

***Aspergillus flavus* expressed sequence tags and microarray as tools in understanding aflatoxin biosynthesis**

J. Yu¹, T. E. Cleveland¹, J. R. Wilkinson¹, B. C. Campbell², J. H. Kim², H. S. Kim³, D. Bhatnagar¹, G. A. Payne⁴, W. C. Nierman^{3,5}

¹Food and Feed Safety Research, USDA-ARS Southern Regional Research Center, 1100 Robert E. Lee Blvd., New Orleans, LA 70124, USA

²Plant Mycotoxin Research, USDA-ARS Western Regional Research Center, 800 Buchanan St., Albany, CA 94710, USA

³The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, USA

⁴Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695, USA

⁵Department of Biochemistry and Molecular Biology, The George Washington University School of Medicine, 2300 Eye Street NW, Washington, DC 20037, USA

Abstract

Aflatoxins are the most toxic and carcinogenic naturally occurring mycotoxins. They are produced primarily by *Aspergillus flavus* and *A. parasiticus*. In order to better understand the molecular mechanisms that control aflatoxin production, identification of genes using *A. flavus* expressed sequence tags (ESTs) and microarrays is currently being performed. Sequencing and annotation of *A. flavus* ESTs from a normalized *A. flavus* cDNA library identified 7,218 unique EST sequences. Genes that are putatively involved in aflatoxin biosynthesis, regulation and signal transduction, fungal virulence or pathogenicity, stress response or antioxidation, and fungal development were identified from these ESTs. Microarrays containing over 5,000 unique *A. flavus* gene amplicons were constructed at The Institute for Genomic Research. Gene expression profiling under aflatoxin-producing and non-producing conditions using this microarray has identified hundreds of genes that are potentially involved in aflatoxin production. Further investigations on the functions of these genes by gene knockout experiments are underway. This research is expected to provide information for developing new strategies for controlling aflatoxin contamination of agricultural commodities.

Keywords: *Aspergillus flavus*, aflatoxin biosynthesis, mycotoxin, microarray, genomics, pathogenesis

Introduction

Aspergillus flavus produces the most potent mycotoxins, aflatoxins B₁ and B₂, while *A. parasiticus* produces aflatoxins G₁ and G₂ in addition to B₁ and B₂. These compounds contaminate pre-harvest field crops when these fungi infect corn, cotton, peanut and tree nuts. They also cause spoilage of post-harvest grains during storage (1-5). Due to the significant health and economic impacts of aflatoxin contaminations, the chemistry, enzymology, and genetics of aflatoxin biosynthesis in *A. flavus* and *A. parasiticus* are being actively studied (1-14).

Genetic studies on aflatoxin biosynthesis in *A. flavus* and *A. parasiticus* led to the cloning of 25 clustered genes within a 75kb DNA region responsible for the enzymatic conversions in the aflatoxin pathway (15, 16). In order to identify genes involved in global regulation of aflatoxin formation and those involved in fungal invasion, *A. flavus* EST (17) and microarray projects have been carried out. Studies indicated that many factors affect aflatoxin formation (5, 18-20). The microarray experiments reported here were focused on profiling genes involved in aflatoxin formation in response to nutritional factors, carbon source in particular. A better understanding of the mechanism of aflatoxin biosynthesis and its regulation at the genomic scale would be helpful in the development of new control strategies to eliminate pre-harvest aflatoxin contamination. Gene expression profiling by EST and microarray technology in *A. flavus* provides a tool for achieving such goal (17, 21-23).

Presented at the EU-USA Bilateral Workshop on Toxigenic Fungi & Mycotoxins, New Orleans, USA, July 5-7, 2005

Correspondence: Jiujiang Yu, 1100 Robert E. Lee Boulevard, New Orleans, LA 70124, USA (jiuyu@srrc.ars.usda.gov)

Financial support: USDA-ARS Food Safety (Animal and Plant Products) CRIS project NP 108

