# Fungal Genomics Program

**Igor Grigoriev**[1] * (ivgrigoriev@lbl.gov)

[1]DOE Joint Genome Institute, Lawrence Berkeley National Laboratory

March 2012

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**

Doubling the number of sequenced and annotated genomes every year, the JGI Fungal Program (jgi.doe.gov/fungi) is moving towards the new large scale initiatives aligned with the 2010 Grand Challenges for Biological and Environmental Research: a long term vision. One of the initiatives, the *1000 Fungal Genomes* project, is aimed to explore fungal diversity across the Fungal Tree of Life in order to provide references for research on plant-microbe interactions and environmental metagenomics. Another initiative, the *Genomic Encyclopedia of Fungi*, is focused on diversity among DOE relevant fungi in the areas of plant health and biorefinery parts lists, which will help us to explore the interactions of bioenergy crop species with symbionts and pathogens as well as to catalog industrially relevant genes, pathways, and hosts for biotechnology applications. In addition to broad exploration of fungal diversity, we will also focus on the functional analysis of several *fungal systems* of varying complexity: new model organisms, symbiotic systems such as lichens, and metagenomes of complex communities

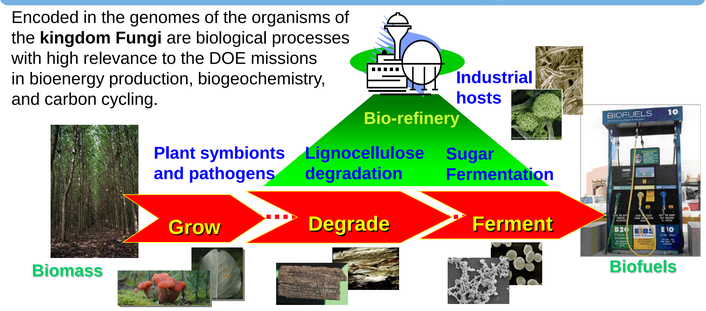
# Fungal Genomics Program


## Igor Grigoriev

DOE JOINT GENOME INSTITUTE
US DEPARTMENT OF ENERGY
OFFICE OF SCIENCE

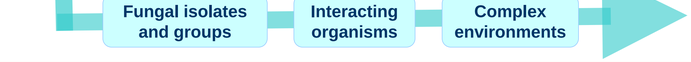

## Mission and Strategy

Encoded in the genomes of the organisms of the **kingdom Fungi** are biological processes with high relevance to the DOE missions in bioenergy production, biogeochemistry, and carbon cycling.




**The JGI Fungal Genomics Program** aims to scale up sequencing and analysis of fungal genomes to explore the diversity of fungi important for energy and the environment, and to promote functional studies on a system level. Combining new sequencing technologies and comparative genomics tools, JGI is now leading the world in fungal genome sequencing and analysis. Over 120 sequenced fungal genomes with analytical tools are available via MycoCosm (www.jgi.doe.gov/fungi), a web-portal for fungal biologists. Our model of interacting with user communities, unique among other sequencing centers, helps organize these communities, improves genome annotation and analysis work, and facilitates new larger-scale genomic projects. This resulted in 20 high-profile papers published in 2011 alone and contributing to the Genomics Encyclopedia of Fungi, which targets fungi related to plant health (symbionts, pathogens, and biocontrol agents) and biorefinery processes (cellulose degradation, sugar fermentation, industrial hosts). Our next grand challenges include larger scale exploration of fungal diversity (1000 fungal genomes), developing molecular tools for DOE-relevant model organisms, and analysis of complex systems and metagenomes.

### Next steps

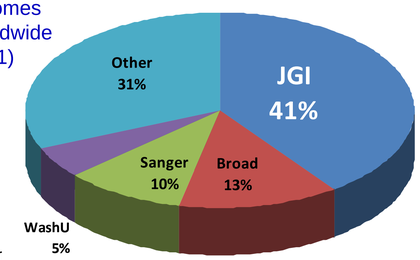



1. **1000 fungal genomes**
   sample fungal diversity
2. **Model fungi**
   deep transcriptome and resequencing
3. **Fungal ecosystems:**
   - Bioenergy crop fungi
   - Population genomics
   - Fungal metagenomes

## Accomplishments

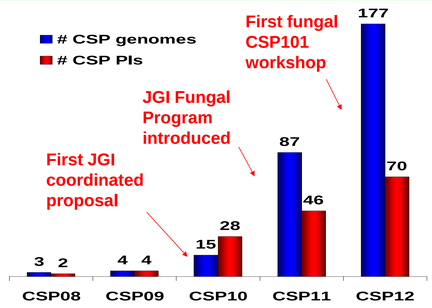
### Lead the World in Fungal Genomics

758 fungal genomes sequenced worldwide (GOLD:Oct 2011)

