Assembly, Annotation, and Analysis of Multiple Mycorrhizal Fungal Genomes

Alan Kuo[1], Igor Grigoriev[1], Annegret Kohler[2], Francis Martin[2], Mycorrhizal Genomics Initiative Consortium

[1]DOE Joint Genome Institute, USA
[2]INRA-Nancy, France

March 2013

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.

Assembly, Annotation, and Analysis of Multiple Mycorrhizal Fungal Genomes

Alan Kuo[1*], Igor Grigoriev[1], Annegret Kohler[2], Francis Martin[2], Mycorrhizal Genomics Initiative Consortium

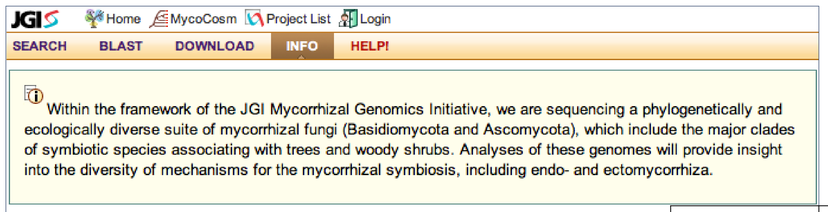
[1]DOE Joint Genome Institute, USA, [2]INRA-Nancy, France. *akuo@lbl.gov

DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

Abstract

Mycorrhizal fungi play critical roles in host plant health, soil community structure and chemistry, and carbon and nutrient cycling, all areas of intense interest to the US Dept. of Energy (DOE) Joint Genome Institute (JGI). To this end we are building on our earlier sequencing of the *Laccaria bicolor* genome by partnering with INRA-Nancy and the mycorrhizal research community in the MGI to sequence and analyze dozens of mycorrhizal genomes of all Basidiomycota and Ascomycota orders and multiple ecological types (ericoid, orchid, and ectomycorrhizal). JGI has developed and deployed high-throughput sequencing techniques, and Assembly, RNASeq, and Annotation Pipelines. In 2012 alone we sequenced, assembled, and annotated 12 draft or improved genomes of mycorrhizae, and predicted ~232831 genes and ~15011 multigene families, All of this data is publicly available on JGI MycoCosm (http://jgi.doe.gov/fungi/), which provides access to both the genome data and tools with which to analyze the data. Preliminary comparisons of the current total of 14 public mycorrhizal genomes suggest that 1) short secreted proteins potentially involved in symbiosis are more enriched in some orders than in others amongst the mycorrhizal Agaricomycetes, 2) there are wide ranges of numbers of genes involved in certain functional categories, such as signal transduction and post-translational modification, and 3) novel gene families are specific to some ecological types.

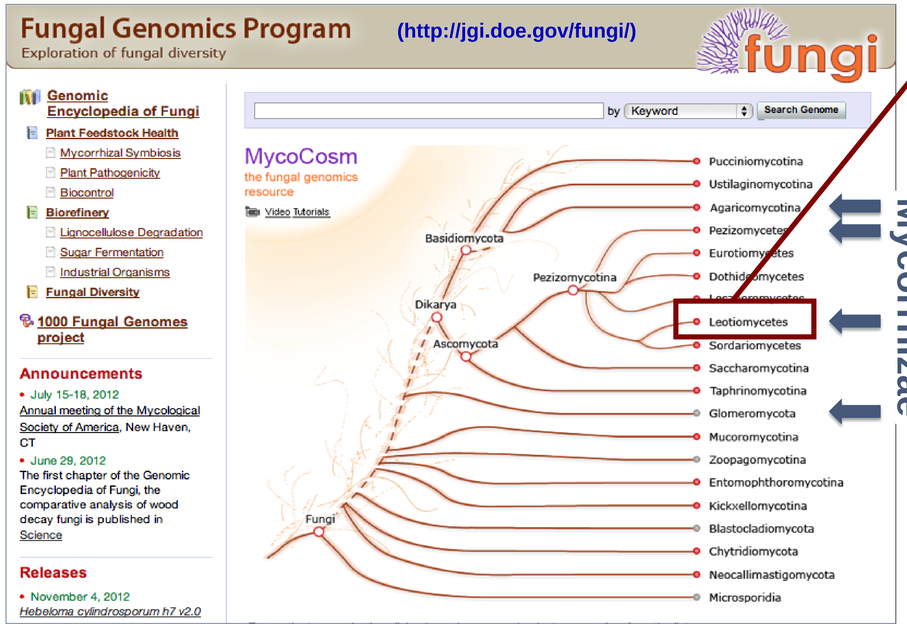
Mycorrhizal Fungi Eco-Group Portal (http://jgi.doe.gov/Mycorrhizal_fungi/)