Materials and Methods

Fungal strain, media and culture conditions

The wild-type strains of *A. flavus* NRRL 3357 (ATCC# 20026), *A. flavus* #13, and *A. parasiticus* SU-1 (SRRC 143, NRRL 5862, ATCC# 56775) were selected for the genomic studies. These strains are stable aflatoxin producers and are widely used in laboratory experiments and field tests. Fungal cultures were maintained on potato dextrose agar (PDA, Difco, Detroit, MI). Fungal conidia were collected from PDA plates using a 0.05% Triton X-100 solution. The liquid cultures were incubated at 30 °C with constant shaking at 150 RPM. The four basic culture media, yeast extract (YE, Difco, Detroit, MI), yeast extract sucrose (YES), peptone minimal salt (PMS), and glucose minimal salts (GMS) (24), were used in microarray experiments. The mycelia were harvested by filtration through miracloth at 48 and 96 hours after inoculation. The harvested mycelial samples were frozen in liquid nitrogen for RNA purification.

RNA isolation, Sequencing and Annotation

The frozen fungal mycelia were ground to a fine powder using a pre-chilled mortar and pestle. Total RNA was purified from the fungal mycelia using a Qiagen RNeasy Plant Mini Kit (Qiagen Inc., Valencia, CA) for cDNA expression library construction. Single pass, unidirectional (5'end) sequencing of the EST clones and annotation were performed at The Institute for Genomic Research (TIGR, Rockville, MD, USA).

Microarray construction

According to the EST sequence information, amplicon microarray was constructed at TIGR following the TIGR Microarray Protocols and Standard Operating Procedures. A total of 9,445 pairs of sequence specific primers to the known unique ESTs were made. Two EST sequences were selected from each Tentative Consensus (TC) sequence group and one from each singleton for primer design. The PCR amplification was performed using *A. flavus* genomic DNA as template in Hard-Shell Skirted 96 well microplate (Bio-Rad, Waltham, MA). The successful amplicons plus 32 aflatoxin pathway genes, representing 5,002 unique genes, were printed in triplicates onto aminosilane coated microscope slides (Corning Inc. Alton, MA) using a high precision spott-

ing robot, TIGR Intelligent Automation Systems (IAS).

Probe Labeling and microarray hybridization:

For generating probes in microarray experiments, total RNA was extracted from 100 mg of the fungal tissue using TRIzol® Reagent (Invitrogen Corporation, Carlsbad, CA) according to manufactures instructions. cDNA synthesis, labeling, and hybridization were performed using one of the three separate protocols. They were Aminoallyl RNA labeling and hybridization protocol (TIGR); 3DNA Array 900 protocol (Genisphere Inc. Hatfield, PA), and 3DNA Array 900 MPX protocol (Genisphere Inc. Hatfield, PA). The TIGR and the Array 900 protocol used 10 µg total RNA, while the Array 900 MPX protocol used 3 µg total RNA in labeling reaction.

Data acquisition and statistic analysis

Following hybridization, the fluorescent images were acquired by scanning slides with array reader (ScanArray Express or a ScanArray Lite 5000, PerkinElmer Life Sciences, Boston, MA). Normalization is necessary to adjust for differences in labeling and detection efficiencies of the fluorescent labels and for differences in the quantity of starting RNA. Data was normalized using a local regression techniques, LOWESS (LOcally WEighted Scatterplot Smoothing), using the MIDAS software tool (TIGR) and the resulting data were averaged over triplicate genes on each array and over duplicate arrays (dye-flip) for each experiment. The data set comparison of the resulting microarray hybridization image was examined by cross comparison between experiments using TIGR MultiExperiment Viewer software (MeV, TIGR, <http://www.tigr.org>).

Results

Genes of interest identified by EST

A total of 26,110 normalized *A. flavus* cDNA clones were sequenced and 22,037 high quality usable sequences were obtained. After annotation, 7,218 unique EST sequences were assembled (17). Among those ESTs that are potentially involved directly or indirectly in aflatoxin production have been identified based on their putative functional classifications (17). These genes were classified into five categories: 1, Genes directly involved in

the aflatoxin biosynthesis such as structural and regulatory genes in the known aflatoxin pathway gene cluster (15, 16); 2, Genes putatively involved in global regulation and signal transduction such as mitogen-activated protein kinase (MAPK), MAPK kinase (MAPKK) and MAPKK kinase (MAPKKK) (25), *laeA* for loss of *aflR* expression (NAGEM53TV) (26); 3, Genes possibly involved in virulence/pathogenicity such as genes encoding hydrolytic enzymes; 4, Genes involved in stress response and antioxidation (18, 25, 27, 28); and 5, Genes putatively involved in fungal development and sporulation.

