

# Lawrence Berkeley National Laboratory

## LBL Publications

### Title

Comparative Genome Analysis of Basidiomycete Fungi

### Permalink

<https://escholarship.org/uc/item/3v02b9z3>

### Authors

Riley, Robert  
Salamov, Asaf  
Morin, Emmanuelle  
et al.

### Publication Date

2012-03-20

## **Comparative genome analysis of Basidiomycete fungi**

Robert Riley<sup>\*1</sup>, Asaf Salamov<sup>1</sup>, Emmanuelle Morin<sup>2</sup>, Laszlo Nagy<sup>3</sup>, Gerard Manning<sup>4</sup>, Scott Baker<sup>5</sup>, Daren Brown<sup>6</sup>, Bernard Henrissat<sup>7</sup>, Anthony Levasseur<sup>7</sup>, David Hibbett<sup>3</sup>, Francis Martin<sup>2</sup>, Igor Grigoriev<sup>1</sup>

<sup>1</sup>DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598

<sup>2</sup>Institut National de la Recherche Agronomique

<sup>3</sup>Department of Biology, Clark University, 950 Main Street, Worcester, MA 01610

<sup>4</sup>Razavi-Newman Center for Bioinformatics, Salk Institute for Biological Sciences, San Diego, CA 92186

<sup>5</sup>Pacific Northwest National Laboratory, PO Box 999, Richland, WA 99352

<sup>6</sup>USDA, ARS, MWA, NCAUR, BFPM, 1815 N University Street, Peoria, IL, 61604

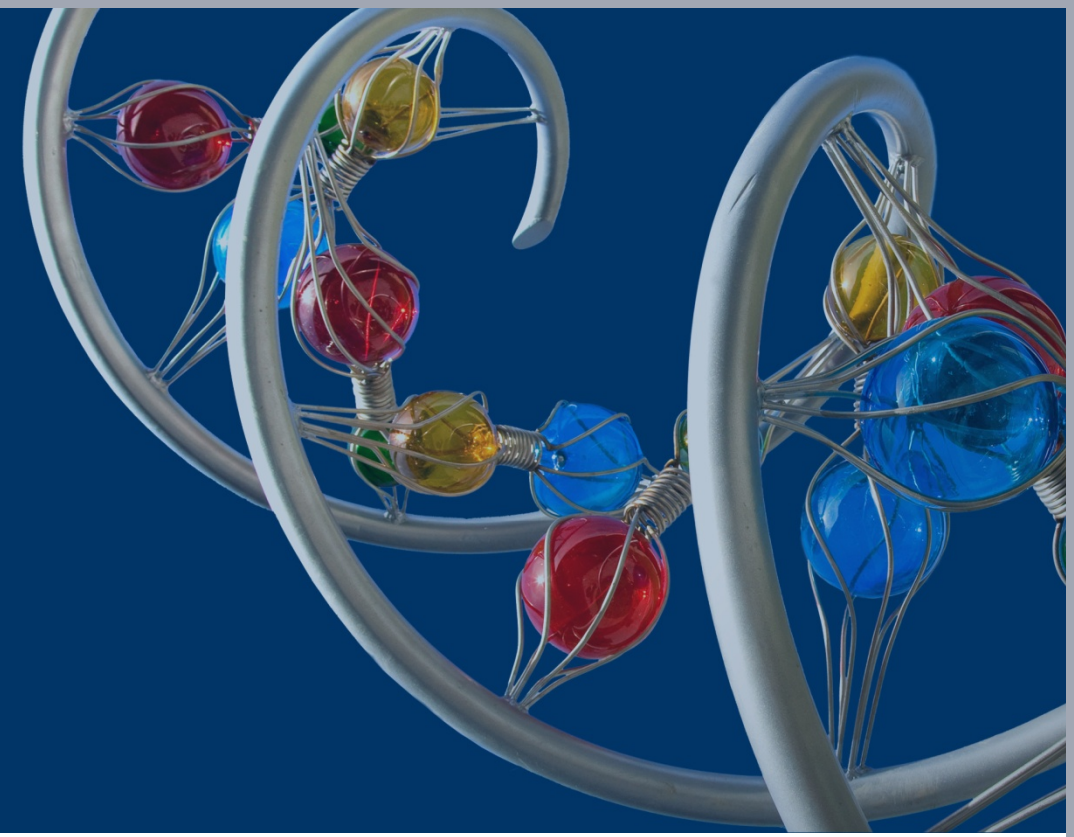
<sup>7</sup>AFMB UMR 6098 CNRS/UI/UII, Case 932,163 Avenue de Luminy,13288 Marseille cedex 9, France

