

# Sequencing the Black *Aspergilli* species complex

Alan Kuo<sup>1</sup>, Asaf Salamov,<sup>1</sup> Kemin Zhou<sup>1</sup>, Robert Olliar<sup>1</sup>, Scott Baker<sup>2</sup> and Igor Grigoriev<sup>1</sup>

<sup>1</sup>DOE Joint Genome Institute and <sup>2</sup>Pacific Northwest National Lab

<sup>1</sup>*To whom correspondence may be addressed. E-mail: akuo@lbl.gov.*

March 14, 2011

## **ACKNOWLEDGMENTS:**

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

## **DISCLAIMER:**

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.



## Abstract

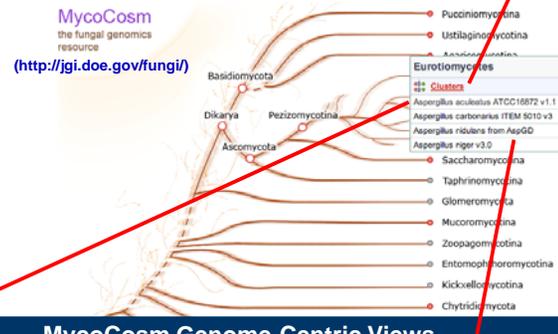
The ~15 members of the *Aspergillus* section *Nigri* species complex (the "Black Aspergilli") are significant as platforms for bioenergy and biotechnological technology, as members of soil microbial communities and players in the global carbon cycle, and as food processing and spoilage agents and agricultural toxigenes. Despite their utility and ubiquity, the morphological and metabolic distinctiveness of the complex's members, and thus their taxonomy, is poorly defined. We are using short read pyrosequencing technology (Roche/454 and Illumina/Solexa) to rapidly scale up genomic and transcriptomic analysis of this species complex. To date we predict 11197 genes in *Aspergillus niger*, 11624 genes in *A. carbonarius*, and 10845 genes in *A. aculeatus*. *A. aculeatus* is our most recent genome, and was assembled primarily from 454-sequenced reads and annotated with the aid of > 2 million 454 ESTs and > 300 million Solexa ESTs. To most effectively deploy these very large numbers of ESTs we developed 2 novel methods for clustering the ESTs into assemblies. We have also developed a pipeline to propose orthologies and paralogies among genes in the species complex. In the near future we will apply these methods to additional species of Black Aspergilli that are currently in our sequencing pipeline.

## Aspergilli in MycoCosm: Genome statistics

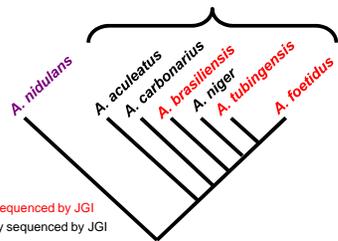
	<i>aculeatus</i>	<i>carbonarius</i>	<i>niger</i>	<i>nidulans</i> (AspGD)
Assembly size (Mbp)	35.4	36.3	34.9	30.5
# of ESTs	338596947	2466582	15360	NA
# of genes	10845	11624	11197	10680
w/ alternative transcripts	6333 (58%)	ND	ND	ND
w/ paralogs	6758 (62%)	7032 (59%)	7045 (63%)	6182 (58%)
# of multigene families	1561	1726	1599	1461

## MycoCosm Eurotiomycetes Node

(<http://genome.jgi-psf.org/eurotiomycetes/>)



## The Black Aspergilli (*Nigri* complex)



- to be sequenced by JGI
- already sequenced by JGI
- already sequenced by Broad

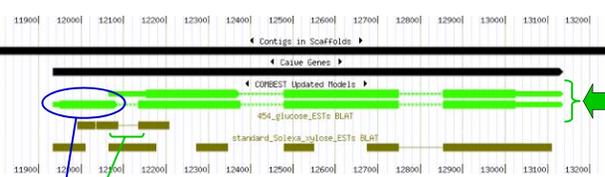
non-*Nigri* Eurotiomycetes:  
*A. sojae*  
*A. wentii*  
*A. versicolor*  
*A. sydowii*  
*A. glaucus*  
*A. zonatus*  
*Penicillium roqueforti*

Black Aspergilli are of interest to:  
 Bioenergy  
 Biotechnological technology  
 Soil ecology  
 Carbon cycle  
 Food processing  
 Food spoilage  
 Food toxicity

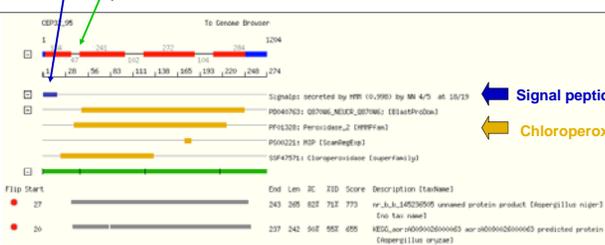
## MycoCosm Genome-Centric Views

### Alternative splicing may affect *A. aculeatus* chloroperoxidase secretion

*A. aculeatus* locus view

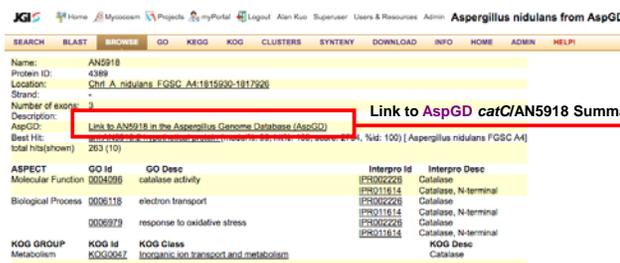


*A. aculeatus* protein view



### *Aspergillus nidulans* catalase gene

*A. nidulans* protein view



Link to AspGD catCAN5918 Summary

### AspGD catCAN5918 Summary



## Acknowledgments

*A. aculeatus* collaborators: Adrian Tsang, Randy Berka,  
*A. carbonarius* collaborators: Giancarlo Perrone, Antonia Gallo  
*A. nidulans* genome: Arnaud, MB et al. at AspGD  
 (<http://www.aspergillusgenome.org>)

## MycoCosm Comparative Genomics View

Many *Nigri* complex gene families are expanded relative to *A. nidulans*

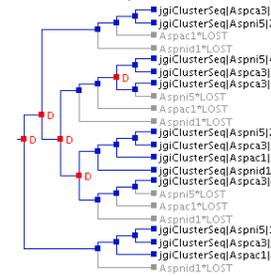
Gene families summary view

Cluster	<i>A. aculeatus</i> v1.0 FilteredModels	<i>A. nidulans</i> from AspGD AspGD_genes	<i>A. niger</i> v3.0 FrozenGeneCatalog071608	<i>Aspergillus carbonarius</i> v3.0 FilteredModels	Total
Totals	5282	4875	5433	5617	21207
78	9	1		5	25
144	4	1		6	18
177	5	1		6	16
205	4	1		5	15
207	1			3	15
216	3			6	14
224	3	1		5	14
229	4	1		6	14
236	4	1		4	13
251	2	1		6	13
256	3	1		5	13

Filter for a single *A. nidulans* gene

### Catalase gene family has multiple duplications and losses

Catalase family tree view



Catalase family synteny view

