Medical Mycology September 2006, **44**, S9–S11

Taylor & Francis
Taylor & Francis Group

Whole genome comparison of *Aspergillus flavus* and *A. oryzae*

G. A. PAYNE*, W. C. NIERMAN†, J. R. WORTMAN†, B. L. PRITCHARD*, D. BROWN*, R. A. DEAN*, D BHATNAGAR‡, T. E. CLEVELAND‡, MASAYUKI MACHIDA§ & J. YU‡

**Department of Plant Pathology, North Carolina State University, Raleigh, NC, USA, †The Institute for Genomic Research, Rockville, MD, USA, ‡USDA/ARS/SRRC, New Orleans, LA, USA and §Institute for Biological Resources and Functions, AIST, Japan*

Aspergillus flavus is a plant and animal pathogen that also produces the potent carcinogen aflatoxin. *Aspergillus oryzae* is a closely related species that has been used for centuries in the food fermentation industry and is Generally Regarded As Safe (GRAS). Whole genome sequences for these two fungi are now complete, providing us with the opportunity to examine any genomic differences that may explain the different ecological niches of these two fungi, and perhaps to identify pathogenicity factors in *A. flavus*. These two fungi are very similar in genome size and number of predicted genes. The estimated genome size (36·8 Mb) and predicted number of genes (12 197) for *A. flavus* is similar to that of *A. oryzae* (36·7 Mb and 12 079, respectively). These two fungi have significantly larger genomes than *Aspergillus nidulans* (30·1) and *Aspergillus fumigatus* (29·4). The *A. flavus* and *A. oryzae* genomes are enriched in genes for secondary metabolism, but do not differ greatly from one another in the predicted number of polyketide synthases, nonribosomal peptide synthases or the number of genes coding for cytochrome P450 enzymes. A micro-scale analysis of the two fungi did show differences in DNA correspondence between the two species and in the number of transposable elements. Each species has approximately 350 unique genes. The high degree of sequence similarity between the two fungi suggests that they may be ecotypes of the same species and that *A. oryzae* has resulted from the domestication of *A. flavus*.

Keywords aflatoxin, secondary metabolism, koji moulds

Introduction

Aspergillus flavus is an opportunistic pathogen of both plants and animals, and it produces one of the most potent naturally occurring carcinogens known. Developing seeds of corn, peanuts, cotton and tree nuts and many other crop plants are susceptible to infection by *A. flavus* under certain environmental conditions. Contamination also may occur in storage, but in the United States the primary concern for aflatoxin contamination occurs before harvest. Aflatoxins have been associated with liver cancer and many veterinary toxic syndromes [1–3]. In developing countries, where detection and decontamination policies are impractical, it is predominately a food safety issue. Last year Kenya reported a minimum of 125 people killed by consuming aflatoxin-contaminated maize (http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5334a4.htm). In the United States, aflatoxin remains predominately an economic problem resulting from inability to sell the crop nationally or internationally. A guideline of 20 parts aflatoxin per billion parts of food or feed substrate (ppb) is the maximum allowable limit imposed by the US Food and Drug Administration for interstate shipment.

Correspondence: Gary A. Payne, Department of Plant Pathology, North Carolina State University, Raleigh, NC 27695-7567, USA. Tel: +919 515 6994, Fax: 919 513 0024; E-mail: gary_payne@ncsu.edu

DOI: 10.1080/13693780600835716

Human infections caused by members of the genus *Aspergillus* are another emerging health problem in developed countries. Immunocompromised individuals are at higher risk of succumbing to aspergillosis, a condition that encompasses a variety of diseases caused by members of the genus *Aspergillus*, including invasive aspergillosis, pulmonary aspergilloma, allergic bronchopulmonary aspergillosis, and others [4]. *Aspergillus flavus* is the second leading cause of aspergillosis in humans and the leading causative agent of chronic indolent invasive sinonasal infection in immunocompetent patients [4]. Mortality from *Aspergillus* infections is high because only a limited number of antifungal drugs are available and resistant strains have been identified [5,6].

