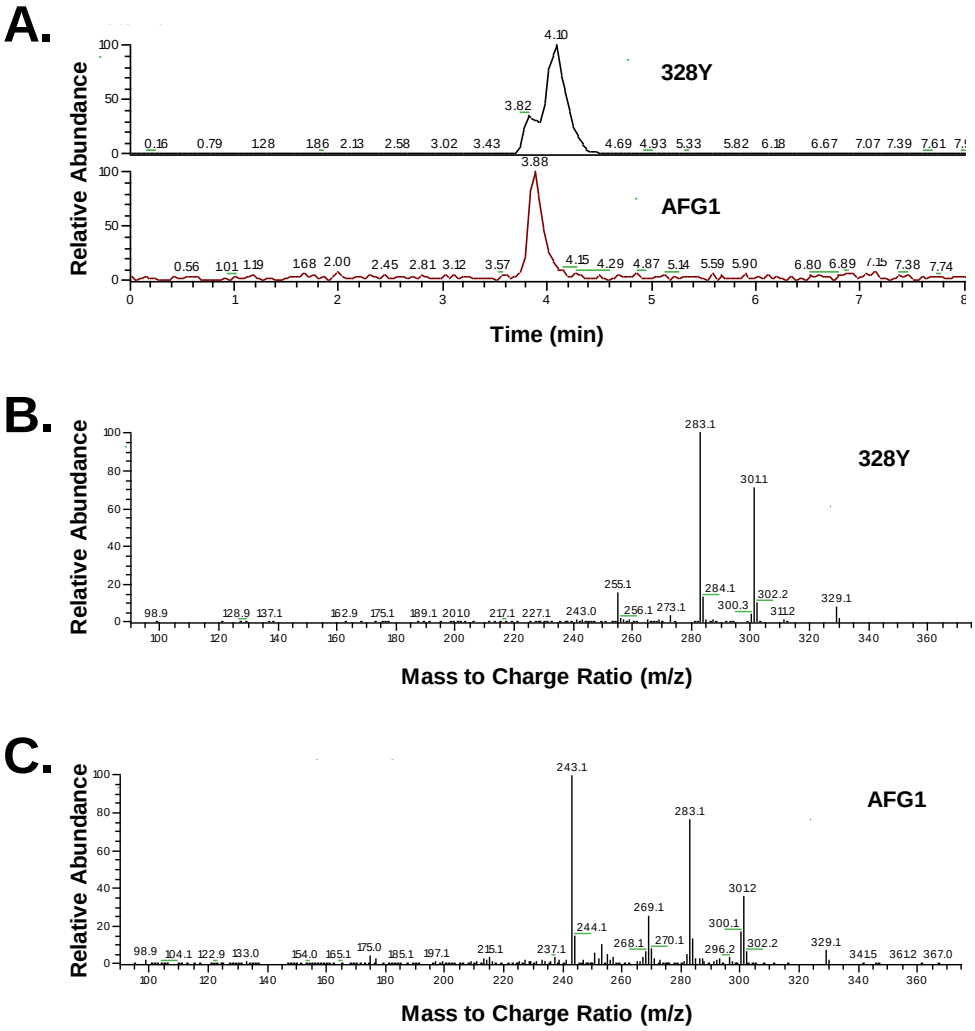


Figure 6. Comparative LC-MS/MS analysis of 328Y and AFG1. (A) Reverse-phase C_{18} profile of 328Y and AFG1 with scanning for ions with a mass to charge ratio (m/z) of 329. Retention times are shown above peaks. (B) MS/MS fragmentation pattern of 328Y. (C) MS/MS fragmentation pattern of AFG1. Daughter ions with m/z 283 and m/z 301 are shared.

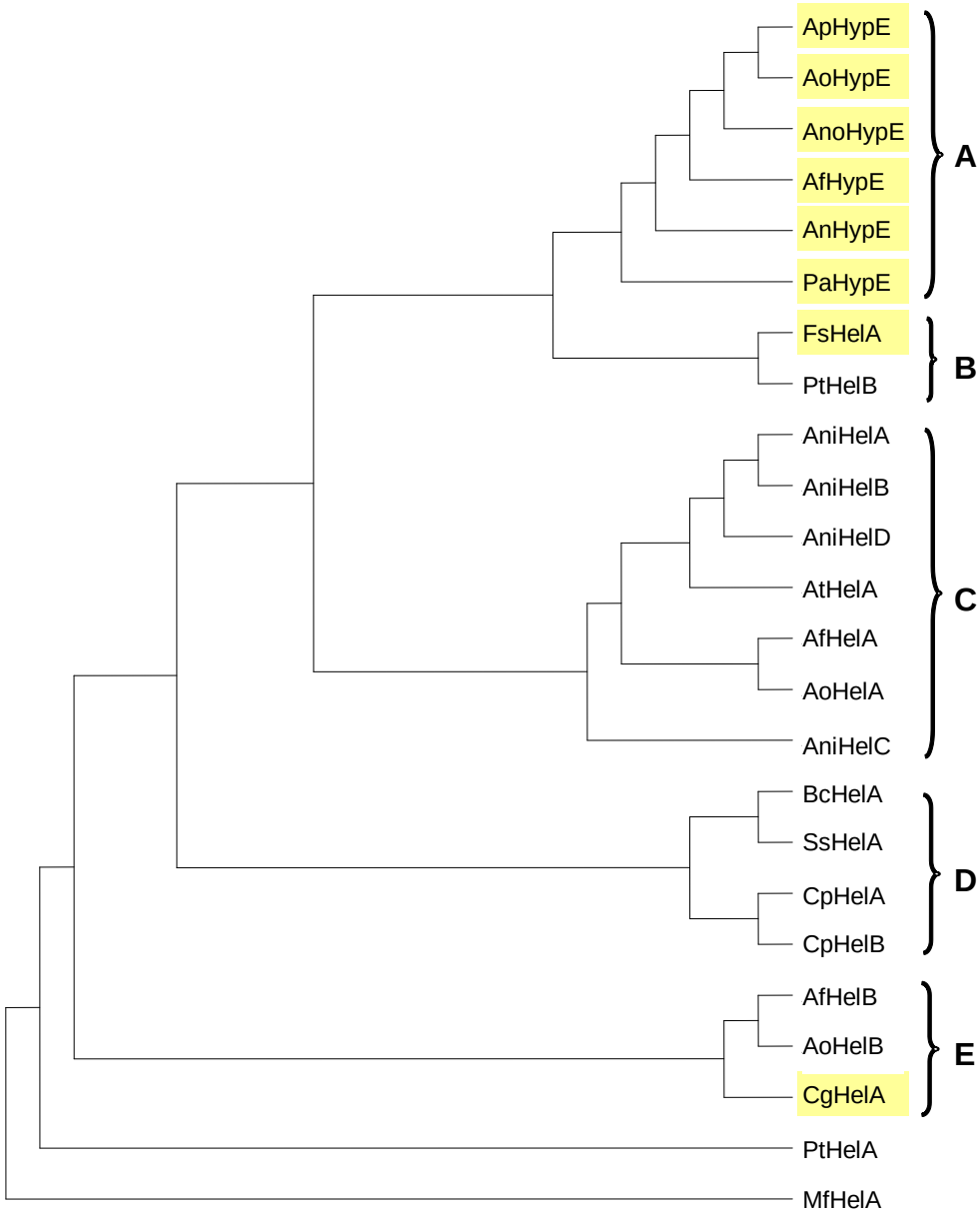


Figure 7**.** Maximum likelihood tree of *hypE-like* protein sequences. Only sequences with e-value ≤ 2e-10 based on BLASTP alignment with AfHypE were included. Clade A represents the *hypE* ortholgous clade which is sister to clade B. Clade C represents the *Aspergillus*-specific clade. Protein sequences containing an EthD domain are highlighted in yellow. The tree was rooted with MfHelA because it is from a member of the Dothideomycetes which are basal to the other classes represented in this tree. Protein designations listed in Table 1.