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

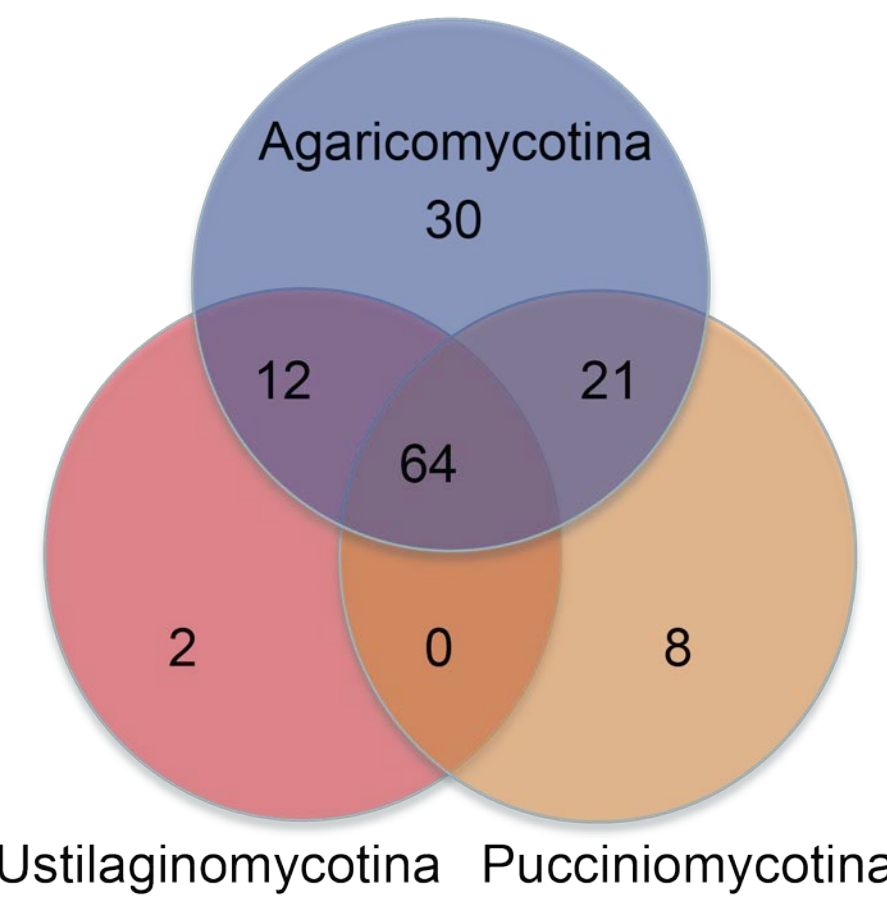
## **DISCLAIMER**

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.



