www.sciencemag.org/cgi/content/full/336/6089/1715/DC1

Supplementary Materials for

The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes

Dimitrios Floudas, Manfred Binder, Robert Riley, Kerrie Barry, Robert A. Blanchette, Bernard Henrissat, Angel T. Martínez, Robert Otillar, Joseph W. Spatafora, Jagjit S. Yadav, Andrea Aerts, Isabelle Benoit, Alex Boyd, Alexis Carlson, Alex Copeland, Pedro M. Coutinho, Ronald P. de Vries, Patricia Ferreira, Keisha Findley, Brian Foster, Jill Gaskell, Dylan Glotzer, Paweł Górecki, Joseph Heitman, Cedar Hesse, Chiaki Hori, Kiyohiko Igarashi, Joel A. Jurgens, Nathan Kallen, Phil Kersten, Annegret Kohler, Ursula Kües, T. K. Arun Kumar, Alan Kuo, Kurt LaButti, Luis F. Larrondo, Erika Lindquist, Albee Ling, Vincent Lombard, Susan Lucas, Taina Lundell, Rachael Martin, David J. McLaughlin, Ingo Morgenstern, Emanuelle Morin, Claude Murat, Laszlo G. Nagy, Matt Nolan, Robin A. Ohm, Aleksandrina Patyshakuliyeva, Antonis Rokas, Francisco J. Ruiz-Dueñas, Grzegorz Sabat, Asaf Salamov, Masahiro Samejima, Jeremy Schmutz, Jason C. Slot, Franz St. John, Jan Stenlid, Hui Sun, Sheng Sun, Khajamohiddin Syed, Adrian Tsang, Ad Wiebenga, Darcy Young, Antonio Pisabarro, Daniel C. Eastwood, Francis Martin, Dan Cullen, Igor V. Grigoriev,* David S. Hibbett*

*To whom correspondence should be addressed. E-mail: ivgrigoriev@lbl.gov; dhibbett@clarku.edu

Published 29 June 2012, *Science* **336**, 1715 (2012)
DOI: 10.1126/science.1221748

This PDF file includes:

SUPPORTING ONLINE MATERIAL

Table of Contents

I. Supplementary Methods, Results and Discussion

1. Sources of strains, culture conditions, and extraction of nucleic acids.

1.1. Methods.

Fungal cultures, except the strains of *Gloeophyllum trabeum* and *Tremella mesenterica* were cultivated in the Hibbett laboratory at Clark University. The Clark cultures were routinely grown under ambient laboratory conditions at 23 C in daylight and simultaneously in a temperature controlled incubator (Precision, GCA, Thermo Scientific, Asheville NC) at 28 C in the dark. Various liquid nutrient media were used to determine the optimal growth conditions for the different fungal isolates, including modified vitamin (VIT) medium (*1*), potato-dextrose (PD) medium (24 g potato-dextrose [EMD] /L), malt extract (ME) medium (20 g malt extract, 0.5 g yeast extract /L), minimal (MM) medium (0.25 g ammonium tartrate, 0.5 g glucose, 0.5 g yeast extract /L), and Avicel medium (40 g Avicel PH-101 from Fluka analytical, 5 g ammonium tartrate,1 g KH_2PO_4, 0.5 g $MgSO_4$ [x $7H_2O$], 0.001 g $CaCl_2$ [x $2H_2O$], 0.1 g yeast extract, 0.88 mg $ZnSO_4$ [x $7H_2O$], 0.81 mg $MnSO_4$ [x $4H_2O$], 0.80 mg $FeCl_3$ [x $6H_2O$] /L). Mycelium used for DNA extractions was grown over a time period of 1-2 weeks and was harvested using a vacuum system including a Buchner funnel and Whatman 4 filter disks. Up to 10 g (wet) of ground mycelium powder were loaded on Qiagen (Valencia CA) Genomic 500/G tips and processed following the lysis protocol for tissue in the Qiagen Blood & Cell Culture DNA Kit, including

the RNase and Proteinase K steps. Materials for RNA extraction were filtered after 3-5 days of growth. The Qiagen RNeasy Midi Kit was used to process up to 1 g (wet) of ground mycelium at a time. The extraction followed the protocol for isolation of total RNA from animal tissues (Qiagen) including on-column DNase digestion and final standard LiCl purification.

The *Tremella mesenterica* strain was maintained on yeast-peptone sugar (YPD) medium (10 g yeast extract, 20 g peptone, 20 g glucose/dextrose /L) in the Heitman laboratory at Duke University Medical Center. For genomic DNA isolation, cells were first harvested from YPD liquid medium after shaking at 24 C overnight. The cell pellet was then lyophilized, and a CTAB DNA isolation protocol was followed to extract genomic DNA (*2*). For the production of RNA, the strain was grown on YPD plates at 24 C for 3 days. The cells were then collected and lyophilized. Total RNA was extracted with TRIZOL (Invitrogen, Grand Island NY) following the manufacturer's instructions.

Gloeophyllum trabeum was grown in the Tsang laboratory at the Centre for Functinal and Structural Genomics, Concordia University. The mycelium for DNA extraction was cultivated at 34° C on 10x *Trametes* defined medium (*3*) (TDM) with 2% glucose as carbon source. Two types of ESTs were produced at Concordia University. The culture conditions for the mRNA-Seq (Illumina/Solexa platform) ESTs were: 34 C in 10X TDM using a mixture of 2% barley and alfalfa straws (ground to 0.5 cm lengths) as carbon source. The culture conditions used for Sangers' ESTs were the same except for the carbon sources. Mycelia in this case were grown independently in glucose, sucrose, xylose, mannose, citric pectin and pulp. The RNAs were pooled to construct a cDNA library. All ESTs were sequenced at the McGill University/Genome Quebec Innovation Centre with financial support from Genome Canada and Genome Quebec.

All cultures are available from the Forest Products Laboratory (FPL, USDA Forest Service) with the exception of *T. mesenterica*, which is available from the American Type Culture Collection (ATCC). The following paragraphs summarize the individual strain information, nutritional modes of species, and media used to produce samples of nucleic acids for sequencing.

1.2. Strains.

Auricularia delicata (Mont.) Henn. 1893
Mycobank: MB122319
Acronym AURDE, strain TFB10046, SS-5 (single spore isolate).
Collection information: 03/14/1999; leg./det. J.L. Mata, R.E. Halling & R.H. Petersen; Lat./Long.: 102609N / 0844221W; Location: Arenal Observatory Lodge, Trail to Lava flow, Fortuna San Carlos Alajuela, Costa Rica. On wodden stump. Voucher deposited at the University of Tennessee (TENN).
Ecology: white rot on dead and decaying wood.
Media used for the production of nucleic acids: three conditions for DNA (ME, PD, VIT), three pooled conditions for RNA (VIT+PD+MM).

Coniophora puteana (Schumach.) P. Karst. 1868
Mycobank: MB204571
Acronym CONPU, strain: RWD-64-598, SS-2 (single spore isolate).
Collection information: 08/04/1964; leg./det. Ross W. Davidson; Location: McNary, Arizona, USA. Substrate is a conifer. Voucher deposited at the Forest Products Laboratory (FPL).
Ecology: brown rot on conifers and sometimes deciduous trees in its natural habitat. Causes structural damage in buildings with unrepaired leaks (the “cellar” fungus).
Media used for the production of nucleic acids: one condition for DNA (PD), two pooled conditions for RNA (PD+VIT).

Dacryopinax sp.
Acronym DACSP, strain: DJM 731, SSP-1 (single-spore isolate); alternative collection number MIN862738.

Collection information: 08/14/1991; leg./det. David J. McLaughlin, Esther G. McLaughlin; Location: La Amistad International Park, Stanford University Biological Station, Las Alturas, Costa Rica. In deciduous woods, clustered on twig. Voucher deposited at the University of Minnesota (MIN).
Ecology: extensive brown rot decay in dead wood.
Media used for the production of nucleic acids: two conditions for DNA (PD, VIT), three pooled conditions for RNA (VIT+PD+MM).

Dichomitus squalens (P. Karst.) D.A. Reid 1965
Mycobank: MB312964
Acronym DICSQ, strain: LYAD-421, SS1 (single spore isolate).
Collection information: 05/17/1967; leg./det. Alix David; Location: Drôme (26), Bordeaux, France. Forest *Pinus* (*sylvestris* or *nigra*?). Voucher deposited at the Forest Products Laboratory (FPL).
Ecology: extensive white rot (commonly called western red rot) on *Pinus ponderosa*.
Media used for the production of nucleic acids: one condition for DNA sample (PD), two conditions for RNA (PD+VIT, Avicel).

Fomitiporia mediterranea M. Fisch. 2002
Mycobank: MB384943
Acronym FOMME, strain: MF3/22 #7 (single spore isolate); alternative collection number AFTOL-ID 688.
Collection information: 08/2001; leg./det. M. Fischer; Location: Kaiserstuhl area near Freiburg, Germany, in a local vineyard on *Vitis vinifera*. Voucher deposited at the University of Regensburg (REG).
Ecology: white rot in grapevine (*Vitis vinifera*). It is one symptom of the destructive grapevine disease complex esca, which may involve a consortium of pathogens.
Media used for the production of nucleic acids: three conditions for DNA (ME, PD, VIT), three pooled conditions for RNA (VIT+PD+MM).

Fomitopsis pinicola (Sw.) P. Karst. 1881
Mycobank: MB101927
Acronym FOMPI, strain: FP-58527 SS1 (single spore isolate).
Collection information: 1933; leg./det. R.W. Davidson & R.L. Gilbertson; Location: South Dakota; Black Hills National Forest, USA. On *Pinus ponderosa*. Voucher deposited at the Forest Products Laboratory (FPL).
Ecology: brown rot on dead conifer trees or occasionally live trees in boreal forests of the northern hemisphere.
Media used for the production of nucleic acids: one condition for DNA (PD), two pooled conditions for RNA (PD+VIT).

Gloeophyllum trabeum (Pers.) Murrill 1908
Mycobank: MB356811
Acronym GLOTR, strain: ATCC 11539 (single spore isolate).
Collection information: 1926; leg./det. J.R. Weir; Location: Wisconsin, Madison, USA. On *Thuja plicata* utility pole. Voucher deposited at the Forest Products Laboratory (FPL).
Ecology: widespread brown rot on hardwoods and conifers, decays structural timbers.
Media used for the production of nucleic acids: one condition for DNA (TDM and glucose), various conditions for RNA (TDM and various carbon sources).

Punctularia strigosozonata (Schwein.) P.H.B. Talbot 1958
Mycobank: MB304837
Acronym PUNST, strain: HHB-11173 SS5 (single spore isolate).
Collection information: 1981; leg./det. H.H. Burdsall; Location: Minnesota, Clearwater, Schoolcraft Trail, USA. On *Populus tremuloides*. Voucher deposited at the Forest Products Laboratory (FPL).

Ecology: widespread white rot on hardwood, especially poplar trees (*Populus* spp.).
Media used for the production of nucleic acids: one condition for DNA (PD), two conditions for RNA (PD+VIT, Avicel).

Stereum hirsutum (Willd.) Pers. 1800
Mycobank: MB189826
Acronym STEHI, strain: FP-91666, SS1 (single spore isolate).
Collection information: 1954; leg./det. T.W. Childs & P.L. Lentz; Location: Oregon, Zig Zag, USA. On *Castanopsis chrysophylla*. Voucher deposited at the Forest Products Laboratory (FPL).
Ecology: pioneer white rot on living hardwood trees and occasionally on conifers.
Media used for the production of nucleic acids: three conditions for DNA (PD, ME, VIT), two conditions for RNA (PD+VIT, Avicel).

Trametes versicolor (L.) Lloyd 1920
Mycobank: MB281625
Acronym TRAVE, strain: FP-101664, SS1 (single spore isolate).
Collection information: 1978; leg./det. M.J. Larsen; Location: Wisconsin, Dane, Madison, University of Wisconsin Arboretum, USA. On *Quercus*. Voucher deposited at the Forest Products Laboratory (FPL).
Ecology: white rot, prime wood decomposer of dead, standing or fallen hardwood trees.
Media used for the production of nucleic acids: one condition for DNA (PD), two conditions for RNA (PD+VIT, Avicel).

Tremella mesenterica (Schaeff.) Retz. 1769
Mycobank: MB154470
Acronym TREME, strain ATCC24925 (single-spore isolate); alternative collection number RJB 2259-6.
Collection information: leg./det. R.J. Bandoni, Vancouver, Canada. On *Alnus rubra*.
Ecology: common on dead wood, especially hardwood. It specializes on parasitizing white rot fungi in the genus *Peniophora*.
Media used for the production of nucleic acids: one condition for DNA (YPD), one condition for RNA (YPD).

Wolfiporia cocos (F.A. Wolf) Ryvarden & Gilb. 1984
Mycobank: MB107372
Acronym WOLCO, strain: MD-104, SS10 (single spore isolate).
Collection information: 1956; leg./det. W.R. Smith & F.F. Lombard; Location: Florida, Alachua, Gainesville, USA. On *Pinus* fence post, copper-naphthanate treated. Voucher deposited at the Forest Products Laboratory (FPL).
Ecology: causes brown rot and is a root parasite of numerous hardwood trees and conifers.
Media used for the production of nucleic acids: one condition for DNA (PD), two pooled conditions for RNA (PD+VIT).

2. Genome and EST sequencing, assembly, and annotation and secretome and transcriptome analyses.

2.1. Genome sequencing and assembly.

Eleven of the 12 genomes were sequenced using a hybrid whole-genome shotgun approach using a combination of ABI3730 (fosmids) (Applied Biosystems, Foster City CA), 454-Titanium (454 Life Sciences, Branford Ct) and Illumina GAII (Illumina Inc., San Diego CA) sequencing platforms. Roche 454 sequence was from one or more standard libraries and typically one 4 kb and one 8 kb paired-end library. Illumina reads were collected from nominal 300 bp fragment libraries, sequenced to 76 bp,

and assembled using Velvet (0.7.55 - REF). The resulting velvet contigs were shredded into overlapping 800 bp chunks with a 200 bp overlap and used by Newbler (2.5-internal-10Apr08-1) together with 454 standard and paired-end reads and Sanger-sequenced fosmids ends. Gaps were closed *in silico* using Gap Resolution (*4*).

Tremella mesenterica was sequenced using only Sanger technology. Three different sized libraries (3.1 kb sequenced at 2.26x sequence read coverage, 6.8 kb at 3.32x, and 36.0 kb fosmids at .76x) were used as templates for the plasmid subclone sequencing process and both ends were sequenced. A total of 340,800 sequence reads (7.44x final sequence coverage) were assembled using Arachne v.20071016 (*5*). Mitochondrial reads were removed from the input data prior to assembly. Final assemblies were checked for contamination using the standard NCBI contaminant database. Genome assembly statistics are summarized in Supplementary Table S1.

2.2. EST sequencing and assembly.

For 10 out of 12 genomes, cDNA libraries were constructed using the methods as outlined in the Roche cDNA Rapid Library Preparation Method Manual (Roche). The 454 libraries were then sequenced using the Genome Sequencer FLX Instrument (Roche). Ribosomal RNA, low quality and low complexity reads were filtered out, then the remaining reads were assembled using either a JGI specific assembly process or Newbler (v2.3-PreRelease-6/30/2009, Roche) with default parameters (see Table S2). In cases where there were multiple libraries for an organism, all the reads from both libraries were combined in the final assembly. For four white rot species (*T. versicolor, S. hirsutum, D. squalens*, and *P. strigosozonata*), EST counts were assessed separately for RNAs obtained from cultures grown in basal salts shake flasks in the presence of microcrystalline cellulose (Avicel) as sole carbon souce or more complex media such as PD and VIT (defined above).

Tremella mesenterica Fries cDNA clone libraries were constructed and sequenced as described previously with minor differences to include: the size ranges of cDNA which were 0.6k-2kb and >2kb (See Table S2) (*6*). *G. trabeum* ESTs were produced at Concordia University and sequenced at the McGill University/Genome Quebec Innovation Centre. Two types of ESTs were produced for *G. trabeum*. In one, cDNA libraries were constructed and sequenced as described by Semova et al. (*7*). In another, the RNA extracted from *G. trabeum* cultured on a mixture of barley and alfalfa straws was sequenced using the RNA-Seq platform of Ilumina/Solexa. The 76-bp reads were assembled de novo (*8*) and used to support gene prediction.

2.3. Genome annotation.

Each of the JGI-sequenced genomes introduced here were annotated using the JGI annotation pipeline, which takes multiple inputs (scaffolds, ESTs, and known genes) and runs several analytical tools for gene prediction and annotation, and deposits the results in MycoCosm (http://jgi.doe.gov/fungi) for further analysis and manual curation. Genomic assembly scaffolds of each genomes were masked using RepeatScout (*9*) and RepeatMasker (*10*) and the RepBase (*11*) library of fungal repeats. tRNAs were predicted using tRNAscan-SE (*12*). Using repeat-masked assemblies, several gene prediction programs falling into three general categories were used: 1) *ab initio* – FGENESH (*12*); GeneMark (*13*), 2) *homology-based* - FGENESH+; Genewise (*14*) seeded by BLASTx alignments against GenBank's database of non-redundant proteins (NR: http://www.ncbi.nlm.nih.gov/BLAST/), and 3) *EST-based* - EST_map (http://www.softberry.com/) seeded by EST contigs. Genewise models were extended where possible using scaffold data to find start and stop codons. EST BLAT (*15*) alignments were used to extend, verify, and complete the predicted gene of models were then filtered for the best models, based on EST and homology models. Homology-based gene models with frameshifts that may correspond to potential pseudeogenes or assembly inaccuracies were removed from the final gene set. The resulting sets support, to produce non-redundant representative sets, subject to further comparative analysis. Measures of model quality include proportions of the models complete with start and stop codons,

consistent with ESTs, supported by similarity with proteins from the NCBI NR database. Quality metrics for gene models are summarized in Supplementary Table S3.

For all genomes, all gene models were functionally annotated using SignalP (*16*), TMHMM (*17*), InterProScan (*18*), BLASTp (*19*) against nr, and hardware-accelerated double-affine Smith-Waterman alignments (deCypherSW; http://www.timelogic.com/decypher_sw.html) against SwissProt (http://www.expasy.org/sprot/), KEGG (*20*), and KOG (*21*). KEGG hits were used to assign EC numbers (http://www.expasy.org/enzyme/), and Interpro and SwissProt hits were used to map GO terms (http://www.geneontology.org/). Multigene families were predicted with the Markov clustering algorithm (MCL (*22*)) to cluster the proteins, using BLASTp alignment scores between proteins as a similarity metric. Functional annotations are summarized in Supplementary Table S3. Manual curation of the automated annotations was performed by using the web-based interactive editing tools of the JGI Genome Portal to assess predicted gene structures, assign gene functions, and report supporting evidence.

2.4 Secretome and transcriptome analyses.

Mass spectrometric analysis of soluble extracellular proteins was performed as described for *P. chrysosporium* and *P. placenta* (*23*) using a medium that facilitates rapid growth of a wide range of wood decay fungi. Briefly, filtrates from Highley's salt medium (*24*) containing 0.5% (w/v) wiley-milled aspen were harvested after 5 days on a rotary shake. Under these conditions plant cell walls are substantially compromised and enzyme accessibility is enhanced. Typically, production of lignin peroxidases and manganese peroxidase is minimal relative to nutrient starved medium (*25*), but CAZYs and accessory enzymes are secreted at high levels (*26-28*). Secretome data were compared to EST counts for four white rot species (*T. versicolor, S. hirsutum, D. squalens, and P. strigosozonata*) grown in basal salts shake flasks in the presence of microcrystalline cellulose (Avicel) as sole carbon souce or more complex media such as PD and VIT (defined above). All secretome and transcriptome data are presented in a Supplementary Electronic File (All_Pro_annotated.xls).

The filtrates are precipitated by direct addition of solid trichloroacetic acid (TCA) to 10% (wt/vol). Following overnight storage at -20°C, the precipitate was centrifuged and the pellet washed several times with cold acetone. Pelleted proteins (estimated at ~300µg) were re-solubilized and denatured in 40µl of 8M Urea / 100mM NH_4HCO_3 for 10 minutes then diluted to 200µl for tryptic digestion with: 10µl of 25mM DTT, 10µl ACN, 110µl 25mM NH_4HCO_3 pH8.2 and 30µl trypsin solution (100ng/µl *Trypsin Gold* from PROMEGA Corp. in 25mM NH_4HCO_3). Digestion was conducted overnight at 37°C then terminated by acidification with 2.5% TFA (Trifluoroacetic Acid) to 0.3% final and 8 µl loaded for nanoLC-MS/MS analysis.

Peptides were analyzed by nanoLC-MS/MS using the Agilent 1100 nanoflow system (Agilent Technologies, Palo Alto, CA) connected to a hybrid linear ion trap-orbitrap mass spectrometer (LTQ-Orbitrap XL, Thermo Fisher Scientific, San Jose, CA) equipped with a nanoelectrospray ion source. Capillary high-performance liquid chromatography (HPLC) was performed using an in-house fabricated column with integrated electrospray emitter essentially as described (*29*) but using 360µm x 75µm fused silica tubing. The column was packed with Jupiter 4µm C_{12} particles (Phenomenex Inc., Torrance, CA) to approximately 12cm. Sample loading (8µl) and desalting were achieved using a trapping column in line with the autosampler (Zorbax 300SB-C18, 5µM, 5x0.3mm; Agilent). HPLC solvents were as follows: Isocratic loading: 1% (v/v) ACN, 0.1% Formic acid; Gradient elution: Buffer A: 0.1% formic acid in water, and Buffer B: 95% (v/v) acetonitrile, 0.1% formic acid in water. Sample loading and desalting were done at 10uL/min whereas gradient elution was performed at 200nL/min and increasing %B from A of 0 to 40% in 75 minutes, 40 to 60% in 20 minutes, and 60 to 100% in 5 minutes. The LTQ-Orbitrap was set to acquire MS/MS spectra in data-dependent mode as follows: MS survey scans from m/z 300 to 2000 were collected in centroid mode at a resolving power of 100,000. MS/MS spectra were collected on the 5 most-abundant signals in each survey scan. Dynamic exclusion was employed to increase dynamic range and maximize peptide identifications. This feature excluded precursors up to 0.55 m/z below and 1.05 m/z above previously selected precursors. Precursors remained on the exclusion list for 15 sec. Singly-charged ions and ions for

which the charge state could not be assigned were rejected from consideration for MS/MS. Raw MS/MS data were searched against the total of number of filtered proteins available for each species (Table S3) using in-house *Mascot* search engine with methionine oxidation and glutamine, asparagine deamidation as variable modifications, peptide mass tolerance was set at 20 ppm and fragment mass at 0.6 Da.

Protein annotation and significance of identification was done with help of Scaffold software (version 3.00.07, Proteome Software Inc., Portland, OR). Peptide identifications were accepted within 0.9% False Discovery Rate as specified by the Peptide Prophet algorithm. Protein identifications were accepted if they could be established at greater than 95.0% probability, as assigned by the Protein Prophet algorithm (*30*), and contained at least 2 identified peptides.

A total of 1989 proteins were unambiguously detected. CAZys included 29 carbohydrate esterases (CEs), 9 polysaccharide lyases (PLs) and, most strikingly, 506 glycoside hydrolases (GHs). Other broad categories included 436 uncharacterized proteins and 225 peptidases. Extracellular proteins were not consistently associated with classical secretion signals and many of these are likely due to hyphal lysis (ribosomal proteins) or to cell wall associated proteins (certain methanol oxidases (*31*)) or to inaccurate 5' termini of some models. It should be emphasized that the absence of detectable peptide in these experiments is equivocal; molecular weight, solubility, and trypsin cleavage sites are critical factors. On the other hand, positive identifications are clear-cut.

3. Analyses of repeated sequences.

3.1. Microsatellites.

Microsatellites are specific DNA sequences composed of 1 to 6 repeated nucleotides in tandem. They play a role in the genome organization, phenotypic diversity and are used as molecular markers for population genetics. The 31 fungal genomes analyzed in this study were surveyed for the presence and richness of microsatellites. MISA (http://pgrc.ipk-gatersleben.de/misa/download/misa.pl) was used to identify mono- to hexa-nucleotide microsatellite motifs using the default parameters (mono- with at least 10 repeats; di- with at least 6 repeats; tri-, tetra-, penta- and hexa- with at least 5 repeats; the maximum number of bases between two SSRs was 100 bp). In the 31 genomes analyzed, the total number of microsatellites varied from 237 for *Batrachochytrium dendrobatidis* to 58,892 for *Phycomyces blakesleeanus* (Table S4 and Fig. S1). Murat *et al.* (*32*) have identified such a large variability in the number of microsatellite in fungi. In saprotrophic Agaricomycotina, the density of microsatellites was constant for the brown rot (BR) species, whereas it was more variable for white rot (WR) (Figs. S1-3). The median number of microsatellites seems higher for WR than BR (Fig. S2).

3.2. Transposable elements.

Transposable elements (TE) are ubiquitous in eukaryotic genomes. The invasion of a genome by TE likely plays an important role in genome plasticity and evolution of species.

TEs were identified in the 31 SAP genomes *de novo* by two approaches. Firstly, RepeatScout (*33*) was used with the default parameters (l=15). RepeatScout generated libraries of consensus sequences. These libraries were then filtered as follows: 1) all the sequences less than 100 bp were discarded; 2) repeats having less than 10 copies in the genome were removed (as they may correspond to protein-coding gene families); and, 3) repeats having significant hits to known proteins in Uniprot (UniProt Consortium, 2010) other than proteins known as belonging to TE were removed. The consensus sequences remaining were annotated manually by a tBLASTx search (*19*) against RepBase (http://www.girinst.org/repbase/index.html). To identify full length LTR retrotransposons, a second *de novo* search was performed with LTR_STRUC (*34*). The full-length candidates LTR retrotransposon sequences were checked for their homology using the t*BLASTx* algorithm against the sequences coming from the RepBase database. The number of TE occurrences and the percent of genome coverage were assessed by masking the genome assemblies using RepeatMasker (*10*) with

the consensus sequences coming from the RepeatScout pipeline and the full-length candidates LTR retrotransposon.

The number of TE sequences identified by RepeatScout varied from 0 for *Malassezia globosa* to 2560 for *Melampsora larici-populina* and the TE genome coverage varied from 0 % for *M. globosa* to 41.4 % for the WR *Fomitiporia mediterranea* (Table S5 and Fig. S1). For all the 31 genomes a positive correlation between genome size and TE richness was found (Fig. S4). The different families of TE contributed differently to genome size according to the species (Table S5 and Fig. S1). Interestingly, most TE-rich species are involved in mutualistic or parasitic interactions with plants (Fig. 1). According to the present results and previous data on plant symbiotic and pathogenic fungi and oomycetes it seems that most TE-rich fungi interact with living plants (*35-38*).

The proportion of TE in the saprotrophic Agaricomycotina genomes was highly variable, ranging from 4.2 % (*Coniophora puteana*) to 29.8 % (*Serpula lacrymans*) for the BR, and from 3.1 % (*Stereum hirsutum) to* 41.4 % (*F. mediterranea)* for the WR. The means of genome size was not different in both groups (Fig. S3). The number of full-length LTR *gypsy* TEs was higher for WR than BR, although the proportion of *gypsy* LTRs was correlated with BR genome size (r2=0.33), but not with WR (r2=0.06). Indeed, the number of full-length *gypsy* is not always related to the genome coverage of this family of TE. For example, 111 full-length *gypsy* elements were identified for *H. annosum* for a genome TE coverage of 8.8 %, whereas only 74 full-length *gypsy* elements were identified for *F. mediterranea* for a genome TE coverage of 21.3 %. This is likely related to size of the current *gypsy* sequences which reflects the different date of the genome invasion by these TEs. The identification of full-length TE sequences relies on the conservation of TE sequences and only the most recent TEs can be identified. For *F. mediterranea,* the TE invasion was older as most of the full-length TEs have invaded the genome >4 millions years ago (Fig. S5). On the other hand, the full-length *gypsy* TE colonized the *H. annosum* genome more recently (Fig. S5). Based on their higher number of full-length TEs, we assumed that the *gypsy* TEs colonized the genomes of the WR species more recently. General conclusions from these analyses are that genome size is related to transposable element proliferation in many species, and that TE proliferation is rampant in plant-associated fungi, such as *Fomitiporia mediterranea and Wolfiporia cocos.*

4. Organismal phylogenetic analyses.

Seventy-one proteins were selected for phylogenetic analyses. Twenty-six were selected based on their inclusion in broad-scale eukaryotic phylogenetic studies (*21, 39*) and other projects (*40, 41*). An additional 45 single-copy proteins with gene models of at least 1 kb that were distributed across a set of 53 core fungal genomes were identified using the phyloinformatic pipeline *Hal* (*42*). A list of the 71 proteins with links to their corresponding *Saccharomyces cerevisiae* protein summaries (www.yeastgenome.org/) is provided in Table S6. Using a core set of 60 well-annotated genomes, alignments for each of the individual 71 proteins were generated using MUSCLE (*43*) and Hidden Markov Models (HMM) were produced from each protein alignment using HMMER 3.0 (hmmer.org). The 71 proteins were subsequently mined from the predicted protein sequences for all genomes in this study using the HMM models. Putative orthologs were identified as the highest full sequence HMM bit score with a minimum E-value of 1E-50.

Phylogenetic quality controls were conducted on each protein alignment to detect erroneous inclusion of potential paralogs. Ambiguous regions of alignments were masked using Gblocks (*44*) with the following settings: maximum number of contiguous nonconserved positions allowed = 4; minimum length of a block allowed = 10. Individual RAxML protein trees were then produced for each protein alignment using the PROTGAMMAWAG model with 100 rapid bootstrap partitions (BP). Individual proteins trees were manually examined for contradictory placement of taxa as measured by 70% BP. Proteins detected as potential paralogs were removed from alignments for subsequent phylogenetic analyses. The individual 71 QC protein alignments (after masking with Gblocks) were concatenated into a single superalignment named 71G using a customized Perl script, exported in a Phylip format, and analyzed using RaxML (*45*) v.7.2.6, with the PROTGAMMAWAG model and 100

rapid BP. The 71G dataset had a length of 34257 amino acids (aa) with a proportion of roughly 3% missing data.

To explore the effect of missing data and fast evolving sites, three additional datasets were assembled (26G, 26GA, 71GA). Alignments used in all phylogenetic analyses are provided as Supplementary Electronic Files. The 71 original full-length protein alignments assembled by *Hal* were subjected to Gblocks runs under more stringent conditions. The minimum number of sequences for a conserved region was defined as 50% of the species plus one and the threshold value for included flanking regions was set to be represented in 85% of the species. The maximum number of contiguous nonconserved positions was fixed to eight positions. Gapped positions were not permitted. In addition, datasets with one or more complete sequence missing or retaining less than 20% of the original amino acids post Gblocks analyses were omitted from further analyses (see Table S7). The concatenated final dataset (26G) included 13371 amino acid positions obtained from 26 proteins (ADO1, AMD1, CCT2, CCT4, CDC60, CHC1, DIP2, ERB1, FAL1, GLN4, HSH155, KRR1, LOS1, LST8, MCM3, MCM7, NBP135, NEW1, RAD3, RBG2, RET1, RPO31, SYF1, TFB2, TPI1, ZWF1) and had no gapped positions.

The impact of evolutionary site-rate variation in the multigene alignments was assessed using the AIR software package (*46*) through the Bioportal at the University of Oslo (http://www.bioportal.uio.no). The 71G and 26G datasets and the inferred optimal RAxML trees were analyzed with AIR-Identifier, calculating the posterior mean values of site rates for eight rate categories under the WAG model. The two fastest evolving rate categories were subsequently excluded using AIR-Remover, resulting in two additional datasets named 71GA (25483 aa, 2.6% missing data) and 26GA (10002 aa). Bootstrap support values (BS) from all four datasets were estimated with RAxML running 500 replicates under the PROTMIXWAGF model. Posterior probability values (PP) were estimated using PhyloBayes (*47*) v.3.2. The runs were performed using the CAT mixed model and were stopped after three million generations when the largest and mean discrepancy observed across all bipartitions became zero (maxdiff, meandiff = 0). A phylogram with branch lengths inferred from the 26G dataset with PhyloBayes and support values for nodes with less than maximal support is shown in Fig. S6.

5. Analyses of decay-related gene families.

5.1. Assembly and phylogenetic analysis of gene family datasets.

Protein models for 27 gene families, including 17 families encoding carbohydrate-active enzymes (CAZymes) and 10 families encoding oxidoreductases (Table 1), were assembled from all 31 species using a combination of IPR domain searches, previously characterized proteins, and the JGI cluster pipeline (http://genome.jgi.doe.gov/31_SAP).

The CAFÉ program (*48*) was used to analyze gene family expansions and contractions in the 27 families of CAZymes and oxidoreductases listed above, as well as in carbohydrate binding modules (CBM1) associated with eleven CAZyme families (see below). CAFÉ uses a stochastic model of gene birth and death to infer statistically significant gain and loss in gene families, given an ultrametric phylogenetic tree (which was generated with BEAST, as described below) and a table of gene copy numbers in each organism. A family-wide significance threshold of 0.01 was used. The search for λ values was run with the –s option, which specifies an optimization algorithm that finds the values of λ that maximize the log likelihood of the data for all families. Fifteen gene families were found to have significant departures from a random birth-death process of gene family evolution ($p<0.01$; Table 1), suggesting that there have been lineage-specific shifts in rates of gene duplication and/or loss. For all gene families with significant family-wide *p*-values, CAFÉ was also used to run a Viterbi algorithm that assigns branch-specific *p*-values for expansions and contractions of each family at each branch of the tree, with a significance threshold of 0.05 employed for the branch-specific changes. A summary of significant branch-specific expansions and contractions for all 15 non-randomly evolving gene families is shown in Fig. S7 and gene copy numbers at internal nodes in the organismal phylogeny estimated with CAFÉ are presented in Table S8.