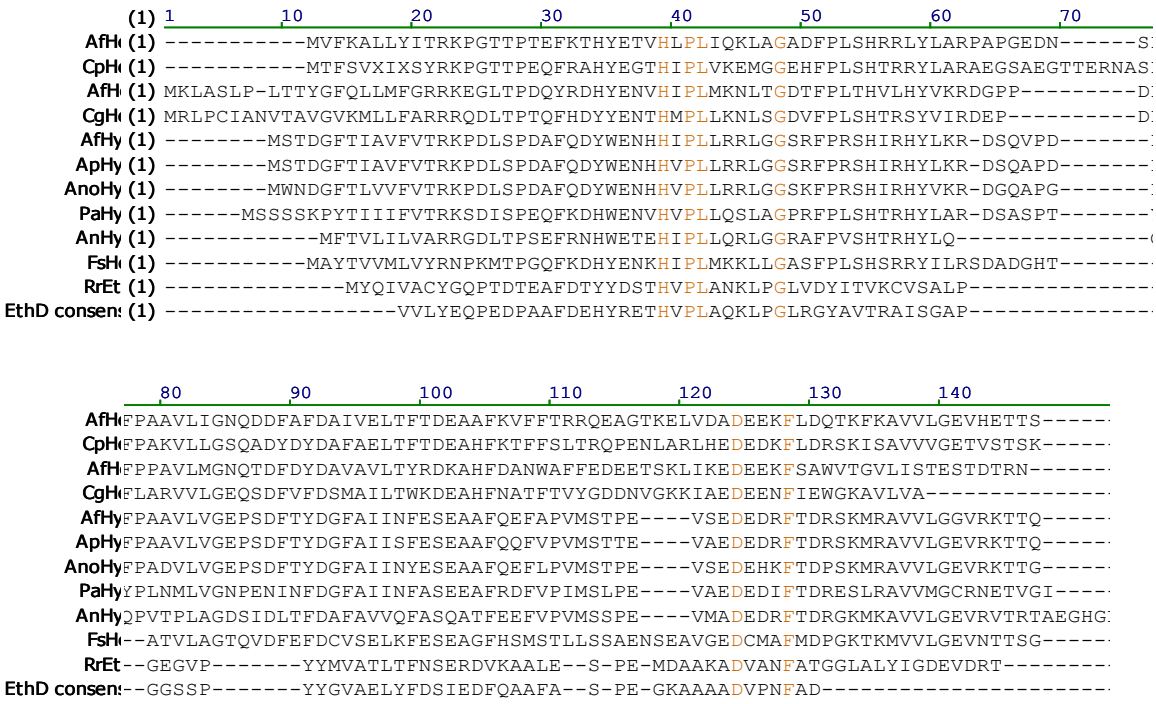


Figure 8. ClustalW alignment of *hypE*-like protein sequences (e-value ≤ 1e-16) with RrEthD and EthD domain consensus sequence (Pfam). Conserved residues are highlighted in red.

TABLE 1. Properties of *hypE* related sequences from the Pezizomycotina[a]

Gene	Species[b]	Predicted Protein Length	EthD Domain	Nucleotide Accession	mRNA/ EST support	Protein Accession	e-value
Eurotiomycetes							
AfhypE	*A. flavus*	127	+	AY510451[f]	+	2911.m00886 [g]	1e-68
AfhelA	*A. flavus*	130	-	afl1.assembly.2541[g]	-	2541.m00688 [g]	9e-20
AfhelB	*A. flavus*	136	-	afl1.assembly.2504[g]	-	2504.m00203 [g]	3e-19
AfhelC	*A. flavus*	167	-	afl1.assembly.2043[g]	+	2043.m00187 [g]	4e-07
AnhypE	*A. nidulans*	124	+	U34740[f]	-		5e-32
AnihelA	*A. niger*	141	-	NW_001594281[f]	-		7e-15
AnihelB	*A. niger*	137	-	XM_001394174	+	XP_001394211	2e-12
AnihelC	*A. niger*	574[d]	-	XM_001389796	+	XP_001389833	2e-15
AnihelD	*A. niger*	126		DR706229	+		3e-12
AnohypE	*A. nomius*	127	+	AY510454[f]	-		6e-59
AohypE	*A. oryzae*	127	+	AB071288[f]	-		4e-65
AohelA.1	*A. oryzae*	119	-	XM_001823675	+	XP_001823727	1e-15
AohelA.2	*A. oryzae*	130	-	AB224297	+		9e-20
AohelB	*A. oryzae*	142	-	XM_001817240	+	XP_001817292	3e-20
AohelC	*A. oryzae*	143	-	NW_001884679[f]	-		2e-07
AphypE	*A. parasiticus*	127	+	AY371490[f]	-	ABD63564[h]	4e-65
AthelA[c]	*A. terreus*	139	-	XM_001216758	+	XP_001216758	4e-14
Sordariomycetes							
CghelA	*Co. graminicola*	128[e]	+	FE740371	+		6e-17
CphelA	*Cr. parasitica*	136[e]	-	CB686521	+		3e-20
CphelB	*Cr. parasitica*	125[e]	-	ES360751	+		8e-14
FshelA	*F. solani*	127	+	sca_22_chr7_11_0	-	Fgenesh1_pg.sca_22_chr_11_0000067	2e-17
PahelA	*Po. anserina*	127[e]	+	CU868858	+		2e-36
Dothideomycetes							
MfhelA	*M. fijiensis*	129	-	FD672562	+	estExt_fgenesh1_pg.C_80135	1e-10
PthelA	*P. tritici-repentis*	143	-	XM_001935445	+	XP_001935480	2e-12
PthelB	*P. tritici-repentis*	162	-	XM_001938622	+	XP_001938657	2e-10
SnhelA	*St. nodorum*	164	-	XM_001798587	+	XP_001798639	1e-09
SnhelB	*St. nodorum*	124	-	XM_001803140	+	XP_001803192	2e-06
SnhelC	*St. nodorum*	136	-	XM_001803435	+	XP_001803487	6e-05
Leotiomycetes							
BchelA	*B. cinerea*	140	-	AL115116	+	XP_001555847	2e-14
SshelA	*Sc. sclerotiorum*	140	-	XM_001591963	+	XP_001592013	1e-13

[a] Only sequences with e-value <1e-04 included. [b] Genus abbreviations are as follows: A, *Aspergillus*; B, *Botrytis*; Cr, *Cryponectria*; Co, *Colletotrichum*; F, *Fusarium*; Po, *Podospera*; Py, *Pyrenophora*; Sc, *Sclerotinia*; St, *Stagonospora*. [c] Genbank description probably mis-annotated. [d] Large predicted protein with a distinct *hypE*-like domain. [e] Length based on EST data which may not contain complete predicted protein. [f] Predicted gene unannotated within larger nucleotide sequence. [g] Sequence available at *A. flavus* genome browser (http://gaplabg5.cfr.ncsu.edu/cgi-bin/gbrowse/final_assembly/). [h] Protein prediction different from our prediction.

REFERENCES

1. **Cai, J., H. Zeng, Y. Shima, H. Hatabayashi, H. Nakagawa, Y. Ito, Y. Adachi, H. Nakajima, and K. Yabe.** 2008. Involvement of the *nadA* gene in formation of G-group aflatoxins in *Aspergillus parasiticus*. Fungal Genet Biol **45:**1081-1093.

2. **Carbone, I., J. H. Ramirez-Prado, J. L. Jakobek, and B. W. Horn.** 2007. Gene duplication, modularity and adaptation in the evolution of the aflatoxin gene cluster. BMC Evol Biol **7:**111-122.

3. **Cary, J. W., K. C. Ehrlich, J. M. Bland, and B. G. Montalbano.** 2006. The aflatoxin biosynthesis cluster gene, *aflX*, encodes an oxidoreductase involved in conversion of versicolorin A to demethylsterigmatocystin. Appl Environ Microbiol **72:**1096-1101.