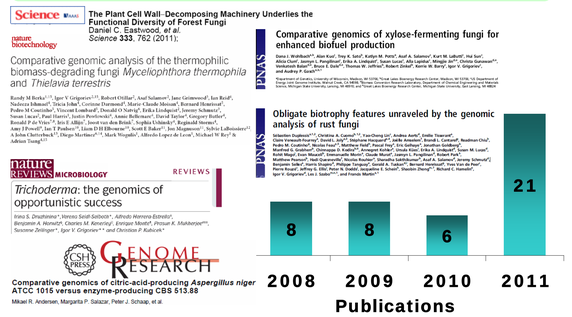



For the second year in a row, we have doubled the number of sequenced and annotated fungal genomes

### Organized User Communities




### Published High-profile Science

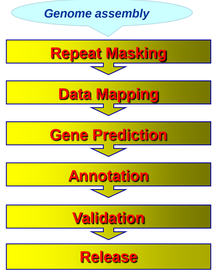



## Tools

### The Annotation Pipeline

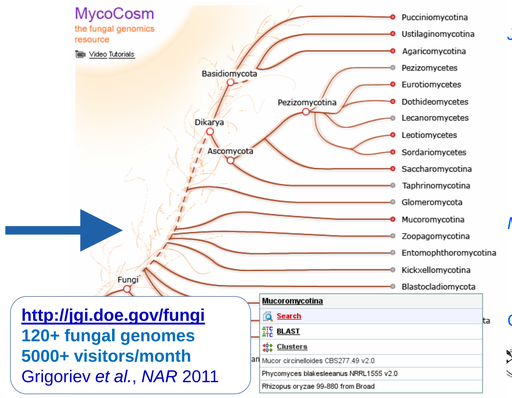
Automated ***Annotation Pipeline*** tuned for high-throughput annotation of fungal genomes runs on compute clusters at NERSC and is equipped with graphical monitoring tools. ***MycoCosm*** provides interactive web-based tools for functional and comparative genomics. ***Quality assurance*** is implemented at every stage of the annotation process: built-in pipeline controls, benchmarked tools, two-tier pre-release data/portal QC, user feedback, and community annotation.

| Pipelines Comparison | MAKER | JGI pipeline | JGI R&D |
|---|---|---|---|
| # of predicted gene models | 9,940 | 12,290 | 12,802 |
| with Swissprot hits | 6,521 | 7,356 | 7,900 |
| with PFAM domains | 5,365 | 6,010 | 6,353 |
| with EST support | 9,252 | 10,796 | 11,105 |
| | | | |
| # of unique PFAM domains | 2,207 | 2,245 | 2,322 |
| EST coverage per gene | 93.0% | 93.3% | 93.3% |
| # EST-supported splice sites | 99,627 | 102,200 | 104,246 |




Pipeline benchmarks, development, and release to production

### MycoCosm: 100+ fungal genomes




http://jgi.doe.gov/fungi
120+ fungal genomes
5000+ visitors/month
Grigoriev *et al.*, *NAR* 2011

### Community Outreach

*Jamborees* to analyze genomes

**Dothideomycetes,** Pacific Grove, CA Mar 15, 2011



**Basidiomycota,** Walnut Creek, CA Mar 21, 2011



**Symbiosis,** INRA-Nancy, France Sep 28-29, 2011

*MycoCosm Tutorials* to introduce the program and tools

MGM (3x@JGI)

MycoCosm Online video tutorial

Satellite Meetings

Close interactions with users to develop strategy and coordination

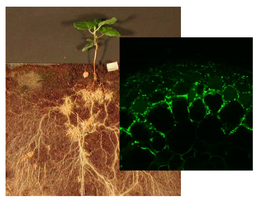
FGC-Asilomar and ECFG Meetings

JGI User Meetings

MSA Meetings

## Science: Genomic Encyclopedia of Fungi

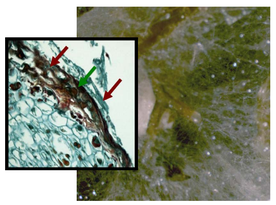
### Plant Health: Symbionts



*Laccaria bicolor* (ECM)

Small secreted proteins are expanded in poplar symbiont and overexpressed in ECM tissue.