##	Organism	Name	Assembly length	# genes
1	Amamu1	Amanita muscaria Koide v1.0	40,699,759	18,153
2	Hebcy2	Hebeloma cylindrosporum h7	38,226,047	15,382
3	Lacam1	Laccaria amethystina LaAM-08-1 v1.0	52,197,432	21,066
4	Lacbi2	Laccaria bicolor v2.0	60,707,050	23,132
5	Oidma1	Oidiodendron maius	46,426,256	16,703
6	Paxin1	Paxillus involutus ATCC 200175	58,301,126	17,968
7	Paxru1	Paxillus rubicundulus Ve08.2h10 v1.0	53,011,005	22,065
8	Pilcr1	Piloderma croceum F 1598	59,326,866	21,583
9	Pismi1	Pisolithus microcarpus 441 v1.0	53,027,657	21,064
10	Pisti1	Pisolithus tinctorius Marx 270 v1.0	71,007,534	22,701
11	Sclci1	Scleroderma citrinum Foug A	56,144,862	21,012
12	Sebve1	Sebacina vermifera MAFF 305830	38,094,242	15,312
13	Suilu1	Suillus luteus UH-Slu-Lm8-n1	37,014,302	18,316
14	Tubme1	Tuber melanosporum from Genoscope	124,945,702	7,496
15	Tulca1	Tulasnella calospora AL13/4D	62,392,858	19,659

Phylogeny	Ecology	transcripts per gene	genes per family
Agaricales	ECM	ND	2.05
Agaricales	ECM	3.22	1.82
Agaricales	ECM	2.99	2.09
Agaricales	ECM	ND	2.58
Leotiomycetes	ericoid	ND	2.15
Boletales	ECM	ND	2.78
Boletales	ECM	3.05	2.38
Atheliales	ECM	ND	2.01
Boletales	ECM	ND	2.24
Boletales	ECM	2.92	2.38
Boletales	ECM	3.23	2.20
Sebacinales	orchid	2.73	2.15
Boletales	ECM	3.82	2.00
Pezizomycetes	ECM	ND	1.42
Cantharellales	orchid	1.94	2.35

Fungal Genomics Program (http://jgi.doe.gov/fungi/)

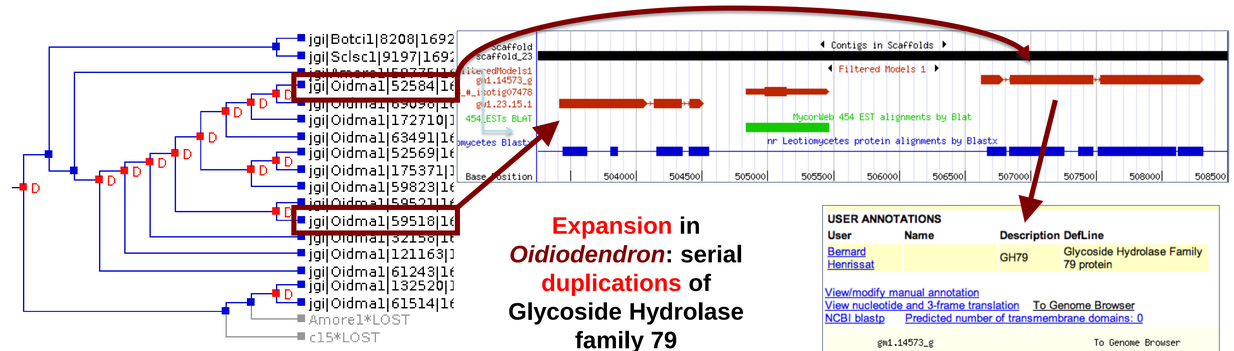


Gene families of Leotiomycetes: mycorrhizal *Oidiodendron* with hydrocarbon-phagic *Amorphotheca*, and pathogenic *Botrytis* and *Sclerotinia*

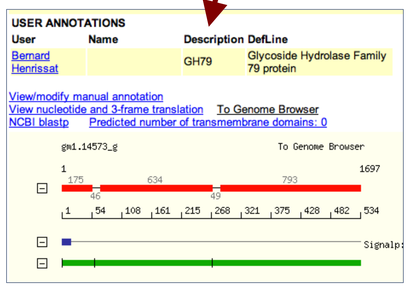
Cluster	A. resinae FilteredModels1	B. cinerea ExternalModels	S. sclerotiorum ExternalModels	O. maius FilteredModels1
	Totals: 9642	Totals: 16447	Totals: 14503	Totals: 16703
176	5	3	2	7
177	3	7	6	1
178	3	3	3	
179	3	4	3	
180	1	4	4	8
181	2	6	4	5
182	3	5	4	5
183	4	3	3	7
184	2	3	3	9
185	1	5	3	8
186	1	5	2	9
187	3	5	5	4
188	3	5	2	7
189	2	3	3	9
190	1	1	1	14
191	3	3	3	