Aflatoxin pathway gene expression

A. parasiticus SU-1 grown in GMS vs. PMS
GMS medium supports aflatoxin production, while PMS does not. When *A. parasiticus* SU-1 grown in GMS medium was compared with that grown in PMS medium for 96 hours, the aflatoxin pathway cluster genes were expressed at a high level under GMS medium condition, but not in PMS medium. A 2 to 9-fold higher expression level (\log_2 value ranged from 1 to 3.5) was scored in *aflC* (*pksA*) (15), *aflD* (*nor-1*), *aflE* (*norA*), *aflG* (*avnA*), *aflH* (*adhA*), *aflI* (*avfA*), *aflJ* (*estA*), *aflK* (*vbs*), *aflL* (*verB*), *aflO* (*omtB*), *aflS* (*aflJ*), *aflV* (*cypX*), and *aflY* (*hypA*). While other aflatoxin pathway genes, *aflR*, *aflU* (*cypA*), *aflB* (*fas-1*), *aflQ* (*ordA*), *aflW* (*moxY*), *aflX* (*ordB*) and the hexose transporter gene, *hxtA*, were expressed but at relatively lower level (0.5 to 1-fold). The sugar utilization genes, *glcA* and *sugR*, the putative aflatoxin transporter gene, *aflT*, were expressed only in the PMS medium (negative \log_2 value). However, some of the aflatoxin pathway genes, *aflF* (*norB*), *aflM* (*ver-1*), and *aflN* (*verA*), were not detected.

Gene expression profiles

in A. flavus NRRL 3357

Using the TIGR protocol we compared 48 hours *A. flavus* NRRL 3357 grown in GMS vs. PMS medium. A total of 3,977 spots gave scorable signals including 20 of the 25 aflatoxin pathway genes. These represent about 50% of the amplicons spotted on the array. When the data were evaluated at 95% confidence level, a total of 263 spots were shown significant. Among them 169 were up-regulated and 94 were down-regulated. All of the aflatoxin pathway genes were scored below 95% significant level (Table 1, experiment 1).

Gene expression profiles in A. flavus #13

YES medium supports aflatoxin production, while YE does not. Gene expression was compared in *A. flavus* #13 grown for 48 hours in YES vs. YE medium using the 3DNA Array 900 protocol (Genisphere). A total of 3,661 amplicon spots gave scorable signals, including most of the known aflatoxin pathway genes. When the results were evaluated at a 95% confidence level, a total of 229 spots were shown significant. Among them, 131 were up-regulated and 98 were down-regulated (Table 1, experiment 2).

Gene expression profiles

in A. parasiticus SU-1

A. parasiticus SU-1 were grown 96 hours in YES vs. YE and GMS vs. PMS. YES was compared to YE and GMS was compared to PMS using the 3DNA Array 900 and Array 900MPX protocols (Genisphere). The four experiments (Table 1, experiments 3, 4, 5, 6) are considered biological replicates with a combination of media (YES vs. YE and GMS vs. PMS) and probe labeling protocols (Array 900 and Array 900MPX). Using both Genisphere protocols, the expressed genes were readily detectable with scorable spots from 6,108 to 7,802 including most of the aflatoxin pathway genes, which represent upto 97% of total spots on array. When the results were evaluated at a 95% confidence level, the significant spots were 428, 335, 396, and 434 respectively. The up- and down-regulated genes were 237 and 191, 165 and 170, 189 and 207, 285 and 149 respectively (Table 1, experiments 3, 4, 5, 6).

Gene expression profiles across experiments

Using TIGR MeV software program, genes expressed across several experiments have been identified. The up- and down-regulated genes in the experimental groups are shown in Table 2. Analysis indicated that there were only 21 genes up-regulated in addition to the aflatoxin pathway genes consistently expressed over the two *A. flavus* strains and there were 21 down-regulated genes. Combining the four *A. parasiticus* replicates, there were 23 up-regulated and 4 down-regulated genes. When comparing all of the six experiments, only 7 genes were shown up-regulated and 3 genes down-regulated across the experiments (Table 2).