Phylogenetic analyses and gene tree/species tree reconciliation analyses were conducted for the 15 non-randomly evolving gene families (except cytochrome P450s, which have up to 250 copies per genome and therefore are not tractable for reconciliation analyses), as well as GH6 and CDH. Preliminary alignments were examined manually and very diverse datasets were subdivided. An unrooted maximum likelihood analysis was performed for each alignment using the RaxML (*49*) server (http://phylobench.vital-it.ch/raxml-bb/) with 100 rapid bootstrap replicates and a final ML search under the GTRGAMMA model, with the WAG substitution matrix. The optimal gene phylogenies for each family are presented in Fig. S8a-q. The best scoring ML tree for each gene family, with bootstrap values, was used in Notung (*50, 51*), which performs a parsimony-based reconciliation analysis in which weakly supported branches in the gene tree (as determined by a user-specified edge weight threshold, EWT) can be rearranged to minimize penalty scores for gene duplications and losses. Default costs for gains and losses (1.5 and 1.0, respectively) were used. Two Notung analyses were performed for each gene family, with the EWT set to 75% or 90% bootstrap support. Graphical summaries of changes in copy number in all 27 families of CAZymes and oxidoreductases inferred with Notung are presented in Figs. S9-11, and the reconciled gene phylogeny for class II fungal peroxidases (with EWT set to 75%) is presented in Fig. S12. Gene copy numbers at internal nodes in the organismal phylogeny estimated with Notung are presented in Table S8 (these include only the gene families found to have significant departures from random diversification in CAFÉ analyses, and GH6 and CDH; for divergent gene families that were subdivided, these figures are the sums of results of analyses of independent gene phylogenies).

To supplement Notung and CAFÉ analyses, the program DrML (*52*) (vers. 0.91) was used to perform gene tree/species tree reconciliations for fungal class II peroxidases only. Given a species tree and a gene tree, DrML program produces a DLS tree (*53*), which is an evolutionary scenario including duplications and losses. DrML seeks the maximum likelihood DLS tree based on an evolutionary model of gene duplication and loss (unlike Notung, DrML does not perform rearrangements in the gene tree). DrML was used to calculate DLS trees from 100 bootstrapped gene trees and the bootstrap consensus tree (generated as described above) on one species tree (as in Figs. 1, S6-7). A custom visualization tool was used to depict an embedding of the gene tree inside the species tree, as well as the distributions of duplications, losses, and copy numbers at each internal node obtained with the 100 bootstrapped gene trees. The optimal reconciliation of the bootstrap consensus gene tree in the species phylogeny, and the distributions of copy numbers and numbers of gains and losses at each internal node of the phylogeny are shown in Figs. S13-S15.

Ancestral states of residues corresponding to Glu35, Glu39 and Asp175 in the manganese binding site of *P. chrysosporium* MnP1 and the Trp171 residue of *P. chrysosporium* LiP-H8 were reconstructed using BayesTraits 1.0 (*54*). Trees were sampled after stationarity in BEAST analyses (see Molecular Clock Analyses) and the posterior sample was thinned to 1000 trees by subsampling in Mesquite (*55*). For the Mn-binding site, genes with 0-1, 2, or 3 of the key residues present were scored as “absent”, “atypical”, or “typical” respectively, while the Trp171 was scored as present or absent in a separate, binary character. Following preliminary reconstructions in Mesquite, ancestral states were reconstructed for candidate nodes in all 1000 trees, while taking into account phylogenetic uncertainty (addmrca command). Hypothesis tests were performed by fixing a node in one state at a time and monitoring the likelihood values. A difference of two log likelihood values was considered significant support for the better fitting state (Fig. 1).

Rates of duplication and loss of class II peroxidase genes in white rot and non-white rot lineages were analyzed with BISSE in the R package Diversitree (*56*) on the same ultrametric Bayesian tree sample used for BayesTraits analyses. In the unconstrained six-parameter, the rate of gene duplication was significantly greater in white rot than in non-white rot lineages (Fig. S16). Seven constrained models with three to five free parameters were investigated; the only models that could not be rejected (using two units of log likelihood as a criterion for significance) were those in which the rate of duplication was allowed to vary in white rot vs. non-white rot lineages (Table S9).

5.2. Overview of CAFÉ and gene tree/species tree reconciliation results.

Eighteen of the 27 decay-related CAZyme and oxidoreductase gene families surveyed are present in the genomes of Basidiomycota and Ascomycota, but absent from those of *Batachochytrium dendrobatidis* (Chytridiomycota) and *Phycomyces blakesleeanus* (Mucoromycotina), including GH7, GH10, GH61, GH43, CE5, POD, GLP, DyP, HTP, GH6, GH11, GH12, GH74, CE12, CE15, OXO, Cytb652, and CDH (Table 1). Only two gene families (DyP and GLP) were limited to Basidiomycota. The common ancestor of the Dikarya therefore possessed a diverse array of enzymes for decay of plant biomass, including multiple gene families that may have arisen in the lineage leading to Dikarya. CAFÉ and Notung with EWT set to 75% or 90% yielded similar estimates of the total number of copies of genes encoding decay-related oxidoreductases at internal nodes across the phylogeny (Table S8, Fig. S11). However, there was significant variation between CAFÉ and Notung in estimates of the number of copies in the fifteen CAZyme gene families selected for reconciliation, particularly in the common ancestor of the Dikarya and the most ancient nodes within the Basidiomycota (nodes 3-6). Inspection of unreconciled gene phylogenies (Fig. S8) suggests that this discrepancy is influenced by large gene families in which there has been significant diversification prior to the origin of Dikarya, including GH3, GH26, GH43, and GH61 (which contain numerous strongly supported groups containing genes from both Basidiomycota and Ascomycota). In contrast, estimates for gene families that coalesce to a single ancestor within the ingroup are more consistently estimated by CAFÉ and Notung. For example, POD genes of Ascomycota and Basidiomycota form strongly supported monophyletic groups in unreconciled phylogenies (Fig. S8l), and both CAFÉ and Notung suggest that a single copy was present in the common ancestor of Dikarya. DrML estimates of shifts in copy number of class II fungal peroxidases largely parallel those obtained with CAFÉ and Notung, with the principal difference being that DrML reconstructs a greater expansion in the "backbone" nodes of the Agaricomycete phylogeny (nodes 9, 10, 11) and the earliest divergences in Polyporales (nodes 18, 19) than does either CAFÉ or Notung (Fig. S11). Thus, DrML analyses imply more dynamic shifts in gene content than either CAFÉ or Notung.

Pucciniomycotina, Ustilaginomycotina, Tremellomycetes and the yeast *Pichia stipitis* are strikingly reduced in genes encoding decay-related CAZYmes and oxidoreductases compared to Agaricomycetes, Dacrymycetes, and filamentous Ascomycota (Table 1). Excluding the rust *Melampsora laricis-populina* (Pucciniomycotina, which has 103 genes in 16 families) these depauperate genomes contain only 12-29 (avg. 16.5) gene copies from 4-13 (avg. 6.7) gene families, while the Agaricomycetes, Dacrymycetes and filamentous Ascomycota have 57-307 genes (avg. 116.5) from 17-27 (avg. 21.6) gene families. Notung reconciliations at EWT 75% or 90% suggest that there have been reductions in from thirteen to sixteen different gene families in each of the lineages leading to Pucciniomycotina, Ustilaginomycotina, Tremellomycetes and *P. stipitis*. (Figs. S9-S10). Notung reconciliations further suggest that nine to eleven gene families have undergone expansions along the internode leading to the most recent common ancestor of the Agaricomycetes (between nodes 7 and 8), with seven gene families expanding between nodes 8 and 9 (Figs. S9-S10). Thus, Notung analyses suggest that the origin of the Agaricomycetes was characterized by expansions in multiple gene families that were present in the ancestor of the Dikarya, but that had remained at low copy number (or were lost entirely) during the early diversification of Pucciniomycotina, Ustilaginomycotina, and Agaricomycotina. Changes in gene content within the Agaricomycete-Dacrymycete clade and in specific gene families are described below and in Figs. S9-11.

5.3. Notes on CAZymes.

5.3.1. GH6, GH7, GH61 (cellobiohydrolases, endoglucanases, "oxidohydrolases").

GH6 and GH7 include cellobiohydrolases and endoglucanases in Ascomycota and mainly cellobiohydrolases in Basidiomycota (*57-59*). Cellobiohydrolases participate in the degradation of crystalline cellulose acting on the non-reducing end (GH6) and the reducing end (GH7) of the cellulose chain. Thus, their activity along with endoglucanases and β-glucosidases represents a

complete system for the degradation of crystalline cellulose. Recently, the role of GH6 as cellobiohydrolases in *C. cinerea* has been questioned (*60*), with only one recognized cellobiohydrolase out of the 5 GH6 genes that the genome possesses (the cellobiohydrolase model is Copci1_959 and includes a CBM1).

CAFÉ analyses returned a family-wide *p*-value of <0.001 for GH7 but results for GH6 were nonsignificant (Table 1). The common ancestor of Agaricomycetes and Dacrymycetes (node 7, Fig. S7) is estimated to have had 3 copies of GH6 based on Notung results (reconstructions were not performed with CAFÉ for GH6). Additionally, both CAFÉ and Notung analyses estimate1-2 copies of GH7 genes present in node 7 (Table 1, Table S8, Fig. S8d). All white rot species have retained one GH6 copy except *A. delicata*, which has two copies. In addition, *S. commune* has one copy and *C. cinerea* has five copies. Both GH6 and GH7 have been lost in the brown rot species and *L. bicolor*. Exceptions include *C. puteana* and *S. lacrymans*, with the former possessing two copies each of GH6 and GH7 genes and the latter one copy of GH6. One copy of GH7 in *C. puteana* and the GH6 of *S. lacrymans* have been shown to carry a cellulose binding module (*61*). Independent expansions of from 3-6 copies in GH7 are reconstructed in terminal lineages leading to three white rot species, including *A. delicata*, *P. strigosozonata*, and *P. chrysosporium* (Figs. S7, S9-10).

The GH61 family also shows significant expansions in white rot lineages, with a family-wide *p*-value of <0.001 in CAFÉ analyses (Table 1, Fig. S7). Formerly recognized as a family with weak endoglucanase activity (*62, 63*), recent data support the view that GH61 comprises proteins that lack hydrolytic and endoglucanase activity (*64, 65*), but that may perform an oxidative attack of crystalline cellulose (*66*). GH61 proteins have been identified in cultures of white rot species growing on wood (*23*) and transcripts of GH61 genes were among the most abundant in cultures of *Phanerochaete carnosa* on wood (*67*). Genes of the family are present in all the Agaricomycetes (but absent from *Dacryopinax* sp.), and the common ancestor on node 7 is estimated to have had between 3 (CAFÉ) and up to 13-14 copies (Notung), suggesting that the diversification of the family predates the separation of Dikarya (*65*). This view is supported by the GH61 phylogeny (Fig. S8k), in which *A. delicata* sequences are scattered in at least in 8 different subclades. Additionally, the family appears to have undergone expansions between nodes 7 and 9, after the divergence of the Dacrymycetes, while expansions are implied higher on the tree for many white rot species. At the same time, both CAFÉ and Notung analyses support losses for *L. bicolor* (9-10 losses) and the brown rot lineages (7-10 losses in the base of the Boletales, supported only in the reconciliation and additional losses for *S. lacrymans*, 7-8 for the *Antrodia* clade, 7 for *G. trabeum* and 3-14 for *Dacryopinax*) (Figs. S7, S9-10).

Overall, results of GH6, GH7 and GH61 analyses suggest that the common ancestor of Agaricomycotina was able to attack crystalline cellulose in a fashion that resembles the approach of contemporary white rot species and soil saprotrophs. Additionally, the gene losses in *L. bicolor* and brown rot lineages support the view that the mycorrhizal lifestyle and the brown rot mechanism are derived nutritional strategies in Agaricomycotina (*68, 69*). Finally, the potential oxidative role of the GH61 family members and their reduction in brown rot species is consistent with the presence of an alternative oxidative system operating through a Fenton reaction and facilitating the oxidative attack of cellulose, which would make the maintenance of multiple GH61 genes redundant.

Expression data generally support the importance of GH6s, GH7s and GH61s in white rot decay. Specifically, LC-MS/MS analyses of aspen culture filtrates revealed proteins corresponding to GH6s (6), GH61s (12), and GH7s (6) were identified in cultures of the white rot fungi *A. delicata, T. versicolor, S. squalens, S. hirsutum,* and *P. strigosozonata* (see Supplementary Electronic Files). Our analysis also allowed examination of EST counts for these species when grown on media containing microcrystalline cellulose as the sole carbon source and, in all cases, umambiguous protein identification were associated with high transcript levels. In these expression studies, a lone brown rot fungus, *G. trabeum*, was able to produce a predicted GH61 protein (Glotr163531). Given the number of unique peptides and the large number of spectra, this identification was unequivocal.

5.3.2. GH43, GH74, CE1, CE12 (CAZymes involved in degradation of hemicellulose, amorphous cellulose, and oligosaccharides).

Fourteen of the CAZyme gene families analyzed are involved in degradation of amorphous cellulose, hemicellulose or oligosaccharides (*70-74*), but only four families (GH43, GH74, CE1, CE12) show patterns of differentiation in copy numbers between white rot fungi and *C. cinerea* vs. brown rot fungi and *L. bicolor*, and only GH43 has a significantly nonrandom pattern of diversification according to CAFÉ analyses (p<0.001). GH43 is a diverse family consisting of proteins with six different recognized enzymatic specificities, all related to hemicellulose degradation (*75*) (http://www.cazy.org/), but studies related to Agaricomycotina are limited (*76, 77*). White rot genomes possess 3-28 (avg. 8.9) copies of GH43 genes, while brown rot and ECM genomes possess 1-7 (avg. 3.1) copies (Table 1). Proteomic analysis of avicel culture filtrates identified 16 different GH43 proteins distributed among both white and brown rot fungi (see Supplementary Electronic Files). Both CAFÉ and Notung analyses suggest independent expansions in GH43 in several white rot species, including *A. delicata* (6-21 gains), *S. hirsutum* (1-5 gains) and *D. squalens* (2-3 gains) and losses for some brown rot species such as *S. lacrymans* (1-4 losses), *W. cocos* and *P. placenta* (4 losses), and the ectomycorrhizal *L. bicolor* (4 losses). Notung results suggest the presence of 12-20 copies of GH43 for nodes 7 and 8, while CAFÉ reconstructs 4-5 copies at these nodes. The CAFÉ result appears to be an underestimate, based on the presence of numerous strongly supported clades that span the root node of the Agaricomycetes in the unreconciled GH43 phylogeny (Fig. S8h).

On average, white rot genomes possess 1-4 (avg. 1.9) copies of GH74, 0-4 (avg. 1.4) copies of CE1, and 0-3 (avg. 1.3) copies of CE12. In contrast, genomes of brown rot fungi and the ectomycorrhizal *L. bicolor* have lost all copies of these gene families, except for *G. trabeum*, which has one copy each of GH74 and CE1, and *S. lacrymans,* which has one copy of GH74. (Among white rot fungi, only *F. mediterranea* and *D. squalens* lack CE1, while *P. strigosozonata, T. versicolor*, and *P. chrysosporium* lack CE12.) The ancestor at node 7 (Agaricomycetes plus Dacrymycetes) is reconstructed with Notung as having 1-2 copies for each of the families GH74, CE1, and CE12. Thus, it appears that white rot species (as well as *S. commune* and *C. cinerea*) have largely retained the plesiomorphic complement of GH74, CE1 and CE12, while brown rot and ectomycorrhizal fungi have undergone parallel losses in these gene families.

5.3.3. GH3, GH5, GH10, GH12, GH28, CE8, CE15, CE16 (endoglucanases, xylanases, mannanases, pectinases, carbohydrate esterases, etc.).

Representatives of these families are widely dispersed and corresponding proteins were unambiguously identified in both white rot and brown rot culture filtrates (see Supplementary Electronic Files). The eight gene families discussed here are widespread across Agaricomycotina and participate in various aspects of carbohydrate degradation as endoglucanases (GH12 (*78, 79*), GH5 in part (*80, 81*)), xylanases (GH10 (*82, 83*)), mannanases (GH5, in part), pectin related enzymes (pectin methylesterases of family CE8 (*84*); pectinases of family GH28 (*85*)), carbohydrate esterases (CEs (*85, 86*)), and enzymes involved in degradation of cellobiose and xylobiose (GH3 (*87*)). CAFÉ analyses suggested significantly non-random diversification in four of the families, including GH3 (p<0.001), GH10 (p=0.004), GH28 (p<0.001), and CE16 (p<0.001). Notung reconstructions for these gene families suggest that the common ancestors of Agaricomycetes and Dacrymycetes (nodes 7 and 8) possessed 14-19 copies of GH3, 2-7 copies of GH10, 16-19 copies of GH28, and 3-8 copies of CE12, while CAFÉ analyses suggest minimum copy numbers of 8, 3, 6, and 4 copies for each family (respectively). In general, there are not large differences in the numbers of gene copies in these families between different nutritional strategies in the sampled genomes. An exception is found in *L. bicolor*, which has completely lost GH10 and has also undergone reductions in GH3. Overall, the maintenance of high numbers of gene copies in extant genomes as well as their representation in the common ancestor of Agaricomycetes and Dacrymycetes indicates their general importance in nutrient acquisition independent of ecological mode.

5.3.4. GH11, CE5 (cutinases, etc.).

Another type of xylanase is found in glycoside hydrolase family 11 (*82, 88*). GH11 is prominent in the genomes of filamentous Ascomycota, which have 3-7 copies each, but the family is rare in the sampled Basidiomycota (17 species lack the family entirely, five species have one copy, but *A. delicata* and *C. cinerea* have three and six copies, respectively), and CAFÉ analyses do not return a significant family-wide *p*-value. In contrast, CAFÉ suggested that CE5 has undergone lineage-specific shifts in diversification rate (p=0.002). CE5 includes cutinases, which are involved in degradation of cutin, the waxy polymer that covers external plant tissues. Like GH11, CE5 is rare across Basidiomycota; 14 species have no copies, four species have one copy, and four species have two to six copies, but the poplar rust *M. larici-populina* has 14 copies. Genomes of filamentous Ascomycota possess four to 15 copies of CE5, with the maximal value in the chestnut blight fungus *Cryphonectria parasitica* which indicates the importance of this gene family for plant pathogens. Consistent with the rarity of Basidiomycota GH11s and CE15s, a single GH11 (Punst160181) was identified by LC-MS/MS. No CE15 were observed.

5.3.5. CBM1 (cellulose binding modules).

Cellulose binding modules (CBMs) are C- or N-terminal functional domains that are found in a large number of CAZY families (http://www.cazy.org/Carbohydrate-Binding-Modules.html) as well as certain oxidoreductases (e.g. cellobiose dehydrogenase). CBM1s enhance the localization of the enzymes that bear them on the surface of crystalline cellulose (*89*) and they may also act to disrupt the crystallinity of cellulose (*90*).

A search of the Pfam domain PF00734 indicated that the fungal genomes analyzed here have from zero to 44 CBM1s (Table S10). The top twelve species, with no fewer than ten CBM1s, include all of the white rot wood decayers as well as *C. cinerea* (44 copies) and the Ascomycota *T. reesei, C. parasitica,* and *S. nodorum* (10-14 copies). Brown rot species and the ectomycorrhizal *L. bicolor* all have zero copies or one copy, except the brown rot Boletales, *S. lacrymans* and *C. puteana*, with eight and three copies, respectively. The presence of CBM1s in filamentous Ascomycota implies that their complete absence from many brown rot species, all Pucciniomycotina, Ustilaginomycotina, and Tremellomycetes, and the yeast *P. stipitis* reflects parallel losses. The presence of CBM1 in *Phycomyces blakesleeanus* (Mucoromycotina) indicates that the origin of this domain predates the origin of the Dikarya. Evolution of CBM1 was estimated using CAFÉ only (because these modules are only about 40 amino acids in length, they contain insufficient phylogenetic information for topology-based gene tree/species tree reconciliation analyses). CAFÉ analyses suggest that there have been significant expansions in CBM1s between nodes 6 and 8 (Fig. S7) and additional expansions for most of the white rot species and *C. cinerea*, with losses for the brown rot species and *L. bicolor*. Eleven copies of CBM1 are reconstructed at node 8 (Table S8), which is greater than any contemporary brown rot species or *L. bicolor,* which further suggests the similarities of carbohydrate degradation of that ancestral species to contemporary soil saprotrophs or white rot species.

The reduced number of CBM1s in brown rot and ectomycorrhizal species compared to white rot species and *C. cinerea* could reflect losses of CBM1 domains from certain gene copies, wholesale losses of genes bearing CBM1s, or a combination of such events. The number and proportion of genes bearing CBM1s was investigated in ten selected families of CAZymes, including GH6, GH7, GH10, GH5, GH61, GH74, CE1, and CE15 (Table S11, Fig. S17). The top eleven species, with at least 25% of the candidate gene copies bearing a CBM1, include all of the white rot wood decayers, *C. cinerea*, *T. reesei,* and *Aspergillus nidulans.* Except for *S. lacrymans* (which has CBM1s associated with 22% of the candidate genes), all brown rot fungi and the ectomycorrhizal L. bicolor have CBM1s associated with 0% to 9% of the focal CAZyme families. Thus, both the proportion and absolute number of CAZymes bearing CBM1s have been reduced in brown rot species, presumably reflecting a shift toward Fenton chemistry for disruption of crystalline cellulose. *Schizophyllum commune* has CBM1s associated with only 7% of the candidate CAZyme families, and in this regard (among others) it more closely resembles a brown rot fungus than a white rot fungus. The distribution

of CBM1s was mapped onto the gene phylogeny for GH5 and GH10 (Figs. S8c,g), suggesting parallel losses of CBM1s in brown rot and ectomycorrhizal lineages and *S. commune*.

The paucity of CBM1s in brown rot was reflected in expression studies where the predicted GH61 model, Glotr1140289, was the only CBM1-containing protein detected in filtrates. In contrast, 28 CBM1-containing proteins were observed in the white rot culture filtrates (see Supplementary Electronic files).

5.4. Notes on oxidoreductases.

5.4.1. Class I peroxidases.

Two genes belonging to Class-I of prokaryotic-origin peroxidases (*91*) were found in the Agaricomycotina genomes analyzed, corresponding to cytochrome-*c* peroxidase (CCP; EC 1.11.1.5), one gene per genome (with the only exception of *T. mesenterica,* which has two different genes) and a new gene found in *D. squalens* that encodes a CCP-ascorbate peroxidase (APX; EC 1.11.1.11) hybrid enzyme.

5.4.2. Class II fungal peroxidases.

Based on structure-functional considerations, class II fungal peroxidases are classified into four main groups: **(i)** lignin peroxidase (LiP; EC 1.11.1.14) proteins are defined as having an exposed catalytic tryptophan homologous to Trp171 in *Phanerochaete chrysosporium* LiP-H8 (encoded by *lipA*) and Trp164 of *Pleurotus eryngii* VPL, **(ii)** manganese peroxidase (MnP; EC 1.11.1.13) proteins are defined as possessing a Mn(II)-oxidation site near the internal propionate of heme formed by three acidic residues homologous to *P. chrysosporium* MnP1 Glu35, Glu39 and Asp179, and *P. eryngii* VPL Glu36, Glu40 and Asp175, and **(iii)** Versatile peroxidase (VP; EC 1.11.1.16) proteins are defined by the presence of both the catalytic tryptophan and the Mn(II) oxidation site of LiP and MnP, respectively. A similar structure-functional classification of these and some other basidiomycete peroxidase genes has been already used by Ruiz-Dueñas et al. (*91-93*).

In agreement with previous results (*95-97*) the LiP, MnP and VP families involved in lignin degradation are distributed in the white rot species, and are absent from the brown rot species. The MnP genes are present in all the white rot fungal genomes (Table S12) and LC-MS/MS analysis of aspen culture filtrates unambiguously identified MnPs in the white rot species *P. strigoso-zonata*, *F. mediterranea*, and *T. versicolor*. The analysis of MnP genes shows the existence of three new MnP "sub-families" well-represented in the SAP genomes (Table S12), in addition to the classical long-MnP described in *P. chrysosporium* (the latter with mature protein sequences 13-16 residues longer than those of typical LiP and up to 27 residues longer than VP) (*98*). The first new MnP sub-family is identified here by their short amino-acid sequences (*99*), similar to those of typical LiP and VP. Their catalytic properties are expected to be similar to those reported for the *Phlebia radiata* short-MnP (*99, 100*) that has Mn-independent activity, similar to generic peroxidases defined below, in addition to its Mn(II) oxidizing activity. In contrast, the second new sub-family is characterized by its "extra-long" sequences (*101*), with a polar C-terminal tail that binds metals providing unusual stability properties to the enzyme, as reported for the *Dichomitus squalens* extralong-MnP (*102*) that is 11 residues longer than the above long-MnP. Finally, several peroxidase genes present in the *A. delicata* and *S. hirsutum* genomes correspond to an atypical MnP with only two acidic residues at the Mn-oxidation site. There are three reasons suggesting that these genes correspond to functional MnPs: **i)** Mn-oxidizing peroxidases have been described in *Ganoderma* species whose sequences (GenBank ABB77243, ABB77244 and BAA88392) also have only 2 acidic residues at the Mn-oxidation site; **ii)** removal of one acidic residue from some Mn-oxidizing sites reduces activity but results in still active variants (*103*); and **iii)** Mn-oxidizing activity has been reported in *S. hirsutum* (*104*) and these are the only *mnp*-type genes present in its genome.

VP genes were found in two genomes (those of *D. squalens* and *T. versicolor*) while LiP genes were found only in one of the newly sequenced genomes, that of *T. versicolor*. Analogously to the

previously reported atypical MnP genes, an atypical VP gene has been found in the *T. versicolor* genome, showing the catalytic tryptophan in addition to an atypical Mn-oxidation site formed by only two acidic residues. No LiP or VP proteins were detected by LC-MS/MS.

In addition to the above ligninolytic peroxidase families, a fourth Class-II peroxidase type was identified in both white rot (*A. delicata, F. mediterranea, S. hirsutum* and *P. strigoso-zonata*) and brown rot (*F. pinicola* and *W. cocos*) fungal genomes, being characterized by the absence of both of the two catalytic sites mentioned above. These enzymes are expected to be low redox-potential peroxidases oxidizing substrates at the main heme channel, with catalytic properties similar to those of the well-known peroxidase of *Coprinopsis cinerea* or the product of the *nop*A gene of *P. chrysosporium* (*105*). Therefore, they are defined here as generic peroxidases (GP; EC 1.11.1.7).

5.4.3. Heme-thiolate peroxidases and dye-decolorizing peroxidases.

The above class I and class II peroxidases belong to the superfamily of Microbial, Fungal and Plant Peroxidases. The present screening of new basidiomycete genomes revealed a considerable number of sequences corresponding to peroxidase groups that were previously known only by a few representatives, and now appear as two largely-distributed new peroxidase (super)families. These new (super)families correspond to heme-thiolate peroxidases (HTP) and dye-decolorizing peroxidases (DyP; EC 1.11.1.19), which were identified by their general folding and typical heme environments.

The HTP (super)family includes both chloroperoxidase (EC 1.11.1.10) type sequences, first described in the ascomycete *Leptoxyphium fumago*, and the peroxygenases (EC 1.11.2.1) recently described in *Agrocybe aegerita* and other basidiomycetes that oxidize both aromatic and aliphatic compounds (*100, 106*). HTP genes have the widest distribution among the different peroxidase genes being present in all the white rot and brown rot genomes analyzed. HTPs are not clearly associated with wood degradation. The *A. aegerita* HTP has been shown to hydroxylate aromatic rings, but the specific role of the enzyme and its action on lignin or humic substances, or its possible role in the production of secondary metabolites are unclear (*107*). The reconstructions of HTP evolution suggest that the common ancestor at node 7 had 5 (CAFÉ) or from 3-7(Notung) gene copies, while both results resolve either 2 expansions between nodes 6 and 7 or 3 expansions between nodes 7 and 9, suggesting that there were expansions in the family early in the evolution of Agaricomycotina. There is no clear transition in copy number among species with different nutritional strategies. Significant expansions ($p<0.01$) are supported in CAFÉ for *A, delicata*, *C. cinerea* and *S. hirsutum* while losses are supported only for the Boletales and *S. commune*.

In contrast to HTPs, DyP genes are widespread among the white rot genomes but are absent from most of the brown rot genomes, with *A. delicata* having the largest number of DyPs among the genomes sampled (16 copies; Table 1). Exceptions to this pattern include the brown rot *P. placenta*, which possesses two DyP-type genes, and *P. chrysosporium* and *S. commune*, which lack DyP genes. CAFÉ and Notung results suggest the presence of one DyP gene at node 7, with expansions in several white rot species, as well as *C. cinerea* and *P. placenta* (Figs. S8m, S9-10; Table S8). Conversely, the reconciliation results support losses for most brown rot species (from 1 to 3 losses). These results along with the recent findings on the activity of DyP for *A. auricula-judae (92)* suggest a potential role of DyP in white rot (*104*, *108*), even though the presence of DyP genes in *L. bicolor* and *C. cinerea* could indicate involvement of the enzyme in additional functions. In this connection, we also detected DyP-like proteins in filtrates from *D. squalens, P. strigosozonata*, and *A. delicata*. The *D. squalens* and *P. strigosozonata* genes had EST support, although their transcripts did not accumulate in Avicel-containing medium relative to complete media (see Supplementary Electronic Files).

5.4.4. Genes of H_2O_2-generating (flavo and Cu-radical) oxidoreductases.

As mentioned above, some secreted peroxidases play a central role in the attack of wood lignin by white rot fungi. However, extracellular oxidases reducing molecular oxygen to hydrogen peroxide play

a more general role in decay since they are involved in wood degradation by both white rot and brown rot fungi. In the first case, peroxide is the required oxidizing-substrate of ligninolytic peroxidases enabling the subsequent oxidation of lignin (*109*), while in the second case peroxide is reduced by ferrous iron to hydroxyl radical (Fenton-type reaction), which acts as a strong diffusible oxidizer initiating cellulose depolymerization without the previous degradation of lignin in brown rot decay (*110, 111*). Genes encoding H_2O_2-generating enzymes in the basidiomycete genomes belong to two different oxidoreductase types: the glucose-methanol-choline (GMC) oxidoreductases and the copper radical oxidase (CRO) gene groups discussed below (oxalate oxidase discussed in the last section can also contribute to peroxide generation).

Concerning GMC flavoproteins, genes (only 1-2 models) of cellobiose dehydrogenase that can also use molecular oxygen as electron acceptor (CDH; EC 1.1.99.18) are present in all the white rot genomes, as well as in those of the brown rot *C. puteana, S. lacrymans* and *G. trabeum* (but absent from other brown rot genomes; Table 1), in agreement with previous reports on its presence in different basidiomycetes including one brown rot species (*112*). The precise role of CDH remains uncertain although recent studies suggest cellulolytic boosting in conjunction with the GH61 metalloproteins (*113*). Along these lines, we noted that 4 of the 5 CDH producing cultures (*S. hirsutum, C. puteana, T. versicolor, D. squalens*) also secreted aldose epimerase. The latter enzyme has been suggested to enhance cellulose degradation via interconversion of glucose anomers (*114*). The reconstruction results using Notung indicate one gene present for the common ancestor on node 7, while 1-2 copies are supported higher on the tree with losses for *Dacryopinax*, the brown rot polypores and *L. bicolor* (Figs. S9-10; Table S8). The identification of other GMC flavooxidases, which were not included in the ancestral reconstruction analyses, revealed that only methanol oxidase (MOX; EC 1.1.3.13) genes are found in all of the new white rot and brown rot genomes analyzed (Table S12). This enzyme has been suggested as the main source of peroxide in brown rot fungi (*115*) but its contribution to white rot decay is less documented. The presence of one copy of methanol oxidase in *Dacryopinax* implies the ancient presence of the gene at least on node 7. Additionally, genes of aryl-alcohol oxidase (AAO; EC 1.1.3.7), which has been reported as the main source of peroxide in white rot *Pleurotus* and *Bjerkandera* species (where it acts synergistically with VP) (*116*), were found in all the white rot genomes (with the only exception of that of *A. delicata*) being more rare in the brown rot genomes. Among sugar oxidoreductase genes, those of glucose oxidase (GOX; EC 1.1.3.4) and pyranose dehydrogenase (PDH; EC 1.1.99.29) were absent from all the genomes analyzed, and that of pyranose-2 oxidase (POX; EC 1.1.3.10), another oxidase whose involvement in wood decay has been suggested (*117*), was found only in the *T. versicolor* and *A. delicata* genomes.

Glyoxal oxidase (GLX; EC 1.2.3.-) is the only well-characterized CRO in wood-rotting basidiomycetes, after its description in *P. chrysosporium*, where it is physiologically coupled with LiP (*118*). GLX is the only extracellular oxidase reported in *P. chrysosporium* and, together with the products of related genes (*cro1* to *cro6*) found in its genome (*119*), they are the main candidates for H_2O_2 production in this fungus. The phylogenetic analysis of CRO without WSC domains (Fig. S8n) recognized 4 distinct clades with GLX (glyoxal oxidases) being the sister group to CRO6, and CRO1 being the sister group to CRO2. CRO3-5, which contain WSC domains, were analyzed separately and the phylogenetic tree indicates that all the related sequences form one clade. The *glx* genes are present in most of the white rot genomes examined here (with the only exception of *F. mediterranea* and *H. annosum*) but are absent from the brown rot genomes and the genomes of *C. cinerea* and *L. bicolor*. However, the other groups of CRO (*cro1- cro6*) have a more widespread distribution across Agaricomycotina (cro 3-5 are not phylogenetically distinct). Additionally, the phylogenetic analysis indicates that homologs from all the groups of CROs were at least present in the common ancestor of *Dacryopinax* and Agaricomycetes or even earlier but the clade of GLX indicates losses for all brown rot fungi. The presence of two *glx* in the genome of *A. delicata* indicates that probably the common ancestor of Agaricomycetes was able to generate H_2O_2 using a glyoxal oxidase as many white rot species do. Accordingly the reconstruction results indicate that there were 4 to 5 copies of CROs present on node 7 with subsequent significant changes (CAFÉ results) in copy number for certain

lineages (expansions in the core polypores, *S. hirsutum*, *P. strigosozonata*, *A. delicata*, *L. bicolor* and contractions for *S. lacrymans* and *S. commune*).

