



Phyloinformatics of Fungi in the Built (and other) Environment(s)

David Hibbett
Clark University, Worcester MA USA



Phyloinformatics of Fungi in the Built (and other) Environment(s)

David Hibbett
Clark University, Worcester MA USA

Goals: ...*The goal of the workshop is to pinpoint specific areas of inquiry in which basic research could most impact society's understanding of fungi in the built environment.*

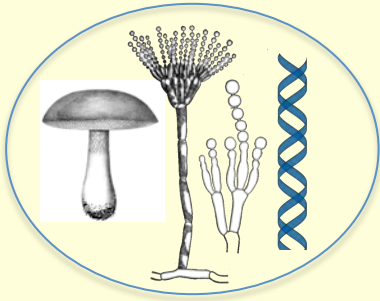
I. Phylogeny and taxonomy.

II. Three attempts:

- ❖ *mor* (past)
- ❖ Open Tree of Life (now)
- ❖ ~~Community Phylogenetics of Agaricomycotina~~

I. Phylogeny and taxonomy. workflow

Environmental Sampling



**Collections,
cultures**

Experimental
resources,
vouchers

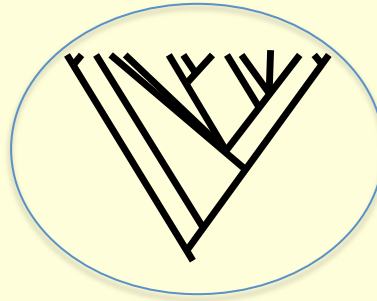
acgatcaccgaagatgtgcatgc
acgatcaccgaagatgatgc
acgatcattacgatcaccgaaga
gaagagagcatgcattacgatca
gaagaattacgatctgattgcat
acgatcaccgaagatgatgc
gaagatgatgcagatgcagatc
gaagtgatgcagatgcagatc
gaagtgatgcagatgcagatc

Sequences

*MOTHUR,
*UNITE, etc

*Ecological tools/community comparison

Phylogenetic Inference



Trees with branch
lengths, support.

*UniFrac,
comparative analyses,
molecular clock,
biogeography, etc.

Classification

Polyporales
..... Polyporaceae,
..... *Trametes*
..... *T. versicolor*
..... *T. ochracea*
..... *T. ectypa*
..... *Ganoderma*
..... *G. lucidum*
..... *G. tsugae*

**Hierarchical
class., names**

*Naïve Bayesian Class.
Communication with:
clinicians, quarantine
officials, plant pathologists,
legislators, educators,
informed citizens

I. Phylogeny and taxonomy. challenges

Construct maximally inclusive, continuously updated phylogenetic trees, with branch lengths and support values.

Translate phylogenetic trees into classifications on an automated basis (with community input).

Disseminate products, and recruit non-specialists (clinicians, pathologists, amateur mycologists, et al.) to provide materials.

II. Three attempts:

- ❖ *mor* AFTOL1 (NSF) 2002-2006.

Concept:

Continuously updated automated phylogenetic reconstruction, using nuc-lsu rRNA genes with backbone constraint.

Automated tree parsing using phylogenetic taxon definitions (least common ancestor statements).

Hibbett, D. S., R. H. Nilsson, M. Snyder, M. Fonseca, J. Costanzo, and M. Shonfeld. 2005.

Automated phylogenetic taxonomy: an example in the homobasidiomycetes (mushroom-forming Fungi). *Systematic Biology* 54: 660-668).



mor

Cladesys Login

E-mail: Password:

explore

Mor Home
Mor Flow Charts
Download Mor

trees

Parsimony Tree
Jackknife
Neighbor-Joining Tree
Archived Trees and Alignments
Recent Changes in the Tree
Tree Visualization

sequences

Current Sequences
Rejected Sequences

interact

Comments and Suggestions
Register for an Account
Login to Create/Edit Clades

miscellaneous

Links

development

The mor Team
Development Blog

welcome to mor

Welcome. *mor* is an ongoing project based in the [Hibbett lab](#) of [Clark University](#) that seeks to produce continually updated phylogenetic trees and classifications for the homobasidiomycetes (mushroom-forming fungi) on an automated basis. In brief, *mor* retrieves, screens, aligns, and analyzes nuc-lsu rDNA sequences of homobasidiomycetes from [GenBank](#), and then parses out the contents of individual clades using node-based phylogenetic taxon definitions. Thus, *mor* represents a step toward automated phylogenetic taxonomy. For more information about homobasidiomycete phylogeny, visit the [Tree of Life Web Project](#) page for [homobasidiomycetes](#).

mor creates two phylogenetic trees each week: an unconstrained jackknife neighbor-joining tree, and a topologically constrained consensus maximum parsimony tree. The parsimony tree is used to delimit the clades.

On the following pages, you can view the NJ and MP trees, link to GenBank, search for images of fungi on the internet, view individual clade pages, download alignments and treefiles (for the entire dataset or individual clades), post comments, and create new clade definitions. Please note that some files are quite large (about 6 MB) and can take some time to load.

sequences in the database

View List of Current...