Table 1. Gene expression profiles of individual experiment

Species, productive vs. non-productive condition (protocol used)	# of observed	% of total entries	Number significant	# up regulated	# down regulated
1. <i>A. flavus</i> NRRL 3357 GMS vs. PMS 48 hours growth (TIGR)	3,977	49.5	263	169	94
2. <i>A. flavus</i> #13 YES vs. YE 48 hours growth (Array 900)	3,661	45.6	229	131	98
3. <i>A. parasiticus</i> SU-1 GMS vs. PMS 96 hours growth (Array 900)	6,995	87.1	428	237	191
4. <i>A. parasiticus</i> SU-1 YES vs. YE 96 hours growth (Array 900)	6,108	76.0	335	165	170
5. <i>A. parasiticus</i> SU-1 GMS vs. PMS 96 hours growth (Array 900MPX)	7,368	91.7	396	189	207
6. <i>A. parasiticus</i> SU-1 YES vs. YE 96 hours growth (Array 900MPX)	7,802	97.1	434	285	149

Note: Total possible amplicons on array: 8,032, which represent a total of 5,002 unique ESTs (3,030 TC, 1,941 singleton plus 31 aflatoxin biosynthetic pathway cluster genes)

Table 2. Gene expression profiles across experiments

Experimental group	Total observed	# signif. across exps.	# up regulated	# down regulated	other*
Experiments: 1, 2	2,664	45	21	21	3
Experiments: 3, 4, 5, 6	4,906	27	23	4	0
Experiments: 1, 2, 3, 4, 5, 6	2,362	10	7	3	0

* Genes neither up- nor down-regulated

Discussion

Employing EST and microarray technologies for profiling expression of the genes potentially involved in aflatoxin contamination of crops by *A. flavus* is a rapid and effective strategy in understanding the mechanism of aflatoxin biosynthesis and its regulation. Experiments on the genes involved in the secondary metabolism, such as those required for aflatoxin biosynthesis, demonstrated that the genes of the aflatoxin pathway are expressed at a much lower level than the primary metabolic pathway genes. This could help to explain why not all the known aflatoxin pathway genes were identified in the EST clones sequenced. This is especially true when using microarrays to detect the low expression aflatoxin pathway genes. In the microarray experiments, only the high expression aflatoxin pathway genes, such as *aflC* (*pksA*), *aflD* (*nor-1*), *aflE* (*norA*), *aflJ* (*estA*), *aflK* (*vbs*), *aflO* (*omtB*), *aflS* (*aflJ*), and *aflY* (*hypA*) were scored, while the genes that were expressed at low level, such as *aflR*, were difficult to detect. It should be noted that during bioinformatics analysis several parameters could be arti-

ficially adjusted to achieve the different significance cutoff level. The significantly up-regulated genes can be controlled to a manageable level for down-stream analysis by targeted mutagenesis. We used an *A. flavus* genome based microarray to study both *A. flavus* and *A. parasiticus*. Both species, though closely related, contain certain differences in their genomes and also in the way aflatoxin biosynthesis is regulated. Some genes with low sequence homology between the two species may not be identified. Therefore, response to environmental factors (nutritional factors in this study) may differ between the species by more than what was demonstrated here. However, this can only be studied in detail when the *A. parasiticus* genome based microarray is available. The genes identified in this study are the putative candidates for further investigation. Genes responsible for the biosynthesis of secondary metabolites such as aflatoxins are those encoding polyketide synthases, fatty acid synthases, carboxylases, dehydrogenases, reductases, oxidases, oxidoreductases, epoxide hydrolases, mono- or di-oxygenases, cytochrome P450 monooxygenases, and methyltransferases (5, 13, 14, 15). In the *A. flavus*

genome as identified by EST and whole genome annotation, numerous genes fall into the categories of those enzymes. Without additional biological evidence it is difficult to predict whether these genes are involved in the primary or secondary metabolisms. In order to identify the global regulatory genes to devise strategy to control aflatoxin contamination, profiling those genes in *A. flavus*, using microarray followed by analysis of targeted mutagenesis is necessary.