4. **Chang, P. K., B. W. Horn, and J. W. Dorner.** 2005. Sequence breakpoints in the aflatoxin biosynthesis gene cluster and flanking regions in nonaflatoxigenic *Aspergillus flavus* isolates. Fungal Genet Biol **42:**914-923.

5. **Cole, R. J., and R. H. Cox.** 1981. Sterigmatocystins, p. 67-93. *In* R.J. Cole and R. H. Cox (ed.), Handbook of toxic fungal metabolites. Academic Press, New York.

6. **Davidson, R. C., J. R. Blankenship, P. R. Kraus, M. de Jesus Berrios, C. M. Hull, C. D'Souza, P. Wang, and J. Heitman.** 2002. A PCR-based strategy to generate integrative targeting alleles with large regions of homology. Microbiology **148:**2607-2615.

7. **Ehrlich, K. C., B. Montalbano, S. M. Boue, and D. Bhatnagar.** 2005. An aflatoxin biosynthesis cluster gene encodes a novel oxidase required for conversion of versicolorin A to sterigmatocystin. Appl Environ Microbiol **71:**8963-8965.

8. **Fetzner, S.** 2002. Oxygenases without requirement for cofactors or metal ions. Appl Microbiol Biotechnol **60:**243-257.

9. **He, Z. M., M. S. Price, G. R. OBrian, D. R. Georgianna, and G. A. Payne.** 2007. Improved protocols for functional analysis in the pathogenic fungus *Aspergillus flavus*. BMC Microbiol **7:**104-114.

10. **Henry, K. M., and C. A. Townsend.** 2005. Ordering the reductive and cytochrome P450 oxidative steps in demethylsterigmatocystin formation yields general insights into the biosynthesis of aflatoxin and related fungal metabolites. J Am Chem Soc **127:**3724-3733.

11. **Huelsenbeck, J. P., and F. Ronquist.** 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17:**754-755.

12. **Keller, N. P., N. J. Kantz, and T. H. Adams.** 1994. *Aspergillus nidulans verA* is required for production of the mycotoxin sterigmatocystin. Appl Environ Microbiol **60:**1444-1450.

13. **Keller, N. P., S. Segner, D. Bhatnagar, and T. H. Adams.** 1995. *stcS*, a putative P-450 monooxygenase, is required for the conversion of versicolorin A to sterigmatocystin in *Aspergillus nidulans*. Appl Environ Microbiol **61:**3628-3632.

14. **Kroken, S., N. L. Glass, J. W. Taylor, O. C. Yoder, and B. G. Turgeon.** 2003. Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. Proc Natl Acad Sci U S A **100:**15670-15675.

15. **Lee, L. W., C. H. Chiou, and J. E. Linz.** 2002. Function of native OmtA in vivo and expression and distribution of this protein in colonies of *Aspergillus parasiticus*. Appl Environ Microbiol **68:**5718-5727.

16. **OBrian, G. R., D. R. Georgianna, J. R. Wilkinson, J. Yu, H. K. Abbas, D. Bhatnagar, T. E. Cleveland, W. Nierman, and G. A. Payne.** 2007. The effect of elevated temperature on gene transcription and aflatoxin biosynthesis. Mycologia **99:**232-239.

17. **Price, E. W., and I. Carbone.** 2005. SNAP: workbench management tool for evolutionary population genetic analysis. Bioinformatics **21:**402-404.

18. **Price, M. S., J. Yu, W. C. Nierman, H. S. Kim, B. Pritchard, C. A. Jacobus, D. Bhatnagar, T. E. Cleveland, and G. A. Payne.** 2006. The aflatoxin pathway regulator AflR induces gene transcription inside and outside of the aflatoxin biosynthetic cluster. FEMS Microbiol Lett **255:**275-279.

19. **Prieto, R., and C. P. Woloshuk.** 1997. ord1, an oxidoreductase gene responsible for conversion of O-methylsterigmatocystin to aflatoxin in *Aspergillus flavus*. Appl Environ Microbiol **63:**1661-1666.

20. **Roberts, I. N., R. P. Oliver, P. J. Punt, and C. A. van den Hondel.** 1989. Expression of the *Escherichia coli* beta-glucuronidase gene in industrial and phytopathogenic filamentous fungi. Curr Genet **15:**177-180.

21. **Skory, C. D., P. K. Chang, J. Cary, and J. E. Linz.** 1992. Isolation and characterization of a gene from *Aspergillus parasiticus* associated with the conversion

of versicolorin A to sterigmatocystin in aflatoxin biosynthesis. Appl Environ Microbiol **58:**3527-3537.

22. **Udwary, D. W., L. K. Casillas, and C. A. Townsend.** 2002. Synthesis of 11-hydroxyl O-methylsterigmatocystin and the role of a cytochrome P-450 in the final step of aflatoxin biosynthesis. J Am Chem Soc **124:**5294-5303.

23. **Watanabe, C. M. H., and C. A. Townsend.** 2002. Initial characterization of a type I fatty acid synthase and polyketide synthase multienzyme complex NorS in the biosynthesis of aflatoxin B1. Chem Biol **9:**981-988.

24. **Yabe, K., Y. Ando, J. Hashimoto, and T. Hamasaki.** 1989. Two distinct O-methyltransferases in aflatoxin biosynthesis. Appl Environ Microbiol **55:**2172-2177.

25. **Yabe, K., and H. Nakajima.** 2004. Enzyme reactions and genes in aflatoxin biosynthesis. Appl Microbiol Biotechnol **64:**745-755.

26. **Yang, Z.** 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Bioinformatics **13:**555-556.

27. **Yu, J., P. K. Chang, K. C. Ehrlich, J. W. Cary, D. Bhatnagar, T. E. Cleveland, G. A. Payne, J. E. Linz, C. P. Woloshuk, and J. W. Bennett.** 2004. Clustered pathway genes in aflatoxin biosynthesis. Appl Environ Microbiol **70:**1253-1262.

CHAPTER 4

Future Considerations

Each chapter of this dissertation explored the problem of aflatoxin production by the filamentous fungus *A. flavus* from a different perspective. In this section I summarize the contributions of each chapter and address unresolved questions that may be fruitful areas for future research.

In Chapter 1 I reported on known inhibitors of aflatoxin biosynthesis and addressed important similarities and differences among these compounds. I made particular care to focus on compounds that inhibit aflatoxin biosynthesis without an accompanying inhibition of growth. The majority of the inhibitors reviewed were plant natural products, many of which have reported antioxidant activity. As the biosynthetic pathways and tools for manipulating these natural product pathways in plants improve, the possibility of engineering crop plants with resistance to *A. flavus* and *A. parasiticus* infection becomes increasingly feasible. In the near term, these diverse compounds present exciting opportunities to probe the regulatory network controlling aflatoxin biosynthesis and other secondary metabolic pathways in the Aspergilli. The availability of microarrays, proteomic and metabolomic tools in *A. flavus* can facilitate this research. For example Kim et al. (1) have already used microarrays to investigate the effect of the antioxidant compound caffeic acid on global gene expression and aflatoxin biosynthetic pathway genes. Integration of this study with future studies that also investigate changes in metabolites and proteins will allow the identification of major pathways and networks that could then be used in hypothesis-driven functional genomics investigations of secondary metabolism.

Chapter 2 reported on an activity present in maize kernels that inhibits fungal growth and aflatoxin production. I showed that this activity likely resulted from the cooperative action of two known classes of antifungal proteins: chitinases and thaumatin-like proteins (TLPs), in this case zeamatin. The active chitinase fraction was highly enriched for the major seed chitinase, ChitA, which belongs to the glycoside hydrolase 19 (GH19) family of chitinases. I also reported the description of the maize GH19 and TLP families and presented a preliminary phylogeny of each family. Finally, in collaboration with Andrea Dolezal, I

reported on the transcriptional response of members of each gene family to *A. flavus* infection in developing maize kernels.