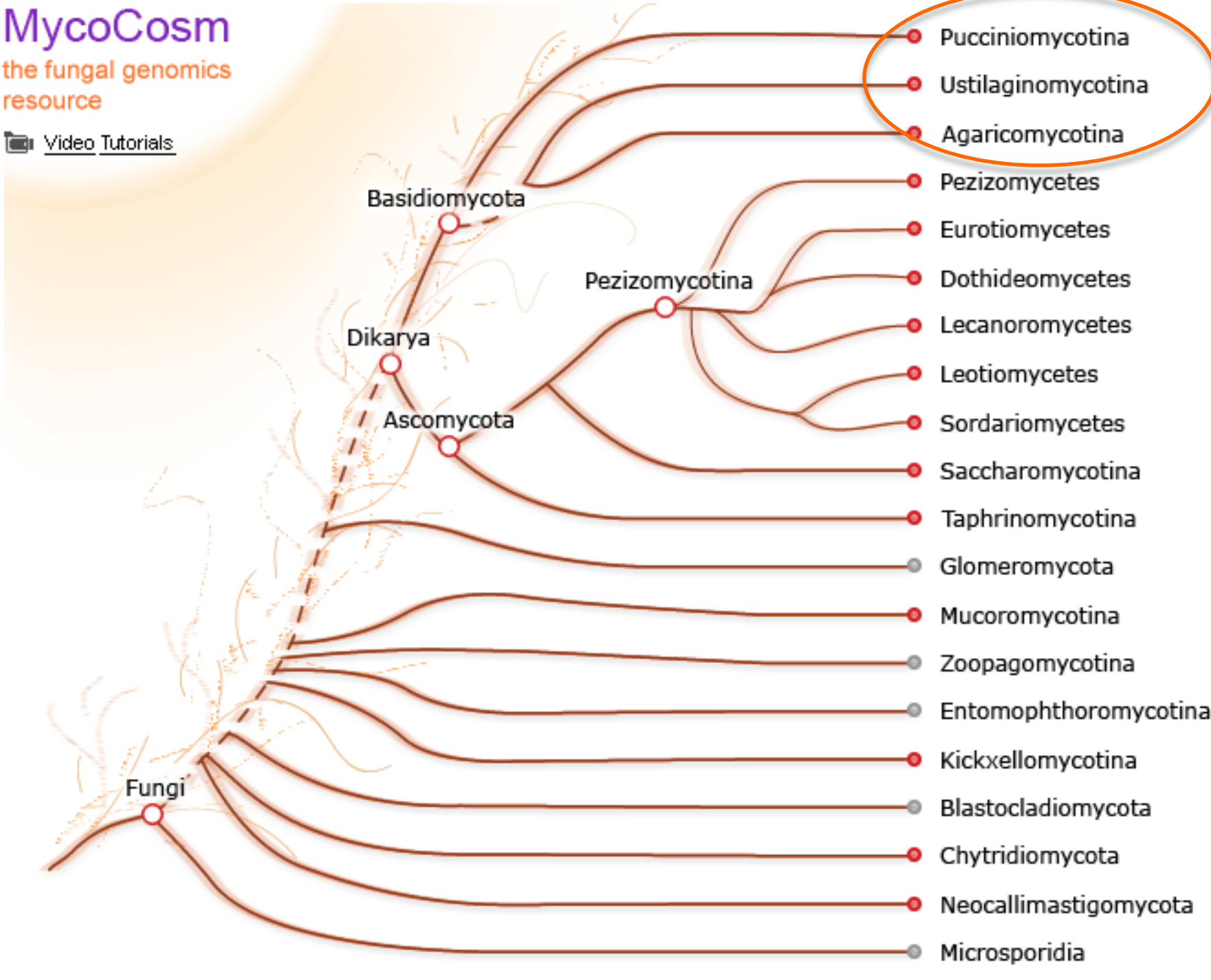
## Abstract

Fungi of the phylum Basidiomycota (basidiomycetes), make up some 37% of the described fungi, and are important in forestry, agriculture, medicine, and bioenergy. This diverse phylum includes the mushrooms, wood rots, symbionts, and plant and animal pathogens. To better understand the diversity of phenotypes in basidiomycetes, we performed a comparative analysis of 35 basidiomycete fungi spanning the diversity of the phylum. Phylogenetic patterns of lignocellulose degrading genes suggest a continuum rather than a sharp dichotomy between the white rot and brown rot modes of wood decay. Patterns of