5.4.5. Laccases and other multicopper oxidases (MCOs).

Laccase activity (EC 1.10.3.2) has traditionally been considered a characteristic of lignin-degrading white rot fungi. The detection of laccase activity in plate cultures has been used as a diagnostic character to identify this (ecological) group of basidiomycetes in pure culture (*120*). However, certain efficient lignin degraders such as *P. chrysosporium* lack any laccases. It is also clear that laccases have insufficient oxidative potential to cleave the major non-phenolic linkages of lignin and involvement of unknown small molecular weight mediators have been proposed. In recent years laccases have been isolated from liquid cultures of several brown rot fungi, and the presence of laccase genes was reported in the first-sequenced brown rot fungal genome (*95*). During wood decay, laccases would oxidize lignin-degradation products or even the whole polymer in the presence of natural mediators (*121*), and would also contribute to hydroxyl radical generation by both white and brown rot fungi (*37, 122*). Our LC-MS/MS analysis of aspen culture filtrates identified 15 laccases including 5 from *F. mediterranea*, 4 from *T. versicolor*, 3 from from *D. squalens*, and 2 from *P. strigosozonata*. A single laccase was detected among the brown rot fungi, Wolco1165246 (see Supplementary Electronic Files).

We were interested to identify MCO and especially laccases in the newly sequenced genomes along with the rest of the genomes included in the study. The data mining of these genomes revealed the presence of laccase "sensu stricto" genes in all the basidiomycete genomes analyzed except *P. chrysosporium*, as reported previously (*123*), *Dacryopinax* and *T. mesenterica*, with the number of genes per genome being higher in the white rot (7-15 models) than in the brown rot species (0-6 models). *Schizophyllum commune* has only 2 laccases, resembling more the brown rot species in this regard. Phylogenetic analyses divide the laccases sensu stricto into two clades, one containing all genes from Agaricomycotina and the other containing all genes from Ascomycota, except that a strongly supported clade of seven genes from *A. delicata* is nested within the ascomycete laccases, with moderate support (Fig. S8o). CAFÉ and Notung reconstructions suggest that the common ancestor of the Dikarya (node 3) possessed from one to three laccase gene copies. Both CAFÉ and Notung reconstructions support the presence of two laccases on node 7 (the common ancestor of Agaricomycetes plus Dacrymycetes) with further expansions either between nodes 6 and 8 (CAFÉ) or between nodes 8 and 9 (Notung) (Table S8). The CAFÉ results support mainly expansions for some white rot species and *C. cinerea* but not for brown rot species, with contractions for *S. commune*, *P. chrysosporium* and *G. trabeum*. On the other hand, the Notung results support more prominent expansions for the white rot species and relative maintenance of gene copy numbers for most of the brown rot species or slight reductions.

In addition to genes encoding laccase "sensu stricto", other MCO genes were assembled (Table S12) but were not included in the reconstruction analysis. The remaining MCOs are split into at least three different types of enzymes: **i)** MCO subfamily ferroxidase/laccase; **ii)** pigment MCOs; and **iii)** fungal ascorbate oxidase. Among them, the Fet3 ferroxidases are relatively conserved across Agaricomycotina with 0-2 copies per genome, while there is more variation in gene numbers for the rest of MCOs (Table S12).

5.4.6. Fe-reducing glycopeptides, quinone reductases, and cytochrome b652 (enzymes involved in Fenton reaction).

The initial decay of wood by brown rot fungi seems to be largely based on the generation of the hydroxyl radical being able to attack cellulose without the previous removal of lignin (*111, 124-126*), and hydroxyl radical can be also generated by white rot fungi (*127, 128*). As mentioned above, hydroxyl radical originates from hydrogen peroxide reduction by easily oxidizable Fe(II) (Fenton-type reaction). Therefore, in order to maintain hydroxyl radical generation, basidiomycetes must operate efficient reducing systems. For this reason Fe(III)-reducing glycopeptides (GLP) (*129*), quinone-

reductases (QR; EC 1.6.5.2) (*130*) and Fe(III)-reducing cytochrome b652 genes were sampled across the 31 genomes. While GLP and cyt b652 reduce Fe(III) directly, QR participates in Fe(III) reduction by quinone redox-cycling. In this mechanism, hydroquinones derived from lignin or synthesized by the fungi are oxidized by laccases (or peroxidases) to semiquinone radicals that quickly reduces Fe(III) to Fe(II) (the role of QR being to maintain redox-cycling by reducing quinones to hydroquinones) (*121, 131*).

Of the three gene families discussed here, only GLP was supported by CAFÉ as having significant family-wide *p*-values, indicating shifts in diversification rates (Table 1). Genes encoding GLP are present in all the Agaricomycetes and *Dacryopinax*, but not in *C. cinerea*, and no copies were identified outside of the Basidiomycota. The ancestral reconstruction results obtained with CAFÉ and Notung suggest 2-3 copies present on node 7, with expansions supported between nodes 6 to 7, 8 to 9, and for some white rot and brown rot species (*C. puteana*, brown rot polypores, *S. hirsutum* and *F. mediterranea*), while losses are supported mostly for some white rot species and less for some brown rot species (*S. lacrymans* and *P. placenta*) (Figs. S7, S8q, S9-10; Table S8). The expansions in some brown rot lineages, along with the generally higher copy number of GLP genes in brown rot genomes (3-11, avg. 6.9) vs. white rot genomes (1-11, avg. 4.8) suggests that these enzymes may play an important role in brown rot.

The Fe(III)-reducing cytochrome b652 genes (which in some cases include a CBM1 domain at the C- terminal region of the protein, and thus could be involved in iron reduction and hydroxyl radical generation in close proximity to crystalline cellulose) were found only in some white rot fungi, while they are missing from all the brown rot fungi except the brown rot Boletales (Table S12). Recently, one of the two copies present in *S. lacrymans* (model 452187) was shown to be highly upregulated during wood degradation (*132*). Of the two copies of cytochrome b652 found in *S. lacrymans* only one has a CBM1 domain. One of the three cytochrome b652 proteins found in *C. puteana* also carries a CBM1 domain. The ancestral reconstruction using Notung implies the presence of two copies of this gene at node 7, with reductions in many lineages.

Quinone reductase genes are widespread across the studied genomes (up to four copies per genome; Table 1) with only *M. globosa* having no copies of QR. The ancestral reconstruction supports the presence of two genes for node 7 with no major changes across Agaricomycetes and Dacrymycetes. The general maintenance of QR genes in fungal species with diverse nutritional strategies could be attributed to the participation of these enzymes in detoxification processes, not limited to wood degradation.

5.4.7. Oxalate oxidases/decarboxylases.

Oxalate seems to have multiple roles in fungal nutritional strategies including decrease of the extracellular pH, which can enhance activity of wood degrading enzymes, metal ion chelation including removal of calcium from pectin, which makes the latter susceptible to pectinases and even production of H_2O_2 through degradation of oxalate by oxalate oxidase (*133*). Thus, the control of accumulation of extracellular oxalate is an important component of fungal nutritional strategies, with oxalate oxidase (EC 1.2.3.4) and oxalate decarboxylase (EC 4.1.1.2) being the two enzymes that control the fate of oxalate. In spite of their different catalytic properties (the oxidase catalyzing O_2 + $HO_2C\text{-}CO_2H \rightarrow H_2O_2 + 2\ CO_2$, and the decarboxylase catalyzing $O_2 + HO_2C\text{-}CO_2H \rightarrow HCO_2H + CO_2$), both oxalate-degrading oxidoreductases (OXOs) share a common fold and structural motifs, as members of the cupin superfamily (*134, 135*). Oxalate decarboxylases have been reported in several basidiomycetes (*136*), but oxalate oxidase has been previously reported only in *Ceriporiopsis subvermispora*, where it could constitute an initial source of hydrogen peroxide for ligninolytic peroxidases (*134*). The screening of OXOs showed that they are present in all the white rot and brown rot basidiomycete genomes analyzed, but absent from that of *T. mesenterica*. Accordingly, the ancestral reconstruction using Notung implies the presence of 4 copies at node 7 and increased numbers of copies for node 10 with a complicated history of expansions and contractions across the Agaricomycetes.

5.4.8. Cytochrome P450s.

Cytochrome P450 monooxygenases are mixed function oxidoreductases ubiquitously distributed throughout living organisms, although their number varies widely among and within the genomes of different kingdoms and phyla. Overall, fungi possess extraordinarily large P450 counts in their genomes, next only to plants. Except for those involved in cell wall ergosterol biosynthesis (CYP51 and CYP61), fungal genomes show no tightly conserved P450 clans/families across the different phylogenetic classes (*137, 138*). Nevertheless, P450 families CYP53 (benzoate p-hydroxylase) and CYP505 (fatty acid oxygenases) are also fairly conserved and distributed across various fungal classes. Previous comparison of the P450omes of previously sequenced basidiomycetes (*Phanerochaete*, *Postia*, *Schizophyllum*) and other fungal classes (ascomycetes, zygomycetes) shows greater expansion of P450 clans and families in basidiomycetes. The sampling across the 31 genomes of P450 domain carrying protein models revealed that while basal lineages in Basidiomycota harbor a low number of these enzymes, including the basal lineage of Tremellomycetes in Agaricomycotina, Dacrymycetes and Agaricomycetes show dramatic expansion in P450 diversity. Reconciliation and CAFÉ analyses were not attempted with P450s, owing to the large number of copies per genome.

The highest number of potential P450s across Pucciniomycotina, Ustilaginomycotina and Tremellales is found in *M. laricis-populina* (28 predicted models). On the other hand, *Dacryopinax* and Agaricomycetes include much higher numbers of potential P450s. Despite the high numbers of P450s for these genomes, there is no distinct pattern between white rot and brown rot genomes. The expansion and reduction of P450 monooxygenase clans/families among the individual SAP genera also do not follow any specific phylogenetic pattern (Fig. S20). The widespread presence of P450s in brown rot and white rot genomes implies their role in the overall wood decay process and reinforces the idea that P450 monooxygenases play a role in downstream events of the lignin degradation process. This is consistent with results from transcriptome and exo-proteome analyses, which demonstrate differential regulation of several P450s under lignocellulolytic culture conditions in selected model basidiomycetes (*95, 137, 139*). In addition, P450s of white rot basidiomycetes play a direct role in the oxidation and biodegradation of xenobiotic compounds including those bearing similarties to lignin substructures, which makes these species prominent candidates for bioremediation studies (*140, 141*). The presence of diverse P450s in *B. dendrobatidis* and *P. blakesleeanus* indicates the ancient presence of these genes in the fungal kingdom.

5.4.9. Summary comments on oxidoreductases.

The overall picture for oxidoreductases suggests that gene families encoding enzymes that directly participate in lignin degradation (class II peroxidases, laccases ‘sensu stricto’, DyP) are under different selection regarding the number of gene copies between white rot species and *C. cinerea* vs. brown rot species and *L. bicolor*. On the other hand, genes that participate in peripheral reactions during wood degradation such as H_2O_2 production (involved in both lignin and cellulose depolymerization), iron reduction and hydroxyl radical formation appear not to be under differential selection for copy number (with few exceptions, such as *glx* genes) with regard to the nutritional strategies referred above. These patterns can be attributed to the complementarity of the systems that can generate the same product such as H_2O_2, implying additionally that such mechanisms are widespread and were present in the common ancestor of the Agaricomycotina, forming a basic system that was modified in white rot and brown rot lineages. In this context, it should be acknowledged that our analysis focused on genes whose function is inferred from structural similarity to known proteins. Our expression analysis lends additional support in many cases. Still, LC-MS/MS identified 436 extracellular proteins for which no function could be determined (see Supplementary Electronic Files). Many of these proteins could be involved directly or indirectly in lignocellulose degradation.

6. Analyses of chromosomal clustering of decay-related genes.

Genes encoding decay-related CAZymes and oxidoreductases (described above) were inferred to be clustered with other genes from the same set if no more than 6 annotated genes separated them on a genomic scaffold. Chromosomal clustering of genes involved in decay processes takes 2 forms: linkage of paralogs (paralog clusters, PC), usually the result of recent local duplications, or associations between different gene families contributing to a common decay path (metabolically related clusters, MRC). A substantial number of decay gene clusters were found among the 31 genome set (see supplementary electronic files). There were on average 7.81 PCs, 4.38 MRCs and 0.74 clusters that are a combination of both types per genome. ClassII peroxidase clusters (Fig. S19) demonstrate gene family expansion patterns, and also suggest convergent origins of associations between peroxidases and glycoside hydrolases. The largest PCs are made up of recent gene family expansions in the white rot Polyporales, *P. chrysosporium* and *T. versicolor.* There is no evidence of conservation or horizontal transfer of MRCs that would suggest strong associations among the constituent gene families.

7. Micromorphological observations of wood decay.

White and brown rot fungi attack wood differently resulting in distinct cell wall changes. Micromorphological observations of alterations in the woody cell walls resulting from these fungal processes have been used to elucidate degradation patterns not only in wood but also in permineralized fossil woods to determine the type of fungal attack (*142-144*). Since the type of decay caused by several fungi used in this study had not been studied, the physical aspects of decay were investigated using scanning electron microscopy on wood wafers (*145*), with six strains, including four white rot species (*A. delicata, F. mediterranea, P. strigosozonata, S. hirsutum*) and two brown rot species (*Dacryopinax* sp., *G. trabeum*). Structural features of wood are different among various species of woody plants and an example of gymnosperm, pine (*Pinus* sp.) and angiosperm wood, aspen (*Populus* sp.) were used to evaluate the type of decay caused by each fungus. Dried wood wafers from the sapwood of aspen (*Populus* sp.) and pine (*Pinus* sp.) were cut to 10 x 10 x 1 mm, hydrated to 80-100% and autoclaved for 60 minutes at 120°C. Ten wood wafers were placed on 7 cm diameter glass filter paper that were in 100 mm x 15 mm Petri plates with 1.5% malt extract agar (15 g malt extract, 15 g agar, 1000 ml water). Six plates of each wood type were inoculated with each fungus by placing six 2 mm plugs of the assay fungus evenly around the glass filter paper 4 days before wafers were added. Sources of strains are listed in Supplementary Notes section 1. Wafers were removed 90 days after inoculation and prepared for scanning electron microscopy as previously described (Blanchette et al. 2010). Samples were examined and photographed using a Hitachi S3500 N (Hitachi, Tokyo, Japan) scanning electron microscope. Patterns of cell wall degradation were consistent with previous reports for white and brown rot fungi (Fig. S20). White rot species completely degraded cell walls in advanced stages of decay, whereas brown rot species caused deformation of cell walls due to cellulose depolymerization, with weakened cell walls fracturing in later stages of decay. See Fig. 20 caption for details.

8. Carbohydrate utilization analyses.

Twenty-two species (19 Basidiomycota and three Ascomycota) were grown in duplicate on minimal medium (Table S13) with the following carbon sources:

- 25 mM D-glucose, D-fructose, D-mannose, D-galactose, L-rhamnose, D-xylose, L-arabinose, D-ribose, D-galacturonic acid, or D-glucuronic acid (monomers)
- 25 mM cellobiose, maltose, sucrose, lactose, or raffinose (oligosaccharides)
- 1% soluble starch, inulin, α-cellulose, beechwood xylan, birchwood xylan, guar gum (galactomannan), arabinogalactan, arabic gum, sodium lignin, apple pectin, or citrus pectin (pure polymers)

- 3% wheat bran, sugar beet pulp, rice bran, citrus pulp, cotton seed pulp, soybean hulls, or alfalfameal (crude biomass)
- No carbon source (control)
- 1% casein (control)

Each species was allowed to grow until the fastest growing colony had almost reached the edge of the plate after which pictures were taken of all the plates. For comparison the growth on plates without a carbon source were used as a negative control. Growth on D-glucose was used as an internal standard as this was the best monomeric carbon source for all fungi. Comparisons between fungi were made by comparing the growth on a specific carbon source relative to growth on D-glucose.

Coprinopsis cinerea medium consists of 2.28 g/l asparagin monohydrate, 1 g/l KH_2PO_4, 2.9 g/l Na_2HPO_4 * H_20, 0.3 g/l Na_2SO_4, 0.5 g/l ammonium tartrate, 0.04 mg/l thiamine hydrochloride (vitamine B1), and 0.25 g/l $MgSO_4$ *$7H_2O$. *Rhizoctonia solani* medium consists of 0.5 g/l $(NH_4)_2$-tartrate, 1 g/l KH_2PO_4, 0.5 g/l $MgSO_4$ *$7H_2O$, 0.1 mg/l thiamine-HCl, and trace elements (Kanieltra) with a pH adjusted to 5.5. *Schizophyllum commune* medium consists of 0.5 g/l MgS04 * 7 H20, 1.5 g/l L-aspargine (monohydrate), 0. 12 mg/l thiaminiumdichloride, 5 mg/l $FeCl_3$ * 6 H_20, 0.06 mg/l HBO_3, 0.04 mg/l $(NH_4)_6Mo_70_{24}$ * $4H_20$, 0.2 mg/l $CuS0_4$ * $5H_20$, 2 mg/l $ZnS0_4$ * $7H_20$, 0.1 mg/l $MnS0_4$ * $4H_20$, 0.4 mg/l $CoCl_2$ * $6H_20$, 1.2 mg/l $Ca(N0_3)_2$ * $4H_20$, 0.46 g/l KH_2PO_4, and 1 g/l K_2HPO_4. *Aspergillus niger* medium consists of 6 g/l $NaNO_3$, 1.5 g/l KH_2PO_4, 0.5 g/l $MgSO_4$ * $7H_2O$, 0.5 g/l KCl, and trace elements (Vishniac) with a pH adjusted to 6.0. *Cryptococcus neoformans* medium consists of 21.5 mM Urea, 10 mM KH_2PO_4, 1.2 mM $MgSO_4$ * $7H_2O$, 1 µM $MnCl_2$ * $4H_2O$, 10 µM $FeSO_4$ * $7H_2O$, 10 µM $ZnSO_4$ * $7H_2O$, 1.2 µM $CuSO_4$ * $5H_2O$, 100 µM $CaCl_2$ * $2H_2O$, 10 µg/l biotin, and 2 mg/l thiamin. Ustilago maydis medium consists of 3 g/l KNO_3, 1 g/l KH_2PO_4, 0.25 g/l Na_2SO_4, 0.5 g/l KCl, 0.13 g/l $MgSO_4$, 0.08 g/l $CaCl_2$ * $2H_2O$, and trace elements with a pH set for 7.0. *Stagonospora nodorum* medium consists of 2.0 g/l $NaN0_3$, 1.0 g/l K_2HP0_4, 0.5 g/l KCl, 0.5 g/l $MgS0_4$ * $7H_20$, 0.01 g/l $ZnS0_4$ * $7H_20$, 10 mg/l $FeS0_4$ * $7H_20$, and 2.5 mg/l $CuS0_4$ * $5H_20$. Yeast medium consists of 10 g/l Yeast nitrogen base (Difco). All fungi were grown at 25°C, except for *A. niger* (30°C) on solid medium containing 1.5% agar. *Trichoderma* medium consists of 2 g/l KH_2PO_4, 1.4 g/l $(NH_4)_2SO_4$, 0.3 g/l $CaCl_2$ * $7 H_2O$, 5 mg/l $FeSO_4$ * $7 H_2O$, 2 mg/l $ZnSO_4$ * $7 H_2O$ and 2 mg/l $MnSO_4$ * H_2O with a pH set for 5.5.

Significant differences in the growth profiles were observed between the species (Table S14). Although all species were able to grow on D-glucose, growth rate and morphology differed strongly with some species growing sparsely on all carbon sources tested. Using growth on D-glucose as a reference, specific difference in their ability to use other carbon sources could be observed. Differences in growth on monomeric carbon sources were most pronounced on D-galactose, L-arabinose and D-galacturonic acid where several species demonstrated no or very poor growth compared to D-glucose.

All species except *C. cinerea*, *D. squalens*, *G. trabeum* and *U. maydis* grew similarly on starch and D-glucose. In contrast, growth on inulin was reduced compared to D-glucose for all basidiomycetes, suggesting this compound is not a common substrate for these fungi. This was also true for growth on sucrose, except for *S. commune*, *S. hirsutum*, *C. neoformans* and *U. maydis*. This does not correlate with the numbers of GH32 enzymes in their genomes, suggesting that regulation of these genes or substrate specificity of the enzymes may play a major role in this. The ability to grow well on sucrose, but poorly on inulin was previously observed in *Aspergillus oryzae* (*146*). Comparison to *A. niger* that grows well on both substrates demonstrated the presence of both endo- and exo-inulinases in this species, while *A. oryzae* only contains exo-inulinases, suggesting that depolymerization of inulin is essential for efficient growth on this substrate. Based on that, the data in this study suggest that the GH32 members of *S. commune*, *S. hirsutum*, *C. neoformans* and *U. maydis* likely encode exo-inulinases, rather than endo-inulinases.

All fungi tested showed poor growth on cellulose, which may not be a true reflection of their ability to degrade cellulose, but could be due to the chemical modification of commercial cellulose during its isolation. Interestingly, for nearly all species (with the exception of *C. cinerea*, *S. hirsutum*, *D. squalens* and *U. maydis*) growth on pectin is similar to or better than growth on D-glucose. While pectin is a minor component of wood, its location around bordered pits may be the reason for this ability in the basidiomycete species, as was suggested previously (*147*). A much larger difference in

growth is observed on xylan, a major component of hardwood, with many species growing poorly on this substrate. In contrast, growth on guar gum, a galactomannan similar in structure to softwood cell wall galactomannan, was better for most basidiomycetes, with the exception of *C. puteana*, *P. placenta*, *F. pinicola* and *P. chrysosporium*. All of the white rot fungi tested were able to grow on lignin-containing medium, except *S. commune*. Four brown rot species and the ectomycorrhizal *L. bicolor* were unable to grow on lignin, but two other brown rot species, *F. pinicola* and *G. trabeum*, were also able to grow on lignin-containing medium.

9. Analysis of symbiosis-related gene families.

The genome of the ectomycorrhizal *Laccaria bicolor* was compared to those of saprotrophic and pathogenic Fungi to gain insight into the mechanisms of evolution of ECM symbiosis. Expansion of protein family sizes in *L. bicolor* was prominent in lineage-specific multigene families (*37*). Marked gene family expansions occurred in genes predicted to have roles in protein–protein interactions (for example, WD40-domain-containing proteins) and signal transduction. Within this set of orphan genes, a large number of genes was found that encode products that have a predicted size of <300 amino acids, which are termed mycorrhiza-induced SSPs (MISSPs). MISSPs may function as effector proteins to manipulate host cell signalling or to suppress defence pathways during infection (*148*), as suggested for pathogenic rusts, smuts and oomycetes. Several transcripts coding for expanded and lineage-specific gene families, including MiSSPs, were upregulated in symbiotic tissues, suggesting a role in ectomycorrhizal tissue differentiation. Many pathogenesis-related proteins, such as virulence factors, evolved very rapidly and show little sequence similarity with proteins from other strains and/or species.

The high number of symbiosis-regulated genes belonging to expanding orphan gene families suggest that the symbiont genomes may possess genes typically lacking in saprotrophic Agaricomycotina. We therefore investigated the occurrence of *L. bicolor* symbiosis-regulated genes in the 31 SAP species. Among the 1,774 *L. bicolor* transcripts up-regulated in ectomycorrhizae (based on NimbleGen array profiling (*37*)), 884 have no homolog in the investigated brown- and white rots (Fig. 21), suggesting that 50% of the SR-genes are coding for symbiosis-related functions that evolved recently. Most *L. bicolor*-specific genes (~90%) were coding for orphan proteins with no homolog in NCBI and/or no known KOG, PFAM or GO functions. Among the 887 symbiosis upregulated-genes that were shared with SAP species, 621, 701, and 630 have no KOG, GO or PFAM function, respectively. The KOG category with the highest number of shared genes coded for components of the transduction pathways.

In contrast, most *L. bicolor* transcripts with a down-regulated expression in ectomycorrhizae were found in the brown- and white rots (905/1,009) (Fig. S22). They code for general metabolic (energy production, protein turnover) and developmental (signal transduction) pathways. In conclusion, these results suggest that most symbiosis-induced genes are specific to the ectomycorrhizal symbiont *Laccaria bicolor* and are not shared by diverse wood decay and pathogenic fungi, suggesting that they play important roles in the ectomycorrhizal lifestyle.

10. Molecular clock analyses.

10.1. Organismal molecular clock analyses.

Fossils representing the fungal tree of life are scarce which complicates molecular clock analyses (*149*). Well characterized fossils were used to calibrate the minimum ages of three different clades in the Dikarya. The 50 mya permineralized suilloid ectomycorrhiza fossil associated with pine roots from the middle Eocene Princeton chert of British Columbia (*150*) was used to calibrate the Boletales (BOL) represented by *Serpula lacrymans* and *Coniophora puteana*. It is the only existing ectomycorrhizal fossil and it has not been associated with fruiting bodies, however, the combination of a pinaceous host and tuberculate ectomycorrhizal morphology places the fossil in an extant suborder of Boletales, the Suillineae. *Archaeomarasmius legetti* from mid-Cretaceous amber already resembled

the morphology of modern agarics (small mushrooms with a stipe and a cap) and it has been placed in the marasmioid clade in the Agaricales accordingly (*151*). It was discovered in a well-characterized layer of clay in New Jersey that dates to 94-90 mya and was used to calibrate the Agaricales (AGA) including *Coprinosis cinerea*, *Laccaria bicolor*, and *Schizophyllum commune*. *Paleopyrenomycites devonicus* is a 400 my old ascomycete fossil from the early Devonian that was associated with the extinct vascular plant *Asteroxylon mackiei* (*152*). Recent studies reject the hypothesis that *P. devonicus* could have represented an early diverging lineage in the Ascomycota, such as the Taphrinomycotina (*153, 154*), while others provide evidence not to calibrate the crown Pezizomycotina (the more derived euascomycetes) with this age (*149, 154, 155*), which would push back the split between Basidiomycota and Ascomycota to almost 850 mya. Consequently, the split between the subphyla Saccharomycotina (yeasts, represented by *Pichia stipidis*) and Pezizomycotina (*Aspergillus niger, Cryphonectria parasitica, Stagonospora nodorum*, and *Trichoderma reesei*) in the Ascomycota (ASC) was calibrated with *P. devonicus*.

The reduced 26G dataset (see section 4, Organismal phylogenetic analyses) was analyzed using Bayesian relaxed molecular clock approaches in BEAST (*156*) v1.6.1 and PhyloBayes (*47, 157*) 3.2. A preliminary xml file was prepared with the aid of BEAUTi (*156*) v1.6.1 to create a starting tree for the BEAST analyses. Twenty-one tMRCA's (time to most recent common ancestors) were defined describing the consensus tree topology obtained from RAxML and PhyloBayes analyses. The three fossil calibrations (nodes 14, 16, and 27 in Figure 1) followed a lognormal distribution, where the minimum ages for the calibrations described above provided the offset values. Secondary estimates for the mean values of the distributions were derived from previous studies (*132, 149*), which used extended taxon sampling for AGA (mean = 124 mya), BOL (mean = 113 mya), and ASC (mean = 452 mya). The standard deviation of all distributions was set to 1.0. Bayesian relaxed uncorrelated lognormal clock analyses with a birth-death tree prior were run for three million generations under the WAG substitution model (Gamma+invariant sites). One of the post-burnin trees in Newick format was added to the final xml file, which was used to launch six independent BEAST MCMC chains running 10 million generations each, sampling data every 1000th generation. The resulting log files were combined using LogCombiner v1.6.1 and inspected with Tracer (*156*) v1.5 to confirm that the estimated sample sizes (ESS) for statistics represent appropriate values. The ESS for Tremellomycetes (node 24) and Ustilaginomycotina (node 25) were slightly below the recommended threshold of effectively independent draws (= 200) from the posterior distribution, but all other statistics had ESS values well in excess of 200 (Table S15). The runs converged to stable likelihood values independently after one million generations and 9000 ultrametric trees from each run were combined and analyzed using TreeAnnotator (*156*) v1.6.1 to estimate the 95% credibility node intervals (CI) or highest posterior densities (HPD), which mark the lower and upper time boundaries. The consensus chronogram including the 95% HPD and the mean age estimates is shown in Fig. 1. Three additional BEAST analyses using the same settings, but omitting one of the fossil calibrations at a time, were run to examine the influence of individual calibration points on the robustness of the age estimates (Table S16).

Soft bound calibrations were used in analyses of the 26G dataset with PhyloBayes, which require a birth-death prior (the alternative option is to use hard bounds). The minimum and maximum boundaries for ages (AGA = 140 – 92 mya; BOL = 130 – 50 mya; ASC = 600 – 452 mya) were assembled using the fossil minimum ages and maximum values were taken from the tail of the lognormal distributions used in BEAST, i.e., PhyloBayes was constrained to estimate the mean ages of the calibrated nodes not to exceed the lower 95% HPD estimates. The CAT model was used to infer a tree with which the tree topology was fixed in the molecular clock analyses, which had overall support of 1.0 posterior probability and was topologically identical to the starting tree used in BEAST analyses. In addition, PhyloBayes ideally requires an age prior for the root, which was not possible in our case because the age of *Batrachochytrium dendrobatidis* is not substantiated by dated fossils. Thus, using a non-optimal prior for the root, PhyloBayes was likely to over estimate the ages of nodes anteceding the ASC calibration. Lognormal autocorrelated relaxed clock (ln) analyses were run using six MCMC chains, sampling data every 100th cycle. The maximum difference in split frequency between runs dropped to zero after seven million generations and the analyses were stopped. The

chains were analyzed using the readpb program, removing 10% of the samples as burn-in. The resulting mean and 95% CI time estimates are shown in Table S16.

10.2. Class II peroxidase gene molecular clock analyses.

The class II peroxidase gene family dataset has 126 amino acid sequences and 307 aligned sites, including lignin peroxidases, manganese peroxidases, versatile peroxidases, and generic peroxidases. To estimate the absolute age of the origin of ligninolytic enzymes, two secondary calibrations taken from the organismal time estimates were applied to this dataset defining the age of the root node of Dikarya (mean = 661.6, offset = 160, stdev = 1) and the node that includes the GPs of Pezizomycotina (mean = 343.6, offset = 100, stdev = 1) using BEAUTi. Six additional tMRCAs (time to most recent common ancestor) were defined to constrain well-supported terminal clades inferred from a RAxML bootstrap analysis (Fig. S8l). Bayesian relaxed uncorrelated lognormal molecular clock analyses were performed using a birth-death tree prior under the WAG substitution model (Gamma+invariant sites). Six independent MCMC BEAST chains were run for 10 million generations each, sampling data every 1000th generation. Average node heights and 95% hpd intervals are shown in Fig. 1.

II. Supplementary Tables

Table S1. Genome assembly summary.

	Genome assembly size (Mbp)	Number of contigs	Contig N50	ContigL50(Kb)	Number of of scaffolds	Scaffold N50	Scaffold L50 (Mbp)	% of scaffold length in gaps	Sanger reads coverage	454 paired-end coverage	454 standard coverage	Illumina reads coverage	Total read coverage
Auricularia delicata	75.1	4884	544	32.1	1531	44	0.49	8.00%	0.20x	8.26x	34.85x	3.26x	46.57x
Coniophora puteana	42.97	622	47	240.5	210	7	2.4	2.60%	-	23.77x	21.22x	4.56x	49.54x
Dacryopinax sp.	29.5	878	117	68.2	99	10	1.2	6.40%	1.49x	11.02x	26.13x	4.55x	43.19x
Dichomitus squalens	42.7	1697	146	73	542	16	0.64	7.70%	-	21.53x	24.99x	4.11x	50.63x
Fomitiporia mediterranea	63.35	3589	287	42.5	1412	6	4.3	10.40%	2.12x	13.31x	20.93x	3.01x	39.37x
Fomitopsis pinicola	46.3	1951	108	101.9	1117	25	0.51	9.20%	1.81x	19.57x	32.92x	4.05x	58.35x
Gloeophyllum trabeum	37.2	4148	174	58.2	443	7	1.9	7.39%	1.71x	17.82x	30.77x	4.18x	54.48x
Punctularia strigosozonata	34.17	761	70	130.4	195	6	2.12	3.20%	1.00x	23.56x	16.5x	4.48x	45.55x
Stereum hirsutum	46.51	669	54	245.3	159	9	1.8	1.90%	1.09x	18.07x	20.97x	4.64x	44.77x
Trametes versicolor	44.79	977	56	204.4	283	7	2.9	4.30%	1.36x	9.48x	24.95x	4.48x	40.27x
Tremella mesenterica	28.6	484	60	123.8	45	5	1.6	2.6	7.44x	-	-	-	7.44x
Wolfiporia cocos	50.5	1288	112	128.2	348	7	2.5	4.40%	1.11x	19.84x	15.81x	3.72x	40.48x

Table S2. EST assembly summary.