This research raised several unresolved questions. First, it is unclear if ChitA is the only chitinase responsible for the cooperative activity observed with the zeamatin fraction. This question can be addressed with two approaches. Production of recombinant ChitA protein would allow us to measure the effect of a pure ChitA preparation alone and in combination with the zeamatin fraction. The second approach would be to perform additional chromatographic separations of ChitA to assess the purity of the chitin-binding fraction. UnoS-1 or MonoS cation exchange and hydroxyapatite chromatography may be sufficient to resolve multiple chitinases present in the chitin binding fraction. Because the maize genome encodes multiple predicted chitinases with similar size and theoretical pI to ChitA it would be beneficial to follow additional purification steps with 2D gel electrophoresis in order to assess purity.

A second unresolved question is whether the chitinase activity of ChitA is unique in interacting with the zeamatin fraction to inhibit *A. flavus* growth and aflatoxin production, or if other chitinases can produce the same effect when combined with the zeamatin fraction. This can be addressed by assaying chitinase enzyme preparations from other organisms alone and in combination with the zeamatin fraction. The corollary question, whether zeamatin is unique or if other TLPs can interact with the chitin-binding fraction, is more difficult to address. TLPs are difficult to produce in heterologous systems. Our attempts to produce soluble zeamatin, TLP2, and TLP4 in *E. coli* were unsuccessful and this has been a persistent problem for production of the model tobacco TLP, osmotin, in *E. coli* and eukaryotic expression systems (M. Narsimhan, personal communication). To address the function of less abundant TLPs and GH19s it may be necessary to co-express members of these families in transgenic plants. This approach is not without its limitations either as endogenous chitinases an TLPs would influence resistance phenotypes.

The evolutionary history of these two gene families in plants has not been described sufficiently. While we presented a preliminary phylogeny on the GH19 and TLP family in

rice, maize and Arabidopsis, it would be instructive to investigate these gene families further and compare their evolutionary histories with those of other gene families involved in plant defense. A particularly intriguing question is the evolution of GH19 genes that do not encode chitin-binding domains. Why these gene lineages have been maintained and what the function of these genes may be raises interesting evolutionary questions about the partitioning of gene functions following gene duplication events.

In Chapter 3 I reported on the characterization of a new aflatoxin biosynthetic cluster gene, *hypE*, involved in production of B aflatoxins. Complementation of a leaky *hypE* mutant and resulting restoration of aflatoxin accumulation to levels near that seen in a wild-type strain showed that *hypE* is important for aflatoxin biosynthesis. We also identified potential aflatoxin precursor compounds that may provide additional insight into the function of the HypE protein. Finally, we described a fungal gene family, the *hypE*-like family, restricted to the Pezizomycotina that contains orthologs of *A. flavus hypE* in other aflatoxin or sterigmatocystin producing Aspergilli and an ortholog in a previously undescribed sterigmatocystin cluster in the Sordariomycete *Podospora anserina*.

Because the *hypE* mutant had leaky expression of the *hypE* transcript it was not possible to conclude if *hypE* is required for aflatoxin biosynthesis. Identification of a *hypE* mutant with no expression of the *hypE* transcript would address this question. In addition to the leaky strain (A16) I have identified two other low aflatoxin accumulating *hypE* gene deletion mutants that accumulate HESUB (strains B8 and B10, data not shown.) Because strains B8 and B10 share the A16 phenotype it is likely that they are also leaky *hypE* mutants. Another *hypE* gene deletion mutant, strain B1, accumulates neither aflatoxin nor VERA and may be a non-leaky *hypE* mutant. Transformation of strain B1 with the *gpdA::hypE* construct would confirm if this phenotype is a result of deletion of *hypE*. Also, RT-PCR (Ch. 3, Fig. 4A) indicated that a VERA accumulating strain (A22) did not accumulate *hypE* transcript. If the VERA accumulation phenotype is due to loss of function in either *aflM* or *aflN*, then it may be possible to complement A22 with *aflN* or *aflM* and generate a non-leaky *hypE* mutant.

Determining the identity of the unknown compounds HESUB and 328Y is critical for understanding the function of HypE. These may represent the same compound or could be different compounds. In either case, they may represent novel aflatoxin precursor compounds that will expand our understanding of the already complex aflatoxin biosynthetic pathway. Identification of these compounds using NMR should be straightforward because HESUB and 328Y are secreted in the growth medium and amenable to purification by reverse phase chromatography. HESUB could also be isolated directly following TLC separation which may make reverse-phase chromatography unnecessary.

The observation that polymorphisms within the *hypE* region correlate with the B/G aflatoxin ratio among populations of *A. parasiticus* suggests that it is important to investigate *hypE* in this organism as well. Because the aflatoxin gene cluster is highly conserved between *A. flavus* and *A. parasiticus* it should be possible to use our existing primers to generate *hypE* gene deletion constructs for *A. parasiticus.* Also, the existing *gpdA::hypE* overexpression construct should function in *A. parasiticus*.

The identification of a sterigmatocystin cluster in *P. anserina* presents exciting possibilities. Production of sterigmatocystin in *P. anserina* has not been reported. The first step in a study on the sterigmatocystin pathway in *P. anserina* would be to verify if sterigmatocystin is produced. Subsequently, it would be interesting to investigate if the sterigmatocystin pathway genes in *P. anserina* are more closely related to ST genes in the Sordariomycete *Chaetomium globosum* or those in *A. nidulans* (a member of the Eurotiomycetes).

REFERENCES

1. **Kim, J.H., J. Yu , N. Mahoney, K.L. Chan, R.J. Molyneux, J. Varga, D. Bhatnagar, T.E. Cleveland, and W.C. Nierman.** 2008. Elucidation of the functional genomics of antioxidant-based inhibition of aflatoxin biosynthesis. Int J Food Microbiol **122**:49-60.

APPENDIX A

Supplemental Data for Chapter 2

This appendix contains supplemental data on the initial purification of ChitA and zeamatin from Tex6 kernels as described in the Materials and Methods section of Chapter 2. We identified an inhibitory activity from Tex6 kernels that was present in a cation exchange binding fraction (Chapter 2, Fig. 1). We performed a linear gradient elution following application of 40 mL Tex6 crude kernel extract to a High S cation exchange column (30 mL 0-0.3 M NaCl in 10 mM sodium phosphate, pH 6.8; Fig. 1A). TLC analysis of aflatoxin produced by *A. flavus* cultured in the presence of protein fractions showed association of the inhibitory activity with the first major peak eluted in the linear gradient (Fig. 1B). Interestingly, the fractions with highest activity were not centered on the peak, but appeared to follow the trailing edge of the peak. To purify the activity further a second linear gradient elution following application of Tex6 crude kernel extract to a High S column was repeated as and fractions corresponding to the first peak and the trailing edge were pooled (Fig. 2A). (For specific purification protocols for chromatograms in Fig. 2 see Chapter 2: Materials and Methods.) This pool was then applied to a t-butyl hydrophobic interaction column and eluted with a linear gradient (Fig. 2B). Because the bound material eluted as a nearly symmetrical single peak this material was pooled and used as starting material for the next cation exchange chromatography step on a UnoS-1 column (Fig. 2C). As in the initial High S separation, shown in Fig. 1A, the inhibitory activity associated with the trailing edge of the major peak. The material from this final separation was used to for the bioassay, SDS-polyacrylamide gel and LC-MS identification of ChitA and zeamatin shown in Chapter 2 (Chapter 2, Fig. 2).