Accepted Sequences

(sorted by...)

name	accession number	date

Rejected Sequences

(sorted by...)

name	accession number	date	reason

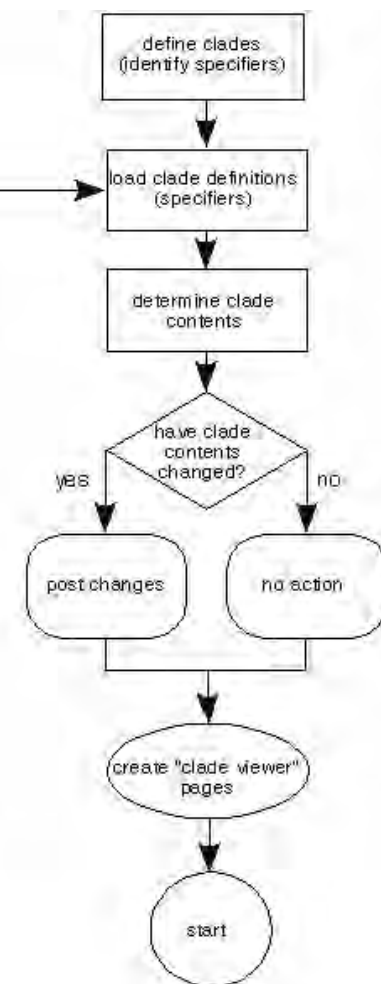
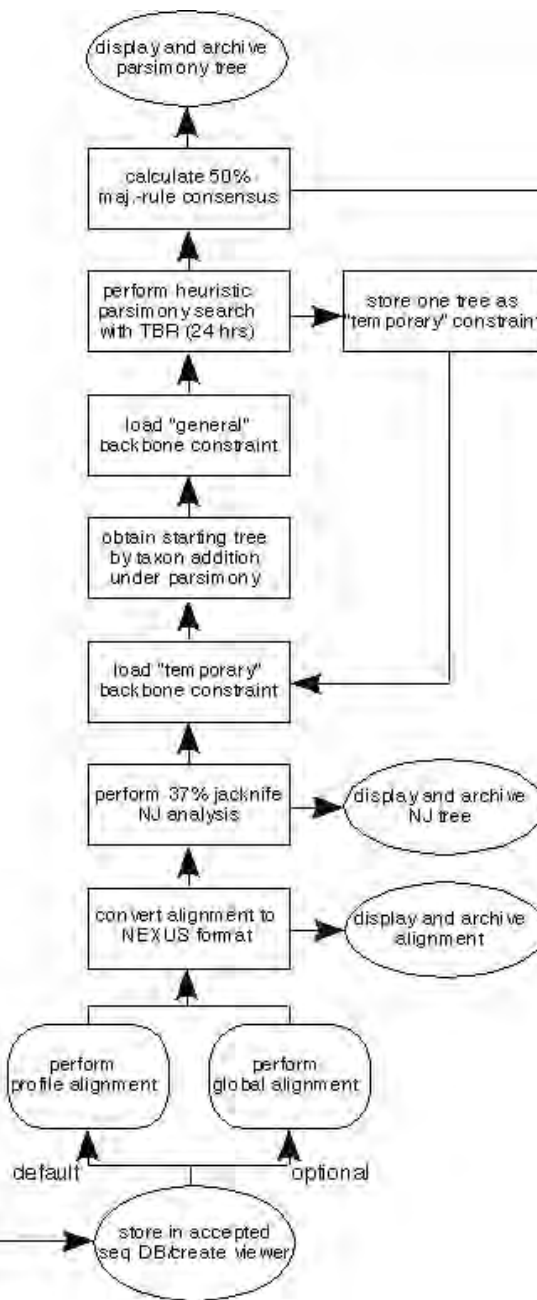
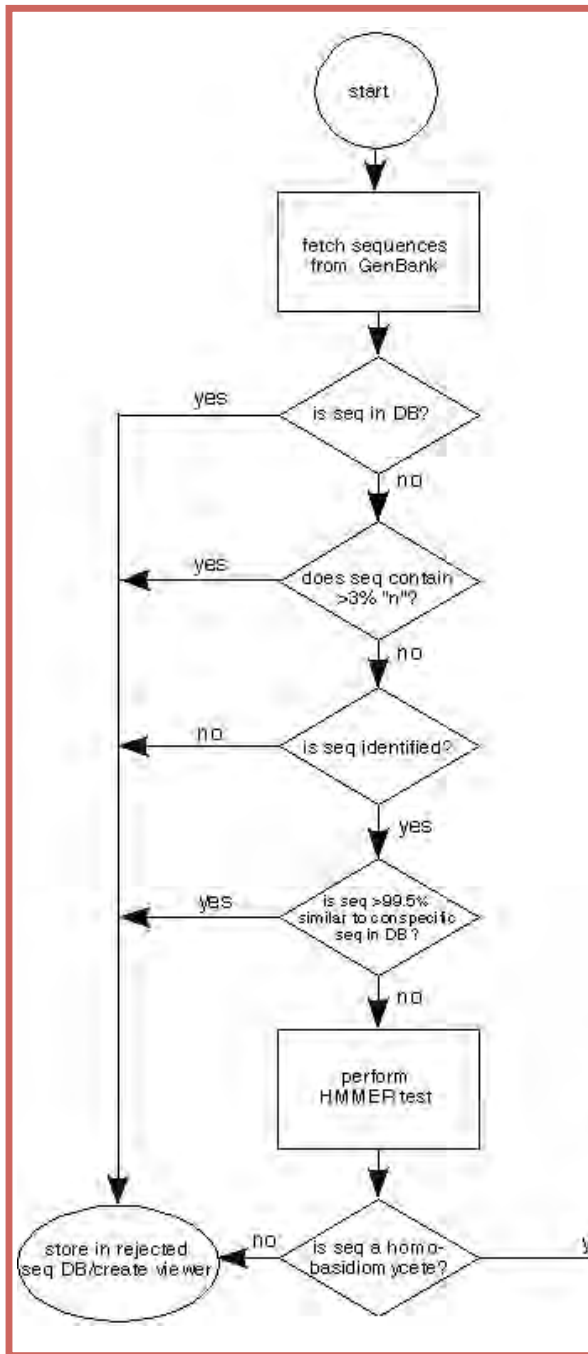
main clades