Conclusion

Using EST and microarray data to identify genes in aflatoxin production and regulation is an effective strategy to narrow down the potential targets for further studies. Since the list of identified genes could be different depending on the bioinformatic parameters used in each experiment, additional validation and functional studies are needed to confirm the involvement of these identified genes in aflatoxin production. Under aflatoxin-producing condition, many genes potentially involved in aflatoxin formation were identified to be up- or down-regulated including those known aflatoxin pathway cluster genes. In *A. flavus*, 21 up- and 21 down-regulated genes were identified. In *A. parasiticus*, 23 up-regulated and 4 down-regulated genes were identified. Further studies on the functional involvement of those genes by targeted mutagenesis are necessary.

Acknowledgement

The authors express their sincere thanks to Dr. Catherine Ronning, The Institute for Genomic Research, for her assistance in bioinformatic analysis and Dr. Charlie Xiang, Director of Microarray Facility, The Venter Institute, Rockville, Maryland, USA, for his helpful suggestions. We are grateful to Dr. Daniel Shelton, Research Leader of the Environmental Microbial Safety Laboratory, USDA-ARS Beltsville Agricultural Research Center, for his support in providing a working environment to continue this project; We are thankful for the secretarial help of Janell Becker.

References

- Bennett JW, Klich M (2003) Mycotoxins. Clin Microbiol Rev 16: 497-516
- Bennett JW, Papa KE (1988) The aflatoxigenic *Aspergillus*. In: Igram DS, Williams PA (eds) Genetics of plant pathology. Academia Press, London, Vol 6, 264-280
- Lancaster MD, Jenkins FP, Phillip JM (1961) Toxicity associated with certain samples of groundnuts. Nature 192: 1095-1096
- Cotty PJ (1997) Aflatoxin-producing potential of communities of *Aspergillus* section Flavi from cotton producing areas in the United States. Mycol Res 101: 698-704
- Payne GA, Brown MP (1998) Genetics and physiology of aflatoxin biosynthesis. Annu Rev Phytopathol 36: 329-362
- Chang P-K, Cary JW, Bhatnagar D, Cleveland TE, Bennett JW, Linz JE, Woloshuk CP, Payne GA (1993) Cloning of the *Aspergillus parasiticus* *apa-2* gene associated with the regulation of aflatoxin biosynthesis. Appl Environ Microbiol 59: 3273-3279
- Chang P-K (2003) The *Aspergillus parasiticus* protein *AFLJ* interacts with the aflatoxin pathway-specific regulator *AFLR*. Mol Genet Genomics 268: 711-719
- Payne GA, Nystrom GJ, Bhatnagar D, Cleveland TE, Woloshuk CP (1993) Cloning of the *afl-2* gene involved in aflatoxin biosynthesis from *Aspergillus flavus*. Appl Environ Microbiol 59: 156-162
- Yu J, Cary JW, Bhatnagar D, Cleveland TE, Keller NP, Chu FS (1993) Cloning and characterization of a cDNA from *Aspergillus parasiticus* encoding an *O*-methyltransferase involved in aflatoxin biosynthesis. Appl Environ Microbiol 59: 3564-3571
- Yu J, Chang P-K, Cary JW, Ehrlich KC, Montalbano B, Dyer JM, Bhatnagar D, Cleveland TE (1998) Characterization of the critical amino acids of an *Aspergillus parasiticus* cytochrome P450 monooxygenase encoded by *ordA* involved in aflatoxin B₁, G₁, B₂, and G₂ biosynthesis. Appl Environ Microbiol 64: 4834-4841
- Yu J, Woloshuk CP, Bhatnagar D, Cleveland TE (2000) Cloning and characterization of *avfA* and *omtB* genes involved in aflatoxin biosynthesis in three *Aspergillus* species. Gene 248: 157-167
- Ehrlich KC, Montalbano BG and Cary JW (1999) Binding of the C6-zinc cluster protein, *AFLR*, to the promoters of aflatoxin pathway biosynthesis genes in *Aspergillus parasiticus*. Gene 230: 249-257
- Yu J, Bhatnagar D, Ehrlich KC (2002) Aflatoxin Biosynthesis. Rev Iberoam Microl (RIAM) Spain. 19: 191-200

