Phylogenomics and Ecogenomics of the Mycorrhizal Symbiosis

Alan Kuo[1], Igor V. Grigoriev[1], Annegret Kohler[2], Francis Martin[2]

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

May 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.

Phylogenomics and Ecogenomics of the Mycorrhizal Symbiosis

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

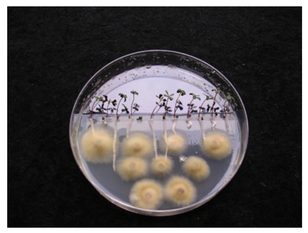
I. Abstract

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 2 dozen mycorrhizal genomes of numerous known mycorrhizal orders and several ecological types (ectomycorrhizal [ECM], ericoid, orchid, and arbuscular). JGI has developed and deployed high-throughput pipelines for genomic, transcriptomic, and re-sequencing, and platforms for assembly, annotation, and analysis. In the last 2 years we have sequenced 21 genomes of mycorrhizal fungi, and resequenced 6 additional strains of *L. bicolor*. Most of this data is publicly available on JGI MycoCosm's Mycorrhizal Fungi Portal (http://jgi.doe.gov/Mycorrhizal_fungi/), which provides access to both the genome data and tools with which to analyze the data. These data allow us to address long-standing issues in mycorrhizal evolution and ecology. For example, a major observation of mycorrhizal evolution is that each of the major ecological types appears to have evolved independently in multiple fungal clades. Using an ecogenomic approach we provide preliminary evidence that 2 clades (Cantharellales and Sebacinales) of a single symbiotic ecotype (orchid) utilize some common regulatory (protein tyrosine kinase) and metabolic (lipase) paths, the latter of which may be the product of HGT. Using a phylogenomic approach we provide preliminary evidence that a particular ecotype (ericoid) may have evolved more than once within a major clade (Leotiomycetes).

II. Mycorrhizae are symbioses between fungi and plants

- 80% of land plants do it.
- Many and diverse fungi do it.
- Including familiar edible mushrooms such as porcini, matsutake, chanterelles.
- Oh, and truffles too.

Simple model is that plant contributes photosynthate and fungus contributes soil nutrients, but can be much more complex.



III. Genomics of mycorrhizal fungi

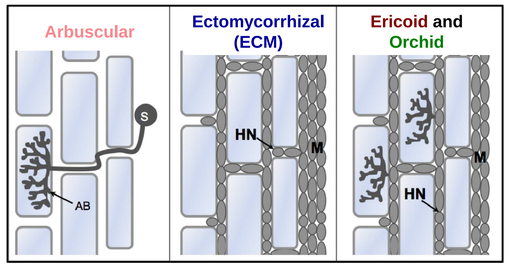
Mycorrhizal Fungi Eco-Group Portal (http://jgi.doe.gov/Mycorrhizal_fungi/)

Within the framework of the JGI Mycorrhizal Genomics Initiative, we are sequencing a phylogenetically and ecologically diverse suite of mycorrhizal fungi (Basidiomycota and Ascomycota), which include the major clades of symbiotic species associating with trees and woody shrubs. Analyses of these genomes will provide insight into the diversity of mechanisms for the mycorrhizal symbiosis, including endo- and ectomycorrhiza.