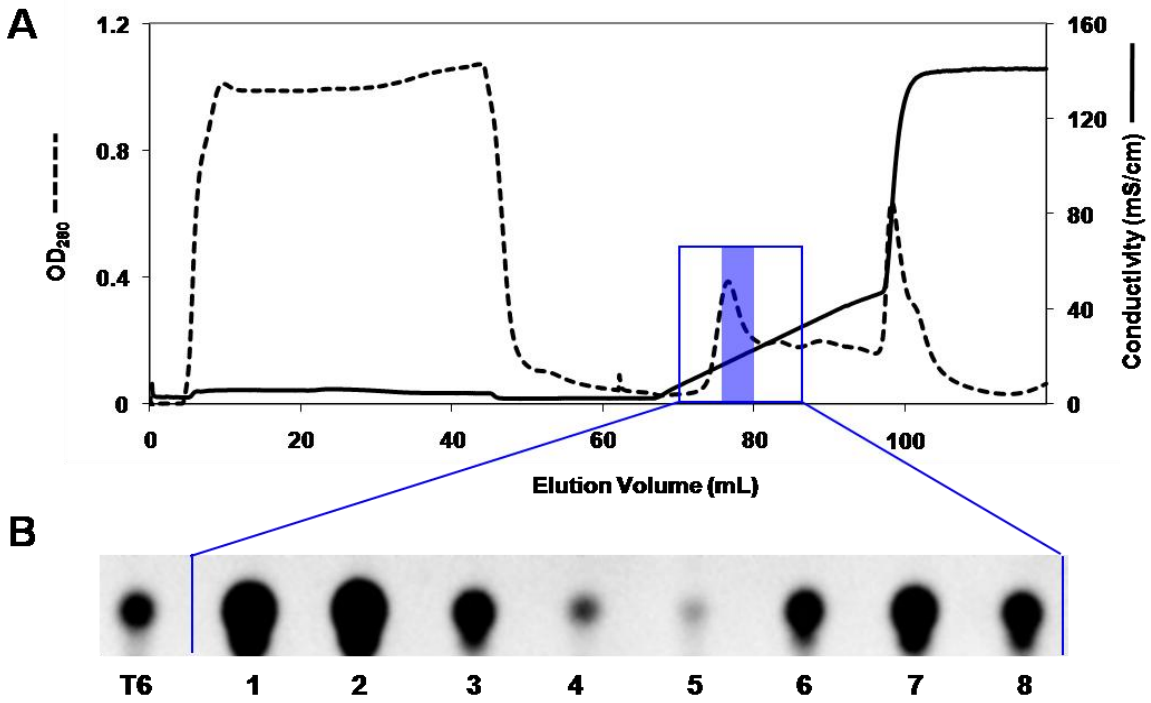
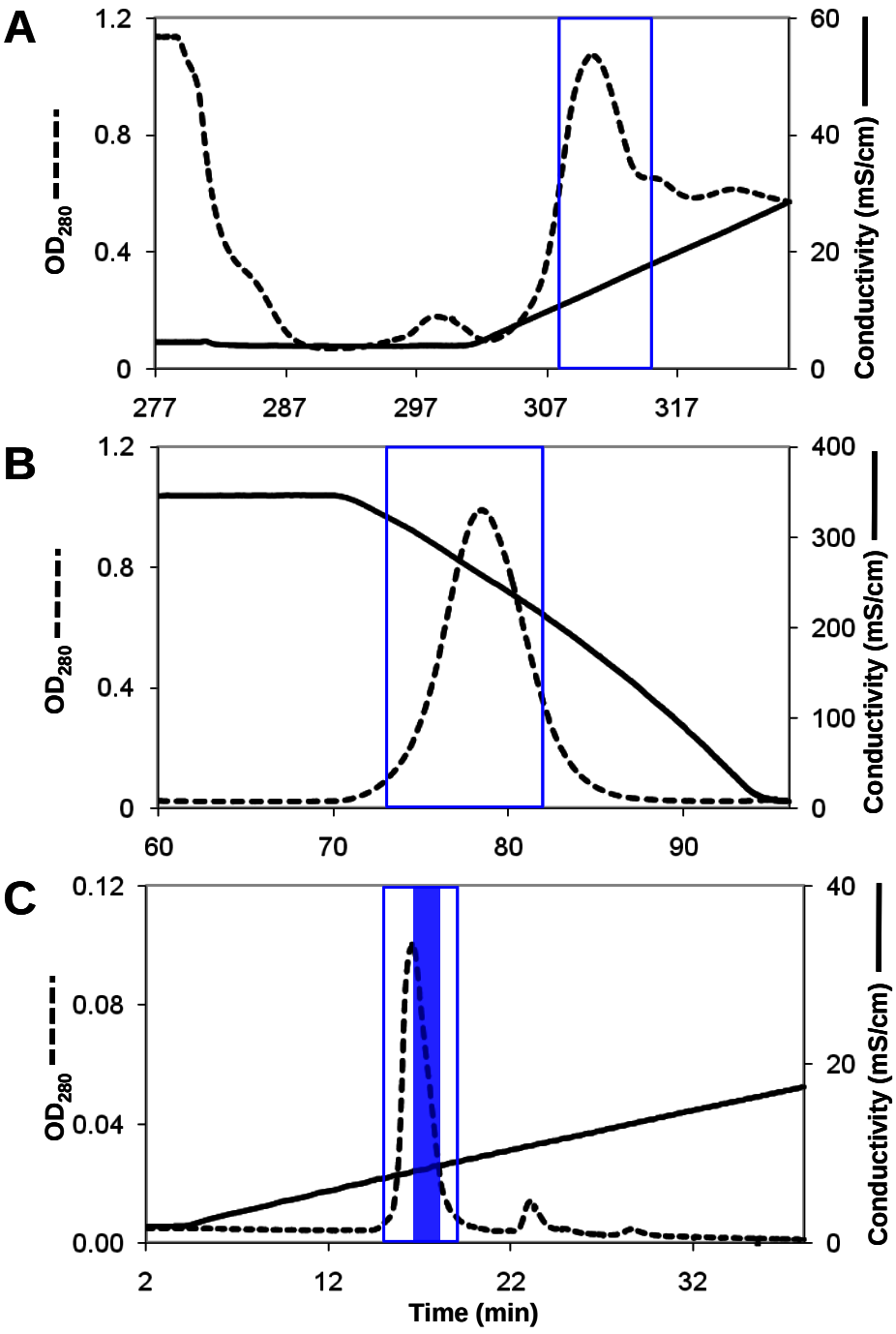


Figure 1. Initial linear elution of cation exchange binding activity from High S column. (A) Chromatograph of protein separation. The region that was assayed for inhibitory activity is represented by the larger box with the light blue line. The region corresponding to the two fractions with the highest inhibitory activity is shaded in light blue. (B) TLC plate of aflatoxin taken from cultures assayed in the presence of Tex6 crude extract (T6) and fractions from the High S column (lanes 1-8).

Figure 2. Fractionation of cation exchange binding activity using High S cation exchange, t-butyl hydrophobic interaction, and UnoS-1 cation exchange columns. (A) High S cation exchange chromatograph showing linear gradient elution following application of Tex6 crude kernel extract. The region boxed in with a light blue line represents fractions pooled for next chromatography step. (B) Tert-butyl hydrophobic interaction chromatograph showing linear gradient elution following application of the High S pool. The region boxed in with a light blue line represents fractions pooled for next chromatography step. (C) UnoS-1 cation exchange chromatograph showing linear gradient elution following application of the tert-butyl hydrophobic interaction pool. The region boxed in with a light blue line represents fractions assayed as shown in Chapter 2 (Chapter 2, Fig. 2B). The region corresponding to the three fractions with the highest inhibitory activity is shaded in light blue.