secondary metabolic enzymes give additional insight into the broad array of phenotypes found in the basidiomycetes. We suggest that the profile of an organism in lignocellulose-targeting genes can be used to predict its nutritional mode, and predict *Dacryopinax sp.* as a brown rot; *Botryobasidium botryosum* and *Jaapia argillacea* as white rots.

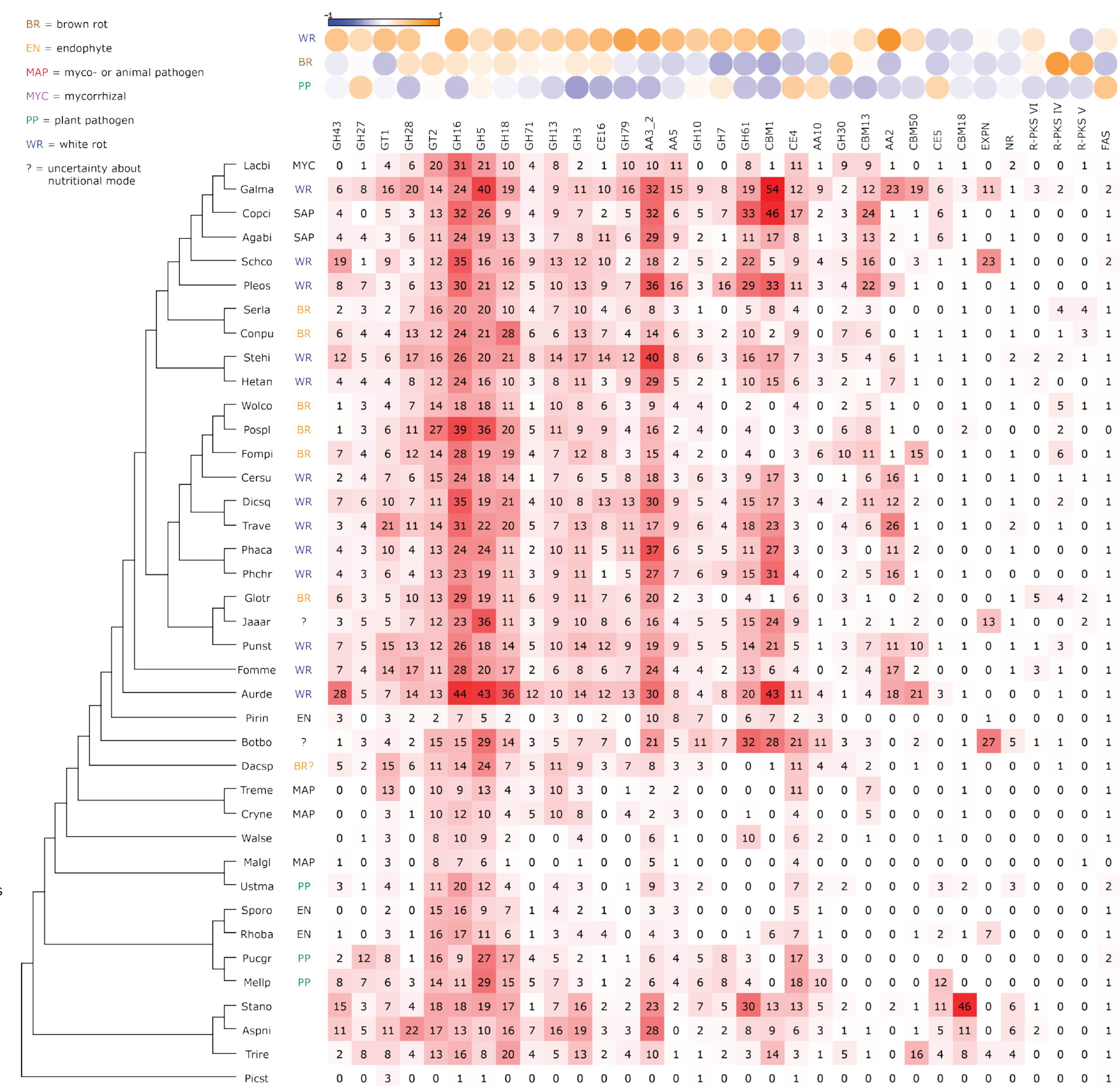


**Distribution of CAZymes in basidiomycetes.** The majority of CAZymes (carbohydrate-active enzymes) are found in Agaricomycotina, with the largest proportion shared by the three subphyla. Pucciniomycotina and Ustilaginomycotina have just a few CAZymes unique to them.