##	Names	Assembly length	# genes	Phylogeny	Ecology	transcripts per gene	genes per family
1	Amanita muscaria Koide v1.0	40,699,759	18,153	Agaricales	ECM	ND	2.05
2	Boletus edulis v1.0	46,637,611	16,933	Boletales	ECM	3.70	2.42
3	Cenococcum geophilum 1.58 v2.0	177,557,160	14,748	Dothideomycetes	ECM	2.17	1.95
4	Hebeloma cylindrosporum h7 v2.0	38,226,047	15,382	Agaricales	ECM	3.22	1.82
5	Laccaria amethystina LaAM-08-1 v1.0	52,197,432	21,066	Agaricales	ECM	2.99	2.09
6	Laccaria bicolor v2.0	60,707,050	23,132	Agaricales	ECM	ND	2.58
7	Meliniomyces bicolor E v2.0	82,384,847	18,619	Leotiomycetes	ECM/ericoid	2.57	2.01
8	Meliniomyces variabilis F v1.0	55,857,776	20,389	Leotiomycetes	ericoid	2.35	2.24
9	Oidiodendron maius Zn v1.0	46,426,256	16,703	Leotiomycetes	ericoid	ND	2.15
10	Paxillus involutus ATCC 200175 v1.0	58,301,126	17,968	Boletales	ECM	ND	2.78
11	Paxillus rubicundulus Ve08.2h10 v1.0	53,011,005	22,065	Boletales	ECM	3.05	2.38
12	Piloderma croceum F 1598 v1.0	59,326,866	21,583	Atheliales	ECM	ND	2.01
13	Pisolithus microcarpus 441 v1.0	53,027,657	21,064	Boletales	ECM	2.92	2.38
14	Pisolithus tinctorius Marx 270 v1.0	71,007,534	22,701	Boletales	ECM	ND	2.24
15	Rhizophagus irregularis DAOM 197198 v1.0	91,083,792	30,282	Glomeromycota	Arbuscular	ND	2.77
16	Scleroderma citrinum Foug A v1.0	56,144,862	21,012	Boletales	ECM	3.23	2.20
17	Sebacina vermifera MAFF 305830 v1.0	38,094,242	15,312	Sebacinales	orchid	2.73	2.15
18	Suillus brevipes v1.0	51,712,595	22,453	Boletales	ECM	3.45	2.74
19	Suillus luteus UH-Slu-Lm8-n1 v1.0	37,014,302	18,316	Boletales	ECM	3.82	2.00
20	Tricholoma matsutake 945 v3.0	175,759,688	22,885	Agaricales	ECM	2.10	2.84
21	Tuber melanosporum from Genoscope	124,945,702	7,496	Pezizomycetes	ECM	ND	1.42
22	Tulasnella calospora AL13/4D v1.0	62,392,858	19,659	Cantharellales	orchid	1.94	2.35

IV. Ecogenomics of mycorrhizal fungi

IV.A. 4 major types of mycorrhizal symbioses

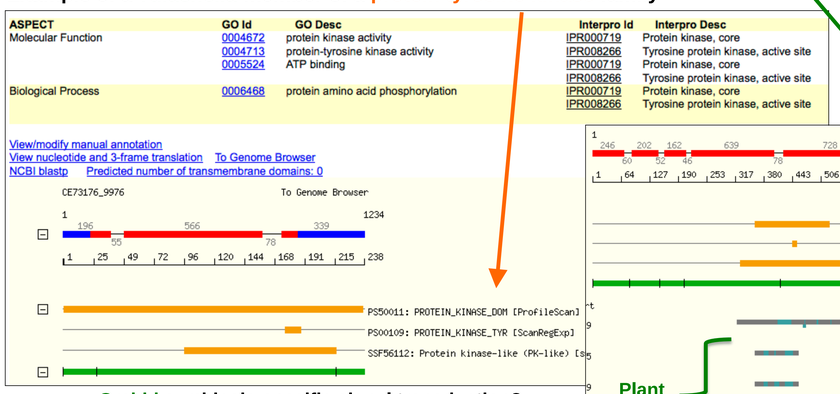


IV.B. Orchid symbioses of Cantharellales and Sebacinales are convergent. Does orchid symbiosis require specific genes?
1.Do Blastp-MCL clustering of all 397639 mycorrhizal proteins in 21 genomes.
2.Filter for genes found in both *Tulasnella* and *Sebacina*, and nowhere else.
3.Found 344 genes in 50 families.

50 clusters, MCL inflation parameter: 2.0, page: 1 GO

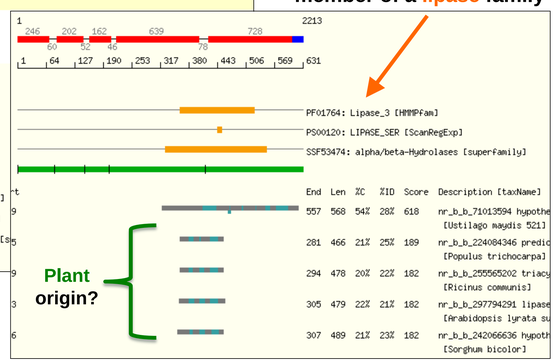
Cluster	T. calospora v1.0 FilteredModels1	T. melanosporum from Genoscope Genoscope_genes	T. matsutake FilteredModels2	S. luteus v1.0 FilteredModels1	S. brevipes FilteredModels1	S. vermifera v1.0 FilteredModels1	S. citrinum v1.0 FilteredModels1
	Totals: 183	Totals: 0	Totals: 0	Totals: 0	Totals: 0	Totals: 161	Totals: 0
1628	7	0	0	0	0	24	0
2786	23	0	0	0	0	1	0
3123	22	0	0	0	0	1	0
5831	4	0	0	0	0	13	0
6442	1	0	0	0	0	15	0
6449	15	0	0	0	0	1	0
7378	12	0	0	0	0	1	0
7651	4	0	0	0	0	8	0
8454	9	0	0	0	0	1	0
8470	6	0	0	0	0	4	0