Aspergillus oryzae, on the other hand, has been used for centuries in the food fermentation industry and is generally regarded as safe [7], and no strains of *A. oryzae* are known to produce aflatoxin. The genes encoding the pathway enzymes for aflatoxin biosynthesis are clustered within a 75 Kb region of DNA in *A. flavus* [8,9]. Strains of *A. oryzae* also have the aflatoxin biosynthetic cluster but it does not appear to be functional. Tominaga *et al.* [10] compared several strains of *A. oryzae* with *A. flavus* and found that compared to the *A. flavus* sequence the *A. oryzae* genes contained deletions, frameshift mutations, and base pair substitutions.

These two fungi are morphologically very similar, and the results of several studies suggest that they in fact may be ecotypes of the same species. One prevailing premise is that *A. oryzae* resulted from the domestication of *A. flavus* after centuries in culture [11]. Now that the genome sequence of *A. oryzae* is available to the public [7] and a whole genome sequence of *A. flavus* is near completion, we have the opportunity to examine these two fungi in detail. A comparison of the genomes of these two fungi will likely reveal information on changes that have occurred during the domestication of *A. oryzae*, and help identify pathogenicity factors in *A. flavus*.

Sequence analysis of *A. flavus*

A whole-genome sequencing project for *A. flavus* was initiated in 2003 from funds obtained from the USDA/NRI Competitive Grants Program and USDA/ARS/SRRC in New Orleans. Sequencing of the genome, which resulted in a 5 × coverage, was done at The Institute of Genomic Research (TIGR). Sequence reads are available at NCBI and at www.aspergillusflavus.org. Automated annotation, which also is being done at TIGR, is near completion. The annotation tools were trained using the genomic sequence of *A. oryzae* and nearly 14,000 unique ESTs from *A. flavus* and *A. oryzae*. Because of these genetic resources that aided in the automated annotation, we feel that this annotation correctly predicted the location of most of the genes of *A. flavus*. Additional manual annotation will be done at North Carolina State University and within the *A. flavus* research community.

A web browser based on the generic browser, GBrowse was developed at North Carolina State University. Currently, it allows Blast matches to genes, proteins and genomic sequence of other *Aspergillus* spp., alignments of Expressed Sequence Tags, and Gene Ontology annotations. Additional functionality will be added to this browser as more sequence information becomes available. Links to the web browser and to other information on the sequencing project can be found at www. aspergillusflavus.org.

Initial comparison of *A. flavus* and *A. oryzae*

Because the final annotation of the *A. flavus* genome is not complete, the data presented here represent a preliminary comparison between these two organisms. The genome of *A. flavus* has been assembled into 79 scaffolds ranging in size from 4·5 Mb to 1·0 Kb. Over 75% of the genome is represented in the ten largest scaffolds, and over 99·6% of the predicted genes are in the largest 16 scaffolds.

The estimated genome size of 36·8 Mb (coding for 12 197 genes) is similar to that for *A. oryzae* (36·7 Mb) with 12 079 genes predicted. These two fungi are enriched in genes for secondary metabolism. *A. flavus*, for example, is predicted to have 35 polyketide synthases, 24 non-ribosomal peptide synthases, 122 p450 enzymes and *A. oryzae* is predicted to have 30 polyketide synthases, 24 nonribosomal peptide synthases and 151 p450 enzymes. By comparison, *A. fumigatus* is predicted to have 14 polyketide synthases, 14 non-ribosomal peptide synthases and 65 p450 enzymes.

A. oryzae scaffolds have been assigned to chromosomes by optical mapping. The high degree of DNA similarity between these two fungi allowed alignment of the *A. flavus* genomic scaffolds to the *A. oryzae* chromosomes predicted by the optical map. The 16 largest genomic scaffolds from *A. flavus* essentially correspond to the 16 arms of the eight predicted chromosomes for *A. oryzae*. In comparison with *A. oryzae*, there has been a translocation event in *A. flavus* between chromosomes II and VI. The break site is associated with a family of uncharacterized repeat elements. Three types of transposable elements have

been predicted in the genome of *A. flavus*. In each case their frequency of occurrence is greater in *A. oryzae* than *A. flavus*.