citation

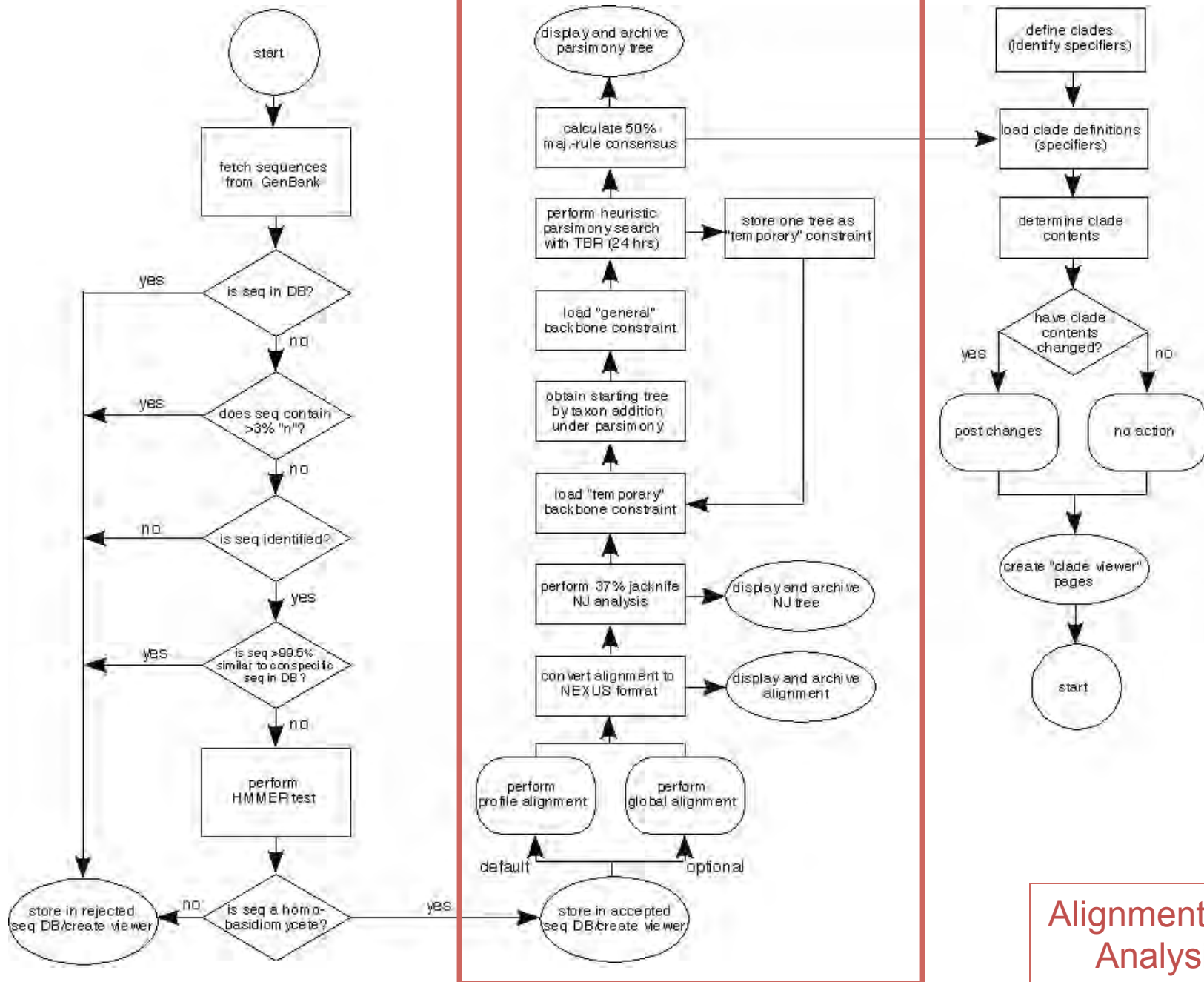
Hibbett, D. S., R. H. Nilsson, M. Snyder, M. Fonseca, J. Costanzo, and M. Shonfeld. 2005. Automated Phylogenetic Taxonomy: An Example in the Homobasidiomycetes (Mushroom-Forming Fungi). *Systematic Biology* 54: 660-668. [Download reprint \(PDF\)](#).

news

Read about mor and WASABI in *Science*: Pennisi, E. 2005. Fungal trees grow faster with computer help. *Science* 309: 374 (meeting report). [Download PDF](#).



Acquisition and Screening



Alignment and Analysis

Total number of sequences: 3009

HRI = n ; the H Rejection Index; n distinct Genbank sequences with the same species name have been tested and found to be over 99.5% identical to this sequence, and hence discarded.