by Keyword Search Genome

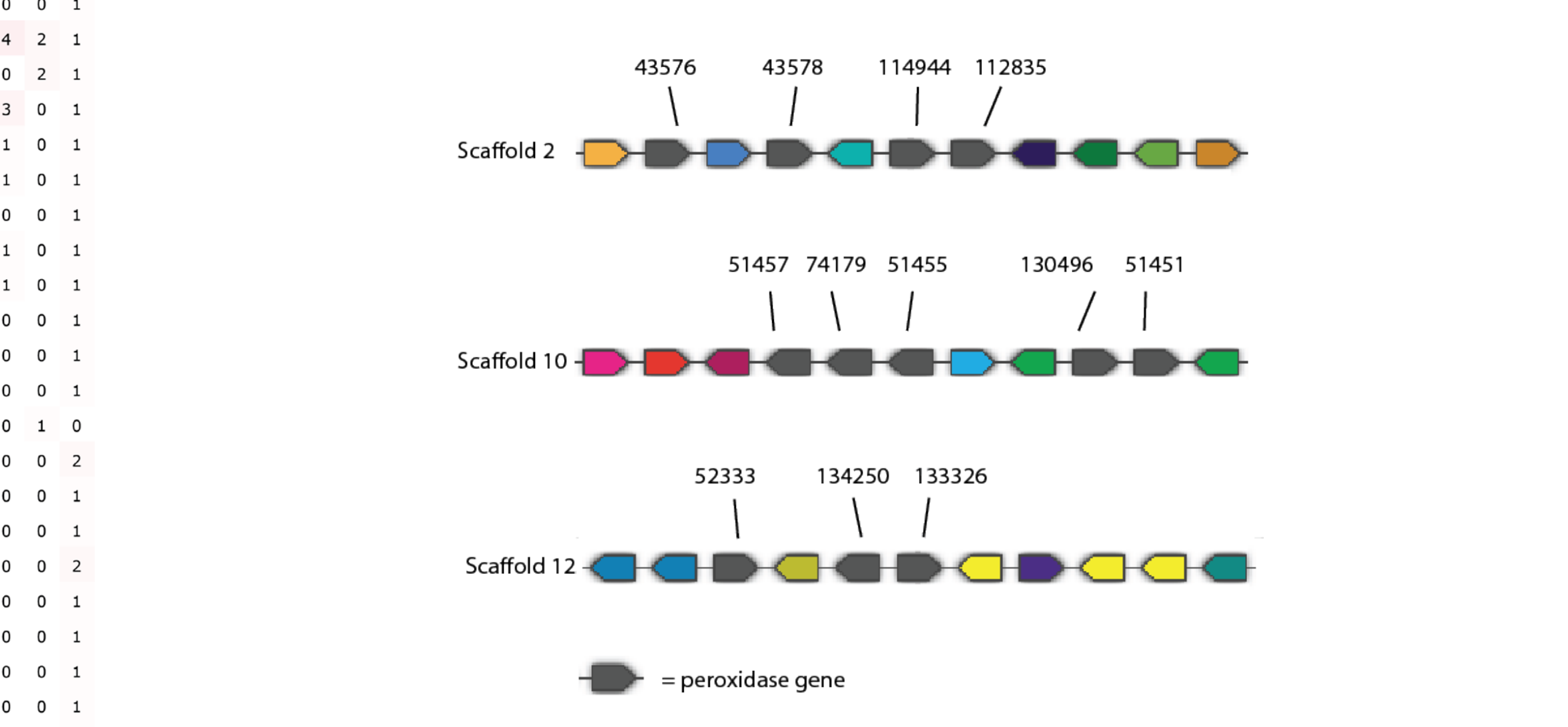
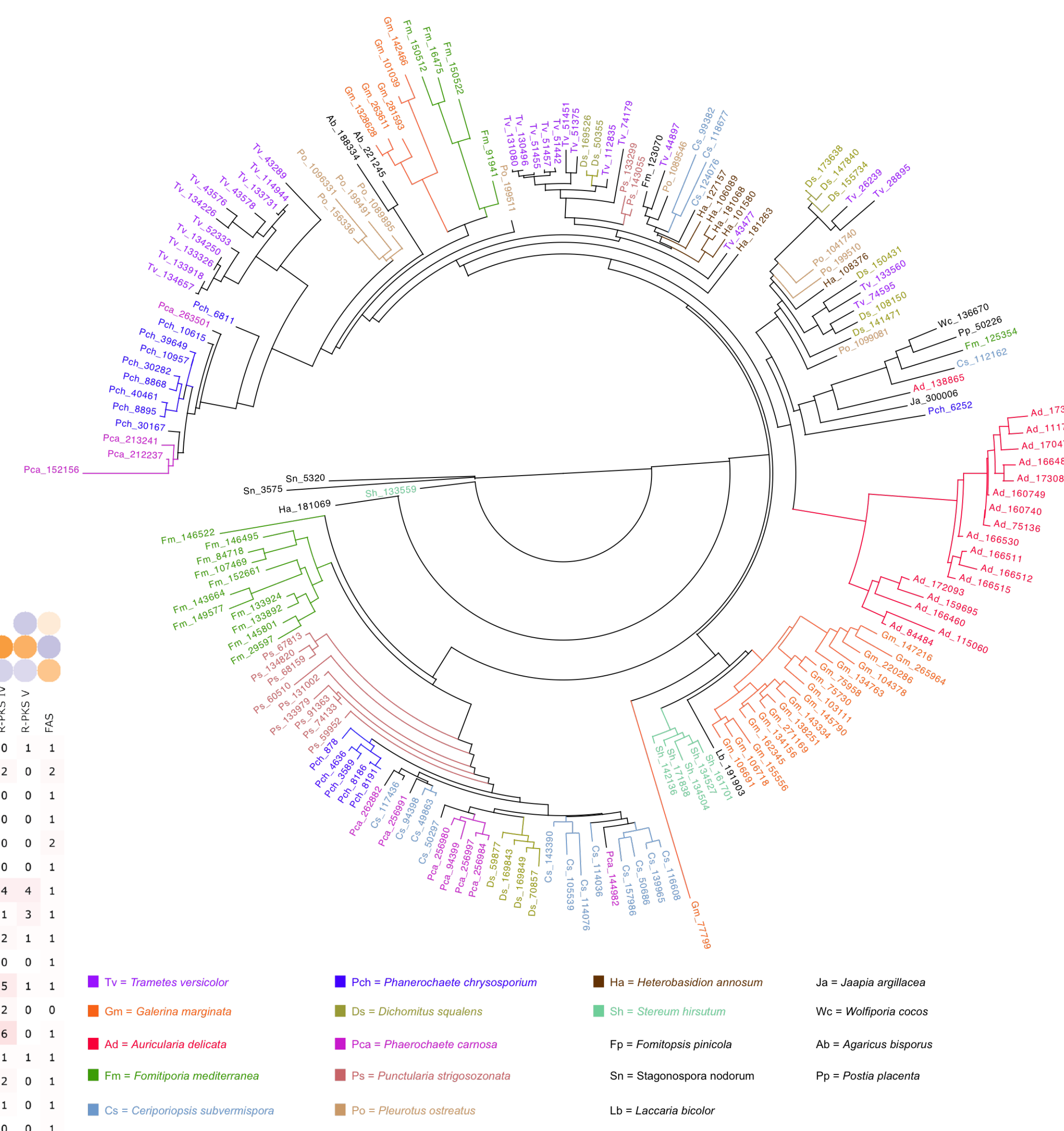