Species (condition)	Library	Library Type	Total Reads	rRNA%	Passing reads	Passing read length, bp
Auricularia delicata SS-5	CGZH	454 Titanium Fam-Y	1,204,720	1.29%	1,184,020	428
Coniophora puteana RWD-64-598 SS-2	CGZG	454 Titanium Fam-Y	1,046,130	1.13%	1,030,680	416
Dacryopinax sp. sp. DJM 731 SSP-1	CHAN	454 Titanium Fam-Y	1,508,140	53.16%	701,007	370
Dichomitus squalens LYAD-421 SS1	CGZN	454 Titanium Fam-Y	957,345	1.80%	935,727	406
Dichomitus squalens LYAD-421 SS1 (avicel)	CHBF	454 Titanium Fam-Y	1,553,618	1.70%	1,519,106	391
Fomitiporia mediterranea MF3 22	CGWI	454 Titanium Fam-Y	1,300,160	0.55%	1,287,880	418
Fomitopsis pinicola FP-58527 SS1	CGZS	454 Titanium Fam-Y	791,470	1.34%	776,365	340
G.trabeum						
Punctularia strigosozonata HHB-11173 SS5	CGZO	454 Titanium Fam-Y	1,091,060	2.00%	1,064,040	282
Punctularia strigosozonata HHB-11173 SS5 (avicel)	CGZP	454 Titanium Fam-Y	1,064,460	1.80%	1,041,050	398
Stereum hirsutum SS1	CGWN	454 Titanium Fam-Y	1,294,117	0.50%	1,282,258	395
Stereum hirsutum SS1 (avicel)	CHBG	454 Titanium Fam-Y	1,626,634	0.60%	1,608,805	362
Trametes versicolor FP-101664 SS1	CGZI	454 Titanium Fam-Y	1,379,640	1.40%	1,352,640	279
Trametes versicolor FP-101664 SS1 (avicel)	CGZT	454 Titanium Fam-Y	833,065	0.90%	821,682	326
Tremella mesenterica Fries	CCON	Sanger (insert 0.6-2kb)	9,984	0.91%	8,495	650
Tremella mesenterica Fries	CCOI	Sanger (insert > 2kb)	14,592	0.89%	12,091	645
Wolfiporia cocos MD-104 SS10	CGZU	454 Titanium Fam-Y	1,324,390	0.53%	1,311,260	387

Table S3. Genome annotation summary.

	Auricularia delicata	*Coniophora puteana*	*Dacryopinax sp.*	*Dichomitus squalens*	*Fomitiporia mediterranea*	*Fomitopsis pinicola*	*Gloeophyllum trabeum*	*Punctularia strigosozonata*	*Stereum hirsutum*	*Trametes versicolor*	*Tremella mesenterica*	*Wolfporia cocos*
Genome size (Mb)	74.92	42.97	29.50	42.75	63.35	46.30	37.18	34.17	46.51	44.79	28.60	50.48
Predicted genes	23577	13761	10242	12290	11333	14724	11846	11538	14072	14296	8313	12746
Models complete with start and stop codons	20619	12482	9338	10924	10469	12505	10513	10322	12973	12634	7515	11581
	87%	91%	91%	89%	92%	85%	89%	89%	92%	88%	90%	91%
Models with exons covered by ESTs ≥75%	9065	7549	6491	9696	8381	8084	6102	7332	8961	9441	1590	8985
	38%	55%	63%	79%	74%	55%	52%	64%	64%	66%	19%	70%
Models with hits to NR database	14114	10937	7703	10031	8725	11524	9792	9546	11254	11300	6408	10730
	60%	79%	75%	82%	77%	78%	83%	83%	80%	79%	77%	84%
Proteins assigned to a KOG	6770	7150	5866	6619	6343	7305	6544	6567	7737	7133	5054	6770
	29%	52%	57%	54%	56%	50%	55%	57%	55%	50%	61%	53%
KOG categories genome-wide	3210	3084	3073	3136	3143	3244	3250	3124	3184	3173	3010	3210
Proteins assigned a GO term	8463	6349	5082	5896	5517	6310	6310	5829	6943	6464	4123	5831
	36%	46%	50%	48%	49%	43%	53%	51%	49%	45%	50%	46%
GO terms genome-wide	2210	2173	2165	2189	2206	2222	2222	2206	2208	2205	2020	2212
Proteins assigned an EC number	3436	2967	2367	2659	2596	2738	2598	2674	3232	2850	1804	2686
	15%	22%	23%	22%	23%	19%	22%	23%	23%	20%	22%	21%
EC numbers genome-wide	750	731	740	735	749	755	765	734	755	743	648	750
Proteins assigned a Pfam domain	8893	6621	5415	6341	5950	6525	6140	6158	7109	6848	4283	6156
	38%	48%	53%	52%	53%	44%	52%	53%	51%	48%	52%	48%
Pfam domains genome wide	2371	2234	2237	2284	2285	2360	2381	2314	2339	2328	2127	2334

Table S4: Microsatellites.

Species		Microsatellites						
Basidiomycota	Life strategy	All	Mono	Di	Tri	Tetra	Penta	Hexa
Coniophora puteana	BR	1658	448	243	828	50	12	77
Gloeophyllum trabeum	BR	1167	423	203	456	35	20	30
Fomitopsis pinicola	BR	1395	219	196	913	34	13	20
Wolfiporia cocos	BR	1451	240	576	516	30	19	70
Dacryopinax sp.	BR	1604	557	205	660	22	28	132
Postia placenta	BR	1816	505	489	642	55	76	49
Serpula lacrymans	BR	2389	1330	308	667	26	12	46
Auricularia delicata	WR	4560	1166	1261	1736	178	161	58
Punctularia strigoso-zonata	WR	2510	967	427	945	104	10	57
Fomitiporia mediterranea	WR	4157	704	2173	892	62	170	156
Dichomitus squalens	WR	1172	282	286	527	21	10	46
Trametes versicolor	WR	1224	228	384	522	45	9	36
Stereum hirsutum	WR	2837	827	440	1209	206	73	82
Schizophyllum commune	WR	1181	336	202	589	23	14	17
Phanerochaete chrysosporium	WR	1393	386	379	598	14	3	13
Heterobasidion annosum	WR	3530	942	740	1237	343	104	164
Cryptococcus neoformans	ANIMAL PATHO	2172	1437	255	397	53	1	29
Malassezia globosa	ANIMAL PATHO	1324	551	440	273	20	15	25
Laccaria bicolor	ECM	5787	3292	718	1469	55	200	53
Tremella mesenterica	MYCOPARASITE	2574	1003	608	785	76	34	68
Ustilago maydis	PLANT PATHO	2462	653	600	861	76	76	196
Melampsora laricis-populina	PLANT PATHO	9972	5224	1612	2580	220	178	158
Coprinopsis cinerea	SAPRO	2050	761	375	843	38	10	23
Sporobolomyces roseus	SAPRO?	9203	1292	4156	3607	59	25	64
Ascomycota								
Stagonospora nodorum	PLANT PATHO	1354	441	302	492	47	23	49
Aspergillus niger	PLANT AND ANIMAL PATHO	4460	1828	1022	1248	289	50	23
Cryphonectria parasitica	PLANT PATHO	18167	9224	4223	4328	286	52	54
Trichoderma reesei	SAPRO	11436	4596	2426	3691	423	156	144
Pichia stipitis	SAPRO?	619	266	97	225	5	6	20
"Basal Fungi" (outgroups)								
30. Batrachochytrium dendrobatidis	ANIMAL PATHO	237	65	14	134	1	19	4
31. Phycomyces blakesleeanus	SAPRO	58892	20342	23850	8544	3330	1624	1202

Table S5: Transposable elements.

Species					No. of Full length LTR****		Gypsy		Copia		
Basidiomycota	Life strategy*	Genome size	Total No. TE**	Total TE coverage %***	Gypsy	Copia	No.	%	No.	%	No.
Coniophora puteana	BR	41.86	255	4.19	1	5	61	0.19	390	0.94	6
Gloeophyllum trabeum	BR	34.43	281	8.08	0	1	276	0.75	324	1	0
Fomitopsis pinicola	BR	42.06	462	9.38	3	0	444	0.95	695	1.42	36
Wolfiporia cocos	BR	48.24	1091	28.01	7	7	3019	4.63	809	0.45	155
Dacryopinax sp.	BR	27.60	101	4.81	3	0	661	1.52	0	0	0
Postia placenta	BR	68.98	753	16.94	14	4	6421	9.69	880	0.72	183
Serpula lacrymans	BR	42.40	494	29.77	21	9	4443	12.33	2796	5.1	15
Auricularia delicata	WR	69.05	897	10.56	28	7	1123	1.97	237	0.33	312
Punctularia strigoso-zonata	WR	33.07	175	5.19	5	1	309	1.46	142	0.46	0
Fomitiporia mediterranea	WR	56.77	316	41.42	74	2	5811	21.33	1660	5.2	0
Dichomitus squalens	WR	39.45	299	9.15	3	2	1163	2.66	743	1.16	58
Trametes versicolor	WR	42.88	233	4.18	5	2	172	0.52	91	0.13	24
Stereum hirsutum	WR	45.64	242	3.09	3	4	81	0.17	150	0.23	0
Schizophyllum commune	WR	37.93	178	5.09	20	7	276	2.04	350	1.1	0
Phanerochaete chrysosporium	WR	32.50	118	6.5	8	5	1005	3.75	357	0.87	0
Heterobasidion annosum	WR	33.64	305	16.03	111	19	1590	8.81	219	0.96	29
Cryptococcus neoformans	AP	18.87	53	5.65	4	2	334	2.2	58	1.22	0
Malassezia globosa	AP	8.95	0	0	0	0	0	0	0	0	0
Laccaria bicolor	ECM	58.68	1843	29.86	17	37	2536	2.85	2126	1.86	153
Tremella mesenterica	MP	27.98	238	19.17	50	6	2521	11.49	760	2.98	
Ustilago maydis	PP	19.68	31	1.74	0	0	0	0	234	0.89	0
Melampsora laricis-populina	PP	97.68	2560	31.97	74	46	7888	6.86	4023	2.82	265
Coprinopsis cinerea	SAP	36.29	182	6.07	39	8	610	3.61	297	0.75	0
Sporobolomyces roseus	SAP?	20.84	31	1.31	3	0	40	0.19	58	0.18	0
Ascomycota, Chytridiomycota, Mucoromycotina											
Cryphonectria parasitica	PP	43.80	104	11.66	37	7	1960	10.15	165	0.62	0
Stagonospora nodorum	PP	37.04	52	5.68	0	0	234	1.12	97	0.71	0
Aspergillus niger	AP/PP	34.85	20	0.25	0	0	0	0	0	0	22
Trichoderma reesei	SAP	33.40	31	0.74	0	2	0	0	44	0.29	0
Pichia stipitis	SAP?	15.44	34	4.56	0	11	0	0	290	2.52	193
Batrachochytrium dendrobatidis	AP	24.13	203	20.22	4	0	71	0.1	55	0.27	43
Phycomyces blakesleeanus	SAP	53.36	527	30.36	22	0	2128	2.47	0	0	745

Table S5, continued.

Species			Other Class I		TIR		DNA Transposon		Helitron		Not Categorized	
Basidiomycota	Life strategy*	*Genome size*	No.	%	No.	%	No.	%	No.	%	No.	%
Coniophora puteana	BR	*41.86*	12	0.05	44	0.08	0	0	0	0	4268	2.93
Gloeophyllum trabeum	BR	*34.43*	0	0	16	0.03	0	0	0	0	6073	6.3
Fomitopsis pinicola	BR	*42.06*	65	0.08	38	0.04	14	0.03	14	0.03	7776	6.74
Wolfiporia cocos	BR	*48.24*	303	0.23	160	0.24	222	0.48	76	0.21	28983	21.47
Dacryopinax spathularia	BR	*27.60*	34	0.04	17	0.04	37	0.17	121	0.43	1932	2.61
Postia placenta	BR	*68.98*	33	0.01	0	0	153	0.1	0	0	16309	6.3
Serpula lacrymans	BR	*42.40*	50	0.13	423	0.7	74	0.12	136	0.35	10514	11.01
Auricularia delicata	WR	*69.05*	352	0.33	194	0.15	179	0.12	0	0	18019	7.55
Punctularia strigoso-zonata	WR	*33.07*	107	0.3	0	0	0	0	0	0	2857	2.97
Fomitiporia mediterranea	WR	*56.77*	448	1.05	951	3.78	668	1.69	15	0.01	8219	8.36
Dichomitus squalens	WR	*39.45*	0	0	0	0	71	0.11	0	0	5737	5.13
Trametes versicolor	WR	*42.88*	25	0.03	10	0.02	22	0.02	59	0.1	4058	3.35
Stereum hirsutum	WR	*45.64*	27	0.01	0	0	0	0	0	0	4012	2.68
Schizophyllum commune	WR	*37.93*	45	0.03	0	0	0	0	0	0	2235	1.92
Phanerochaete chrysosporium	WR	*32.50*	65	0.16	39	0.15	67	0.1	0	0	1148	1.47
Heterobasidion annosum	WR	*33.64*	12	0.01	29	0.05	0	0	38	0.27	3279	5.86
Cryptococcus neoformans	AP	*18.87*	362	1.02	52	0.17	0	0	0	0	306	1.04
Malassezia globosa	AP	*8.95*	0	0	0	0	0	0	0	0	0	0
Laccaria bicolor	ECM	*58.68*	857	1.48	511	0.38	601	0.59	382	0.33	36798	22.03
Tremella mesenterica	MP	*27.98*	6	0.01	131	0.61	54	0.4	24	0.23	3843	3.45
Ustilago maydis	PP	*19.68*	47	0.03	13	0.03	0	0	0	0	426	0.79
Melampsora laricis-populina	PP	*97.68*	1229	0.87	5447	3.14	468	0.32	604	0.28	62914	17.27
Coprinopsis cinerea	SAP	*36.29*	10	0.01	12	0.02	0	0	0	0	1742	1.68
Sporobolomyces roseus	SAP?	*20.84*	13	0.01	0	0	0	0	0	0	357	0.93
Ascomycota, Chytridiomycota, Mucoromycotina												
Cryphonectria parasitica	PP	*43.80*	0	0	173	0.46	22	0.06	0	0	624	0.37
Stagonospora nodorum	PP	*37.04*	54	0.02	238	0.54	0	0	0	0	933	3.29
Aspergillus niger	AP/PP	*34.85*	81	0.03	34	0.07	0	0	0	0	338	0.14
Trichoderma reesei	SAP	*33.40*	55	0.05	0	0	0	0	0	0	603	0.4
Pichia stipitis	SAP?	*15.44*	0	0	0	0	0	0	0	0	226	0.82
Batrachochytrium dendrobatidis	AP	*24.13*	45	0.11	11	0.11	172	0.88	0	0	8789	18.35
Phycomyces blakesleeanus	SAP	*53.36*	150	0.18	2589	3.8	1686	7.39	224	0.6	15576	13.55

*BR = brown rot, WR = white rot, AP = animal pathogen, PP = plant pathogen, SAP = saprotroph, ECM = ectomcyorrhiza, MP = mycoparasite.
**Total number of TE coming from the de novo identification of repeat sequences with RepeatScout.
***Total TE coverage estimated with RepeatMasker using the seqences identified with RepeatScout and LTRSTRUC.

Table S6. Proteins used in phylogenetic analyses. Abbreviations follow the *Saccharomyces* Genome Database, with links to protein summaries.

ABD1[4]	CIA1[4]	GCD6[4]	MCM3[1,2]	RBG2[4]	SPC97[4]
ADO1[4]	COX15[4]	GLN4[4]	MCM7[1,2]	RET1[1,2]	SSL1[4]
AMD1[3]	CTK1[4]	GPI8[1,2]	MIP1[4]	RIO1[4]	SYF1[4]
APC1[4]	DBP3[4]	HAT1[4]	MOT1[1,2]	RPA135[1]	TAP42[4]
BUB2[4]	DIP2[4]	HOS2[3]	MPD1[4]	RPA49[4]	TEF1[3]
CCT2[1,2]	DPB2[4]	HSE1[3]	NBP35[4]	RPB1[3]	TFB2[1]
CCT4[1,2]	EFT2[2]	HSH155[1,2]	NEW1[4]	RPB2[3]	TIM44[4]
CDC5[3]	ELP2[4]	KAE1[1,2]	NMD2[1]	RPO31[1,2]	TPI1[3]
CDC60[3]	ERB1[4]	KRR1[4]	OST1[4]	RRP8[4]	UBP6[4]
CEX1[4]	FAL1[4]	LOS1[4]	PWP1[4]	SEC10[4]	ULA1[4]
CFT2[4]	FRS2[1,2]	LST8[4]	RAD3[1,2]	SEC15[4]	UTP15[4]
CHC1[4]	GCD10[4]	MAK11[4]	RAM1[4]	SGD1[4]	VMS1[4]
				SNU66[4]	ZWF1[3]

[1]Analyzed by Koonin et al. (*21*).
[2]Analyzed by Rokas at el. (*39*).
[3]Analyzed by James et al. (*158*) and Regier et al. (*41*)
[4]Obtained using *Hal* (*42*).

Table S7. Aligned gene regions retained for phylogenetic reconstruction following stringent Gblocks analysis (dataset 26G).

Gene	Presence in # species	Full length	Length post Gblocks	Conserved and flanking regions
ABD1	28	1510	127 (8%)	993-1013, 1056-1070, 1111-1123, 1169-1178, 1216-1249, 1379-1391, 1393-1413.
ADO1	31	388	304 (78%)	9-23, 28-46, 48-63, 65-118, 123-177, 180-240, 247-259, 265-279, 311-340, 342-367.
AMD1	31	1314	325 (24%)	593-613, 621-650, 659-734, 736-790, 872-888, 892-910, 923-968, 1034-1094.
APC1	31	3113	138 (4%)	1953-1968, 1993-2020, 2023-2062, 2080-2110, 2122-2144.
BUB2	29	584	168 (28%)	260-269, 380-396, 398-428, 447-524, 533-564.
CCT2	31	549	453 (82%)	65-123, 136-244, 246-350, 353-485, 500-546.
CCT4	31	608	459 (75%)	88-199, 201-211, 261-439, 450-538, 540-568, 570-608.
CDC5	30	2121	205 (9%)	189-202, 237-270, 272-317, 336-353, 362-379, 383-427, 443-455, 457-473.
CDC6	31	2258	284 (12%)	400-414, 428-444, 480-497, 510-540, 542-551, 558-570, 623-662, 719-749, 791-875, 1263-1286.
CDC60	31	1304	823 (63%)	93-150, 209-297, 300-342, 365-404, 411-425, 447-469, 484-520, 525-635, 643-652, 659-719, 732-752, 772-897, 939-982, 1002-1047, 1050-1081, 1096-1126, 1154-1166, 1196-1218.
CEX1	30	2689	286 (10%)	393-406, 432-445, 479-499, 551-567, 580-649, 655-706, 719-768, 771-818.
CFT2	30	1692	183 (10%)	477-499, 533-551, 611-623, 638-661, 921-931, 933-949, 1064-1109, 1140-1169.
CHC1	31	1802	1443 (80%)	9-18, 34-44, 55-81, 115-161, 176-208, 226-267, 284-441, 444-522, 538-593, 596-696, 700-894, 896-956, 1005-1284, 1294-1365, 1407-1442, 1444-1497, 1503-1517, 1519-1550, 1553-1686.
CIA1	30	1173	158 (13%)	95-111, 116-131, 184-222, 244-264, 283-299, 452-465, 519-528, 547-558, 626-637.
COX15	30	758	242 (31%)	252-276, 278-349, 358-380, 447-465, 511-552, 554-567, 594-611, 676-704.
CTK1	29	1848	221 (11%)	979-1011, 1018-1095, 1118-1165, 1168-1199, 1259-1288.
DBP3	30	785	324 (41%)	296-321, 323-334, 369-389, 399-417, 447-497, 508-538, 553-567, 570-582, 613-714, 718-730, 748-768.
DIP2	31	1319	295 (22%)	233-251, 308-318, 566-577, 600-625, 627-644, 669-688, 713-736, 810-861, 863-911, 929-992.
DPB2	31	4724	124 (2%)	1234-1249, 1348-1375, 1385-1399, 1534-1547, 1573-1582, 1609-1620, 1626-1636, 1714-1731.
EFT2	30	850	797 (93%)	13-83, 95-262, 274-376, 382-486, 491-811, 822-850.
ELP2	30	2280	69 (3%)	1075-1092, 1116-1128, 1336-1345, 1351-1378.
ERB1	31	1051	357 (33%)	342-357, 386-427, 430-473, 510-584, 848-863, 878-926, 933-976, 979-1049.
FAL1	31	448	353 (78%)	33-101, 127-410.
FRS2	31	669	350 (52%)	99-123, 158-176, 196-211, 220-231, 266-347, 397-409, 414-448, 494-548, 552-610, 622-655.
GCD10	28	975	17 (1%)	441-457.
GLN4	31	1401	430 (30%)	750-779, 782-834, 872-908, 936-991, 999-1030, 1035-1105, 1125-1149, 1151-1161, 1163-1181, 1184-1195, 1204-1213, 1227-1242, 1249-1264, 1266-1277, 1337-1349, 1371-1387.
GPI8	30	566	164 (28%)	144-307.
HAT1	29	759	152 (20%)	49-58, 60-75, 229-254, 303-318, 373-430, 646-657, 672-685.
HOS2	30	790	269 (34%)	117-138, 142-165, 199-305, 311-322, 330-433.
HSE1	29	1830	199 (10%)	363-384, 398-452, 483-497, 548-561, 741-768, 770-779, 783-799, 852-889.
HSH155	31	1392	903 (64%)	126-136, 153-172, 211-234, 432-444, 449-465, 508-533, 547-735, 740-822, 828-874, 876-898, 900-954, 957-1037, 1040-1049, 1056-1083, 1088-1214, 1216-1237, 1239-1283, 1287-1368.
KAE1	30	1848	188 (10%)	173-243, 246-271, 398-453, 455-489.
KRR1	31	661	227 (34%)	98-292, 364-395.
LOS1	31		495 (32%)	288-302, 307-317, 337-356, 371-384, 386-398, 443-465, 481-497, 499-518, 545-558, 594-632, 665-676, 725-739, 771-786, 897-907, 909-927,

				954-978, 1036-1049, 1064-1076, 1101-1123, 1158-1205, 1258-1275, 1295-1329, 1342-1360, 1421-1438, 1456-1478.
LST8	31	711	227 (31%)	46-56, 70-85, 157-189, 201-269, 308-329, 338-350, 463-525.
MAK11	27	1559	120 (7%)	184-195, 227-243, 257-271, 380-389, 391-400, 402-416, 456-483, 515-527.
MCM3	31	2215	515 (23%)	1256-1294, 1309-1353, 1367-1376, 1381-1406, 1425-1460, 1478-1493, 1498-1514, 1522-1685, 1687-1754, 1811-1825, 1827-1850, 1859-1913.
MCM7	31	975	549 (56%)	286-300, 309-377, 379-498, 500-720, 732-769, 785-829, 843-857, 928-953.
MIP1	30	3107	613 (20%)	645-654, 677-693, 722-750, 832-842, 844-864, 931-984, 1112-1135, 1160-1234, 1594-1624, 1711-1741, 1763-1789, 1868-2021, 2037-2073, 2094-2185.
MOT1	29	3835	255 (6%)	520-532, 713-749, 803-826, 828-842, 879-898, 972-982, 985-1012, 1020-1031, 1033-1046, 1055-1067, 1078-1112, 1120-1152.
MPD1	24	616	105 (17%)	73-119, 133-153, 189-213, 263-274.
NBP35	31	610	281 (45%)	42-66, 69-111, 152-174, 177-230, 276-301, 304-413.
NEW1	31	1697	707 (41%)	571-582, 585-608, 680-697, 701-710, 713-734, 737-771, 774-839, 996-1045, 1048-1058, 1060-1074, 1086-1100, 1104-1156, 1162-1174, 1188-1280, 1292-1365, 1373-1548, 1640-1659.
NMD2	30	2762	235 (8%)	917-942, 956-993, 996-1095, 1113-1140, 1189-1199, 1635-1666.
OST1	30	1527	171 (11%)	954-963, 1057-1109, 1131-1153, 1199-1218, 1262-1276, 1308-1327, 1330-1346, 1349-1361.
PWP1	31	2082	287 (13%)	758-787, 890-900, 921-930, 956-990, 1021-1041, 1084-1114, 1160-1191, 1268-1283, 1294-1306, 1313-1328, 1364-1389, 1416-1436, 1442-1466.
RAD3	31	1421	636 (44%)	604-621, 665-706, 725-772, 783-856, 859-873, 887-1008, 1010-1083, 1086-1210, 1213-1276, 1279-1316, 1323-1338.
RAM1	30	1002	168 (16%)	350-373, 376-389, 401-411, 470-498, 516-533, 581-597, 651-671, 780-796, 827-843.
RBG2	31	417	228 (54%)	37-47, 57-182, 190-200, 210-237, 239-252, 255-281, 285-295.
RET1	31	1616	1015 (62%)	335-349, 351-400, 420-452, 457-516, 521-530, 547-556, 558-616, 625-640, 666-689, 695-779, 783-857, 859-894, 897-912, 916-979, 982-1031, 1042-1181, 1203-1224, 1261-1285, 1287-1333, 1345-1380, 1408-1449, 1490-1554, 1574-1608.
RIO1	26	991	180 (18%)	305-333, 359-387, 392-435, 477-507, 562-597, 604-614.
RPA49	27	1575	108 (6%)	182-198, 241-265, 326-345, 554-563, 584-599, 605-614, 652-661.
RPA135	31	1911	325 (17%)	61-70, 83-97, 262-303, 326-380, 399-419, 449-461, 474-526, 569-582, 584-593, 595-615, 617-637, 712-724, 727-743, 761-780.
RPO31	31	2304	1082 (46%)	662-678, 700-759, 763-821, 874-885, 887-913, 917-940, 942-968, 973-1001, 1017-1100, 1138-1271, 1286-1326, 1403-1701, 1850-1933, 1950-1967, 1998-2012, 2068-2091, 2111-2163, 2166-2183, 2188-2244.
RRP8	28	828	85 (10%)	323-332, 345-355, 383-392, 467-481, 604-630, 635-646.
SEC10	30	2100	394 (18%)	651-676, 779-789, 894-953, 996-1021, 1023-1046, 1103-1126, 1188-1197, 1200-1215, 1402-1432, 1482-1491, 1512-1564, 1578-1596, 1674-1709, 1717-1734, 1740-1759, 1802-1811.
SEC15	29	921	435 (47%)	78-93, 97-208, 233-271, 403-424, 444-523, 526-548, 610-621, 648-673, 686-717, 737-754, 757-781, 783-812.
SGD1	30	1309	176 (13%)	654-682, 740-771, 795-823, 846-861, 914-935, 987-1005, 1021-1049.
SNU66	29	1613	237 (14%)	442-458, 546-563, 590-599, 654-669, 693-715, 948-962, 1036-1045, 1194-1213, 1217-1228, 1282-1295, 1297-1317, 1325-1365, 1369-1388.
SPC97	30	2108	367 (17%)	366-375, 390-399, 401-465, 471-499, 570-656, 718-742, 811-866, 872-887, 906-931, 1011-1053.
SSL1	31	956	114 (11%)	191-230, 261-272, 382-404, 444-454, 506-533.
SYF1	31	1569	498 (31%)	277-292, 328-358, 366-378, 484-496, 637-702, 733-774, 793-811, 819-831, 862-897, 943-964, 973-1024, 1026-1043, 1050-1060, 1067-1107, 1125-1155, 1159-1198, 1203-1236.
TAP42	26	645	117 (18%)	93-113, 208-220, 305-333, 431-459, 486-496, 502-515.
TFB2	31	695	166 (23%)	265-276, 294-322, 341-358, 390-406, 417-434, 469-527, 532-544.
TIM44	30	827	252 (30%)	193-235, 290-312, 335-352, 543-562, 570-628, 641-674, 686-704, 706-727, 740-753.
TPI1	31	279	224 (80%)	2-15, 17-35, 37-50, 63-112, 120-165, 180-260.
UBP6	30	694	243 (35%)	136-159, 194-230, 249-263, 321-336, 340-384, 401-446, 540-550, 563-588, 644-666.
ULA1	31	1800	223 (12%)	303-328, 352-382, 387-409, 476-485, 539-561, 578-590, 601-610, 677-

				687, 699-726, 788-800, 977-1011.
UTP15	30	712	288 (40%)	98-122, 125-153, 163-181, 185-202, 217-228, 235-253, 292-307, 331-340, 373-387, 437-452, 528-552, 596-638, 641-681.
VMS1	31	1129	143 (12%)	476-487, 542-559, 561-587, 619-628, 638-649, 659-674, 918-942, 954-976.
CIA1	30	1173	158 (13%)	95-111, 116-131, 184-222, 244-264, 283-299, 452-465, 519-528, 547-558, 626-637.
ZWF1	31	560	381 (68%)	76-90, 109-140, 145-172, 181-234, 246-304, 310-332, 342-352, 356-406, 408-451, 454-496, 515-535.

Table S8. Gene copy numbers at ancestral nodes estimated with Notung at 75% edge weight threshold. Includes all gene families found to have significantly non-random diversification according to CAFÉ analyses, as well as GH6 and CDH. For node numbers, see Fig. 1 (main text).

Node	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Oxidoreductases																														
ClassII	-	-	1	1	1	1	1	3	4	4	6	3	2	2	2	0	4	6	6	1	1	8	2	0	0	0	1	1	1	1
Laccases	-	-	3	2	2	2	2	2	5	5	5	5	4	4	4	4	6	4	4	2	2	5	4	0	0	0	2	3	7	8
DYEP	-	-	-	1	1	1	1	1	1	3	4	4	3	3	3	0	1	2	2	1	1	1	3	1	0	1	0	0	0	0
CRO	1	2	3	6	7	7	5	6	6	6	6	6	5	5	6	5	6	5	5	4	4	6	4	3	2	3	2	2	1	1
GLP	-	-	-	1	1	1	2	2	4	5	5	5	4	3	1	4	4	3	6	6	6	6	6	0	0	1	0	0	0	0
HTP	-	-	7	7	8	7	7	9	9	9	9	8	4	5	5	3	7	4	4	3	3	4	8	0	2	3	4	8	7	3
CDH	-	-	1	1	1	1	1	1	1	2	2	2	2	1	1	1	1	1	1	0	0	1	1	0	0	0	1	3	4	3
sum*	1	2	14	18	20	19	18	23	29	32	35	31	22	22	21	16	28	24	27	17	17	30	27	4	4	8	9	14	16	13
CAZymes																														
GH7	-	-	3	3	1	1	1	1	1	1	2	2	1	1	1	1	2	2	1	0	0	3	1	0	0	2	2	3	3	4
GH61	-	-	14	13	13	13	13	14	16	17	17	17	16	18	15	6	11	13	11	4	3	14	11	1	0	0	14	20	14	11
GH3	1	2	22	14	16	17	17	19	19	20	20	17	11	9	6	10	13	12	10	8	7	7	11	6	3	3	19	25	20	17
GH10	-	-	3	3	3	3	3	7	10	11	8	6	4	4	4	2	3	4	4	2	2	5	5	0	2	1	2	2	2	2
GH28	-	3	22	18	18	17	17	19	21	21	15	14	10	7	5	8	10	8	9	11	9	9	16	1	1	1	19	20	19	17
GH43	-	-	23	19	19	16	16	20	20	20	17	15	12	12	4	2	9	8	6	4	0	5	10	0	3	1	16	18	14	12
CE16	-	1	4	4	4	4	4	8	9	11	11	9	7	6	5	5	7	8	9	8	5	9	9	1	0	2	4	4	5	6
CE5	-	-	5	6	6	4	4	4	4	4	4	4	3	3	2	1	1	0	0	0	0	0	1	0	3	2	4	8	8	7
GH6			4	3	3	3	3	3	3	3	3	3	3	3	2	1	1	1	1	0	0	1	1	0	0	0	3	4	2	1
sum*	1	6	96	80	80	75	75	92	100	105	94	84	64	60	42	35	56	55	50	37	26	52	64	9	12	12	80	100	85	76

*excluding CDH and GH6.

Table S8, continued. Gene copy numbers at ancestral nodes estimated with Notung at 90% edge weight threshold

Node	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Oxidoreductases																														
ClassII	-	-	1	1	1	1	1	2	3	3	3	2	2	2	2	0	3	3	3	1	1	7	2	0	0	0	1	1	1	1
Laccases	-	-	2	2	2	2	2	2	4	4	4	4	4	4	4	4	8	4	4	3	3	5	4	0	0	0	2	3	6	7
DYEP	-	-	-	1	1	1	1	2	2	3	3	3	2	2	2	0	1	2	2	1	1	1	3	1	0	1	0	0	0	0
CRO	1	2	3	4	4	4	4	5	5	5	5	5	4	4	6	4	6	5	5	4	4	6	4	3	2	3	2	2	1	1
GLP	-	-	-	1	1	1	2	2	4	4	4	4	4	3	1	4	4	3	6	9	9	6	6	0	0	1	0	0	0	0
HTP	-	-	4	4	4	3	3	5	6	6	6	6	4	4	4	3	6	3	3	3	3	4	6	0	2	3	4	8	8	4
CDH	-	-	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0	0	1	2	4	3
Sum*	1	2	10	13	13	12	13	18	24	25	25	24	20	19	19	15	28	20	23	21	21	29	25	4	4	8	9	14	16	13
CAZymes																														
GH7	-	-	3	3	2	2	2	2	2	2	2	2	2	2	1	1	2	2	2	0	0	3	1	0	0	2	3	3	3	4
GH61	-	-	14	14	14	14	14	15	17	16	15	17	16	18	15	9	13	13	12	4	3	15	11	1	0	0	14	19	15	12
GH3	1	2	16	12	13	14	14	16	16	17	16	14	9	9	6	10	13	10	8	6	5	7	10	6	2	3	14	21	20	16
GH10	-	-	2	2	2	2	2	4	7	7	5	4	4	4	4	2	4	4	3	2	2	5	5	0	2	1	2	2	2	2
GH28	-	3	18	17	17	16	16	18	19	20	15	14	10	7	5	8	10	8	9	8	7	7	16	1	1	1	16	16	19	17
GH43	-	-	15	14	14	12	12	15	16	16	15	13	11	11	4	2	8	8	6	4	0	5	9	0	3	1	13	14	13	13
CE16	-	1	3	3	3	3	3	8	8	7	8	7	6	5	4	6	7	8	8	8	5	8	7	1	0	1	3	3	4	4
CE5	-	-	7	6	5	3	3	3	3	3	3	3	2	2	2	1	1	0	0	0	0	0	1	0	3	4	7	9	10	10
GH6	-	-	3	3	3	3	3	3	3	3	3	3	3	3	2	1	1	1	1	0	0	1	1	0	0	0	3	4	2	1
Sum*	1	6	78	71	70	66	66	81	88	88	79	74	60	58	41	39	58	53	48	32	22	50	60	9	11	13	72	87	86	78

*excluding CDH and GH6.