Length: 61490 CI: 0.070467 RI: 0.842812
RC: 0.059390 HI: 0.929533 GFit: -356.940741

```

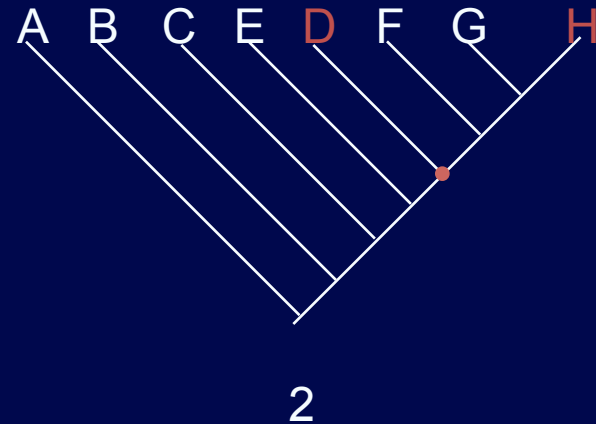
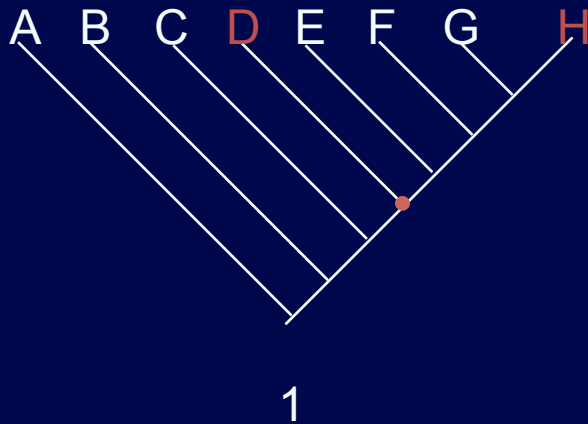
    /- Calvatia sp (AF261482) image? comments? ni(1)
    100
    | \- Calvatia gigantea (AF518603) image? comments? ni(2)
    100
    | \- Lycoperdon sp (AF287873) image? comments? ni(8)
    100
    | \--- Bovista sp (AF261483) image? comments? ni(3)
    |
    | /- Lycoperdon marginatum (AF261485) image? comments? ni(4)
/-----100100
| | \- Lycoperdon perlatum (AJ406583) image? comments? ni(5)
| 100
| | \- Lycoperdon perlatum (AF518630) image? comments? ni(6)
| 100
| \--- Lycoperdon coloratum (AF261484) image? comments? ni(7)
|
| /----- Stagnicola perplexa (AF261509) image? comments? ni(464)
|
| /- Pluteus brunneoradiatus (AF261567) image? comments? ni(2622)
|
| 100- Pluteus cervinus (AF261570) image? comments? ni(2623)
|
| | \- Pluteus pallidus (AF261569) image? comments? ni(2624)
|
| 100
| +- Pluteus pouzarianus (AF261568) image? comments? ni(2625)

```


Parsing a node-based phylogenetic taxon definition:

*“Taxon X is the least-inclusive clade that contains **D** and **H**”.*

D and **H** are “specifiers”



Tree 1 (A(B(C(**D**(E(F(G,**H**))))))) Taxon X = **D**, E, F, G, **H**

Tree 2 (A(B(C(E(**D**(F(G,**H**))))))) Taxon X = **D**, F, G, **H**

II. Three attempts:

- ❖ *mor* AFTOL1 (NSF) 2002-2006.

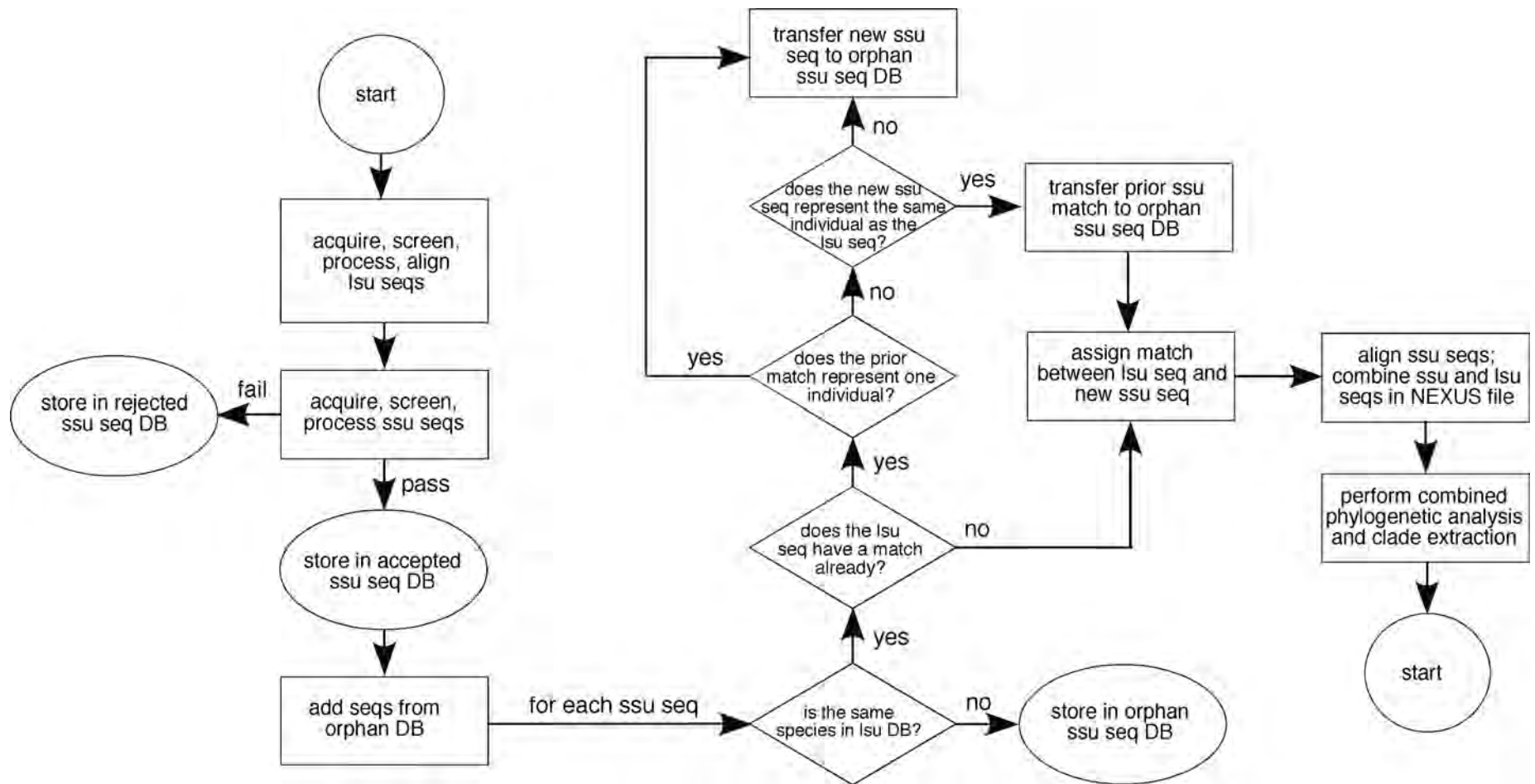
Limitations:

No ITS data; multigene data only included in constraint.

Proposed solution:

Automated supertree analysis.

Matchmaker pipeline to combine LSU sequences to other genes.



NSF BDI (2005) New tools for automated phylogenetic taxonomy as applied to Homobasidiomycetes.

II. Three attempts:

- ❖ **Open Tree of Life** AVAToL (NSF) 2012-2015

Concept:

Synthesize published **phylogenies** and **taxonomy** (Index Fungorum, NCBI, etc) using graph database.

Facilitate community updates of synthetic phylogeny.

Ten PIs, three years.



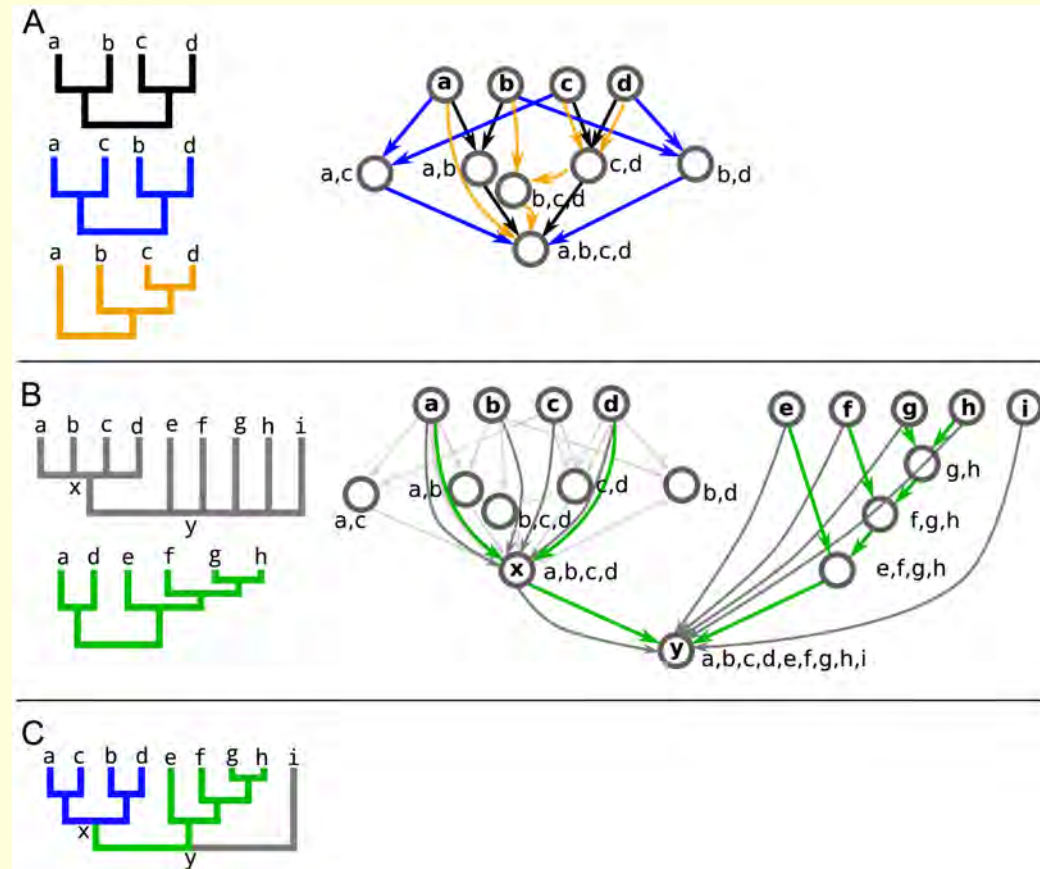
<http://blog.opentreeoflife.org/>





Graph building method

- Alternative method to supermatrix and supertree approaches
- Allows merging of multiple phylogenies with incomplete overlapping taxa (at least one taxa in common)
- Resolution of conflict can be automated based on specific criteria (i.e., year, # of genes, # taxa)



Smith, Brown, Hinchliff. 2013. Analyzing and synthesizing phylogenies using tree alignment graphs. PLOS Computational Biology.