All genome data were downloaded from [jgi.doe.gov/fungi](http://jgi.doe.gov/fungi)



**Correlating genotype with degradative phenotype.** Distribution of 28 CAZy families and lignin-active enzymes (AA) identified as nonrandomly evolving in basidiomycetes, and selected secondary metabolite protein families. Each organism's number of genes from each family is shown in the white/red map. The correlation of each family with the white rot (WR), brown rot (BR), and plant pathogen (PP) phenotype is shown in the blue/orange map.

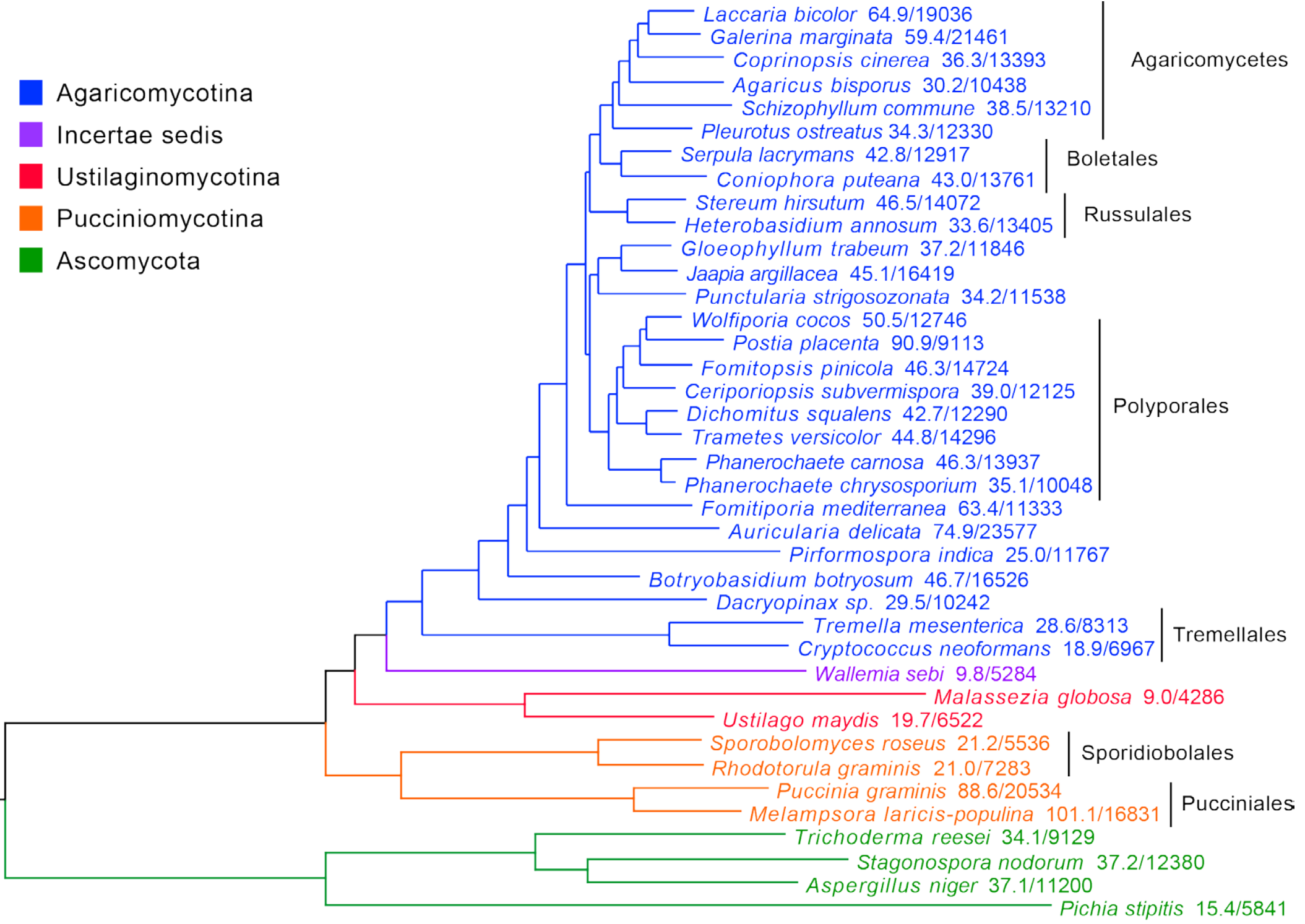
Notice that the AA2 (lignin peroxidase), AA3\_2 (cellobiose dehydrogenase), and several CAZy families IV and V correlate with the BR phenotype.



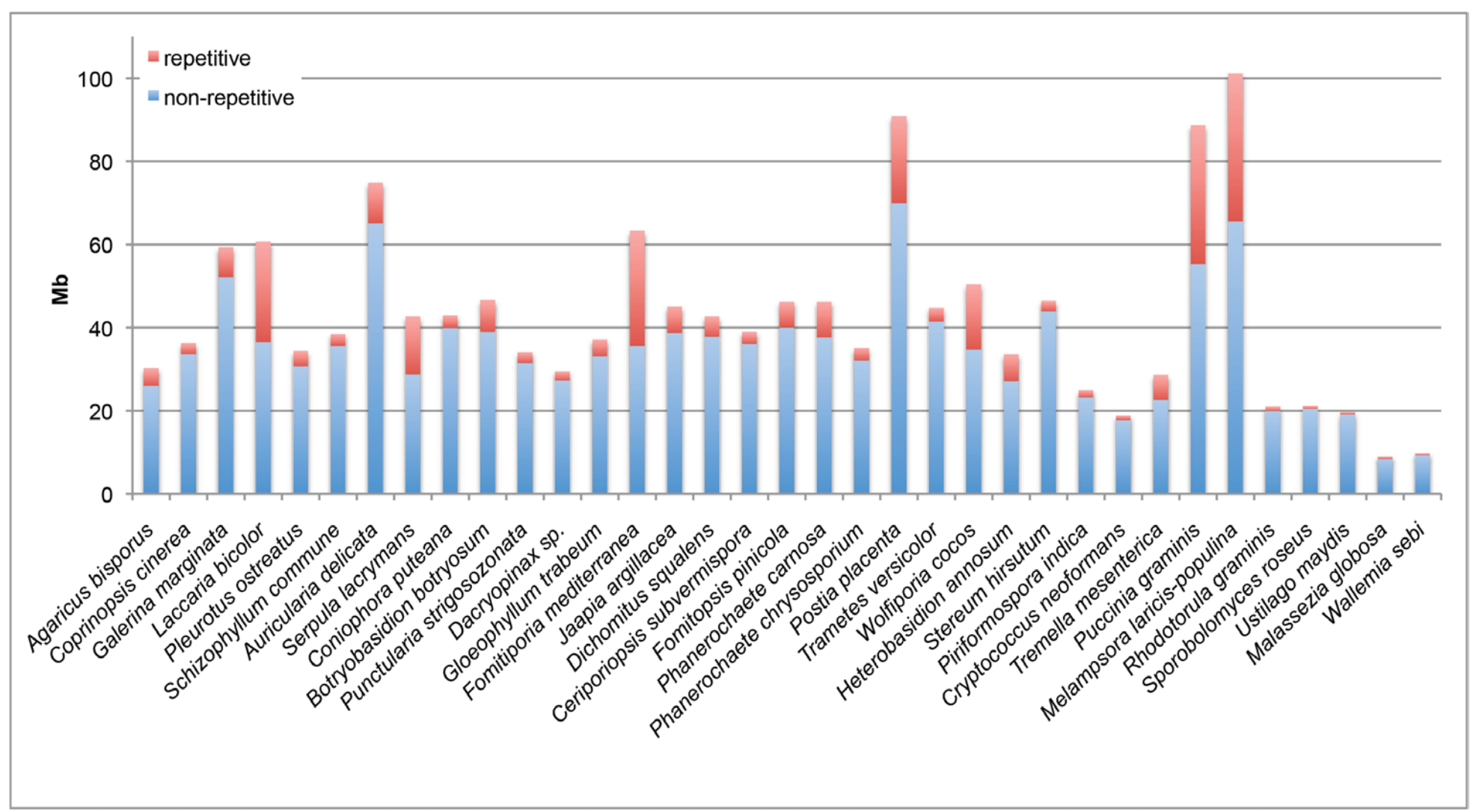
**Phylum-wide expansions in lignin peroxidases.** Maximum likelihood tree constructed from lignin peroxidase protein sequences. Sequences from white rots are colored by organism, with mono- and paraphyletic regions shaded accordingly.