Table S8, continued. Gene copy numbers at ancestral nodes estimated with CAFE.

Node	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Oxidoreductases																														
ClassII	-	1	1	1	1	1	2	7	7	7	7	6	1	1	1	0	61	8	8	1	1	13	7	0	0	0	1	1	1	1
Laccases	-	1	1	1	1	1	2	5	6	6	6	6	6	6	9	6	1	5	5	5	4	7	6	0	0	0	1	3	4	4
DYEP	-	1	1	1	1	1	1	1	1	1	1	1	1	1	2	0	1	1	1	1	1	1	1	1	0	1	0	0	0	0
CRO	-	2	2	3	3	3	4	5	5	5	5	5	5	5	6	5	5	5	5	4	4	8	5	3	3	3	1	1	1	1
GLP	-	1	1	1	1	1	3	3	5	5	5	5	5	5	1	5	5	5	5	7	7	5	5	0	0	1	0	0	0	0
HTP	-	2	3	31	31	31	51	52	52	52	52	52	52	52	52	31	63	42	42	42	42	43	52	0	2	3	3	51	51	41
Sum*	-	8	9	0	0	0	7	6	9	9	9	8	8	8	4	9	4	8	8	2	1	8	9	4	5	8	5	0	1	0
CAZymes																														
GH7	-	1	1	1	1	1	1	2	21	21	21	21	21	21	21	21	21	21	21	0	0	31	21	0	0	1	1	2	2	2
GH61	-	1	1	1	1	1	3	9	1	1	1	1	1	4	4	0	21	1	1	4	3	5	1	1	0	0	2	71	71	71
GH3	-	5	6	5	5	6	8	9	9	9	9	9	9	9	7	9	1	9	9	9	8	9	9	5	3	4	8	2	3	3
GH10	-	2	2	2	2	2	3	4	4	4	4	4	4	4	4	3	41	4	4	3	3	5	4	0	2	2	2	2	2	21
GH28	-	3	3	3	3	3	6	8	9	9	9	9	8	6	6	8	0	9	9	9	9	9	9	1	1	2	3	7	9	0
GH43	-	2	2	2	2	2	4	5	5	5	5	5	5	5	4	5	5	4	4	4	0	4	5	0	2	2	2	6	6	6
CE16	-	1	1	1	1	2	4	6	6	6	6	6	6	6	5	6	6	6	6	6	6	7	6	1	0	1	1	2	2	3
CE5	-	1	1	1	1	1	1	11	11	11	11	11	11	21	21	1	11	01	01	01	01	01	11	0	1	1	2	6	6	71
CBM1	-	11	11	11	11	11	43	14	14	14	14	14	04	04	14	84	65	24	24	03	02	85	14	0	0	01	02	74	84	15
Sum*	-	6	7	6	6	8	0	4	7	7	7	7	6	8	4	4	1	5	5	5	9	2	7	8	9	3	1	4	7	0

*excluding CDH and GH6.

Table S9. Comparison of alternative models of gene diversification in white rot and non-white rot lineages. Models significantly rejected against the unconstrained model are marked by an asterisk. Of the rate parameters, q stands for rate of change in the nutritional mode (state 0 being non-white rot and state 1 being white rot), μ stands for gene loss rate and λ stands for gene birth rate.

Model	Rates in the model	Mean LogLh
Unconstrained model	q_{01}, q_{10}, μ_0, μ_1, λ_0, λ_1	-13.051
Five-parameter model	q_{01}, q_{10}, μ, λ_0, λ_1	-13.051
Four-parameter model	q, μ, λ_0, λ_1	-14.425
Five-parameter model	q, μ_0, μ_1, λ_0, λ_1	-14.425
Five-parameter model	q_{01}, q_{10}, μ_0, μ_1, λ	-17.901*
Four-parameter model	q, μ_0, μ_1, λ	-18.068*
Four-parameter model	q_{01}, q_{10}, μ, λ	-23.873*
Three-parameter model	q, μ, λ	-24.825*

Table S10. Total CBM1 copies in all fungal genomes, based on search with PFAM domain PF00734.

Species	No. CBM1s
COPCI	44
AURDE	42
PHCHR	28
PUNST	27
TRAVE	23
STEHI	22
DICSQ	18
HETAN2	17
TRIRE	14
CRYPA	12
FOMME	10
STANO	10
ASPNI	9
SERLA	8
SCHICO	4
CONPU	3
GLOTR	1
LACBI2	1
DACSP	1
PHYBL	1
POSPL	0
WOLCO	0
FOMPI	0
TREME	0
CRYNE	0
MELLP	0
SPORO	0
USTMA	0
MALGL	0
PICST	0
BATDE	0

Table S11. Distribution of CBM1s in GH6, GH7, GH10, GH5, GH61, GH74, CE1, and CE15 gene families. Rows labeled “ALL” indicate total number of genes in that gene family. Rows labeled “wCBM1” indicate number of gene copies containing CBM1.

Species	GH6	GH7	GH10	GH5 End[1]	GH5 Man[2]	CE15	CE1	CE16 group2	GH61_1	GH61_2	GH74	sum
Copci ALL	5	6	6	1	5	8	3	4	22	13	1	74
Copci wCBM1	1	0	2	1	1	3	3	0	8	0	0	19
Lacbi ALL	0	0	0	1	2	0	0	3	1	4	0	11
Lacbi wCBM1	0	0	0	1	0	0	0	0	0	0	0	1
Schico ALL	1	2	5	2	1	2	4	6	12	10	1	46
Schico wCBM1	0	0	1	0	0	0	2	0	0	0	0	3
Serla ALL	1	0	1	3	5	0	0	2	2	3	1	18
Serla wCBM1	1	0	0	2	1	0	0	0	0	0	0	4
Conpu ALL	2	2	3	5	3	0	0	4	2	8	0	29
Conpu wCBM1	1	0	1	0	0	0	0	0	0	1	0	3
Hetan ALL	1	1	2	3	3	1	1	4	6	4	1	27
Hetan wCBM1	1	0	1	3	1	0	1	1	2	1	0	11
Stehi ALL	1	3	6	3	2	1	1	7	8	8	2	42
Stehi wCBM1	1	1	4	2	1	1	1	1	2	1	1	16
Punstri ALL	1	5	5	3	3	2	2	5	9	5	2	42
Punstri wCBM1	1	5	3	3	1	1	1	1	2	1	1	20
Gloetr ALL	0	0	4	2	3	1	1	4	1	3	1	20
Gloetr wCBM1	0	0	1	0	0	0	0	0	0	0	0	1
Phchr ALL	1	8	6	2	3	2	4	2	8	7	4	47
Phchr wCBM1	1	6	4	2	1	1	2	1	3	3	1	25
Pospl ALL	0	0	2	3	2	1	0	3	0	2	0	13
Pospl wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Wolco ALL	0	0	2	2	2	1	0	4	0	2	0	13
Wolco wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Fompi ALL	0	0	4	3	2	1	0	9	2	2	0	23
Fompi wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Trave ALL	1	4	6	3	2	2	3	6	9	9	1	46
Trave wCBM1	1	0	3	3	1	1	3	1	2	1	1	17
Dicsq ALL	1	3	5	3	2	2	0	6	7	8	1	38
Dicsq wCBM1	1	0	3	3	1	1	0	1	2	1	1	14
Fomme ALL	2	2	4	3	3	1	0	4	6	7	4	36
Fomme wCBM1	1	0	1	1	1	0	0	1	2	1	1	9
Dacsp ALL	0	0	3	3	2	1	0	2	0	0	0	11
Dacsp	0	0	0	1	0	0	0	0	0	0	0	1

wCBM1												
Aurde ALL	2	6	2	4	3	6	3	18	10	9	1	64
Aurde wCBM1	2	5	2	3	1	3	1	1	4	1	1	24
Treme ALL	0	0	0	0	0	0	0	1	0	0	0	1
Treme wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Cryne ALL	0	0	0	0	0	0	0	0	0	1	0	1
Cryne wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Ustma ALL	0	0	2	0	0	0	1	0	0	0	0	3
Ustma wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Malgl ALL	0	0	0	0	0	0	0	0	0	0	0	0
Malgl wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Mellp ALL	0	9	6	4	0	0	0	0	0	0	0	19
Mellp wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Sporo ALL	0	0	0	0	0	0	0	2	0	0	0	2
Sporo wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Crypa ALL	2	5	4	5	2	2	2	5	4	8	2	41
Crypa wCBM1	1	0	2	1	1	0	0	2	0	1	1	9
Trire ALL	1	2	1	2	1	1	0	0	1	2	1	12
Trire wCBM1	1	2	0	1	0	1	0	0	0	1	1	7
Aspni ALL	2	2	2	3	1	0	1	1	3	4	1	20
Aspni wCBM1	1	1	0	0	0	0	0	0	2	1	1	6
Stano ALL	4	5	7	3	1	1	6	1	19	10	0	57
Stano wCBM1	1	0	2	0	0	0	0	0	3	1	0	7
Picst ALL	0	0	1	0	0	0	0	0	0	0	0	1
Picst wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Phybl ALL	0	0	0	0	0	0	0	0	0	0	0	0
Phybl wCBM1	0	0	0	0	0	0	0	0	0	0	0	0
Batde ALL	0	0	0	1	0	0	0	0	0	0	0	1
Batde wCBM1	0	0	0	0	0	0	0	0	0	0	0	0

[1]Endoglucanases
[2]Mannanases

Table S12. Oxidoreductase genes in twelve new Agaricomycotina genomes. AD, *Auricularia delicata*; PST, *Punctularia strigoso-zonata;* FM, *Fomitiporia mediterranea*; DS, *Dichomitus squalens*; TV, *Trametes versicolor*; SH, *Stereum hirsutum*; CP, *Coniophora puteana*; GT, *Gloeophyllum trabeum*; FP, *Fomitopsis pinicola*; WC, *Wolfiporia cocos*; D, *Dacryopinax* sp; and TM, *T. mesenterica*

	White Rot						Brown Rot					Myc.par
	AD	PST	FM	DS	TV	SH	CP	GT	FP	WC	D	T
Class II peroxidases												
short Mn-peroxidases (MnP-short)	0	2	3	5	13	0	0	0	0	0	0	0
long Mn-peroxidases (MnP-long)	0	5	11	0	0	0	0	0	0	0	0	0
extra long Mn-peroxidases (MnP-extra long)	0	3	0	4	0	0	0	0	0	0	0	0
Mn-peroxidases lacking 1 acidic residue (MnP-atypical)	5	0	2	0	0	5	0	0	0	0	0	0
all Mn-peroxidases (all MnP)	5	10	16	9	13	5	0	0	0	0	0	0
lignin peroxidases (LiP)	0	0	0	0	10	0	0	0	0	0	0	0
versatile peroxidases (VP)	0	0	0	3	2	0	0	0	0	0	0	0
versatile peroxidases lacking 1 acidic residue (VP-atypical)	0	0	0	0	1	0	0	0	0	0	0	0
generic peroxidases (GP)	11	1	1	0	0	1	0	0	1	1	0	0
Glucose-methanol-choline oxidoreductases												
cellobiose dehydrogenases (CDH)	1	1	1	1	1	1	2*	1	0	0	0	0
aryl-alcohol oxidases (AAO)	0	4	2	9	3	14	0	2	1	0	0	0
methanol oxidases (MOX)	3	3	2	4	4	7	2	1	4	4	1	0
pyranose-2 oxidases (POX)	2	0	0	0	2	0	0	0	0	0	0	0
pyranose dehydrogenases (PDH)	0	0	0	0	0	0	0	0	0	0	0	0
glucose oxidases (GOX)	0	0	0	0	0	0	0	0	0	0	0	0
Copper radical oxidases												
glyoxal oxidases (GLX)	2	3	0	5	5	3	0	0	0	0	0	0
copper radical oxidases 1 (CRO1)	0	0	1	1	1	0	1	0	1	1	2	2
copper radical oxidases 2 (CRO2)	4*	3	1	1	1	3	4	1	1	1	0	0
copper radical oxidases 3 (CRO3)	0	0	0	0	0	0	0	0	0	0	0	0
copper radical oxidases 4&5 (CRO4/5)	1	1	1	1	1	1	0	0	1	1	0	0
copper radical oxidases 6 (CRO6)	2	2	1	1	1	1	1	1	1	1	1	0
Multicopper oxidases												
Laccases 'sensu stricto' (Lac)	7	12	10	11	7	15	6	3	5	3	0	0
Fet3 ferroxidases (Fet3)	1	1	1	1	2	2	1	1	1	1	2*	1
ferroxidase/laccase	1	0		2	1	0	1	1	0		3	2
ascorbate oxidases	0	0		1	0	0	0	0	0		0	1
melanin synthesis (MCO)	1	0		0	0	0	0	0	0		0	0
cytochrome P450s												
'Authentic P450s'	226	127	115	157	178	189	216	108	154	174	111	8
'Tentative P450s'	23	17	15	30	12	26	22	22	36	32	15	1
Total P450s	249	144	130	187	190	215	238	130	190	206	126	9

Aurde* 1 potential pseudogene
Conpu* 1 could be a non functional gene
Dacsp1*similar to Fet3

Table S13. P450 monooxygenases (P450ome) subclassification (for species abbreviations see Table S12).

	White rot						Brown rot					Myc.par
	AD	PST	FM	DS	TV	SH	CP	GT	FP	WC	D	TM
P450 summary												
No. of Clans	11	11	11	11	11	12	11	11	12	11	12	5
No. of Families	39	36	29	34	34	38	30	35	33	35	31	7
No. of Sub-families	109	64	54	57	55	81	121	51	64	60	86	8
Clan-level classification												
CYP7	4	0	0	0	0	0	1	0	0	0	1	0
CYP51	3	4	1	4	3	3	4	2	3	4	1	1
CYP52	27	10	9	12	14	17	20	15	11	14	7	0
CYP53	14	8	12	6	11	6	5	8	7	12	3	1
CYP54	18	4	13	14	8	10	6	9	16	21	6	0
CYP61	1	2	2	1	1	1	1	1	1	1	1	1
CYP64	107	55	40	62	72	93	128	41	63	70	56	0
CYP67	8	4	9	12	6	7	1	3	3	2	1	0
CYP505	9	2	1	2	2	1	0	2	1	1	6	0
CYP534	33	31	21	33	54	41	39	23	42	42	24	3
CYP547	0	1	3	8	5	6	8	3	3	5	2	2
CYP609	0	0	0	0	0	1	0	0	1	0	0	0
UA	2	6	4	3	2	3	3	1	3	2	3	0

Table S14. Carbon utilization analyses.

		no C source	glc	frc	gal	man	xyl	ara	galA	starch	inulin	sucrose	cellulose	citrus pectin	beechwood xylan	Guar gum	lignin	min medium*	days
A. delicata	WR	±	++	++	±	+	±	±	±	+	±	±	±	+	±	+++	++	Rs	8
D. squalens	WR	±	++	++	±	++	±	±	±	±	±	±	±	±	±	++	±	Rs	8
F. mediterranea	WR	±	++	++	±	++	+	±	±	++	+	±	±	++	++	+++	+	Rs	7
H. annosum	WR	±	±	±	±	±	±	±	±	+++	±	±	-	+	±	+++	++	Rs	7
P. chrysosporium	WR	-	++	+	-	++	±	-	-	++	-	-	-	+++	+	-	±	An	7
P. strigoso-zonata	WR	±	+	+	±	+	+	±	+	+	+	+	±	++	+	++	+	Rs	4
S. commune	WR	-	++	++	+	++	+	-	±	+++	+	++	+	+	±	±	-	Sc	5
S. hirsutum	WR	-	+++	+++	-	+++	+	±	++	+++	+	+++	+	++	+	+	±	Rs	6
C. puteana	BR	-	++	+	+	++	+	+	++	++	±	±	+	+++	++	+	-	Rs	8
F. pinicola	BR	-	++	+++	+	+	++	+	+	+++	±	±	+	++	+++	+++	+	Rs	7
G. trabeum	BR	±	++	+	+	++	+	+	++	+	+	±	±	++	+++	++	++	Rs	5
P. placenta	BR	±	++	++	±	+++	++	++	+++	++	+	±	-	+++	+++	+++	-	Rs	8
S. lacrymans	BR	±	++	++	+	++	++	+	+	++	±	±	±	++	±	++	-	Rs	31
W. cocos	BR	+	++	+	+	++	+	+	+	+	+	±	-	+	±	+++	-	Rs	3
L. bicolor	ECM	+	++	+++	±	±	-	±	-	++	±	++	-	+	-	++	-	Rs	17
C. cinerea	litter	±	++	++	±	+	+	±	-	++	+	+	+	+	++	++	++	Cc	6
A. niger	mold	-	++	++	±	++	++	+	+	++	++	+++	-	+++	++	+++	±	An	7
T. reesei	mold	±	++	++	+	++	++	+	+	++	++	++	+	++	++	++	±	T	3
C. neoformans	path.	-	++	++	++	++	++	+	+	±	+	++	-	+++	+	+	±	Cn	5
S. nodorum	path.	±	++	++	+	++	+	±	±	++	+	++	+	+	++	+	+	Sn	5
U. maydis	path.	+	++	++	-	+++	+	±	±	+	+	++	±	++	+	++	+++	Um	3

*Rs = *Rhizoctonis solani* medium; An = *Aspergillus niger* medium; Sc = *Schizophyllum commune* medium; Cc = *Coprinopsis cinerea* medium; T = *Trichoderma* medium; Sn = *Stagonospora nodorum* medium; Cn = *Cryptococcus neoformans* medium; Um = *Ustilago maydis* medium.

Table S15. Traces: Estimated sample sizes (ESS) from six combined Bayesian relaxed uncorrelated lognormal clock analyses.

Statistic	Mean	ESS
posterior	-3.01E5	10512.627
prior	-240.898	34191.724
likelihood	-3.008E5	9028.238
treeModel.rootHeight	813.204	1197.538
tmrca (node 2)	769.863	1267.956
tmrca (node 3)	661.593	1487.454
tmrca (node 4)	521.343	674.819
tmrca (node 5)	491.411	626.82
tmrca (node 6)	428.968	972.606
tmrca (node 7)	349.335	1113.738
tmrca (node 8)	290.472	1221.578
tmrca (node 9)	237.024	1507.905
tmrca (node 10)	195.633	1790.09
tmrca (node 11)	183.855	1745.801
tmrca (node 12)	168.156	1830.736
tmrca (node 13)	148.939	1687.009
tmrca (node 14)	122.735	1416.225
tmrca (node 16)	104.226	633.085
tmrca (node 17)	100.943	384.971
tmrca (node 18)	140.849	886.673
tmrca (node 23)	170.417	1126.39
tmrca (node 24)	152.802	191.386
tmrca (node 25)	272.879	188.532
tmrca (node 26)	382.036	273.447
tmrca (node 27)	517.548	1352.199
birthDeath.meanGrowthRate	2.987E-3	39286.625
birthDeath.relativeDeathRate	0.199	27464.525
alpha	0.8	1026.913
plnv	0.157	1335.596
ucld.mean	6.558E-4	886.557
ucld.stdev	0.307	298.924
meanRate	5.875E-4	2151.354
coefficientOfVariation	0.277	381.024
covariance	-2.099E-3	886.75
treeLikelihood	-3.008E5	9028.238
speciation	-202.996	39625.201

Table S16. Age estimates inferred from Bayesian relaxed molecular clock analyses. The presence of fossil calibrations is indicated using acromyms: AGA = Agaricales (node 14), BOL = Boletales (node 16), and ASC = Ascomycota (node 27). Node numbering corresponds to Fig. 1.

Meth.	BEAST v1.6.1								PhyloBayes 3.2	
Calib.	AGA BOL ASC		AGA BOL		AGA ASC		BOL ASC		AGA BOL ASC	
Ages (mya)	mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI	mean	95% CI
Node										
1	811.93	621.63-1042.93	739.52	523.52-977.92	815.83	618.05-1065.11	771.14	583.82-1011.03	1610.02	1079.57-2453.8
2	769.86	600.4-977.82	701.22	502.98-920.13	774.72	600.5-999.85	733.24	566.17-958.17	1141.83	892.39-1466.6
3	661.59	519.99-831.05	599.1	432.38-786.67	670.14	513.27-855.63	631.41	490.79-802.03	861.2	700.65-1047.41
4	521.34	402.79-664.85	480.36	361.37-636.94	536.07	410.29-709.85	490.92	360.37-639.67	565.23	461.26-675.77
5	491.41	377.3-626.12	451.88	336.37-594.23	506.94	385.87-673.16	463.65	339.68-605.19	545.8	444.85-652.71
6	428.97	328.59-549.76	394.2	292.33-514.33	440.89	328.05-578.88	405.63	296.28-530.71	468.13	380.81-556.76
7	349.72	266.79-448.73	320.11	237.4-420.92	360.47	265.49-466.62	330.7	237.87-435.41	367.96	298.76-434.44
8	290.47	221.75-371.5	266.72	196.74-347.21	299.32	218.94-388.43	274.12	198.74-359.69	288.05	233-338.2
9	237.02	182.45-301.14	219.32	165.85-284.65	243.56	181.96-312.32	223.27	161.82-293.05	235.61	190.19-274.99
10	195.63	153.2-247.57	182.01	139.62-233.17	201.13	155.31-258.5	184.34	138.32-242.69	191.16	153.84-220.9
11	183.9	143.8-232.05	171.25	132.9-220	189.17	146.63-243.61	173.29	127.67-226.13	183.18	147-5-211.22
12	168.16	131.84-211.58	157.08	122.64-201.06	172.6	131.95-219.19	158.4	117.53-208.38	169.59	136.65-194.82
13	148.94	117.17-187-56	139.6	109.89-176.71	152.95	117.84-194.63	140.42	102.64-185.11	149.72	120.64-170.26
14	122.74	97.46-154.27	115.48	94.6-144.78	125.12	98.12-158.97	114.1	80.37-154.32	123.13	98.98-138.99
15	86.94	62.65-115.18	82.39	59.38-108.4	86.89	61.6-117.1	81.09	51.78-114.07	93.36	74.39-107.97
16	104.23	73.93-138.7	97.62	70.03-129.44	110.18	78.12-147.24	100.35	67.64-136.98	112.47	90.0-128.33
17	100.94	65.06-138.8	92.78	58.4-129.44	103.53	67.87-144.93	91.32	57.37-128.09	106.2	83.62-126.42
18	140.85	104.49-183.4	129.75	92.98-170.87	145.34	109.14-189.2	133.47	97.29-174.67	143.37	114.97-166.69
19	116.14	83.08-152.28	106.96	75.55-143.8	120.9	87.5-157.34	111.55	80.92-149.46	121.9	97.35-143.63
20	80.6	54.44-109.99	73.98	50.06-101.93	84.64	56.85-115.84	78.24	53.6-106	89.86	70.74-108.17
21	65.45	41.63-92.04	60.16	38.76-85.5	68.92	43.39-97.53	64.22	41.14-90.39	77.26	60.27-93.94
22	59.95	35.38-85.65	56.79	34.85-82.41	60.79	36.02-87.74	56.59	34.67-79.17	57.96	44.6-71.82
23	170.42	129.86-220.18	159.47	120.23-208.84	175.67	132.44-227.27	160.78	115.98-212.39	174.16	139.64-202.48
24	152.8	91.2-222.98	144.1	85.98-209.83	161.0	89.65-238.43	144.69	86.32-209.82	115.62	86.11-146.26
25	272.88	165.9-	252.9	152.55-	296.62	191.79-	265.7	167.08-	294.01	224.29-

		378.42		363.32		408.73		372.24		369.6
26	382.04	256.9-513.9	353.44	242.2-485.19	391.93	267.72-538.52	356.14	234.75-489.91	421.05	331.4-516.01
27	517.55	417.14-644.9	451.64	319.03-606.98	517.14	416.32-649.89	496.39	406.27-618.67	524.08	454.7-595.19/
28	343.6	248.44-454.62	299.2	210.94-404.84	340.16	246.4-446.28	306.07	203.22-421.26	309.96	238.83-385.06
29	290.6	203.02-387.07	249.92	170.5-341.83	287.04	191.6-387.98	255.5	150.68-349.42	288.09	221.66-358.93
30	147.27	89.82-211.62	127.61	73.17-183.55	142.56	87.19-206.83	129.89	76.99-189.45	119.56	86.66-158.45

III. List of Supplementary Electronic Files

The following electronic files have been deposited in the DRYAD data repository (http://datadryad.org/): http://dx.doi.org/10.5061/dryad.5k3t47p0.

Aligned datasets for phylogenetic analyses:

- Multigene alignments for phylogenetic and molecular clock analyses:
- 71G.nex = 71-gene alignment, 34257 aa
- 71G.tre =Treefile with optimal topology derived from 71G dataset
- 71GA.nex = 71-gene alignment with positions in two fastest rate categories removed (of eight rate categories), 25483 aa
- 71GA.tre = Treefile with optimal topology derived from 71GA dataset
- 26G.nex = 26-gene alignment, 13371 aa
- 26G.tre = Treefile with optimal topology derived from 26G dataset
- 26GA.nex = 26-gene alignment with positions in two fastest rate categories removed, 10002 aa
- 26GA.tre = Treefile with optimal topology derived from 26GA dataset
- 26G_BEAST.xml = xml file for BEAST analyses

Single-gene alignments that were concatenated for phylogenetic analyses. For each of the following, two FASTA files are available containing the aligned sequences before and after G-blocks analysis.

- ABD1, CIA1, GCD6, MCM3, RBG2, SPC97, ADO1, COX15, GLN4, MCM7, RET1, SSL1, AMD1, CTK1, GPI8, MIP1, RIO1, SYF1, APC1, DBP3, HAT1, MOT1, RPA135, TAP42, BUB2, DIP2, HOS2, MPD1, RPA49, TEF1, CCT2, DPB2, HSE1, NBP35, RPB1, TFB2, CCT4, EFT2, HSH155, NEW1, RPB2, TIM44, CDC5, ELP2, KAE1, NMD2, RPO31, TPI1, CDC60, ERB1, KRR1, OST1, RRP8, UBP6, CEX1, FAL1, LOS1, PWP1, SEC10, ULA1, CFT2, FRS2, LST8, RAD3, SEC15, UTP15, CHC1, GCD10, MAK11, RAM1, SGD1, VMS1, SNU66, ZWF1.

Protein IDs for all decay-related CAZymes and oxidoreductases analyzed, grouped by gene family:

- CDH_cytb652_QRD_OXO_GLP.xlsx
- CE1_5_8_12_15_16.xlsx
- ClassII_DyP_CRO_HTP_MCO.xlsx
- GH5_12_3_28_43.xlsx
- GH6_7_61.xlsx
- GH74_10_11.xlsx

Decay-related protein sequences from all taxa used for phylogenetic analyses. For each gene family, three files are available, including: unaligned sequences, original aligned sequences (all positions retained), and edited alignments (with some positions excluded).

- QRD, CDH, CE1, CE5, CE8, CE12, CE15, CE16, CE16_group1, CE16_group2, cytb652, DyP, GH_6, GH_7, GH_10, GH3, GH3_b_glucosid_group3, GH3_b_N_acetyl_glucosaminid, GH3_b_xylosidases, GH3_group1and2, GH5, GH5_group1_endogluc, GH5_group2_mannan, GH11, GH12, GH28, GH28_group1, Gh28_group2, GH28_group3, GH43, GH43_group1, GH43_group2, GH43_group3, GH61, GH61_group1, GH61_group2, GH74, GLP, HTP, HTP_1, HTP_2, HTP_3, MCO, Fet3_MCOother, L_ascor_oxid, Lacc_Bas_Asco, MCOs related to melanin synthesis, OXO/DEC, POD.

Gene clusters

- Gene clustering.xlsx

Secretome and transcriptome results

- All_Pro_annotated_final.xls

IV. Additional acknowledgements.

The following sources of support are gratefully acknowledged:
National Science Foundation Postdoctoral Fellowship: DBI-0805625
NGI Kluyver Centre for Genomics of Industrial Fermentation (KCII 2.2.27)
Dutch Technology Foundation STW, Applied Science division of NWO
Technology Program of the Ministry of Economic Affairs UGC (11108)
For technical assistance, the authors thank Sakinie Misiedjan and Christine Brussel
PEROXICATS European Union project (KBBE2010-265397, www.peroxicats.org; to ATM)
Spanish projects RAPERO (BIO2008-01533, to ATM) and HIPOP (BIO2011-26694, to FJR-D)
FJR-D is supported by a MINECO Ramón y Cajal contract

V. References

1. M. Binder, K. H. Larsson, P. B. Matheny, D. S. Hibbett, Amylocorticiales ord. nov. and Jaapiales ord. nov.: Early diverging clades of Agaricomycetidae dominated by corticioid forms. *Mycologia* **102**, 865 (2010).
2. A. Velegraki, M. Kambouris, A. Kostourou, G. Chalevelakis, N. J. Legakis, Rapid extraction of fungal DNA from clinical samples for PCR amplification. *Med. Mycol.* **37**, 69 (1999).
3. B. P. Roy, F. Archibald, Effects of kraft pulp and lignin on *Trametes versicolor* carbon metabolism. *Appl. Environ. Microbiol.* **59**, 1855 (1993).
4. S. Trong *et al.*, in *4th Annual Meeting on Sequencing Finishing, Analysis in the Future*. (Santa Fe, NM, 2009), pp. 35.
5. D. B. Jaffe *et al.*, Whole-genome sequence assembly for mammalian genomes: Arachne 2. *Genome Res.* **13**, 91 (2003).
6. T. W. Jeffries *et al.*, Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. *Nat. Biotechnol.* **25**, 319 (2007).
7. N. Semova *et al.*, Generation, annotation, and analysis of an extensive Aspergillus niger EST collection. *BMC Microbiol.* **6**, 7 (2006).
8. M. Qi *et al.*, Snapshot of the eukaryotic gene expression in muskoxen rumen: a metatranscriptomic approach. *PLoS One* **6**, e20521 (2011).
9. A. L. Price, N. C. Jones, P. A. Pevzner, De novo identification of repeat families in large genomes. *Bioinformatics* **21 Suppl 1**, i351 (Jun, 2005).
10. H. R. Smit AFA, Green P, RepeatMasker Open-3.0. (1996-2010).
11. J. Jurka *et al.*, Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* **110**, 462 (2005).
12. T. M. Lowe, S. R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955 (1997).
13. K. Isono, J. D. McIninch, M. Borodovsky, Characteristic features of the nucleotide sequences of yeast mitochondrial ribosomal protein genes as analyzed by computer program GeneMark. *DNA Res.* **1**, 263 (1994).
14. E. Birney, R. Durbin, Using GeneWise in the *Drosophila* annotation experiment. *Genome Res.* **10**, 547 (2000).
15. W. J. Kent, BLAT--the BLAST-like alignment tool. *Genome Res* **12**, 656 (2002).
16. H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne, Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**, 1 (1997).
17. K. Melen, A. Krogh, G. von Heijne, Reliability measures for membrane protein topology prediction algorithms. *J. Mol. Biol.* **327**, 735 (2003).
18. E. M. Zdobnov, R. Apweiler, InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17**, 847 (2001).
19. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403 (1990).