A
1.2
0.8
0.4
0
OD280
60
40
20
0
Conductivity (mS/cm)
277
287
297
307
317
B
1.2
0.8
0.4
0
OD280
400
300
200
100
0
Conductivity (mS/cm)
60
70
80
90
C
0.12
0.08
0.04
0.00
OD280
40
20
0
Conductivity (mS/cm)
2
12
22
32
Time (min)

APPENDIX B

Characterization of additional inhibitory activities from Tex6 maize kernels

We observed that the crude extract from Tex6 kernels contained more than one inhibitory activity against aflatoxin production by *A. flavus* (Chapter 2, Fig. 1). To investigate the activity that did not interact with the High S cation exchange column we performed anion exchange chromatography on the Tex6 crude extract and assayed for inhibition of aflatoxin production. Tex6 crude extract was prepared as described in Chapter 2 and subjected to boiling for 10 minutes. After boiling the extract was subjected to centrifugation for 10 minutes at 14,500 x g. The supernatant fraction was partitioned against an equal volume of ethyl acetate and the aqueous phase was then digested with 50 µg/mL Proteinase K at 37ºC for one hour. Following digestion with Proteinase K the sample was dialyzed against 10 mM sodium phosphate, pH 6.8 overnight at 4ºC. The sample was then applied to a High Q anion exchange column (Bio-Rad, Hercules, CA) and bound material was eluted with two step gradients (0.2 M NaCl and 0.55 M NaCl in 10 mM sodium phosphate, pH 6.8; Fig. 1A). Fractions associated with the peak from the second step gradient were dialyzed and then assayed for inhibition of aflatoxin production (Fig. 1A, inset). Three fractions associated with the peak were inhibitory to aflatoxin production. Coomassie staining of an SDS-polyacrylamide gel loaded with the active fractions showed that most protein had been removed from the active fraction (Fig. 1B). A pool of the active fractions (fractions 2-4) and two side fractions without strong activity (fractions 1 and 5) were compared by LC-MS at the Genomic Sciences Laboratory (NCSU). A heatplot of compounds detected by LC-MS showed that the active pool was highly enriched in compounds with mass/charge ratios (m/z) between 600 and 700 (Fig. 1C). The LC-MS spectrum for compounds with m/z between 600 and 700 showed that most abundant compound among those enriched in the active fraction had an m/z of 677.6 (Fig. 1C). A compound with an m/z of 660.9 was also detected. (Fig. 1C, sideplot). These two compounds correspond to phytic acid (inositol hexakisphosphate or IP_6) and an ammonium-phytic acid adduct (an artifact of the ionization process) based on confirmation with a phytic acid standard. Additional LC-MS analysis of the pooled active fraction showed that it also contained phytic acid precursor compounds (IP_2, IP_3, IP_4, and IP_5 ;data not shown).

To investigate if phytic acid or its precursor compounds were important for the inhibitory activity we compared the inhibitory activity of extracts taken from normal (+) and *low phytic acid 1-1* mutant (*lpa1-1*) B73 maize kernels. The *lpa1-1* mutation is within a gene encoding an ABC transporter (2). Mutant seeds accumulate reduced phytic acid and other inositol polyphosphate precursors. Because phytic acid accumulates primarily in the embryo we enriched for embryo by drilling the embryos out of mature kernels and extracting the powder in 1:8 (w/v) in 20 mM Bis-Tris, pH 6.0. The resulting extracts were then compared for ability to inhibit aflatoxin production (Fig. 2A). Embryo-enriched extracts from normal kernels showed greater inhibition than extracts from *lpa1-1* kernels. To test if a phytic acid precursor may be responsible for the inhibitory activity we used phytase enzyme from *Aspergillus ficuum* (Sigma, St. Louis, MO) to treat crude extract from Tex6 and an enriched fraction (obtained by step elution of Tex6 crude extract from a High Q column). In comparison to untreated controls, the phytase treated fractions showed enhanced inhibitory activity against aflatoxin production (Fig. 2B). However, phytase treatment of pure phytic acid did not yield inhibitory activity (Fig. 2C). Also, pure inositol polyphosphate isomers did not have inhibitory activity against aflatoxin production (Fig. 2D).

Additional chromatographic separation of the activity that interacted with the High Q column showed that the activity was likely not due to an inositol polyphosphate. Tex6 kernels were ground in a coffee mill and extracted in 1:5 (w/v) Bis-Tris, pH 6.0 by magnetic stirring at 4ºC for 2-4 hours. The resulting extract was filtered through cheesecloth and subjected to centrifugation at 4ºC for 20 minutes at 14,500 x g. The resulting supernatant was filtered through two layers of Miracloth (Calbiochem, La Jolla, CA). Extracts were stored at -20º C until use. Prior to chromatography extracts were again subjected to centrifugation (14,500 x g, 20', 4º C) and the supernatant was vacuum-filtered through 0.45 μm filters (Pall, Ann Arbor, MI) to yield the crude extract fraction. Thirty-five mL of Tex6 crude extract were applied to a High Q anion exchange column and eluted with an 80 mL linear gradient (0-0.8 M NaCl in 25 mM Bis-Tris, pH 6.0; Fig. 3A). Fractions were assayed for ability to inhibit aflatoxin production and three fractions with inhibitory activity were identified that

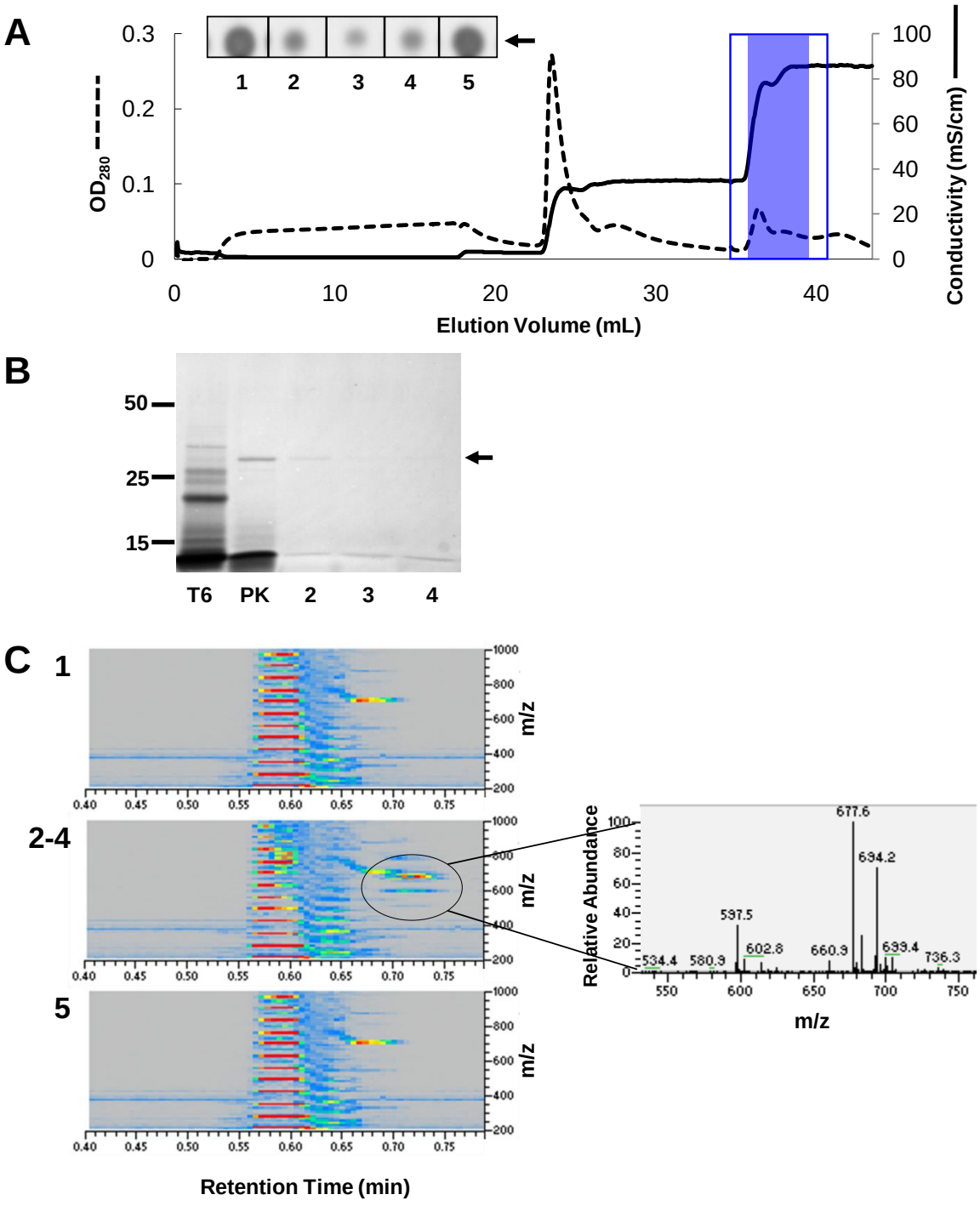
eluted at the front edge of the second major peak in the gradient (Fig 3B). This activity was assigned as the anion exchange binding activity.