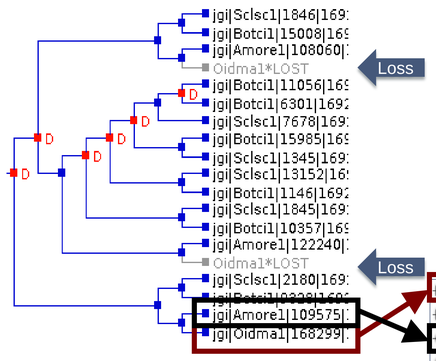
Contraction

Expansion



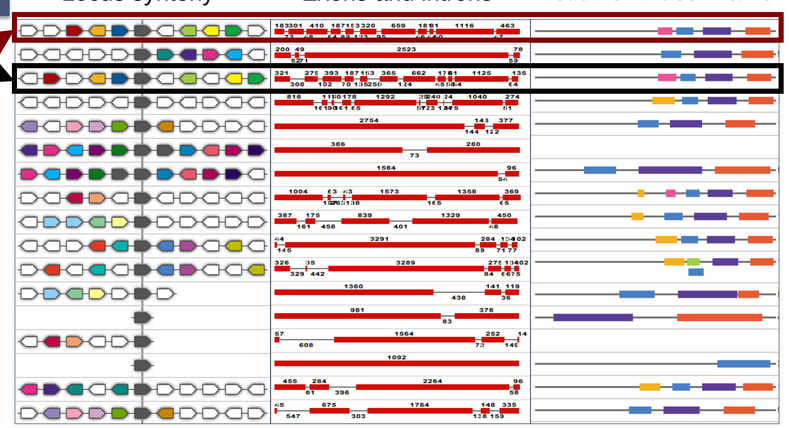
Expansion in *Oidiodendron*: serial **duplications** of Glycoside Hydrolase family 79



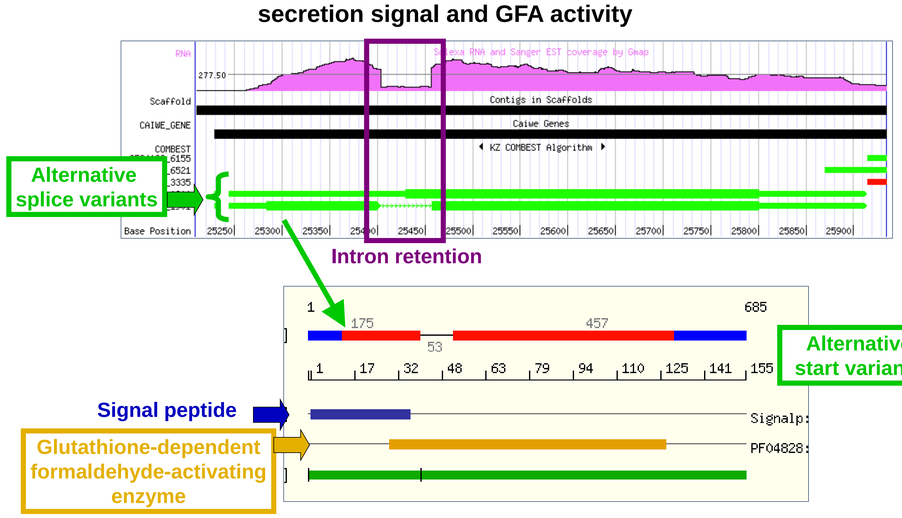


Contraction in *Oidiodendron*: multiple **losses** of histidine kinase response regulator

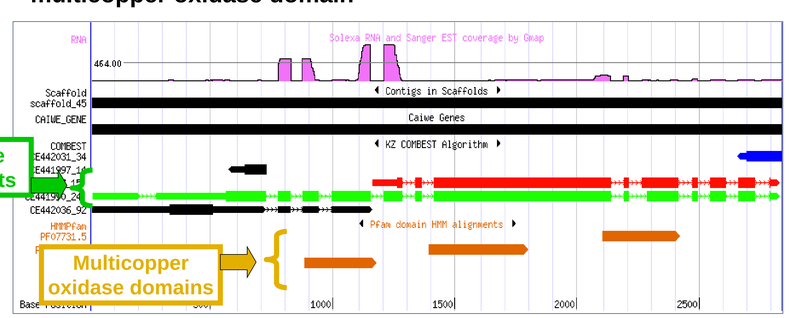
Remaining *Oidiodendron*: histdine kinase has *Amorphotheca* ortholog



Alternative splicing in *Suilus*: intron retention loses secretion signal and GFA activity



Alternative starts in *Suilus*: downstream initiation site loses multicopper oxidase domain



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.