Status of the project

- Synthetic taxonomy with >2.5 million taxa
→ OpenTreeTaxonomy = OTT v. 2.6 (now v. 2.8)
- 1st version of the synthetic tree with all taxonomy taxa plus phylogenetic resolution from > 300 studies (May, 2014)
- 2nd version of the synthetic tree with all taxonomy taxa plus phylogenetic resolution from > 500 studies (July, 2014)
- Several thousand other studies in database in line for synthesis



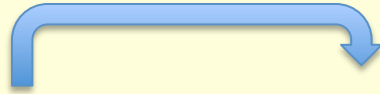
Status for Fungi

- Synthetic taxonomy
 - OTT v. 2.8 – mainly based on Index Fungorum
Tips: 346,480 / 65,153 hidden (excl. incertae sedis, fossils)
- Synthetic tree contains only 12 phylogenetic studies
 - Version 2 = ~60 fungal tree sources
- ~ 1,000 studies waiting to be included in synthesis

Updated here: <http://blog.opentreeoflife.org/>

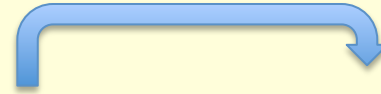


Challenges (Fungi):



DATA availability

Low number of available phylogenies (~17%)

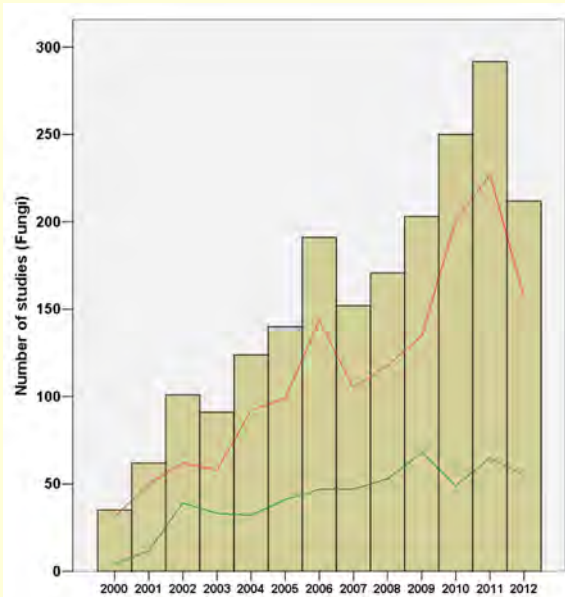


DATA curation

Rooting of trees
Mapping of terminals
Determine ingroup

Real conflict?

Real conflict because of gene history OR artifact of current taxonomy



- Most of the studies needed extra curation (re-rooting, label editing and mapping of terminals)
- Conflicts between taxonomic databases (**IF** / GBIF / NCBI)

- If there is too much conflict with the OTT taxonomy the trees do not pass to synthesis

Drew, Gazis, Cabezas, Swithers, Soltis, Hibbett, Crandall, Katz. 2013. Lost branches on the tree of life. PLOS Biology.



People:



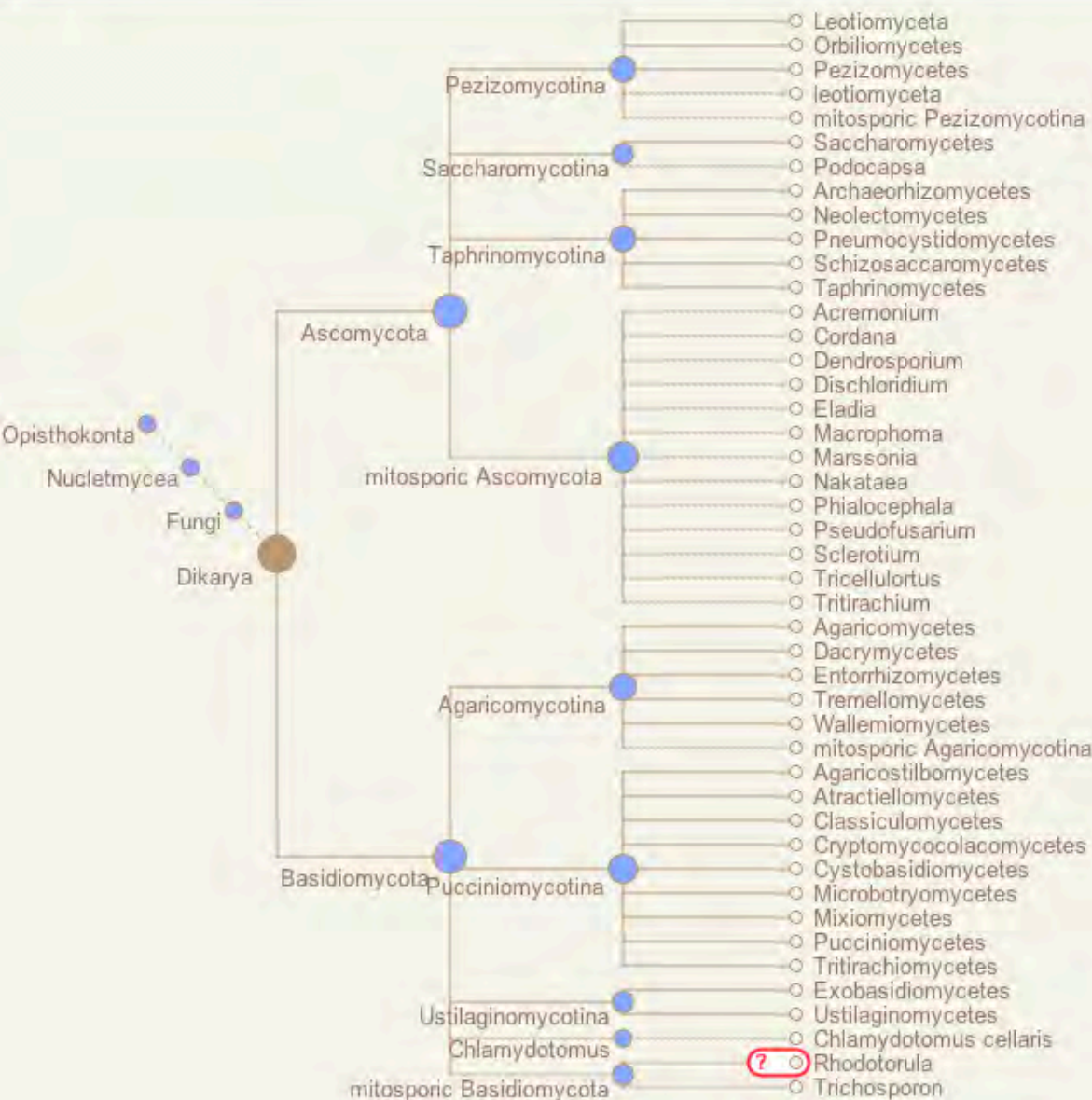
Jiaqi Mei

Romina Gazis



Dikarya

0 Show comments

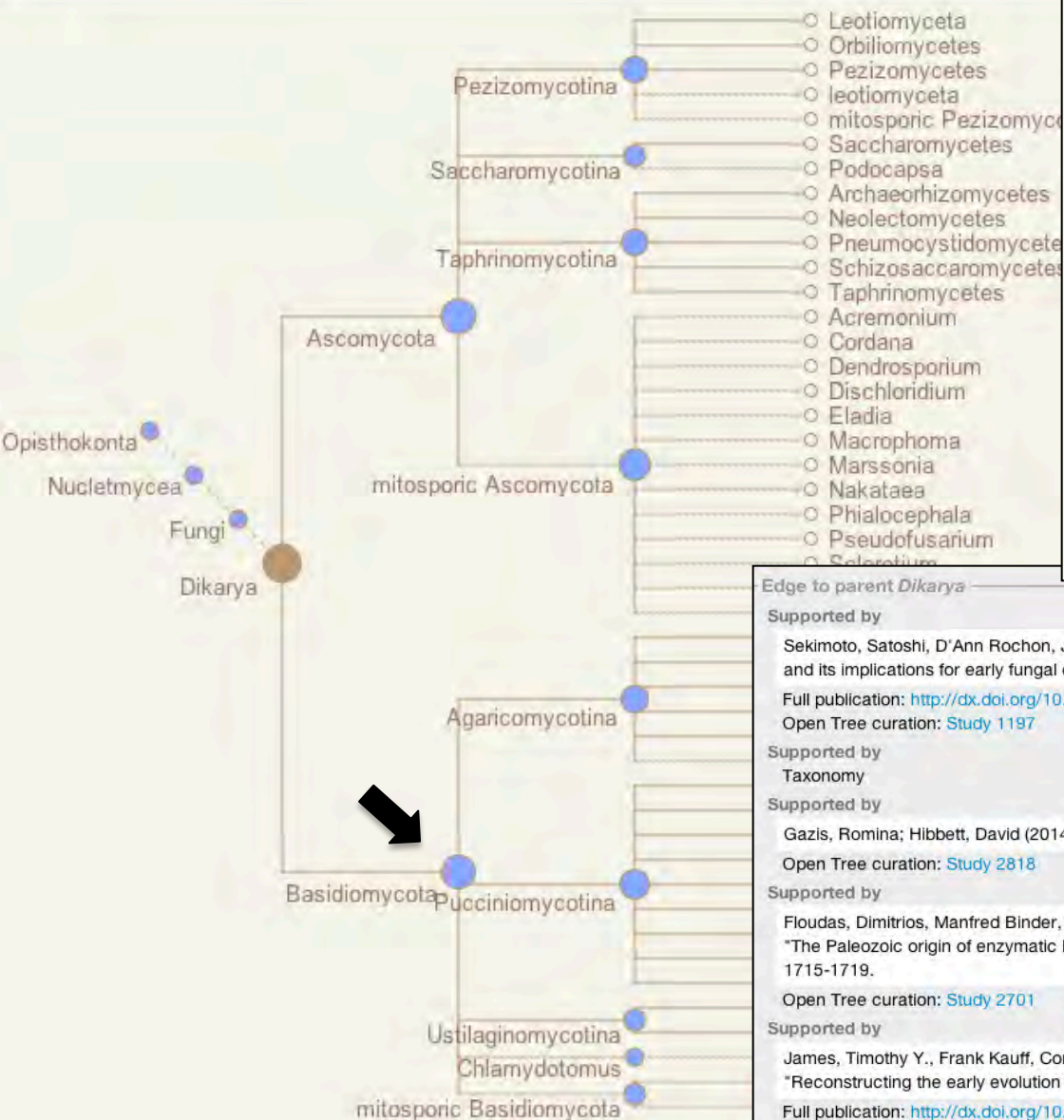




Dikarya

Exploring the current synthetic

Show comments



Basidiomycota

Node properties

Source taxonomy
[Hibbett et al. 2007 updated: 13](#)