We pursued additional purification of the anion exchange binding activity. The same purification protocol was repeated as for Fig. 3 (Fig. 4A). Activity was scored visually and 6 mLs of the pooled active fractions from the High Q column were diluted to a final volume of 50 mL with 25 mM Bis-Tris, pH 6.0. Forty mL of this diluted pool was applied to a DEAE Protein Pak column (Waters, Milford, MA) that had been equilibrated in 25 mM Bis-Tris, pH 6.0 + 30 mM NaCl. Bound material was eluted with two linear NaCl gradients (0.03-0.4 M NaCl, 20 mL and 0.4-1.0 M NaCl, 8 mL) in 25 mM Bis-Tris, pH 6.0 and assayed for inhibitory activity (Fig. 4B, C). Inhibitory activity was observed in fractions 6-8. Fractions were also assayed for the ability to bind iron using the assay described by Phillippy and Bland (Fig. 4B; 1). We assayed for iron binding because inositol polyphosphates chelate iron and are detectable with this assay. Although we detected a peak, presumably due to the presence of a phytic acid precursor, this peak was not associated with the inhibitory activity. We also measured dry weight of fungal cultures (Fig. 2D). Similar to the cation exchange binding activity described in Chapter 2, the anion exchange binding activity was initially detected by screening for reduced aflatoxin production but also inhibited growth when concentrated.

In summary, it is likely that the inhibitory anion exchange-binding activity is due to a compound(s) related to the phytic acid biosynthetic pathway. However, the active compound may not be a direct precursor to phytic acid since individual inositol polyphosphate isomers did not show inhibitory activity and the active fractions from the DEAE Protein Pak column did not associate with the major iron binding peak that is presumably a phytic acid precursor. Because the *lpa1* mutation is in an ABC transporter other pathways besides the inositol polyphosphate pathway may be altered in mutant seeds. Also, enhanced inhibition following treatment with the phytase enzyme may not necessarily be due to phosphatase activity against phytic acid since phytase enzymes may be active against other substrates (B. Phillippy, personal communication). Additional chromatographic purification in combination with metabolic profiling and bioassays comparing other maize seeds with altered phytic acid

content (e.g. transgenic lines overexpressing phytase enzymes or the *lpa2* and *lpa3* mutants) may provide additional clues to the identity of the inhibitory compound(s).

Figure 1. Anion exchange fractionation and LC-MS characterization of activity inhibitory to aflatoxin biosynthesis. (A) Chromatograph showing fractionation of Tex6 kernel extract on a High Q anion exchange column following boiling, organic partitioning and dialysis. The boxed region corresponds to fractions 1-5 and the region corresponding to fractions 2-4 is shaded in blue. Inset: TLC image showing aflatoxin produced in the presence of fractions taken from boxed region. Arrow, AFB1. (B) 12% SDS-polyacrylamide gel of boiled Tex6 crude extract (T6), proteinase K treated extract (PK) and active fractions 2-4. PK was used as the onput for the chromatograph in panel A. Arrow, proteinase K. (C) Heatplot showing LC-MS analysis of the pooled active fractions (2-4) and side fractions 1 and 5. The X and Y axes indicate retention time and mass to charge ration (m/z), respectively. Color indicates relative abundance (red>orange>yellow>green>blue). Compounds enriched in the active fraction are circled. Side plot: mass spectrum of major peak enriched in the active fraction. The most abundant compound is $IP_6NH_3^+$ (m/z 677.6) which is an ammonium adduct of IP_6 and forms as an artifact of the ionization process. IP_6 was also detected (m/z 660.9).


A
0.3
0.2
0.1
0
OD280
1 2 3 4 5
100
80
60
40
20
0
Conductivity (mS/cm)
0 10 20 30 40
Elution Volume (mL)
B
50
25
15
T6 PK 2 3 4
C
1
2-4
5
m/z
0.40 0.45 0.50 0.55 0.60 0.65 0.70 0.75
Retention Time (min)
Relative Abundance
677.6
694.2
597.5
602.8
660.9
693.4
736.3
534.4
580.9
550 600 650 700 750
m/z

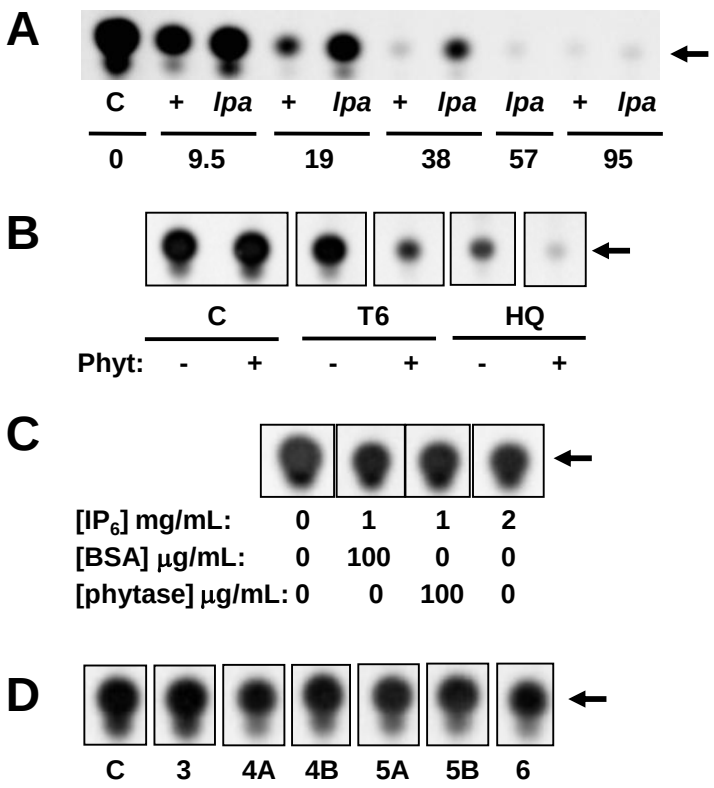


Figure 2. Effect of *lpa1-1* embryo-enriched extracts, phytase treatment and inositol polyphosphates on aflatoxin production. (A) TLC image showing aflatoxin accumulation in the presence of buffer control (C) and increasing concentrations of embryo-enriched extracts from normal (+) and *lpa1-1* (*lpa*) kernels of maize line B73. Numbers below treatment dry indicate dry weight equivalents of starting material in mg/mL. (B) TLC image showing aflatoxin accumulation in the presence of samples treated with *A. ficuum* phytase (+) or BSA (-) at 100 µg/mL. C, buffer control; T6, boiled Tex6 crude extract; HQ, anion exchange fraction. (C) TLC image showing aflatoxin production in the presence of phytic acid (IP_6) treated with *A. ficuum* phytase, an equivalent concentration of BSA, or alone at a higher concentration (2 mg/mL). (D) TLC image showing aflatoxin accumulation in the presence of various inositol polyphosphate isomers at 1 mg/mL. C, buffer control; 3, mixture of IP_3 isomers; 4A, $I(1,2,3,6)P_4$; 4B, $I(1,2,5,6)P_4$; 5A, $I(1,2,3,4,5)P_5$; 5B, $I(1,2,4,5,6)P_5$; 6, IP_6. All TLC images were inverted. Arrows indicate AFB1.