20. M. Kanehisa *et al.*, KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **36**, D480 (2008).
21. E. V. Koonin *et al.*, A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* **5**, R7 (2004).
22. A. J. Enright, S. Van Dongen, C. A. Ouzounis, An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575 (2002).
23. A. Vanden Wymelenberg *et al.*, Significant alteration of gene expression in wood decay fungi *Postia placenta* and *Phanerochaete chrysosporium* by plant species. *Appl. Environ. Microbiol.* **77**, 4499 (2011).
24. T. L. Highley, Influence of carbon source on cellulase activity of white rot and brown rot fungi. *Wood Fiber* **5**, 50 (1973).
25. A. Vanden Wymelenberg *et al.*, Computational analysis of the *Phanerochaete chrysosporium* v2.0 genome database and mass spectrometry identification of peptides in ligninolytic cultures reveals complex mixtures of secreted proteins. *Fungal Genet. Biol.* **43**, 343 (2006).
26. A. Vanden Wymelenberg *et al.*, Significant alteration of gene expression in wood decay fungi Postia placenta and Phanerochaete chrysosporium by plant species. *Appl. Environ. Microbiol.* **77**, 4499 (2011).
27. A. Vanden Wymelenberg *et al.*, Transcriptome and secretome analysis of *Phanerochaete chrysosporium* reveal complex patterns of gene expression. *Appl. Environ. Microbiol.* **75**, 4058 (2009).
28. A. Vanden Wymelenberg *et al.*, Comparative transcriptome and secretome analysis of wood decay fungi *Postia placenta* and *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* **76**, 3599 (2010).
29. S. E. Martin, J. Shabanowitz, D. F. Hunt, J. A. Marto, Subfemtomole MS and MS/MS peptide sequence analysis using nano-HPLC-ESI fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **72**, 4266 (2000).
30. A. I. Nesvizhskii, A. Keller, E. Kolker, R. Aebersold, A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646 (2003).
31. G. Daniel *et al.*, Characteristics of *Gloeophyllum trabeum* alcohol oxidase, an extracellular source of H_2O_2 in brown rot decay of wood. *Appl. Environ. Microbiol.* **73**, 6241 (2007).
32. C. Murat *et al.*, Mining microsatellites in the Perigord black truffle genome and identification of new molecular markers. *Fung. Genet. Biol.* **48**, 592 (2011).
33. A. L. Price, N. C. Jones, P. A. Pevzner, De novo identification of repeat families in large genomes. *Bioinformatics* **21**, i351 (2005).
34. E. M. McCarthy, J. F. McDonald, LTR_STRUC: a novel search and identification program for LTR retrotransposons. *Bioinformatics* **19**, 362 (2003).
35. B. J. Haas *et al.*, Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. *Nature* **461**, 393 (2009).
36. F. Martin *et al.*, Perigord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* **464**, 1033 (2010).
37. F. Martin *et al.*, The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* **452**, 88 (2008).
38. P. D. Spanu *et al.*, Genome expansion and gene loss in powdery mildew fungi reveal tradeoffs in extreme parasitism. *Science* **330**, 1543 (2010).
39. A. Rokas, D. Kruger, S. B. Carroll, Animal evolution and the molecular signature of radiations compressed in time. *Science* **310**, 1933 (2005).
40. T. Y. James *et al.*, Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443**, 818 (2006).
41. J. C. Regier *et al.*, Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**, 1079 (2010).
42. B. Robbertse, R. J. Yoder, A. Boyd, J. Reeves, J. W. Spatafora, Hal: an automated pipeline for phylogenetic analyses of genomic data. *PLoS Currents Tree of Life* **7**, 3 (2011).
43. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792 (2004).

44. J. Castresana, Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540 (2000).
45. A. Stamatakis, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688 (2006).
46. S. Kumar *et al.*, AIR: A batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. *BMC Bioinformatics* **10**, 357 (2009).
47. N. Lartillot, T. Lepage, S. Blanquard, PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* **25**, 2286 (2009).
48. T. De Bie, N. Cristianini, J. P. Demuth, M. W. Hahn, CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269 (2006).
49. A. Stamatakis, P. Hoover, J. Rougemont, A Rapid Bootstrap Algorithm for the RAxML Web Servers. *Syst. Biol.* **57**, 758 (2008).
50. B. Vernot, M. Stolzer, A. Goldman, D. Durand, Reconciliation with Non-Binary Species Trees. *J. Comp. Biol.* **15**, 981 (2008).
51. D. Durand, H. B. V., B. Vernot, A hybrid micro-macroevolutionary approach to gene tree reconstruction. *J. Comp. Biol.* **13**, 320 (2006).
52. P. Górecki, G. J. Burleigh, O. Eulenstein, Maximum likelihood models and algorithms for gene tree evolution with duplications and losses. *BMC Bioinformatics* **12**, S15 (2011).
53. P. Górecki, J. Tiuryn, DLS-trees: A model of evolutionary scenarios. *Theor. Comp. Sci.* **359**, 378 (2006).
54. D. Barker, A. Meade, M. Pagel, Constrained models of evolution lead to improved prediction of functional linkage from correlated gain and loss of genes. *Bioinformatics* **23**, 14 (2007).
55. W. P. Maddison, D. R. Maddison, Mesquite: a modular system for evolutionary analysis, version 2.75 (2011).
56. R. Fitzjohn, W. P. Maddison, S. P. Otto, Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Syst. Biol.* **58**, 595 (2009).
57. C. E. Chow, E. Yague, S. Raguz, D. A. Wood, C. F. Thurston, The *cel3* gene of *Agaricus bisporus* codes for a modular cellulase and is transcriptionally regulated by the carbon source. *Appl. Environ. Microbiol.* **60**, 2779 (1994).
58. Y. Ohnishi, M. Nagase, T. Ichiyanagi, Y. Kitamoto, T. Aimi, Transcriptional regulation of two cellobiohydrolase encoding genes (cel1 and cel2) from the wood-degrading basidiomycete *Polyporus arcularius*. *Appl. Microbiol. Biotechnol.* **76**, 1069 (2007).
59. A. Grassick *et al.*, Three-dimensional structure of a thermostable native cellobiohydrolase, CBH IB, and molecular characterization of the cel7 gene from the filamentous fungus, *Talaromyces emersonii*. *Eur. J. Biochem.* **271**, 4495 (2004).
60. Y. Liu *et al.*, Crystal structure of a glycoside hydrolase family 6 enzyme, CcCel6C, a cellulase constitutively produced by *Coprinopsis cinerea*. *FEBS Journal* **277**, 1532 (2010).
61. T. Kajisa, K. Igarashi, M. Samejima, The genes encoding glycoside hydrolase family 6 and 7 cellulases from the brown-rot fungus *Coniophora puteana*. *J. Wood Sci.* **55**, 376 (2009).
62. J. Karlsson *et al.*, Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei. Eur. J. Biochem.* **268**, 6498 (2001).
63. T. Koseki *et al.*, Biochemical characterization of a glycoside hydrolase family 61 endoglucanase from *Aspergillus kawachii*. *Appl. Microbiol. Biotechnol.* **77**, 1279 (2007).
64. S. Karkehabadi *et al.*, The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. *J. Mol. Biol.* **383**, 144 (2008).
65. P. V. Harris *et al.*, Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry* **49**, 3305 (2010).
66. G. Vaaje-Kolstad *et al.*, An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330**, 219 (2010).
67. J. MacDonald *et al.*, Transcriptomic responses of the softwood-dDegrading white-rot fungus *Phanerochaete carnosa* during growth on coniferous and deciduous wood. *Appl. Environ. Microbiol.* **77**, 3211 (2011).

68. D. S. Hibbett, M. J. Donoghue, Analysis of character correlations among wood decay mechanisms, mating systems, and substrate ranges in homobasidiomycetes. *Syst. Biol.* **50**, 215 (2001).
69. D. S. Hibbett, L. B. Gilbert, M. J. Donoghue, Evolutionary instability of ectomycorrhizal symbioses in basidiomycetes. *Nature* **407**, 506 (2000).
70. A. A. Hasper, E. Dekkers, M. van Mil, P. J. I. van de Vondervoort, L. H. de Graaff, EglC, a new endoglucanase from *Aspergillus niger* with major activity towards xyloglucan. *Appl. Environ. Microbiol.* **68**, 1556 (2002).
71. T. Ishida, K. Yaoi, A. Hiyoshi, K. Igarashi, M. Samejima, Substrate recognition by glycoside hydrolase family 74 xyloglucanase from the basidiomycete *Phanerochaete chrysosporium*. *FEBS Journal* **274**, 5727 (2007).
72. K. Yaoi, Purification, characterization, cDNA cloning, and expression of a xyloglucan endoglucanase from *Geotrichum* sp. M128. *FEBS Letters* **560**, 45 (2004).
73. V. F. Crepin, C. B. Faulds, I. F. Connerton, A non-modular type B feruloyl esterase from Neurospora crassa exhibits concentration-dependent substrate inhibition. *Biochem. J.* **370**, 417 (2003).
74. P. A. Kroon, G. Williamson, N. M. Fish, D. B. Archer, N. J. Belshaw, A modular esterase from *Penicillium funiculosum* which releases ferulic acid from plant cell walls and binds crystalline cellulose contains a carbohydrate binding module *Eur. J. Biochem.* **267**, 6740 (2000).
75. M. Skjùt *et al.*, Functional cloning of an endo-arabinanase from *Aspergillus aculeatus* and its heterologous expression in *A. oryzae* and tobacco. *Mol. Genet. Genomics* **265**, 913 (2001).
76. H. Ichinose, An Exo- -1,3-galactanase Having a Novel -1,3-Galactan-binding Module from Phanerochaete chrysosporium. *J. Biol. Chem.* **280**, 25820 (2005).
77. T. Kotake *et al.*, Molecular cloning and expression in *Pichia pastoris* of a *Irpex lacteus* Exo-β-(1→3)-galactanase gene. *Biosci. Biotechnol. Biochem.* **73**, 2303 (2009).
78. G. Henriksson *et al.*, Endoglucanase 28 (Cel12A), a new *Phanerochaete chrysosporium* cellulase. *Eur. J. Biochem.* **259**, 88 (1999).
79. F. Goedegebuur *et al.*, Cloning and relational analysis of 15 novel fungal endoglucanases from family 12 glycosyl hydrolase. *Curr. Genet.* **41**, 89 (2002).
80. R. Cohen, M. R. Suzuki, K. E. Hammel, Processive endoglucanase active in crystalline cellulose hydrolysis by the brown rot basidiomycete *Gloeophyllum trabeum*. *Appl. Environ. Microbiol.* **71**, 2412 (2005).
81. H. Toda *et al.,* Gene cloning of an endoglucanase from the basidiomycete *Irpex lacteus* and its cDNA expression in *Saccharomyces cerevisiae*. *Biosci. Biotechnol. and Biochem.* **69**, 1262 (2005).
82. B. Decelle, A. Tsang, R. Storms, Cloning, functional expression and characterization of three *Phanerochaete chrysosporium* endo-1,4-beta-xylanases. *Curr. Genet.* **46**, 166 (2004).
83. P. W. J. De Groot *et al.*, An endo-1,4-beta-xylanase-encoding gene from *Agaricus bisporus* is regulated by compost-specific factors. *J. Mol. Biol.* **277**, 273 (1998).
84. S. Christgau *et al.*, Pectin methyl esterase from *Aspergillus aculeatus* : expression cloning in yeast and characterization of the recombinant enzyme. *Biochem. J.* **319**, 705 (1996).
85. O. Markovic, S. Janecek, Pectin methylesterases: sequence-structural features and phylogenetic relationships. *Carbohydr. Res.* **339**, 2281 (2004).
86. X.-L. Li, S. Špániková, R. P. de Vries, P. Biely, Identification of genes encoding microbial glucuronoyl esterases. *FEBS Lett.* **581**, 4029 (2007).
87. D. Faure, The Family-3 Glycoside Hydrolases: from Housekeeping Functions to Host-Microbe Interactions. *Appl. Environ. Microbiol.* **68**, 1485 (2002).
88. C. B. Faulds *et al.*, Synergy between xylanases from glycoside hydrolase family 10 and family 11 and a feruloyl esterase in the release of phenolic acids from cereal arabinoxylan. *Appl. Microbiol. Biotechnol.* **71**, 622 (2005).
89. D. Guillén, S. Sánchez, R. Rodríguez-Sanoja, Carbohydrate-binding domains: multiplicity of biological roles. *Appl. Microbiol. Biotechnol.* **85**, 1241 (2010).
90. L. Wang, Y. Zhang, P. Gao, A novel function for the cellulose binding module of cellobiohydrolase I. *Sci. China C Life Sci.* **51**, 620 (2008).

91. F. J. Ruiz-Dueñas, A. T. Martinez, in *Biocatalysts based on heme peroxidases,* E. Torres, M. Ayala, Eds. (Springer-Verlag, Berlin, 2010), pp. 37-59.
92. F. J. Ruiz-Dueñas, A. T. Martínez, Microbial degradation of lignin: How a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. *Microb. Biotechnol.* **2**, 164 (2009).
93. F. J. Ruiz-Dueñas, E. Fernández-Fueyo, M. J. Martínez, A. T. Martínez, *Pleurotus ostreatus* heme peroxidases: An in silico analysis from the genome sequence to the enzyme molecular structure. *C. R. Biol.* **334**, 795 (2011).
94. K. Tamura, J. Dudley, M. Nei, S. Kumar, MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596 (2007).
95. D. Martinez *et al.*, Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. *Pro. Natl. Acad. Sci. U.S.A.* **106**, 1954 (2009).
96. D. Martinez *et al.*, Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nat. Biotechnol.* **22**, 695 (2004).
97. A. T. Martínez, F. J. Ruiz-Dueñas, M. J. Martínez, J. C. d. Río, A. Gutiérrez, Enzymatic delignification of plant cell wall: from nature to mill. *Curr. Opin. Biotechnol.* **20**, 348 (2009).
98. A. T. Martinez, Molecular biology and structure-function of lignin-degrading heme peroxidases. *Enzyme Microbial Technol.* **30**, 425 (2002).
99. K. Hildén, A. T. Martinez, A. Hatakka, T. Lundell, The two manganese peroxidases Pr-MnP2 and Pr-MnP3 of *Phlebia radiata*, a lignin-degrading basidiomycete, are phylogenetically and structurally divergent. *Fungal Genet. Biol.* **42**, 403 (2005).
100. M. Hofrichter, R. Ullrich, M. Pecyna, C. Liers, T. Lundell, New and classic families of secreted fungal heme peroxidases. *Appl. Microbiol. Biotechnol.* **87**, 871 (2010).
101. D. M. Li, H. L. Youngs, M. H. Gold, Heterologous expression of a thermostable manganese peroxidase from *Dichomitus squalens* in *Phanerochaete chrysosporium*. *Arch. Biochem. Biophys.* **385**, 348 (2001).
102. F. J. Ruiz-Dueñas *et al.*, Manganese oxidation site in *Pleurotus eryngii* versatile peroxidase: A site-directed mutagenesis, kinetic and crystallographic study. *Biochemistry* **46**, 66 (2007).
103. F. Nerud, Z. Zouchova, Z. Misurcová, Ligninolytic properties of different white-rot Fungi. *Biotechnol. Lett.* **13**, 657 (1991).
104. C. Liers, C. Bobeth, M. Pecyna, R. Ullrich, M. Hofrichter, DyP-like peroxidases of the jelly fungus *Auricularia auricula-judae* oxidize nonphenolic lignin model compounds and high-redox potential dyes. *Appl. Microbiol. Biotechnol.* **85**, 1869 (2010).
105. L. Larrondo, A. González, A. T. Pérez, D. Cullen, R. Vicuña, The nop gene from *Phanerochaete chrysosporium* encodes a peroxidase with novel structural features. *Biophys. Chem.* **116**, 167 (2005).
106. A. Gutiérrez *et al.*, Regioselective oxygenation of fatty acids, fatty alcohols and other aliphatic compounds by a basidiomycete heme-thiolate peroxidase. *Arch. Biochem. Biophys.* **514**, 33 (2011).
107. R. Ullrich, M. Hofrichter, The haloperoxidase of the agaric fungus *Agrocybe aegerita* hydroxylates toluene and naphthalene. *FEBS Lett.* **579**, 6247 (2005).
108. T. K. Lundell, M. R. Mäkelä., K. Hildén, Lignin-modifying enzymes in filamentous basidiomycetes--ecological, functional and phylogenetic review. *J. Basic Microbiol.* **50**, 5 (2010)
109. A. T. Martínez *et al.*, Biodegradation of lignocellulosics: Microbiological, chemical and enzymatic aspects of fungal attack to lignin. *Int. Microbiol.* **8**, 195 (2005).
110. P. Baldrian, V. Valášková, Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiol. Rev.* **32**, 501 (2008).
111. G. Daniel *et al.*, Characteristics of *Gloeophyllum trabeum* Alcohol Oxidase, an Extracellular Source of H_2O_2 in Brown Rot Decay of Wood. *Appl. Environ. Microbiol.* **73**, 6241 (2007).
112. P. J. Kersten, Glyoxal oxidase of *Phanerochaete chrysosporium*: Its characterization and activation by lignin peroxidase. *Proc. Natl. Acad. Sci. U.S.A* **87**, 2936 (1990).
113. J. A. Langston *et al.*, Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Appl. Environ. Microbiol*. **77**, 7007 (2011).

114. A. Vanden Wymelenberg *et al.*, The *Phanerochaete chrysosporium* secretome: database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium. *J. Biotechnol.* **118**, 17 (2005).
115. F. Guillén, A. T. Martínez, M. J. Martínez, C. S. Evans, Hydrogen peroxide-producing system of *Pleurotus eryngii* involving the extracellular enzyme aryl-alcohol oxidase. *Appl. Microbiol. Biotechnol.* **41**, 465 (1994).
116. G. Daniel, J. Volc, E. Kubátová, Pyranose oxidase, a major source of H2O2 during wood degradation by *Phanerochaete chrysosporium*, *Trametes versicolor*, and *Oudemansiella mucida*. *Appl. Environ. Microbiol.* **60**, 2524 (1994).
117. T. Kajisa *et al.*, Characterization and molecular cloning of cellobiose dehydrogenase from the brown-rot fungus *Coniophora puteana*. *J. Biosci. Bioeng.* **98**, 57 (2004).
118. A. Vanden Wymelenberg *et al.*, Structure, organization, and transcriptional regulation of a family of copper radical oxidase genes in the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* **72**, 4871 (2006).
119. A. Käärik, The identification of the mycelia of wood-decay fungi by their oxidation reactions with phenolic compounds. *Studia Forestalia Suecica* **31**, 3 (1965).
120. A. I. Cañas, S. Camarero, Laccases and their natural mediators: Biotechnological tools for sustainable eco-friendly processes. *Biotechnol. Adv.* **28**, 694 (2010).
121. F. Guillén, C. Muñoz, V. Gómez-Toribio, A. T. Martínez, M. J. Martínez, Oxygen activation during the oxidation of methoxyhydroquinones by laccase from *Pleurotus eryngii*. *Appl. Environ. Microbiol.* **66**, 170 (2000).
122. S. Kilaru, P. J. Hoegger, U. Kües, The laccase multi-gene family in Coprinopsis cinerea has seventeen different members that divide into two distinct subfamilies. *Curr. Genet.* **50**, 45 (2006).
123. D. Martinez *et al.*, Genome sequence of the lignocellulose degrading fungus Phanerochaete chrysosporium strain RP78. *Nat. Biotechnol.* **22**, 695 (Jun, 2004).
124. H. Tanaka, K. Koike, S. Itakura, A. Enoki, Degradation of wood and enzyme production by Ceriporiopsis subvermispora. *Enz. Mic. Tech.* **45**, 384 (2009).
125. F. Guillén, V. Gómez-Toribio, M. J. Martínez, A. T. Martínez, Production of hydroxyl radical by the synergistic action of fungal laccase and aryl alcohol oxidase. *Arch. Biochem. Biophys.* **383**, 142 (2000).
126. V. Gómez-Toribio, A. B. García-Martín, M. J. Martínez, A. T. Martínez, F. Guillén, Enhancing the production of hydroxyl radicals by Pleurotus eryngii via quinone redox cycling for pollutant removal. *Appl. Environ. Microbiol.* **75**, 3954 (2009).
127. D. Wei *et al.*, Laccase and its role in production of extracellular reactive oxygen species during wood decay by the brown rot basidiomycete *Postia placenta*. *Appl. Environ. Microbiol.* **76**, 2091 (2010).
128. H. Tanaka *et al.*, Characterization of a hydroxyl-radical-producing glycoprotein and its presumptive genes from the white-rot basidiomycete *Phanerochaete chrysosporium*. *J. Biotechnol.* **128**, 500 (2007).
129. L. Akileswaran, B. J. Brock, J. L. Cereghino, M. H. Gold, 1,4-benzoquinone reductase from Phanerochaete chrysosporium: cDNA cloning and regulation of expression. *Appl. Environ. Microbiol.* **65**, 415 (1999).
130. M. R. Suzuki, C. G. Hunt, C. J. Houtman, Z. D. Dalebroux, K. E. Hammel, Fungal hydroquinones contribute to brown rot of wood. *Env. Microbiol.* **8**, 2214 (2006).
131. M. R. Makela, K. Hilden, A. Hatakka, T. K. Lundell, Oxalate decarboxylase of the white-rot fungus Dichomitus squalens demonstrates a novel enzyme primary structure and non-induced expression on wood and in liquid cultures. *Microbiology* **155**, 2726 (2009).
132. D. C. Eastwood *et al.*, The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science* **333**, 762 (2011).
133. M. V. Dutton, C. S. Evans, Oxalate production by fungi: its role in pathogenicity and ecology in the soil environment *Canad. J. Microbiol.* **42**, 881 (1996).
134. S. Khuri, F. T. Bakker, J. M. Dunwell, Phylogeny, function, and evolution of the cupins, a structurally conserved, functionally diverse superfamily of proteins *Mol. Biol. Evol.* **18**, 593 (2001).

135. U. Urzua, P. J. Kersten, R. Vicuna, Manganese peroxidase-dependent oxidation of glyoxlic and oxalic acids synthesized by Ceriporiopsis subvermispora produces extracellular hydrogen peroxide. *Appl. Environ. Microbiol.* **64**, 68 (1998).
136. M. R. Escutia *et al.*, Cloning and Sequencing of Two Ceriporiopsis subvermispora Bicupin Oxalate Oxidase Allelic Isoforms: Implications for the Reaction Specificity of Oxalate Oxidases and Decarboxylases. *Appl. Environ. Microbiol.* **71**, 3608 (2005).
137. H. Doddapaneni, R. Chakraborty, J. Yadav, Genome-wide structural and evolutionary analysis of the P450 monooxygenase genes (P450ome) in the white rot fungus Phanerochaete chrysosporium : Evidence for gene duplications and extensive gene clustering. *BMC Genomics* **6**, 92 (2005).
138. H. Doddapaneni, V. Subramanian, J. Yadav, Physiological regulation, xenobiotic induction and heterologous expression of P450 monooxygenase gene pc-3, a new member of the CYP63 gene cluster in the white rot fungus Phanerochaete chrysosporium. *Curr. Microbiol.*, (2005).
139. H. Doddapaneni, J. S. Yadav, Microarray-based global differential expression profiling of P450 monooxygenases and regulatory proteins for signal transduction pathways in the white rot fungus *Phanerochaete chrysosporium*. *Mol. Genet. Genomics.* **274**, 454 (2005).
140. V. Subramanian, J. S. Yadav, Role of P450 Monooxygenases in the Degradation of the Endocrine-Disrupting Chemical Nonylphenol by the White Rot Fungus Phanerochaete chrysosporium. *Appl. Environ. Microbiol.* **75**, 5570 (2009).
141. K. Syed, H. Doddapaneni, V. Subramanian, Y. W. Lam, J. S. Yadav, Genome-to-function characterization of novel fungal P450 monooxygenases oxidizing polycyclic aromatic hydrocarbons (PAHs). *Biochem. Biophys. Res. Commun.* **399**, 492 (2010).
142. K. Eriksson, R. Blanchette, P. Ander, Biodegradation of Lignin. *Microbial and enzymatic degradation of wood and wood components*, 225 (1990).
143. S. P. Stubblefield, T. N. Taylor, Wood decay in silicified gymnosperms from Antarctica. *Bot. Gaz.* **147**, 116 (1986).
144. S. P. Stubblefield, T. N. Taylor, C. B. Beck, Studies of paleozoic fungi. IV. Wood-decaying fungi in *Callixylon newberryi* from the Upper Devonian. *Am. J. Bot.* **72**, 1765 (1985).
145. R. Blanchette *et al.*, An Antarctic Hot Spot for Fungi at Shackleton's Historic Hut on Cape Royds. *Microb. Ecol.* **60**, 29 (2010).
146. P. M. Coutinho *et al.*, Post-genomic insights into the plant polysaccharide degradation potential of *Aspergillus nidulans* and comparison to *Aspergillus niger* and *Aspergillus oryzae*. *Fungal. Genet. Biol.* **46 Suppl 1**, S161 (Mar, 2009).
147. R. A. Ohm *et al.*, Genome sequence of the model mushroom Schizophyllum commune. *Nat. Biotechnol.*, (Jul 11, 2010).
148. J. M. Plett *et al.*, A Secreted Effector Protein of Laccaria bicolor Is Required for Symbiosis Development. *Curr. Biol.* **21**, 1197 (2011).
149. M. L. Berbee, J. W. Taylor, Dating the molecular clock in fungi – how close are we? *Fung. Biol. Rev.* **24**, 1 (2010).
150. B. A. LePage, R. S. Currah, R. A. Stockey, G. W. Rothwell, Fossil ectomycorrhizae from the Middle Eocene. *Am. J. Bot.* **84**, 470 (1997).
151. D. S. Hibbett, D. Grimaldi, M. J. Donoghue, Fossil mushrooms from Cretaceous and Miocene ambers and the evolution of homobasidiomycetes. *Am. J. Bot.* **84**, 981 (1997).
152. T. N. Taylor, H. Hass, H. Kerp, M. Krings, R. T. Hanlin, Perithecial ascomycetes from the 400 million year old Rhynie chert: an example of ancestral polymorphism. *Mycologia* **97**, 269 (2005).
153. E. J. Douzery, E. A. Snell, E. Bapteste, F. Delsuc, H. Philippe, The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proc. Natl. Acad. Sci.* **101**, 15386 (Oct 26, 2004).
154. R. Lücking, S. Huhndorf, D. H. Pfister, E. R. Plata, H. T. Lumbsch, Fungi evolved right on track. *Mycologia* **101**, 810 (2009).
155. J. W. Taylor, M. L. Berbee, Dating divergences in the Fungal Tree of Life: review and new analyses. *Mycologia* **98**, 838 (Nov-Dec, 2006).
156. A. J. Drummond, A. Rambaut, BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).

157. N. Lartillot, A Bayesian Mixture Model for Across-Site Heterogeneities in the Amino-Acid Replacement Process. *Mol. Biol. Evol.* **21**, 1095 (2004).
158. T. Y. James *et al.*, Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443**, 818 (2006).

Figure S1: Genome size, number of microsatellites and TE genome coverage according to phylogeny of the 31 genomes analysed. The coverage for the different families of TE are indicated as well as the species interacting with living plants

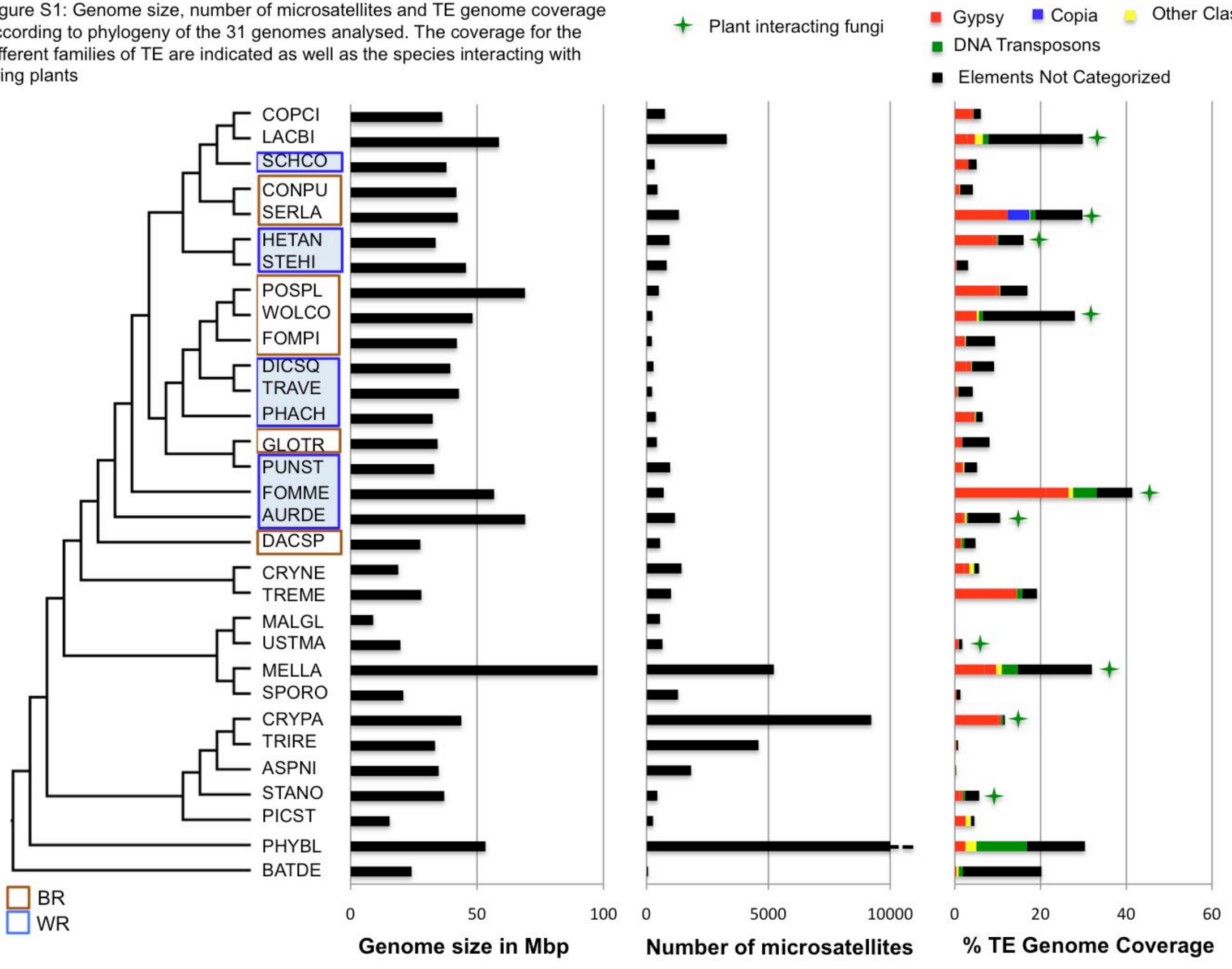


Figure S2: Number of microsatellites in saprotrophic Agaricomycotina presented for each species (A) or as the median for all the brown-rot (BR) and (WR) species (B).

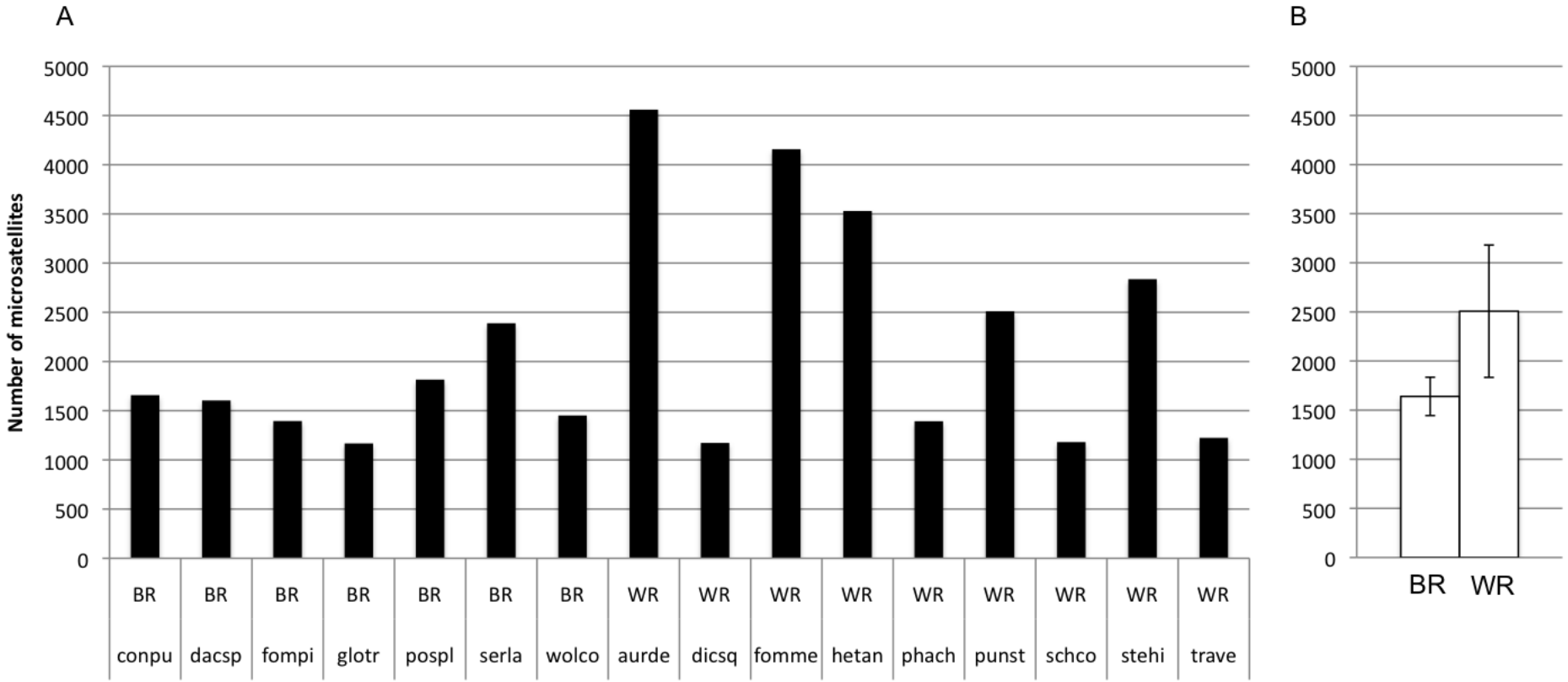


Figure S3: Median values for genome size, TE genome coverage, microsatellites, number of fulll-length Gypsy and copia elements for all the brown-rot (BR) and (WR) species.