Source taxonomy
[Index Fungorum: 90050](#)

Source taxonomy
[NCBI: 5204](#)

Source taxonomy
[GBIF: 34](#)

Source taxonomy
[IRMNG: 236](#)

Taxonomic rank
phylum

Terminal taxa within this clade
20

[Extract subtree](#)
(depth limited to 4 levels)

Edge to parent Dikarya

Supported by

Sekimoto, Satoshi, D'Ann Rochon, Jennifer E. Long, Jaclyn M. Dee, and Mary L. Berbee. "A multigene phylogeny of *Olpidium* and its implications for early fungal evolution." *BMC Evolutionary Biology* 11, no. 1 (2011): 331.

Full publication: <http://dx.doi.org/10.1186/1471-2148-11-331>

Open Tree curation: [Study 1197](#)

Supported by

Taxonomy

Supported by

Gazis, Romina; Hibbett, David (2014): Fungi Classification Update 2014. figshare. <http://dx.doi.org/10.6084/m9.figshare.915439>

Open Tree curation: [Study 2818](#)

Supported by

Floudas, Dimitrios, Manfred Binder, Robert Riley, Kerrie Barry, Robert A. Blanchette, Bernard Henrissat, Angel T. Martinez et al. "The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes." *Science* 336, no. 6089 (2012): 1715-1719.

Open Tree curation: [Study 2701](#)

Supported by

James, Timothy Y., Frank Kauff, Conrad L. Schoch, P. Brandon Matheny, Valérie Hofstetter, Cymon J. Cox, Gail Cello et al. "Reconstructing the early evolution of Fungi using a six-gene phylogeny." *Nature* 443, no. 7113 (2006): 818-822.

Full publication: <http://dx.doi.org/10.1038/nature05110>

Open Tree curation: [Study 1162](#)



Lentinus

0 Show comments

- ☐ Lentinus berteri
- ☐ Lentinus berteri
- ☐ Lentinus bertieri
- ☐ Lentinus bogoriensis
- ☐ Lentinus bonariensis
- ☐ Lentinus braccatus
- ☐ Lentinus brunneofloccosus
- ☐ Lentinus brunneomaculatus
- ☐ Lentinus caelopus

more... (Lentinus caesariatus - Lentinus isabellina)

more... (Lentinus japonicus - Lentinus scyphoides)

more... (Lentinus sericeosquamosus - Pocillaria coriacea)

- ☐ Pocillaria crassiuscula
- ☐ Pocillaria crinita
- ☐ Pocillaria cubae
- ☐ Pocillaria cuneata
- ☐ Pocillaria curtisii
- ☐ Pocillaria dactylophora
- ☐ Pocillaria denticulata
- ☐ Pocillaria dichotoma
- ☐ Pocillaria divisa
- ☐ Pocillaria dorsalis
- ☐ Pocillaria echinopus
- ☐ Pocillaria exasperata
- ☐ Pocillaria exilis



Index Fungorum

Search by:-

Name Epithet Genus Family higher Enter a search term:-



Lentinus

Search

Name, Author, Year, (Current name), Parent taxon

Pages: 1 2 3 4 [Next >>] of 628 records. TofP BofP

[Lentinus abnormis](#) Berk. 1878, (also see Species Fungorum: [Lentinus connatus](#)); Polyporaceae[Lentinus abnormis](#) (Alb. & Schw.) Berk. 1878 (also see Species Fungorum: [Lentinus connatus](#)); Polyporaceae

Pocillaria ianthina

Exploring the cur

Lentinus

Node properties

Source taxonomy

[Index Fungorum: 17925](#)

Source taxonomy

[NCBI: 5357](#)

Source taxonomy

[GBIF: 2526675](#)

Source taxonomy

[IRMNG: 1281628](#)

Taxonomic rank

genus

Terminal taxa within this clade

432

[Extract subtree](#)

(depth limited to 4 levels)



493139 records on-line

[add new record](#)

Polyporales

Lentinus

II. Three attempts:

- ❖ **Open Tree of Life** AVAToL (NSF) 2012-2015

Summary of limitations:

Too few phylogenies in TreeBASE, Dryad, etc.

No branch lengths or support values in synthetic tree.

Errors in taxonomic databases.

Lots of manual curation still required....and time is running out.



Conclusions:

27 years after PCR and 12 years after AFTOL1, there is still **no** web-accessible comprehensive fungal phylogeny integrating specimen/culture-based and environmental sequences. No centralized tree-based taxonomic resource.

Lack of phylogenetic synthesis limits progress in all downstream uses of trees, hierarchical classifications, and names, including: ecology, medical mycology/epidemiology, plant pathology, quarantine management, biogeography, conservation/climate change biology, bioprospecting, and education.

**But what about
species recognition?...**