Example of a *Tulasnella* member of a protein tyrosine kinase family



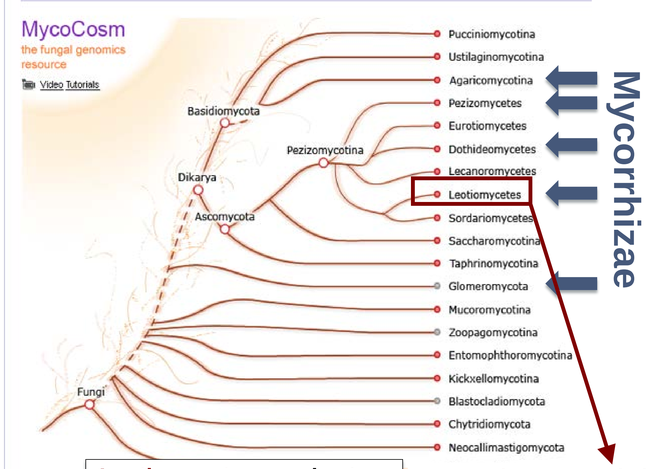
Orchid symbiosis-specific signal transduction?

Example of a *Sebacina* member of a lipase family

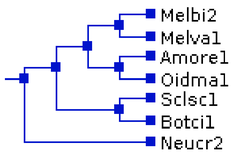


V. Phylogenomics of mycorrhizal fungi

V.A. Mycorrhizal symbiosis arose in multiple clades



Leotiomycetes species tree



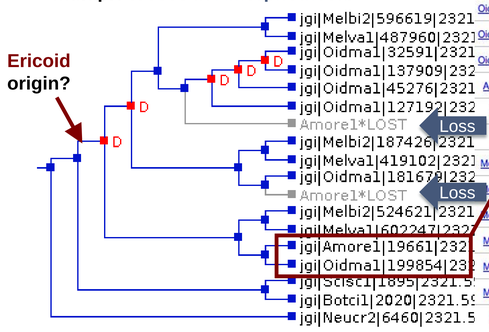
Amorphotheca and *Oidiodendron* branch together

V.B. Leotiomycetes clade has both ericoid symbionts and non-symbionts. Did ericoid symbiosis arise once?
1.Do Blastp-MCL clustering of all 2156 Leotiomycetes proteins in 3 ericoid (*Oidiodendron*, *Meliniomyces*), 2 pathogen, and 1 hydrocarbon-phagic (*Amorphotheca*) genomes.
2.Find orthologs and build clustalW-RaxML species tree, and then build gene trees.
3.Browse for ericoid-enriched gene family expansions.

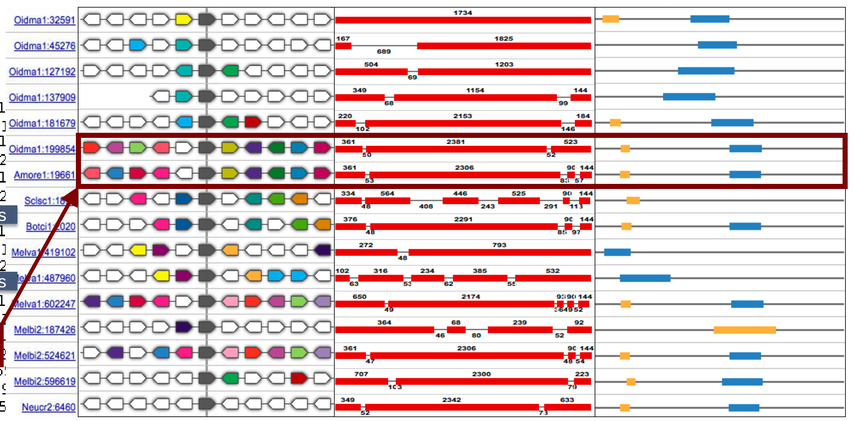
Cluster	S. sclerotiorum ExternalModels	O. maius FilteredModels1	M. variabilis FilteredModels1	M. bicolor FilteredModels2	Neurospora crassa BroadModels	B. cinerea ExternalModels	A. resinae FilteredModels1
	Totals: 14503	Totals: 16703	Totals: 20389	Totals: 18619	Totals: 10785	Totals: 16447	Totals: 9642
594	0	2	10	5	0	0	0
595	1	6	3	3	1	1	1
596	1	7	2	1	1	2	2

Example of an ericoid-enriched gene family:
Serial duplications of a transcription factor in the Amorphotheca-Oidiodendron common ancestor were followed by multiple losses in Amorphotheca.

Ericoid origin?



Locus synteny | **Exons and introns** | **Transcription factor Pfams**



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.