Conclusion

Our initial examination of these two genomes indicates that the two fungi are very similar. Either they are extremely closely related species or ecotypes of the same species. In either case, the two fungi provide a good model to better understand genomic changes that may be associated with different ecological niches. Three areas are of particular interest to us. (i) What changes have occurred in number and organization of genes for secondary metabolism? Information from these studies may shed light on the role of secondary metabolites in fungal ecology. (ii) Can we identify pathogenicity factors in *A. flavus*? Because *A. oryzae* has been found only rarely to be pathogenic on plants or animals, an examination of the unique genes in *A. flavus* may provide gene targets in a search for pathogenicity genes. (iii) What genes are important in survival of a fungus in the natural ecosystem? Presumably some of the genes have been lost in *A. oryzae* and thus may appear as unique genes in *A. flavus*. If unique genes are associated with any of these traits, it should be relatively easy to narrow down the target genes as the two fungi have only approximately 350 unique genes.

Certainly, the presence of a gene does not imply necessarily that the gene is actively expressed. To address differences in gene expression that may occur between these two species, we have begun metabolic profiling of the two strains. Our preliminary experiments indicate that the metabolic profile of *A. flavus* is more complex than that of *A. oryzae*. We plan to complement these studies with gene expression profiling. Funding has been obtained for whole genome affymetrix arrays and oligo-nucleotide arrays for *A. flavus*. These DNA arrays will be employed to monitor gene expression during development, secondary metabolism, saprophytic growth, and pathogensis.

Acknowledgements

We wish to thank Greg OBrian for his assistance with this manuscript.

References

1 Bressac B, Kew M, Wands J, *et al*. Selective G-mutation to T-mutation of P53 gene in hepatocellular-carcinoma from southern Africa. *Nature* 1991; **350**: 429–431.
2 Hsu IC, Metcalf RA, Sun T, *et al*. Mutational hotspot in the P53 gene in human hepatocellular carcinomas. *Nature* 1991; **350**: 427–428.
3 Wogan GN, Househam KC, Hundt HKL. Aflatoxin exposure and its relationship to kwashiorkor in African children. *Cancer Res* 1992; **52**: s2114–s2118.
4 Stevens DA, Kan VL, Judson MA, *et al*. Practice guidelines for diseases caused by *Aspergillus*. *Clin Infect Dis* 2000; **30**: 696–709.
5 Lass-Florl C, Jenkins FP, Philip JM. Toxicity associated with certain samples of groundnuts. *Nature* 1961; **192**: 1095–1096.
6 Seo K, Akiyoshi H, Ohnishi Y. Alteration of cell wall composition leads to amphotericin B resistance in *Aspergillus flavus*. *Microbiol Immunol* 1999; **43**: 1017–1025.
7 Machida M, Asai K, Sano M, *et al*. Genome sequencing and analysis of *Aspergillus oryzae*. *Nature* 2005; **438**: 1157–1161.
8 Payne GA, Brown MP. Genetics and physiology of aflatoxin biosynthesis. *Annu Rev Phytopathol* 1998; **36**: 329–362.
9 Yu JJ, Chang PK, Ehrlich KC, *et al*. Clustered pathway genes in aflatoxin biosynthesis. *Appl Environ Microb* 2004; **70**: 1253–1262.
10 Tominaga M, Lee YH, Hayashi R, *et al*. Molecular analysis of an inactive aflatoxin biosynthesis gene cluster in *Aspergillus oryzae* RIB strains. *Appl Environ Microb* 2006; **72**: 484–490.
11 Kurtzman CP, Smiley MJ, Robnett CJ, Wicklow DT. DNA relatedness among wild and domesticated species in the *Aspergillus flavus* group. *Mycologia* 1986; **78**: 955–959.