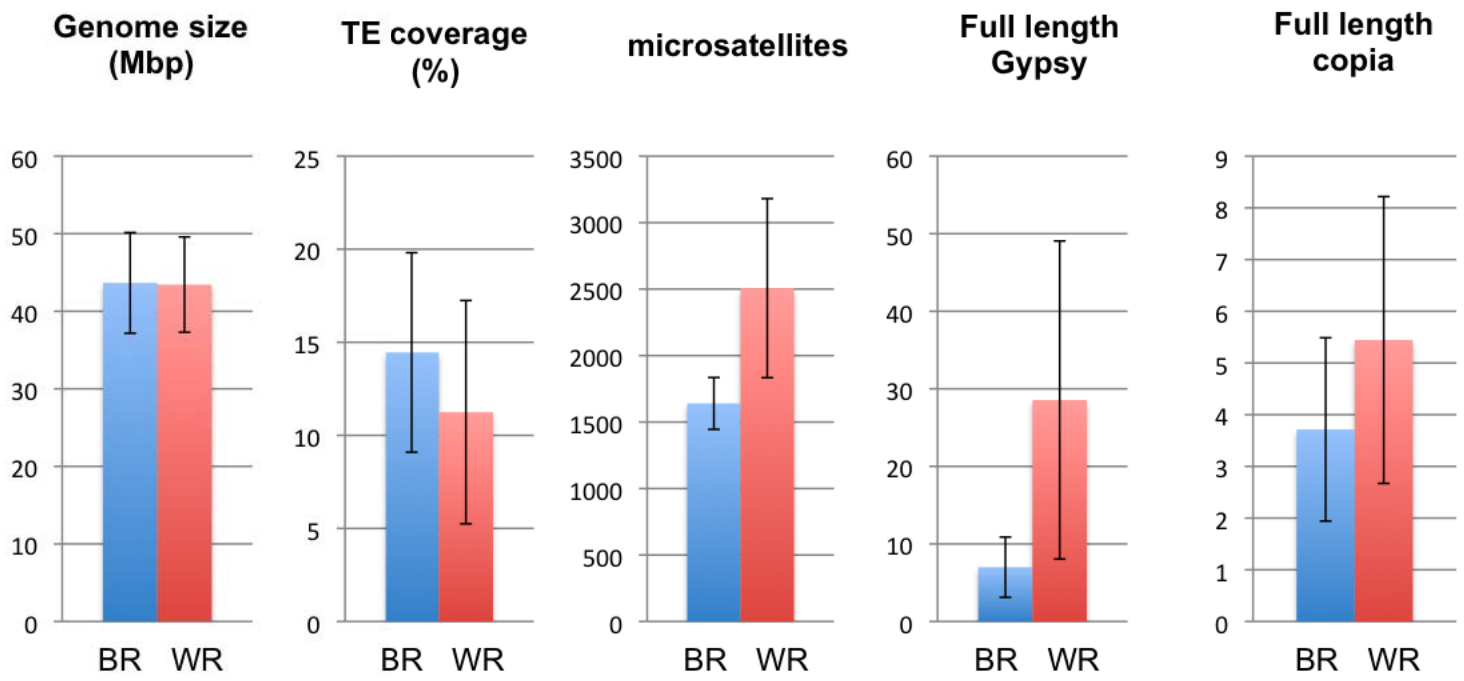


Figure S4: Correlation between genome size and TE coverage for all the 31 genomes analysed.

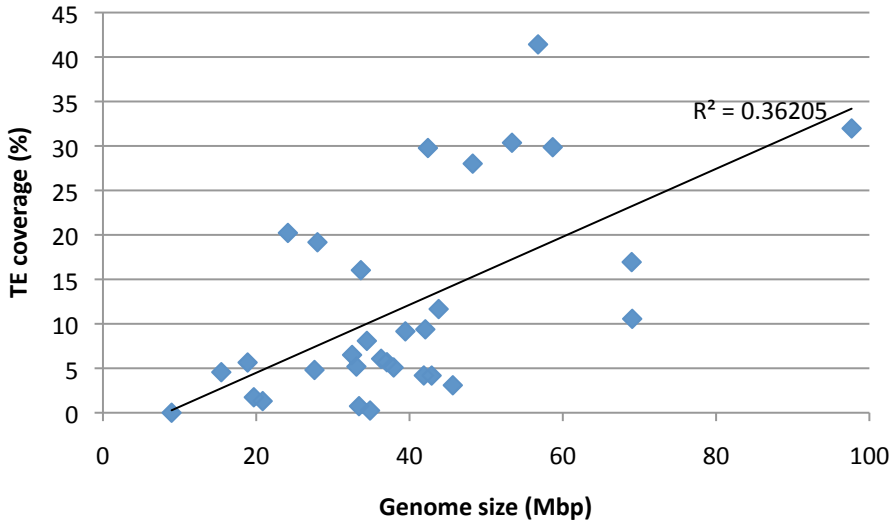


Figure S5: Estimation of insertional age for full length gypsy elements. A substitution mutation rate of 1.3 x 10-8 was used with both solo LTRs of each element.

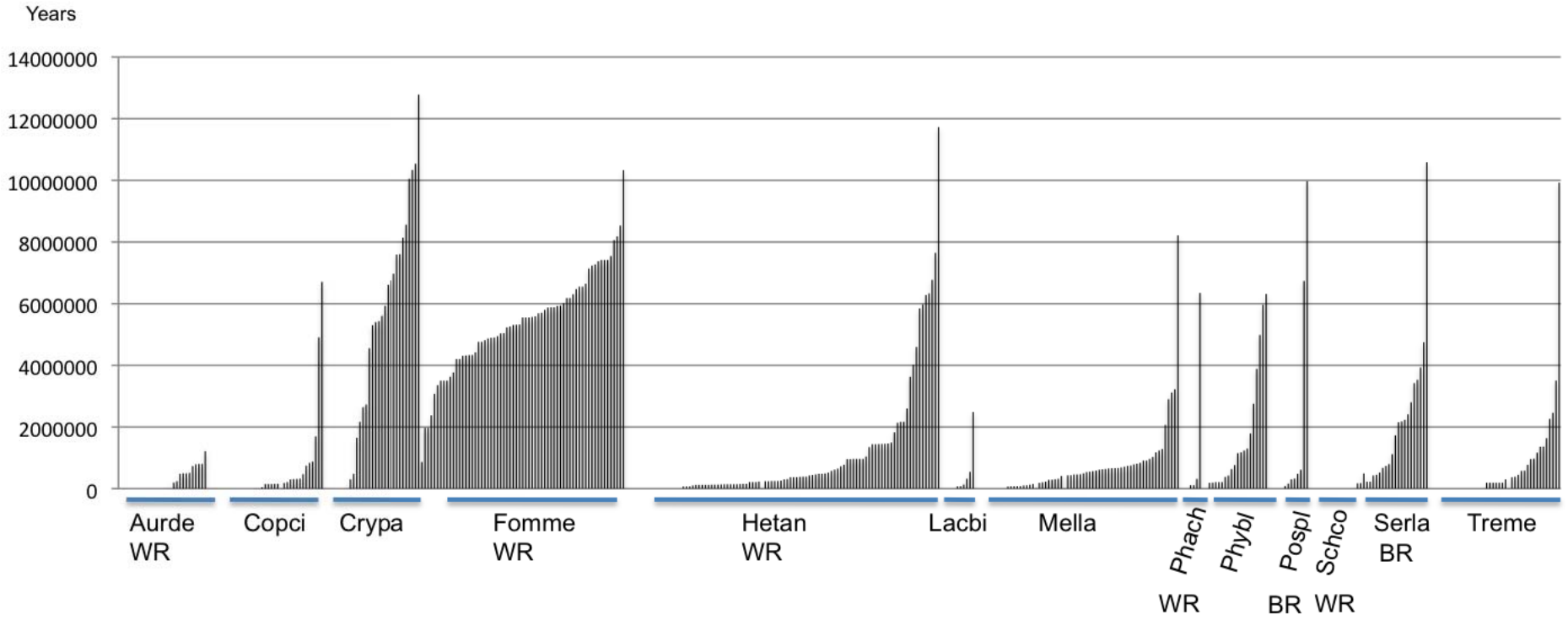


Fig. S6. Phylogenetic analysis of the 26G dataset using PhyloBayes 3.2. The topology of the tree (a single optimal tree shown with branch length) is identical with the tree topology estimated by BEAST v1.61, which was used as a starting tree in the downstream molecular clock analyses. Support values for nodes 4, 10, 12, 23, and 28 are indicated in the inset box for all four datasets (26G, 26GA, 71G, 71GA) inferred using RAxML 7.0.4 and PhyloBayes. All other nodes received maximal support in all analyses. (See Supporting Notes for descriptions of phylogenetic datasets). AURDE = A. delicata, ASPNI = A. niger, BATDE = B. dendrobatidis, COPCI = C. cinerea, CRYNE = C. neoformans, CONPU = C. puteana, CRYPA = C. parasitica, DACSP = Dacryopinax sp., DICSQ = D. squalens, FOMME = F. mediterranea, FOMPI = F. pinicola, GLOTR = G. trabeum, HETAN = H. annosum (has been reclassified as H. irregulare), LACBI = L. bicolor, MALGL = M. globosa, MELLA = M. laricis-populina, PHYBL = P. blakesleeanus., PHACH = P. chrysosporium, POSPL = P. placenta, PICST = P. stipitis, PUNST = P. strigosozonata, SCHCO = S. commune, STEHI = S. hirsutum, SERLA = S. lacrymans, STANO = S. nodorum, SPORO = S. roseus, TREME = T. mesenterica, TRIRE = T. reesei, TRAVE = T. versicolor, USTMA = U. maydis, WOLCO = W. cocos.

Node number	RAxML 71G,71GA; 26G, 26GA	PhyloBayes 71G,71GA; 26G, 26GA
4	91, 93; 100, 100	1.0, 1.0; 0.98, 1.0
10	45, 100; -, -	-, 1.0; 1.0, -
12	29, 59; 96, 66	1.0, 1.0; 1.0, 0.99
21	99, 69; 84, 80	1.0, 0.96; 0.99, 0.99
23	99, 100; 62, 64	1.0, 1.0; 1.0, 0.99
28	91, 93; 100, 100	1.0, 1.0; 0.98, 1.0

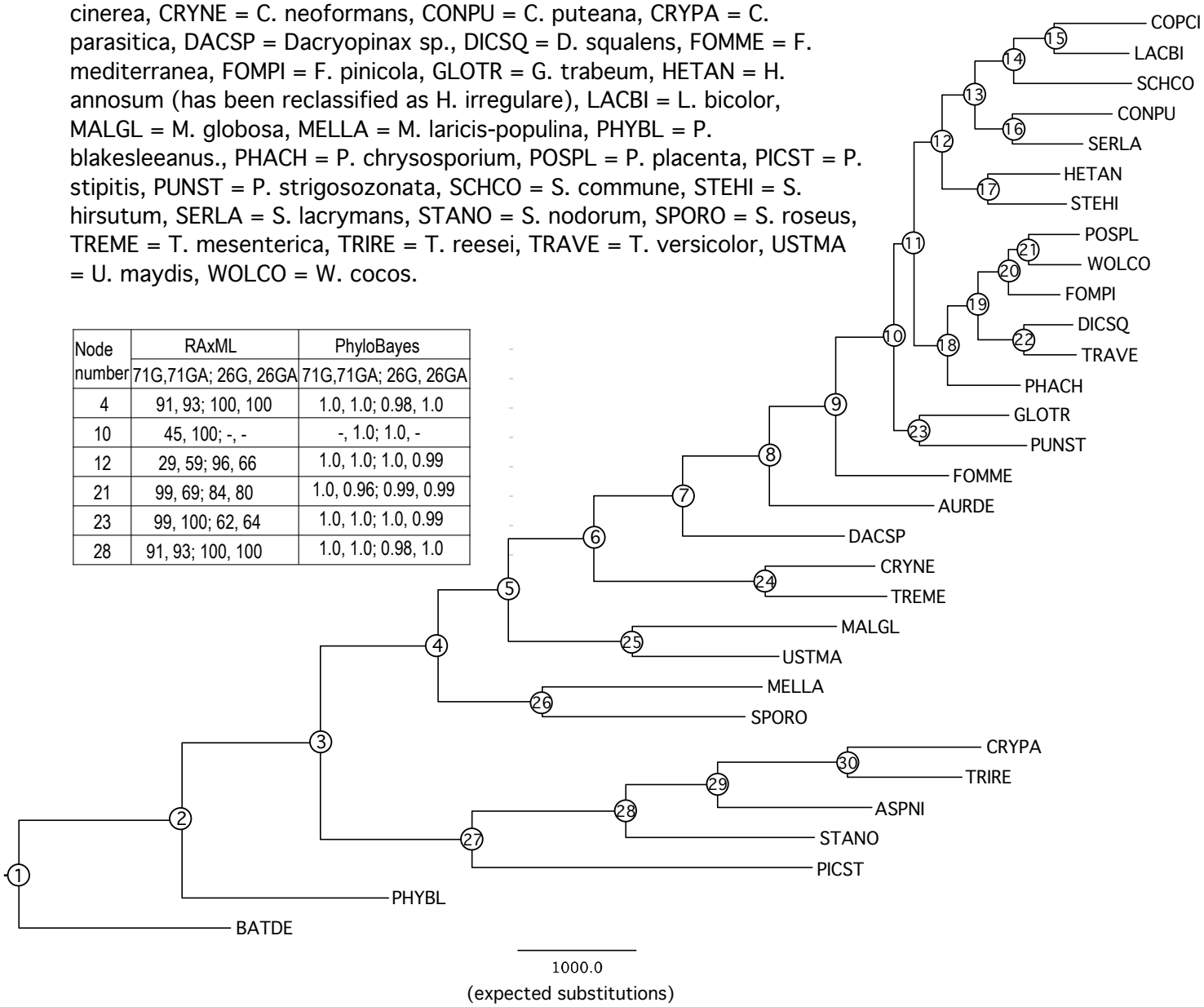


Fig. S7. Changes in gene copy numbers estimated with CAFÉ. Blue font indicates gene expansion, red indicates contraction. Gene and species codes as in Table 1.

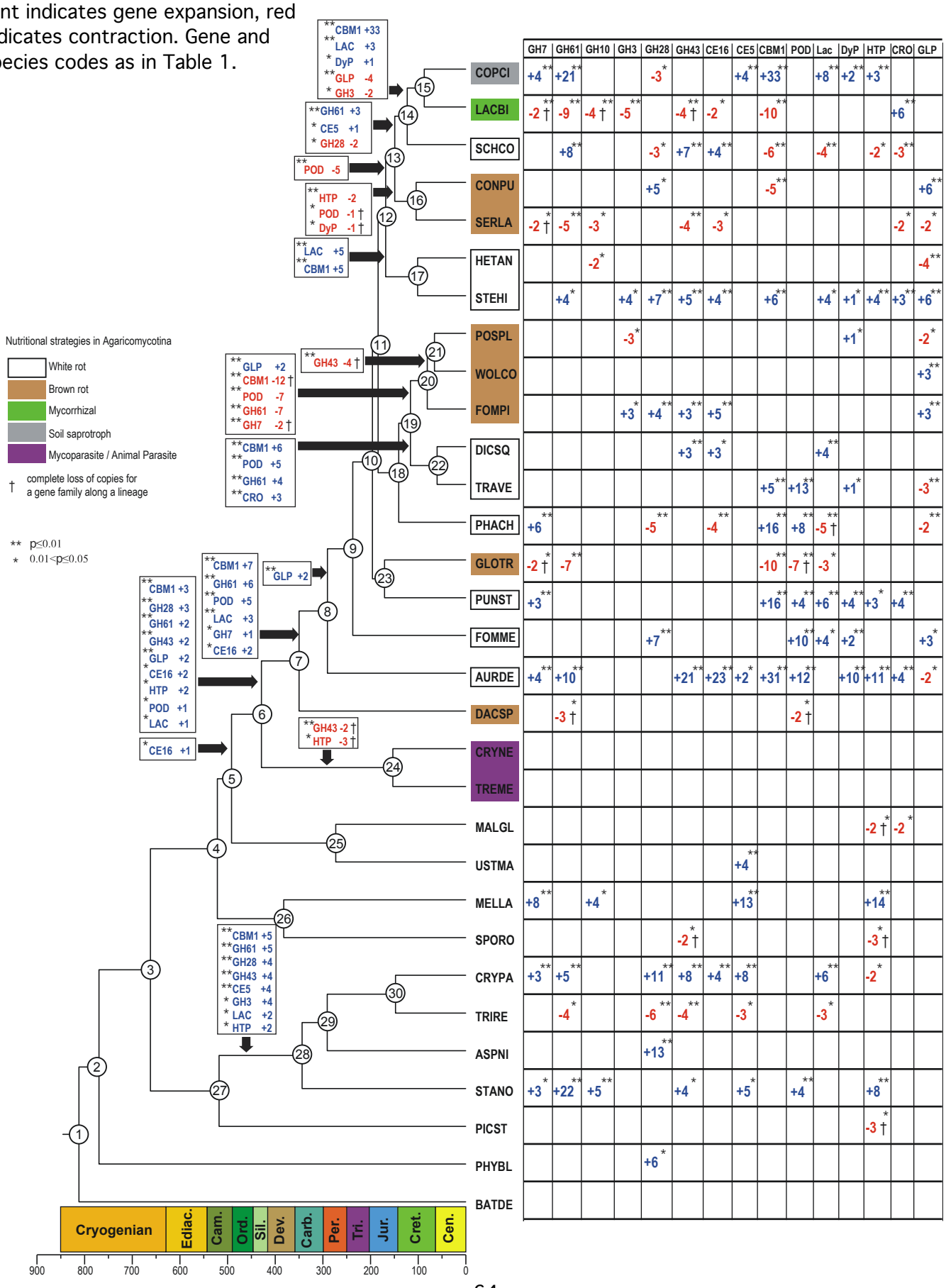


Fig. S8. Best ML trees from RAxML analyses of 14 gene families encoding decay-related CAZymes and oxidoreductases supported in CAFE analysis as having non-random changes in gene copies numbers (and GH5, GH6 and CDH). Bootstrap values greater than 65% are shown.

Fig. S8a. Glycoside Hydrolases family 3 group 1 and 2.

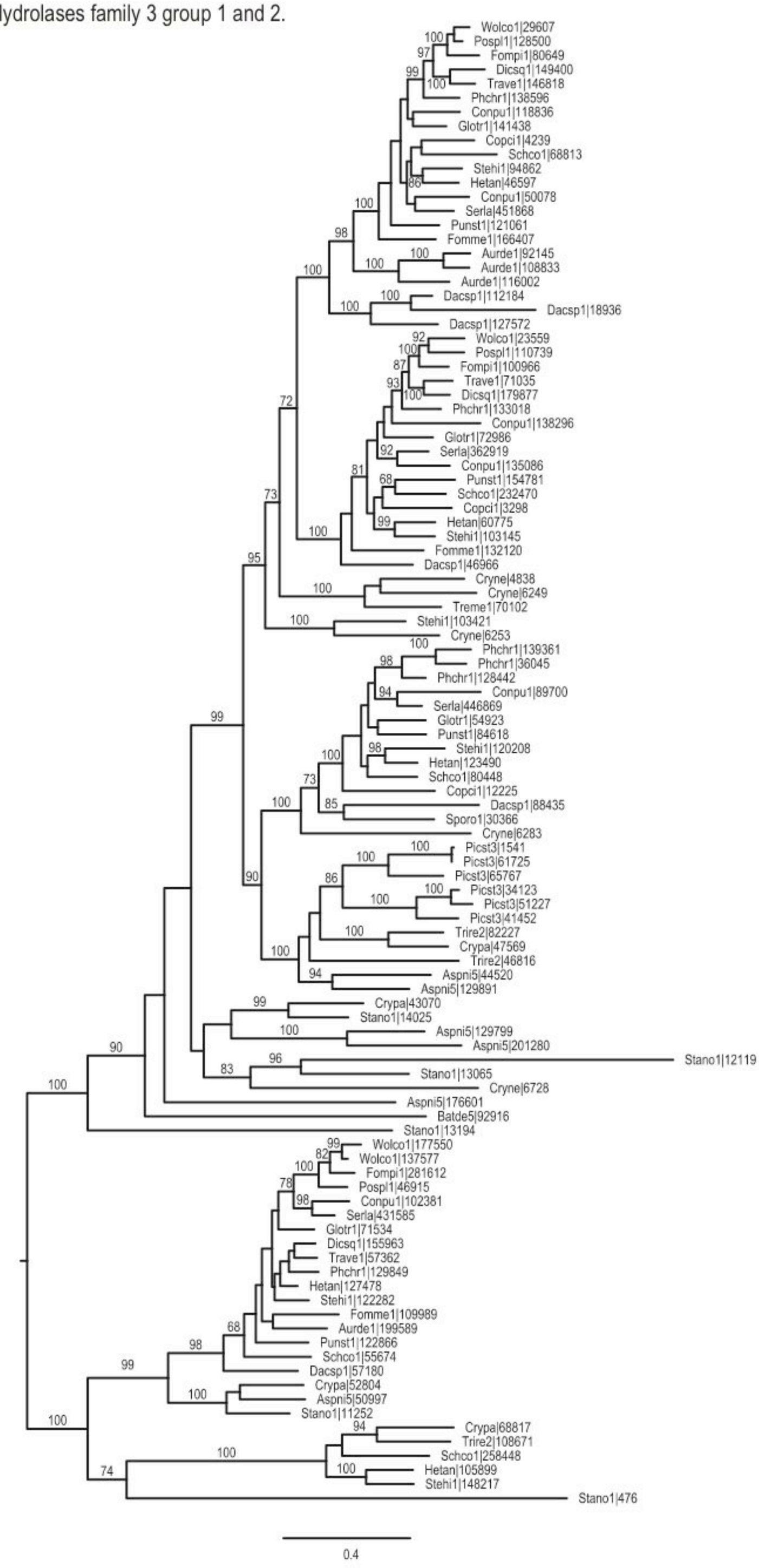


Fig. S8b. Glycoside Hydrolases family 3 group 3 (left) and potential beta-xylosidases (right)

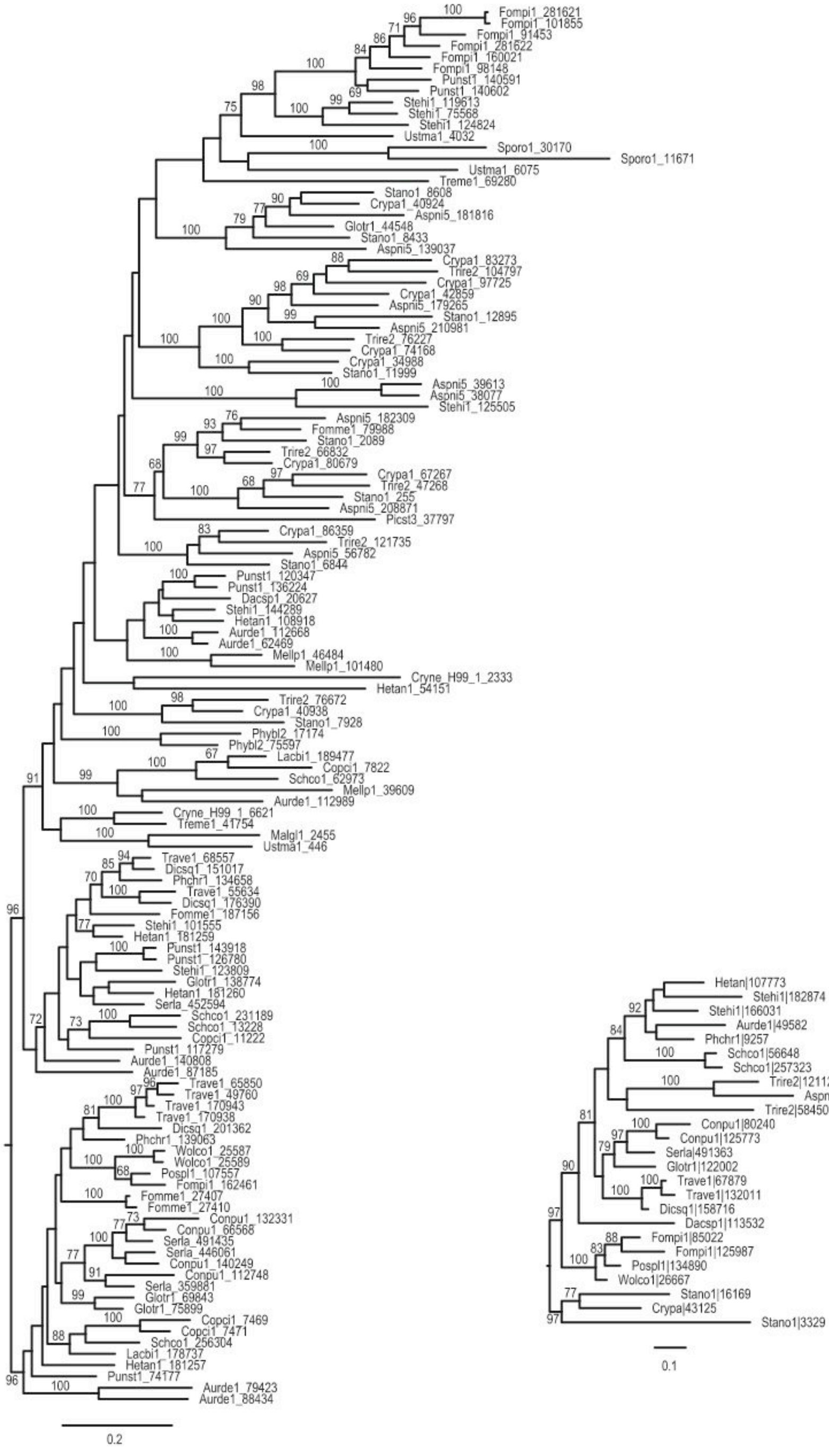


Fig. S8c. Glycoside Hydrolases family 5, candidate mannanases (left) and candidate endoglucanases (right) with CBM1 distribution. Protein models are indicated with green or brown if they are present in white rot species or brown rot species respectively.

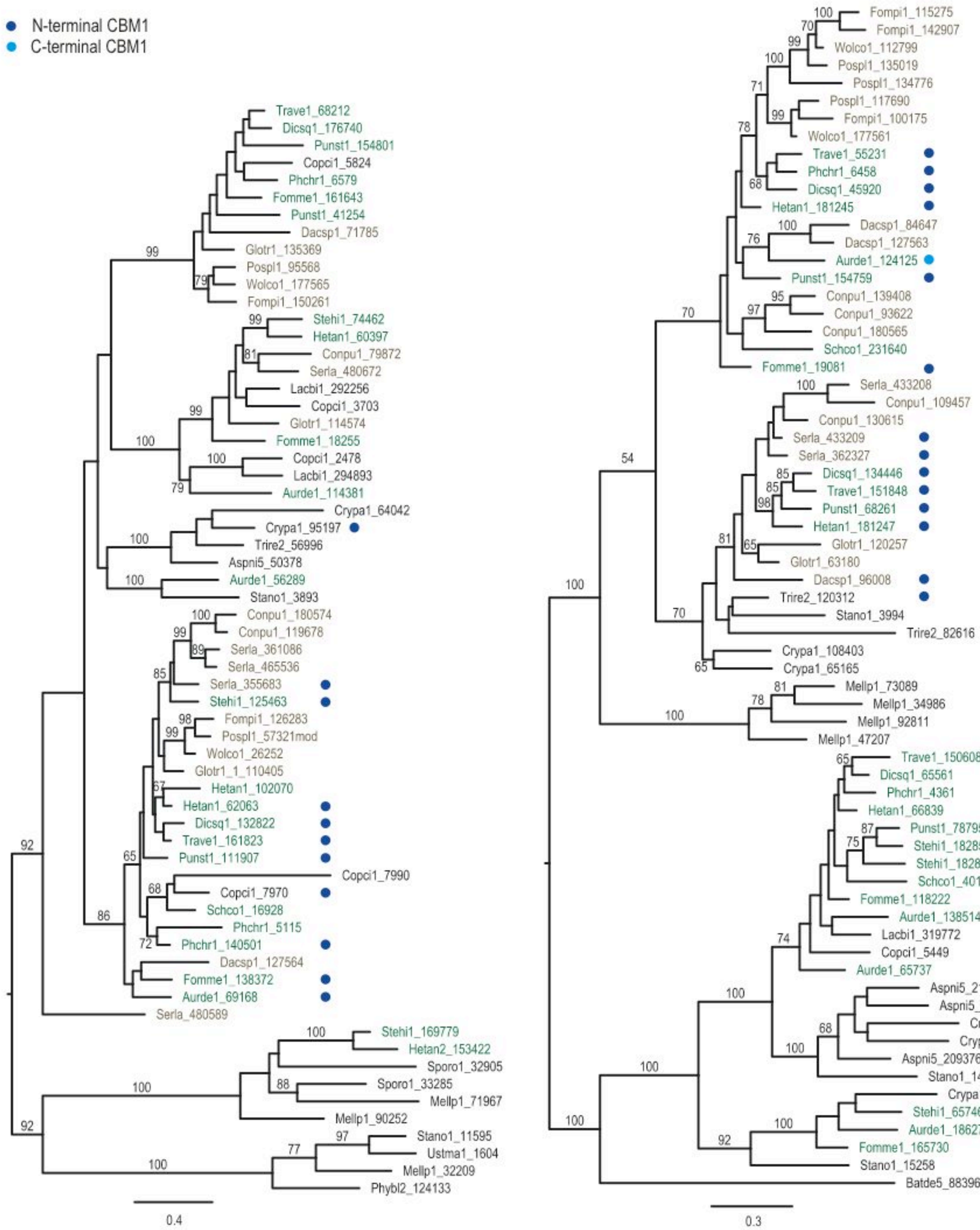


Fig. S8d. Glycoside Hydrolases family 7 (upper tree) and family 6 (lower tree)

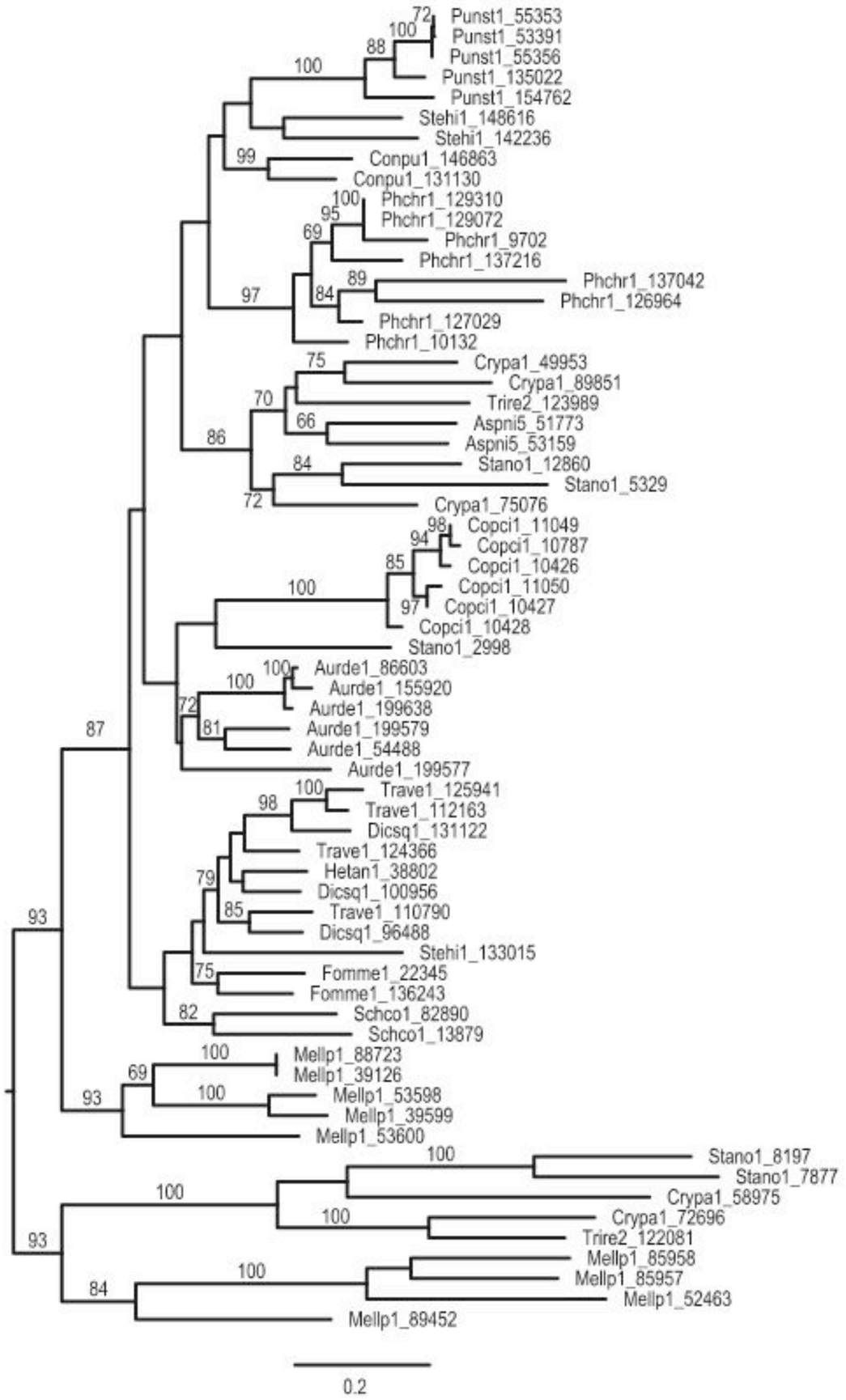


Fig. S8e. Glycoside Hydrolases family 28 group 1 (upper tree) and group 2 (lower tree).