- 14 Yu J (2004) Genetics and Biochemistry of Mycotoxin Synthesis. In: Arora DK (ed). Fungal Biotechnology in Agricultural, Food, and Environmental Applications. Marcel Dekker, New York, Vol. 21, 343-361
- 15 Yu J, Chang PK, Ehrlich KC, Cary JW, Bhatnagar D, Cleveland TE, Payne GA, Linz JE, Woloshuk CP, Bennett JW (2004) The clustered pathway genes in aflatoxin biosynthesis. *Appl Environ Microbiol* 70: 1253-1262
- 16 Yu J, Bhatnagar D, Cleveland TE (2004) Completed sequence of aflatoxin pathway gene cluster in *Aspergillus parasiticus*. *FEBS Lett* 564: 126-130
- 17 Yu J, Whitelaw CA, Nierman WC, Cleveland TE, Bhatnagar D (2004) *Aspergillus flavus* expressed sequence tags for identification of genes with putative roles in aflatoxin contamination of crops. *FEMS Microbiol Lett* 237: 333-340
- 18 Campbell BC, Molyneux, RJ, Schatzki TF (2003) Current research on reducing pre- and post-harvest aflatoxin contamination of U. S. almond, pistachio and walnut. In: Abbas H (ed) Aflatoxin and Food Safety. Part I. *J Toxicol-Toxin Rev* 22: 225-266
- 19 Yu J, Chang P-K, Bhatnagar D, Cleveland TE (2000) Cloning of sugar utilization gene cluster in *Aspergillus parasiticus*. *Biochim Biophys Acta* 1493: 211-21.
- 20 Yu J, Mohawed SM, Bhatnagar D, Cleveland TE (2003) Substrate-induced lipase gene expression and aflatoxin production in *Aspergillus parasiticus* and *Aspergillus flavus*. *J Appl Microbiol* 95: 1334-1342
- 21 O'Brian GR, Fakhoury AM, Payne GA (2003) Identification of genes differentially expressed during aflatoxin biosynthesis in *Aspergillus flavus* and *Aspergillus parasiticus*. *Fungal Genet Biol* 39: 118-127
- 22 Price MS, Conners SB, Tachdjian S, Kelly RM, Payne GA (2005) Aflatoxin conducive and non-conductive growth conditions reveal new gene associations with aflatoxin production. *Fungal Genet Biol* 42: 506-518
- 23 Yu J, Bhatnagar D, Cleveland TE (2004) Genetics and biochemistry of aflatoxin formation and genomics approach for eliminating aflatoxin contamination. In: Romeo JT (ed) Recent Advances in Phytochemistry: *Secondary Metabolism in Model Systems*. Elsevier, Vol 38, 224-242
- 24 Adye J, Mateles RI (1964) Incorporation of labeled compounds into aflatoxins. *Biochim Biophys Acta* 86: 418-420
- 25 Kim JH, Campbell BC, Yu J, Mahoney N, Chan KL, Molyneux RJ, Bhatnagar D, Cleveland TE (2005) Examination of fungal stress response genes using *Saccharomyces cerevisiae* as a model system: Targeting genes affecting aflatoxin biosynthesis by *Aspergillus flavus* Link. *Appl Microbiol Biotech* 67: 807-815
- 26 Bok JW, Keller NP (2004) LaeA, a regulator of secondary metabolism in *Aspergillus* spp. *Eukaryot Cell* 3: 527-535
- 27 Kim JH, Campbell BC, Mahoney NE, Chan KL, Molyneux RJ (2004) Identification of phenolics for control of *Aspergillus flavus* using *Saccharomyces cerevisiae* in a model target-gene bioassay. *J Agric Food Chem* 52: 7814-7821
- 28 Kim JH, Mahoney N, Chan KL, Molyneux RJ, Campbell BC (2006) Controlling food-contaminating fungi by targeting their antioxidative stress-response system with natural phenolic compounds *Appl Microbiol Biotechnol*. In press.