Martin *et al.*, *Nature* 2008

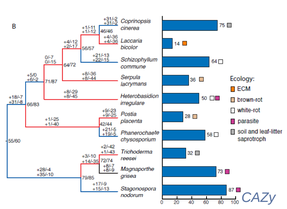


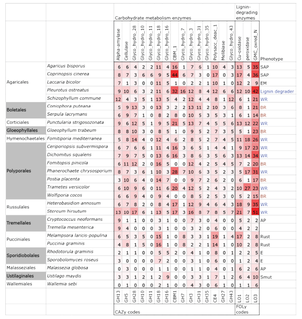
*Serpula lacrymans* (brown rot)

Convergent evolution of brown rot and ECM fungi from white rot ancestor.

Eastwood *et al.*, *Science* 2011

### Biorefinery: Lignocellulose Degradation

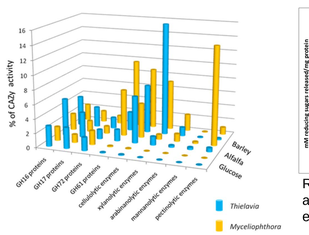


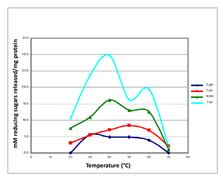




Genomes of diverse fungi of different lifestyles enrich a catalogue of enzymes involved in lignocellulose degradation

Riley *et al.*, *in prep.*

### Biorefinery: Thermophiles



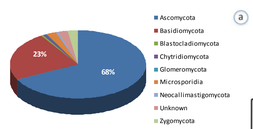



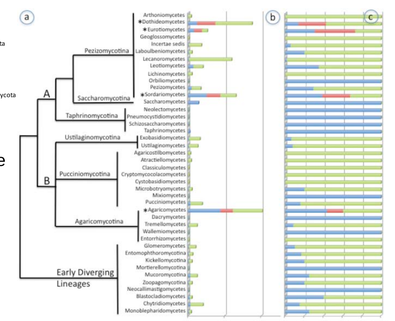

Release of reducing sugars from alfalfa straw by extracellular enzymes

Thermophilic biomass-degrading fungi *M.thermophila* and *Thielavia terrestris,* the first with completely finished genomes, are new candidates for cell factories with secreted thermostable enzymes.

Berka *et al.*, *Nat Biotech* 2011.

### Diversity: 1000 Fungal Genomes

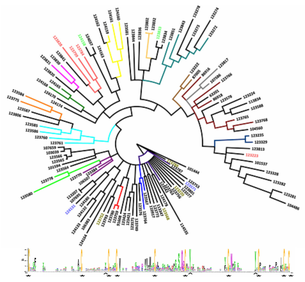


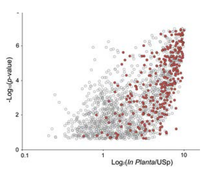




Sequencing 1000 fungal genomes from across the Fungal Tree of Life will provide references for research on plant-microbe interactions and environmental metagenomics.

Spatofora *et al.*, CSP12.

### Plant Health: Pathogens



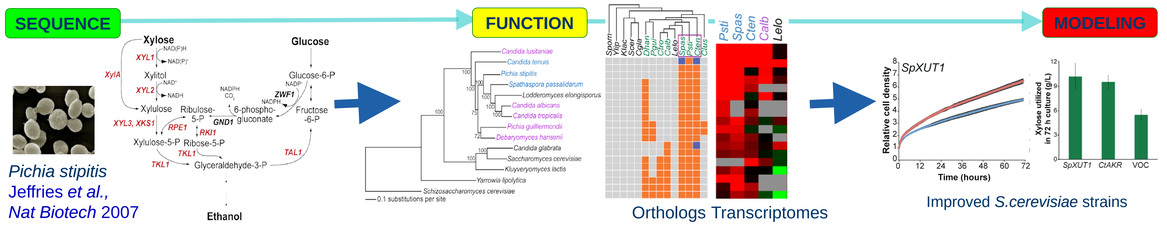



Small secreted proteins are expanded in poplar and wheat rust pathogens and overexpressed *in planta*.

Duplessis *et al.*, *PNAS* 2011

### Biorefinery: Engineering Improved Xylose Fermentation

SEQUENCE → FUNCTION → MODELING




*Pichia stipitis* Jeffries *et al.*, *Nat Biotech* 2007

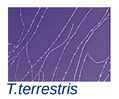
Orthologs Transcriptomes

Improved *S.cerevisiae* strains

Comparative genomics and transcriptomics of *xylose fermenters* and *xylose growers* revealed candidate genes for *S.cerevisiae* strain improvement.

Wohlbach, *et al. PNAS* 2011

### Fungal Systems & Metagenomes



*T.terrestris*

*S.commune*

*Lichen: alga+ fungus*

*ECM: plant+ fungus*

*Forest soil metagenomes*

Model fungi → Simple systems → Complex environments

Understanding fungal systems requires developing molecular and analytical tools for model organisms, simple systems and complex environments


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.