Notice large organism-specific expansions in *Galerina marginata*, *Auricularia delicata*, *Fomitiporia mediterranea*, *Trametes versicolor*, and *Phanerochaete chrysosporium*.

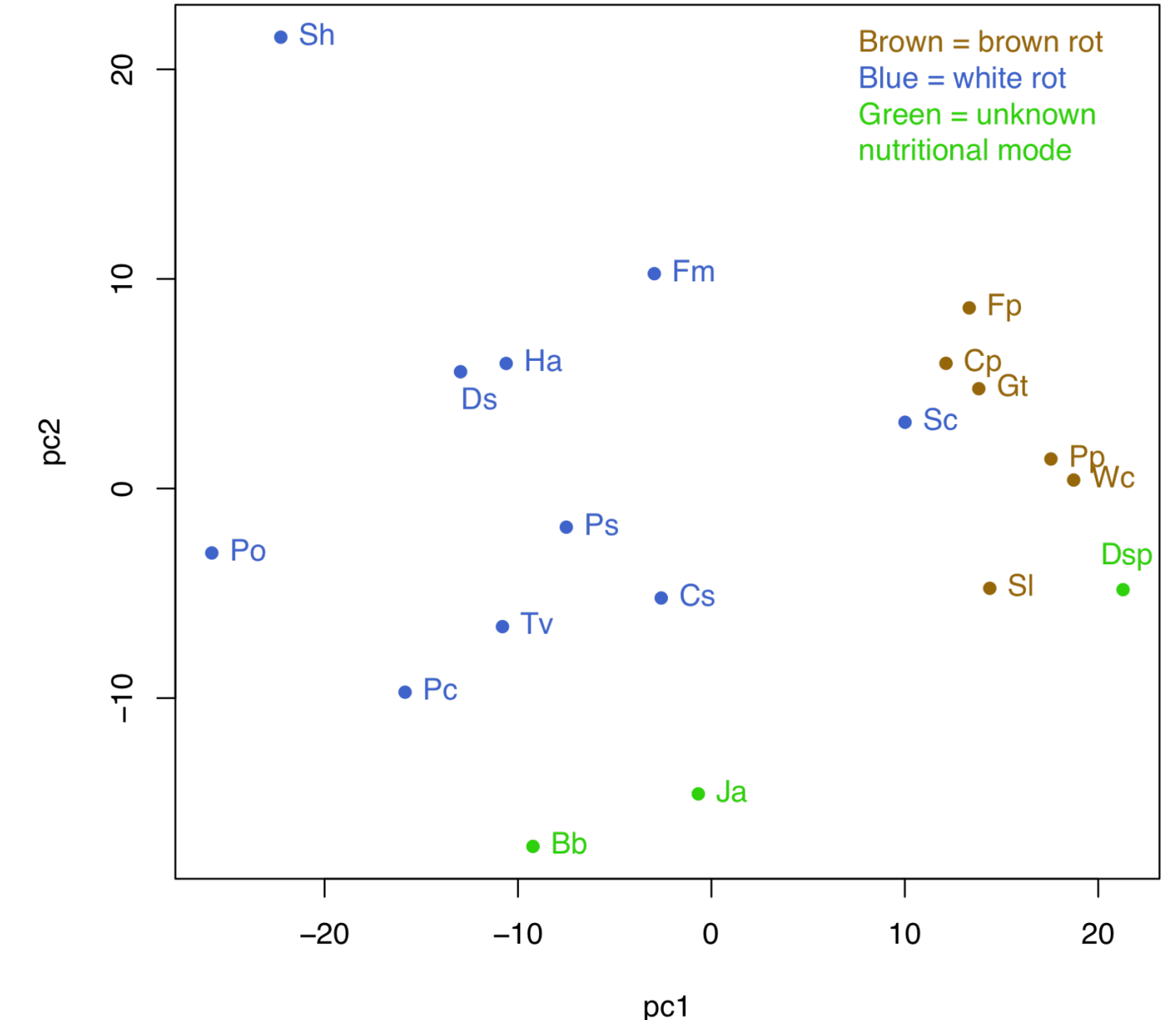
Analysis of chromosomal gene position suggests that lignin peroxidase expansions were driven in part by tandem duplication. Above, three *Trametes versicolor* scaffold regions are shown in which lignin peroxidase genes are found in clusters (implying tandem duplication).