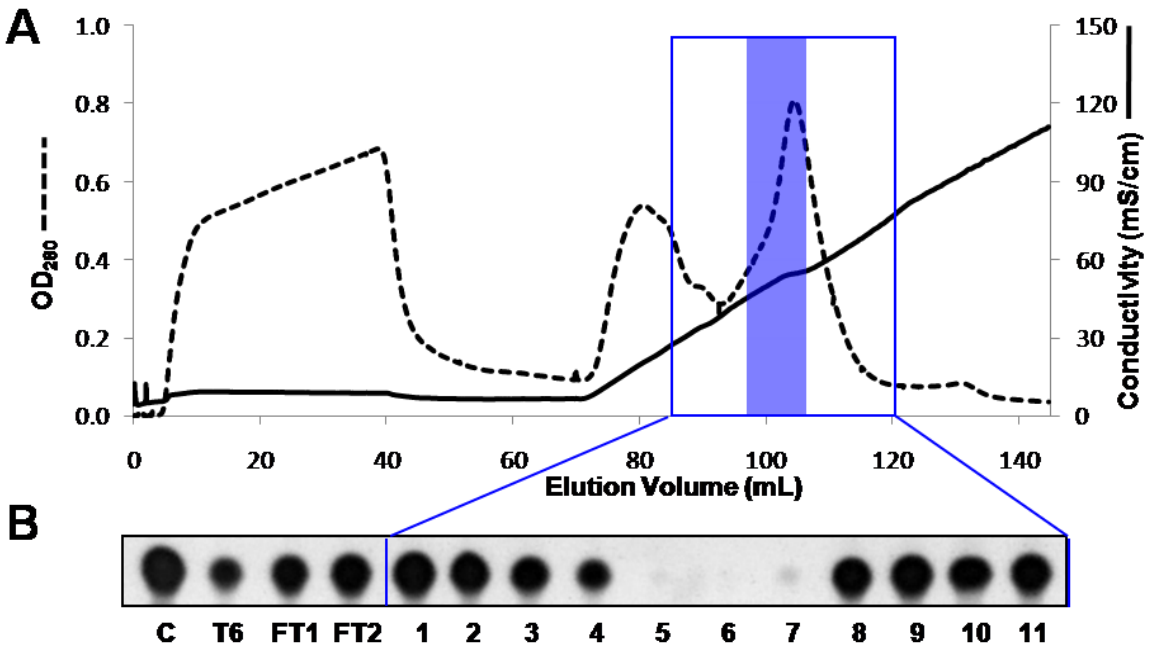
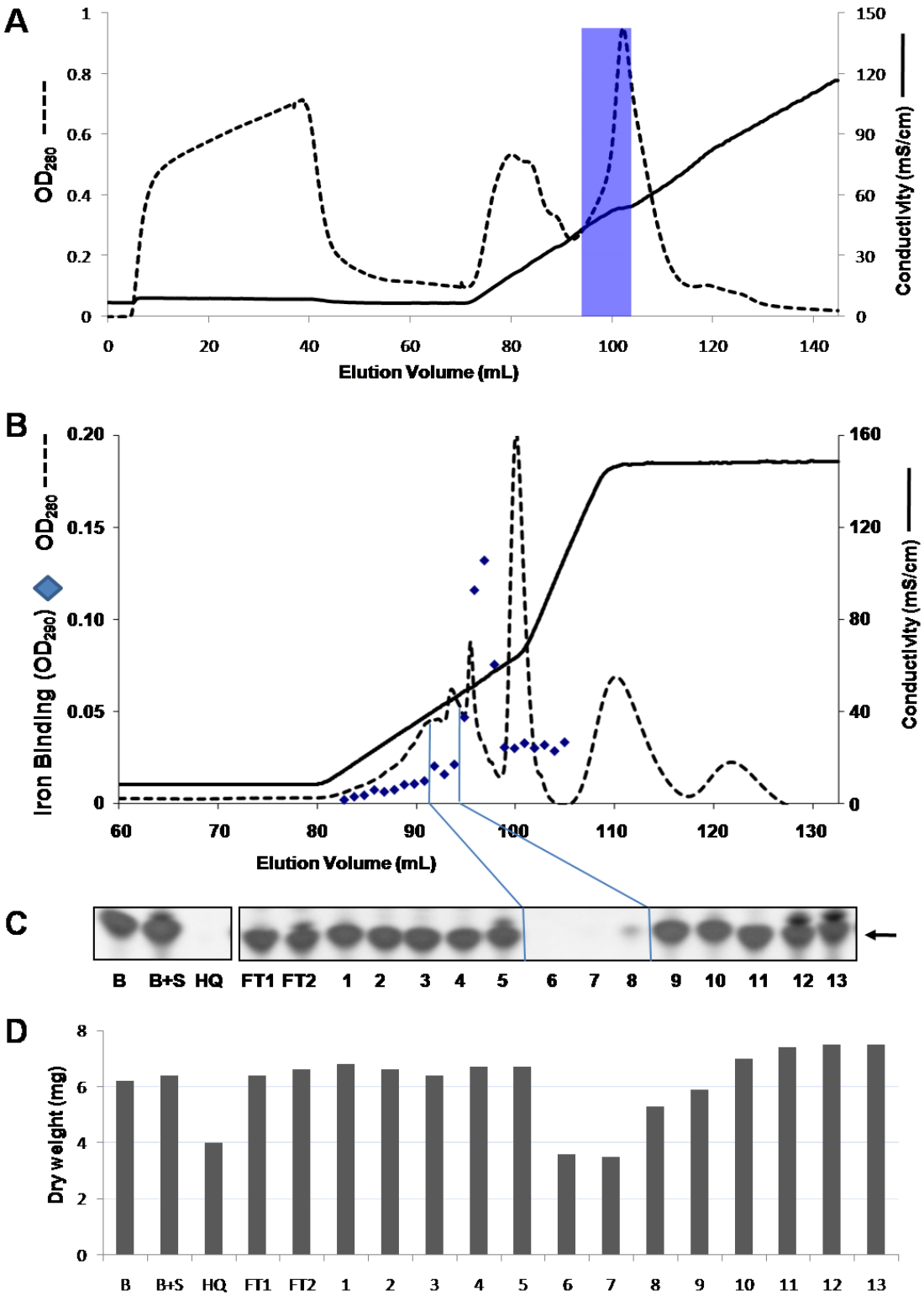


Figure 3. Linear gradient fractionation of Tex6 crude extract by anion exchange chromatography and assay for inhibition of aflatoxin production. (A) Chromatograph showing linear elution of Tex6 crude extract applied to a High Q anion exchange column. (B) Inverted TLC image of aflatoxin production in the presence of buffer control (C), Tex6 crude extract (T6), two fractions from the unbound flow-through (FT1, FT2) and fraction corresponding to the region boxed in with a blue line (lanes 1-11). The inhibitory activity in fractions 5-7 corresponds to the blue shaded region in the chromatograph.

Figure 4. Additional purification and characterization of the anion exchange binding activity. (A) Chromatograph showing linear elution of Tex6 crude extract applied to a High Q anion exchange column. Activity was determined by visual inspection and the region corresponding to the active fractions is shaded in blue. (B) Chromatograph showing linear elution of High Q active pool applied to DEAE Protein Pak column. Iron binding activity is marked with a blue diamond. Absorbances at 280 and 290 nm share the same Y axis. (C) TLC image showing aflatoxin accumulation in the presence of buffer (B), buffer + 1M NaCl (B+S), HQ active fractions prior to dilution, two unbound flow-through fractions (FT1, FT2) and fractions from the linear elution (lanes 1-13). Arrow, aflatoxin B1. (D) Dry weights of fungal cultures. Designations correspond to those in panel C.


A
OD280
Conductivity (mS/cm)
Elution Volume (mL)
B
Iron Binding (OD290)
OD280
Conductivity (mS/cm)
Elution Volume (mL)
C
B
B+S
HQ
FT1
FT2
1
2
3
4
5
6
7
8
9
10
11
12
13
D
Dry weight (mg)
B
B+S
HQ
FT1
FT2
1
2
3
4
5
6
7
8
9
10
11
12
13

REFERENCES

1. **Phillippy, B.Q., and J.M. Bland.** 1988. Gradient ion chromatography of inositol phosphates. Anal Biochem **175**:162-166.

2. **Shi, J., H. Wang, K. Schellin, B. Li, M. Faller, J.M. Stoop, R.B. Meeley, D.S. Ertl, J.P. Ranch, and K. Glassman.** 2007. Embryo-specific silencing of a transporter reduces phytic acid content of maize and soybean seeds. Nat Biotechnol **25**:930-937.