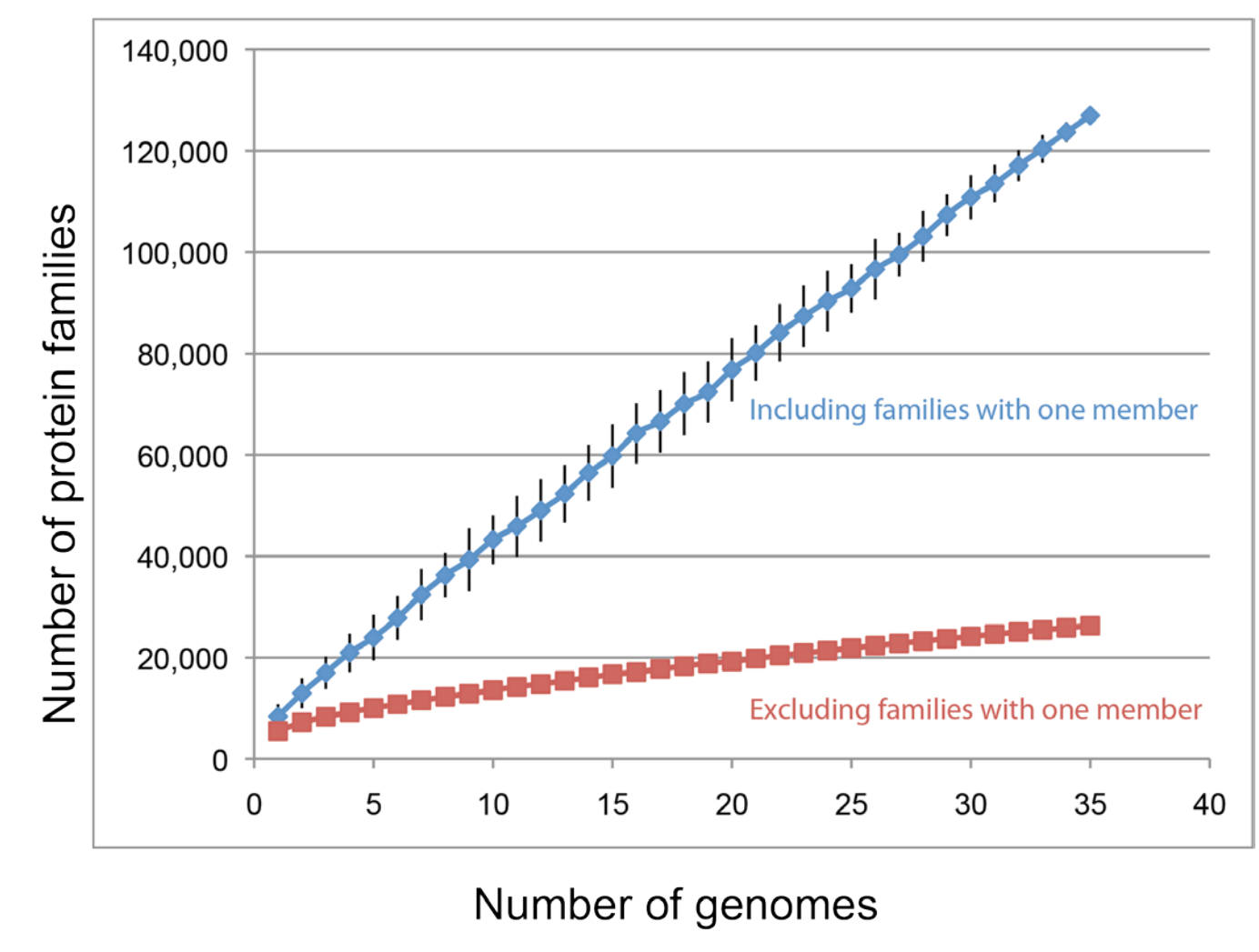
Maximum likelihood phylogeny of 35 basidiomycete and selected ascomycete fungi based on 237 widely conserved genes. Taxonomic orders with more than one organism are noted with black vertical bars



**Genomes overview.** Genome size in basidiomycetes varies over an order of magnitude with *M. globosoa* (9.0 Mb) and *M. laricis-populina* (101 Mb) at the extremes. Repeat content varies from a few percent (*M. globosoa*, *S. roseus*, *U. maydis*, *W. sebi*), to ~40 percent (*L. bicolor*, *F. mediterranea*, *P. graminis*, *M. laricis-populina*)



**Predicting degradative phenotype from lignocellulose genes.** Principal components analysis (PCA) was performed using selected lignocellulose genes as features. Genomes are shown plotted on the first two principal components from PCA of the lignocellulose genes of the organisms. Note the separation of brown rots (brown), white rots (blue). We predict that *B. botryosum* and *J. argillacea* are white rots, and that *Dacryopinax sp.* is a brown rot. Our results also raise questions about the traditional classification of *S. commune* as a white rot.



**Rate of protein family discovery.** MCL clustering of protein sequences was used to identify protein families. For each number of genomes, samples of genomes of a given number were randomly generated, and the protein families were counted. The figure illustrates that, in terms of protein family discovery, good gains are still being made from genome sequencing of basidiomycetes.