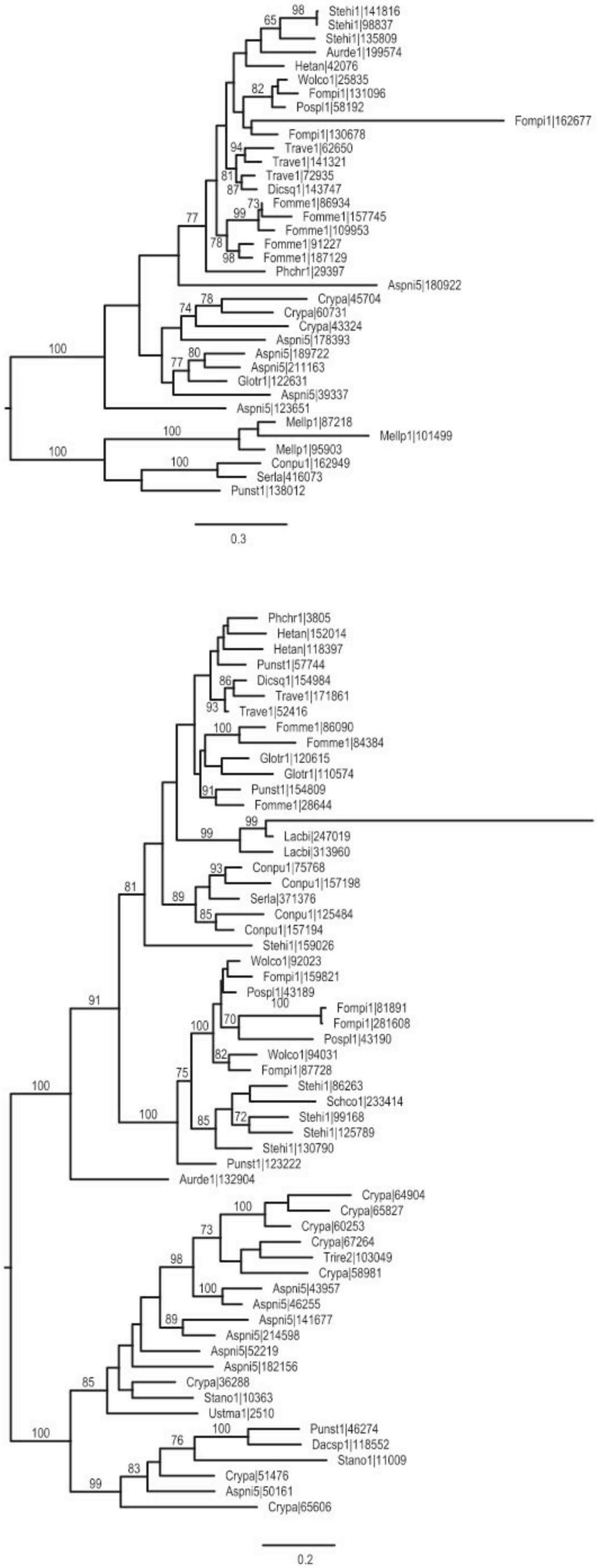


Fig. S8f. Glycoside Hydrolases family 28 group 3.

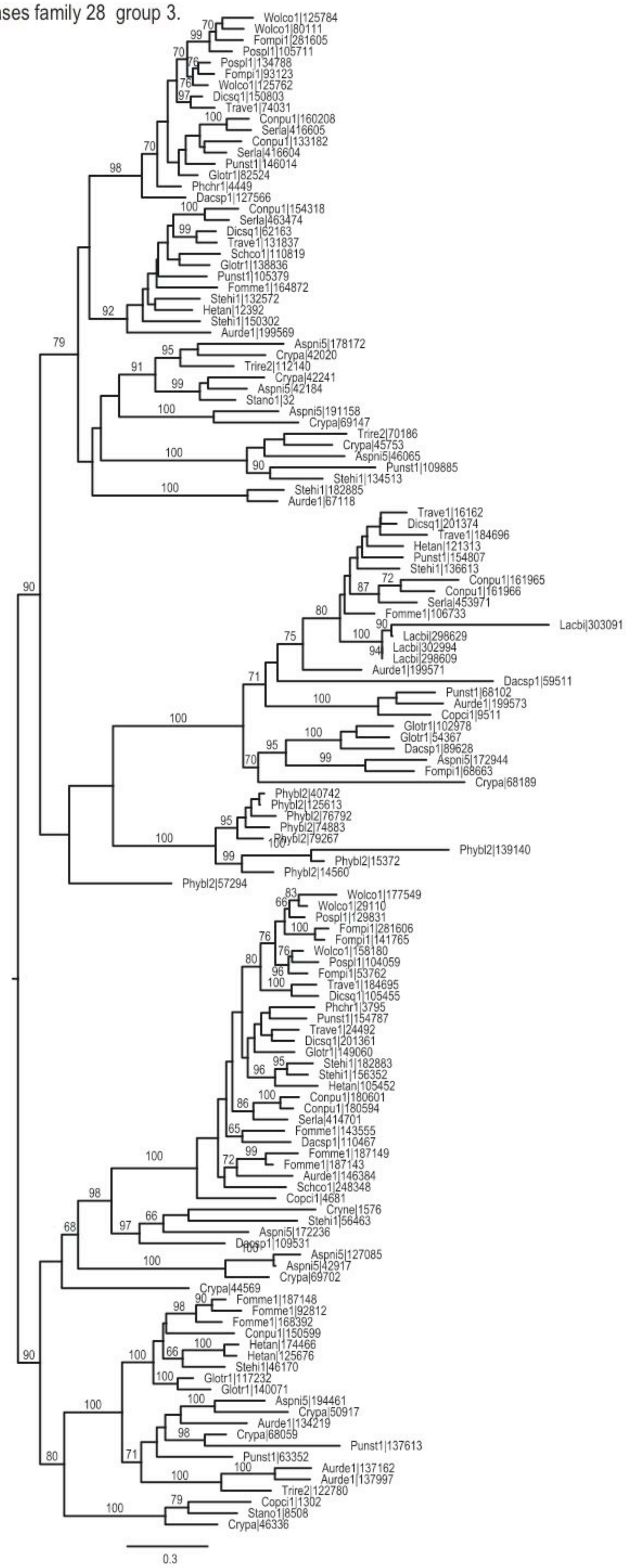


Fig. S8g. Glycoside Hydrolases family 10 with CBM1 distribution. Protein models are indicated with green or brown if they are present in white rot species or brown rot species respectively.

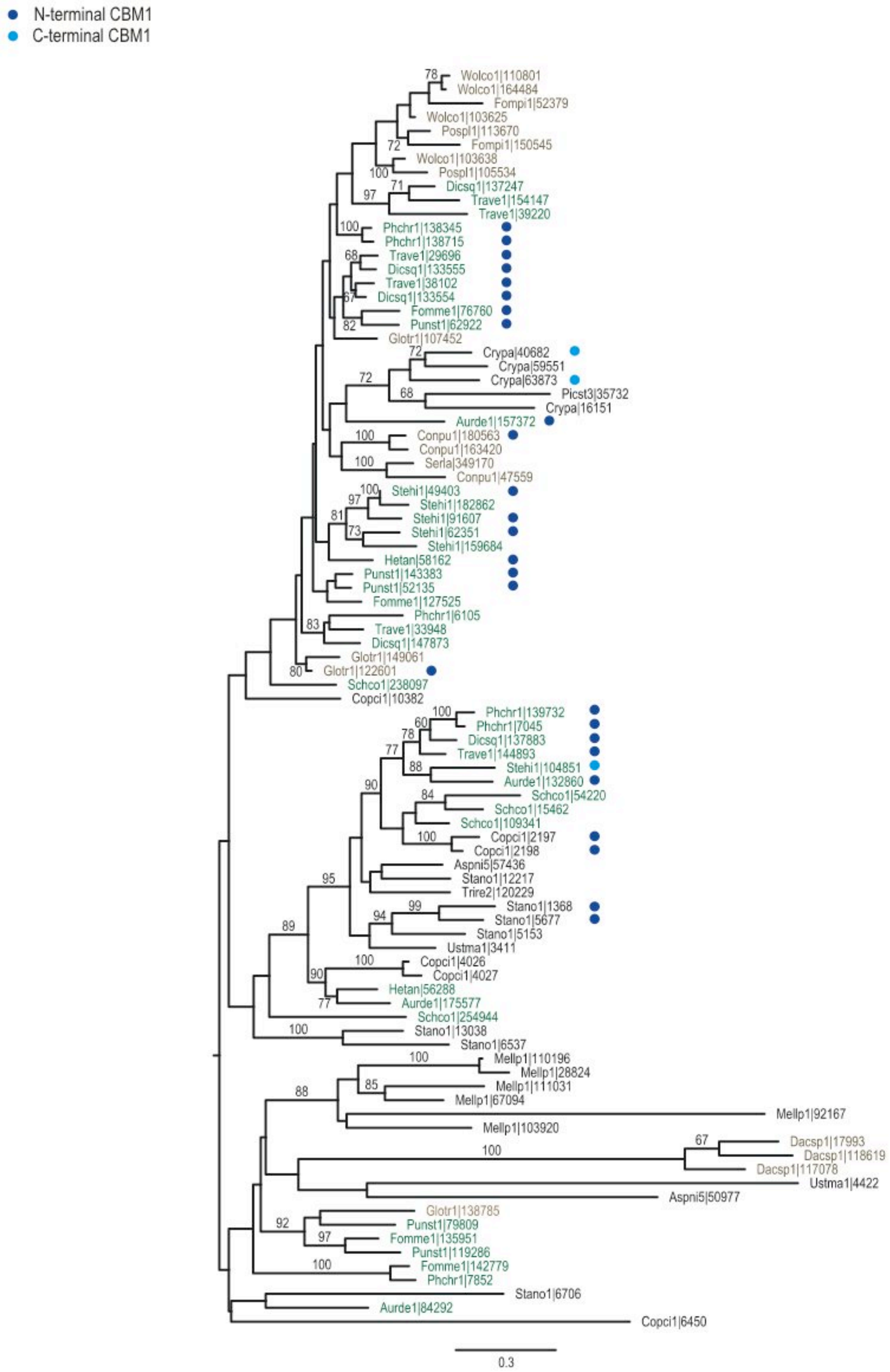
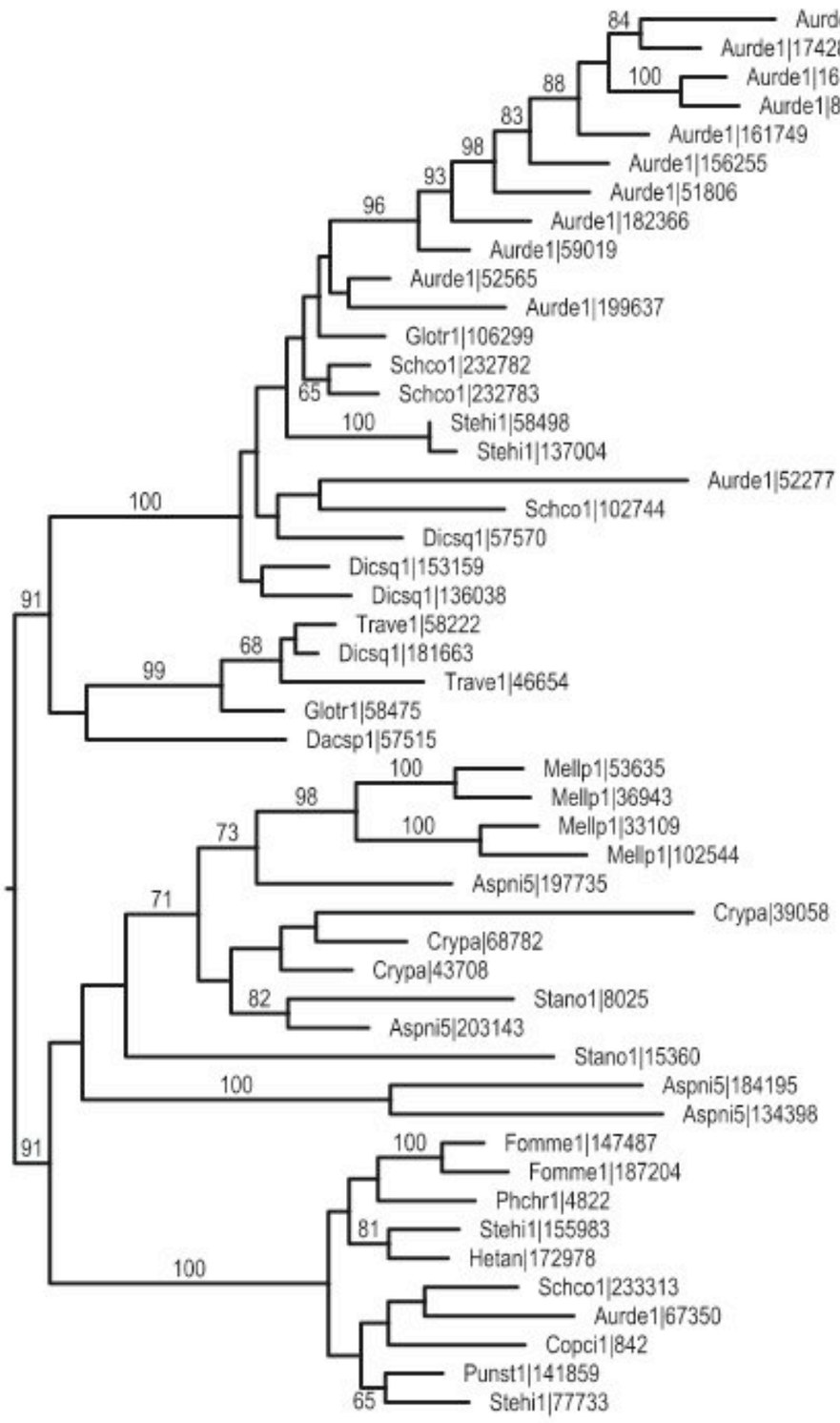
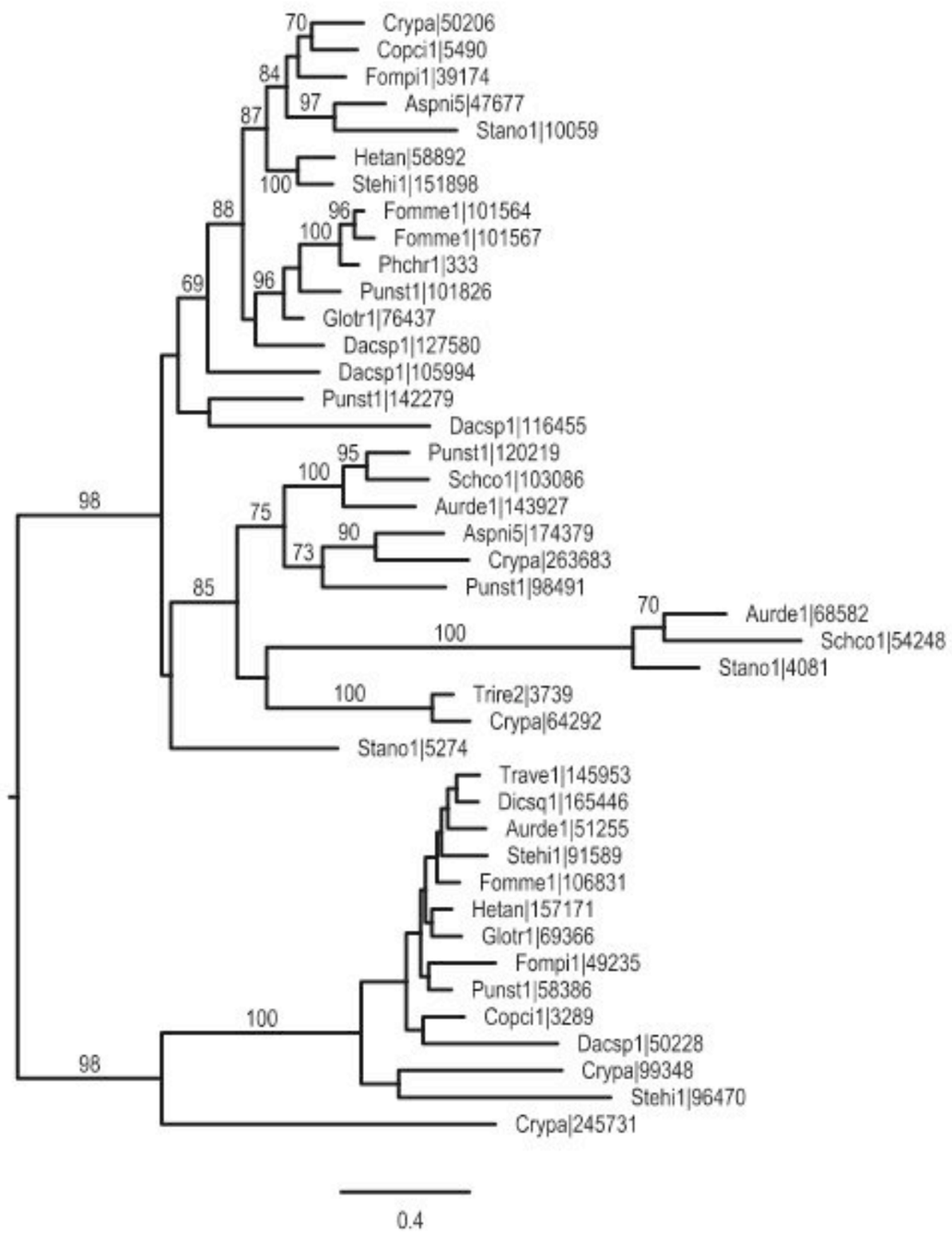


Fig. S8h. Glycoside Hydrolases family 43 group 1 (upper left), group 2 (upper right) and group 3 (lower tree).





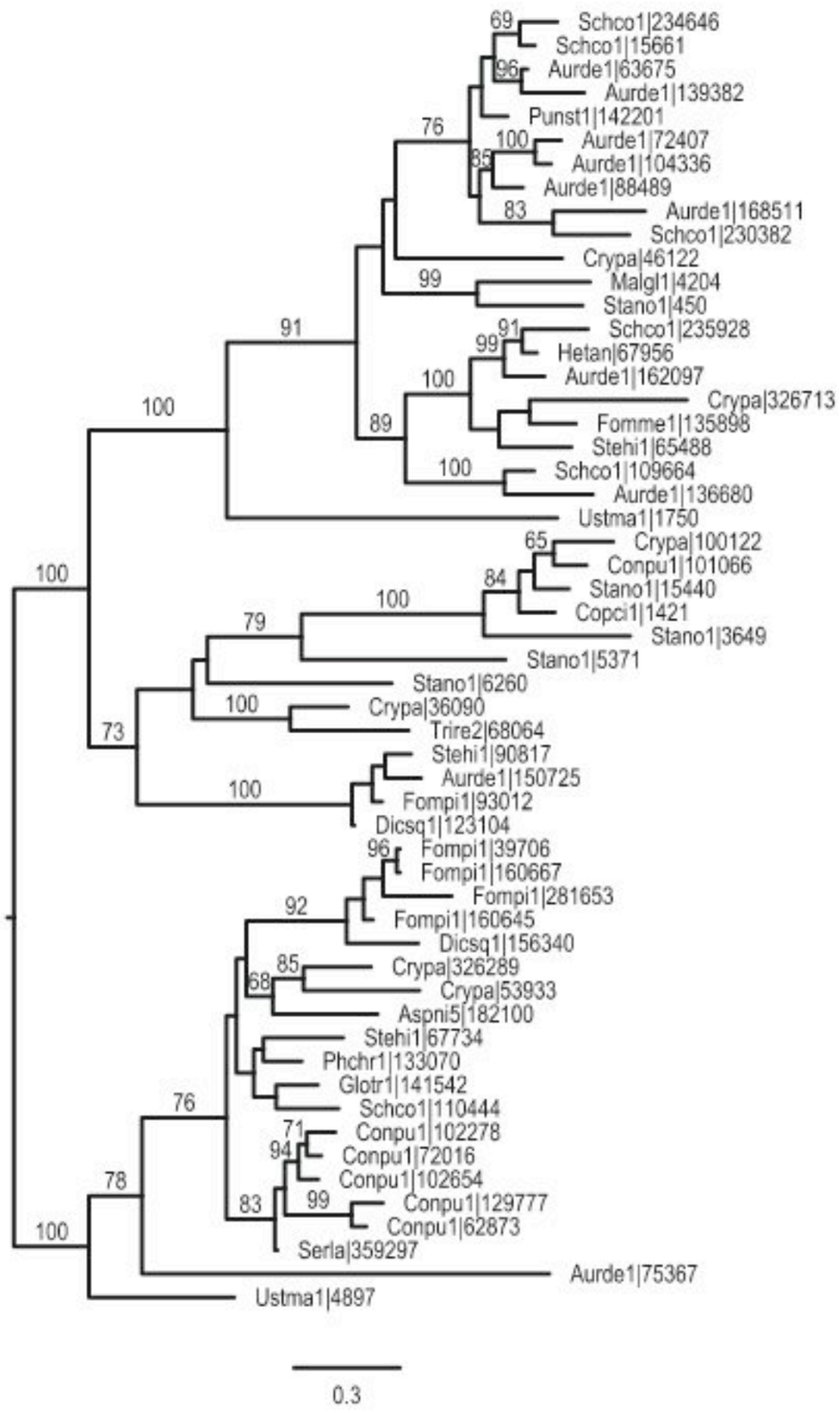


Fig. S8i. Carbohydrate Esterases family 5 (potential cutinases).

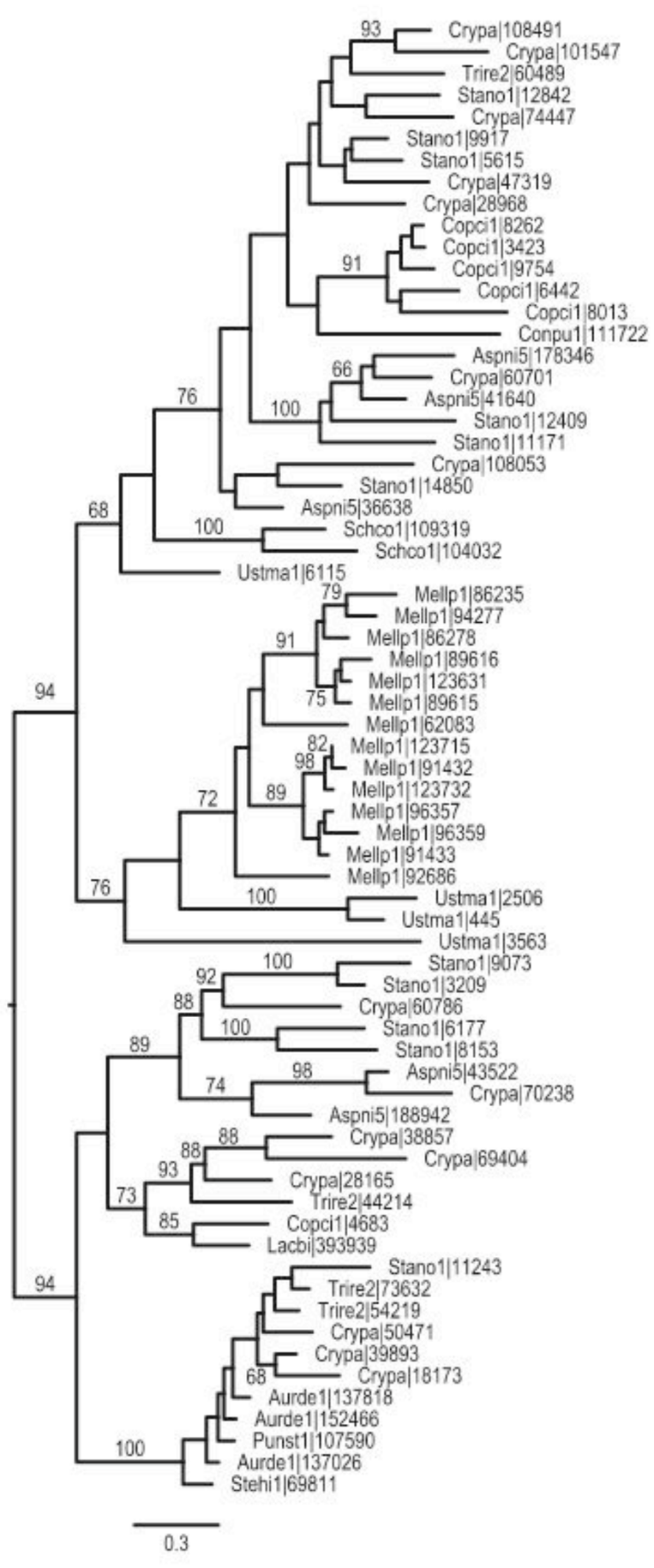
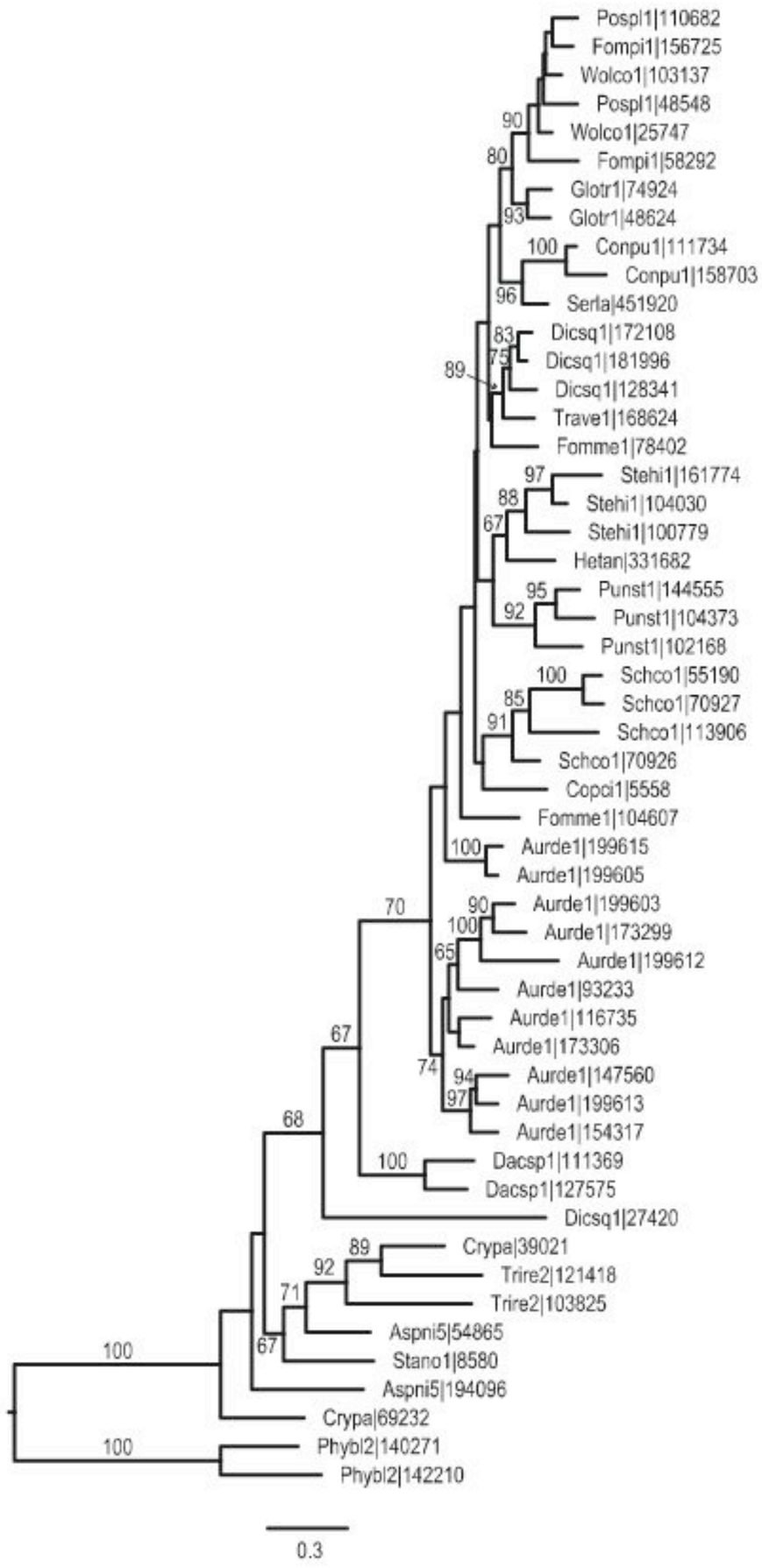


Fig. S8j. Carbohydrate Esterases family 16 group 1 (left) and group 2 (right).



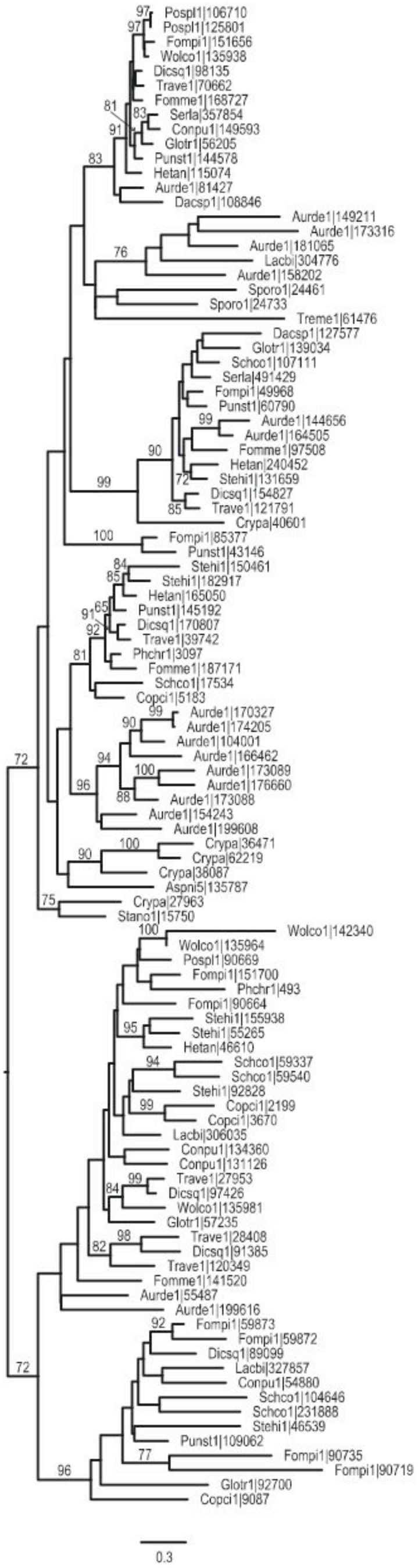


Fig. S8k. 'Glycoside Hydrolases' family 61 group 1 (left) and group 2 (right).

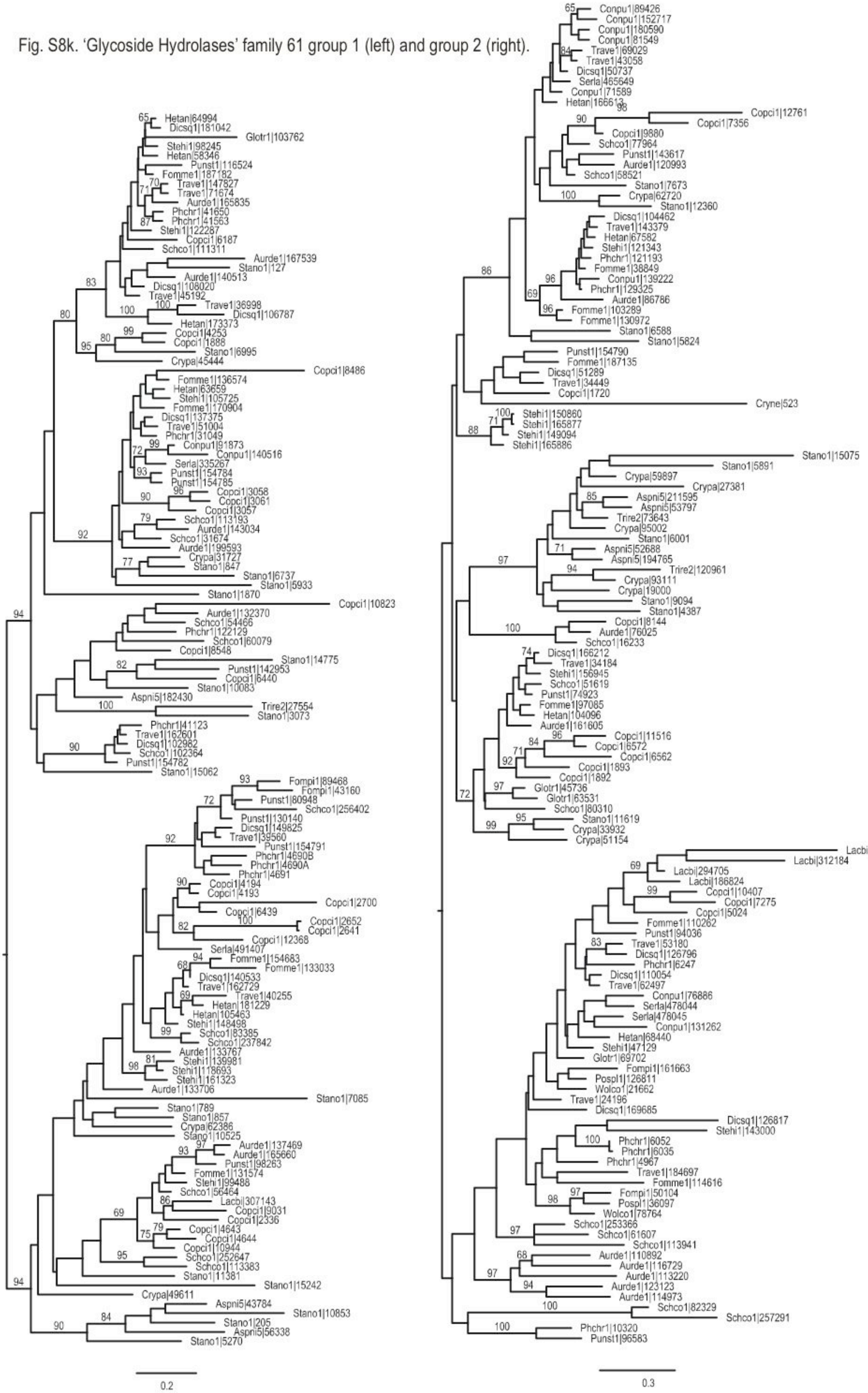


Fig. S8I. Class II peroxidases. Subclassification of the basidiomycete sequences is highlighted as LiP, MnP, VP, GP.

MnP: Manganese peroxidase
GP: generic peroxidase
LiP: Lignin peroxidase
VP: versatile peroxidase

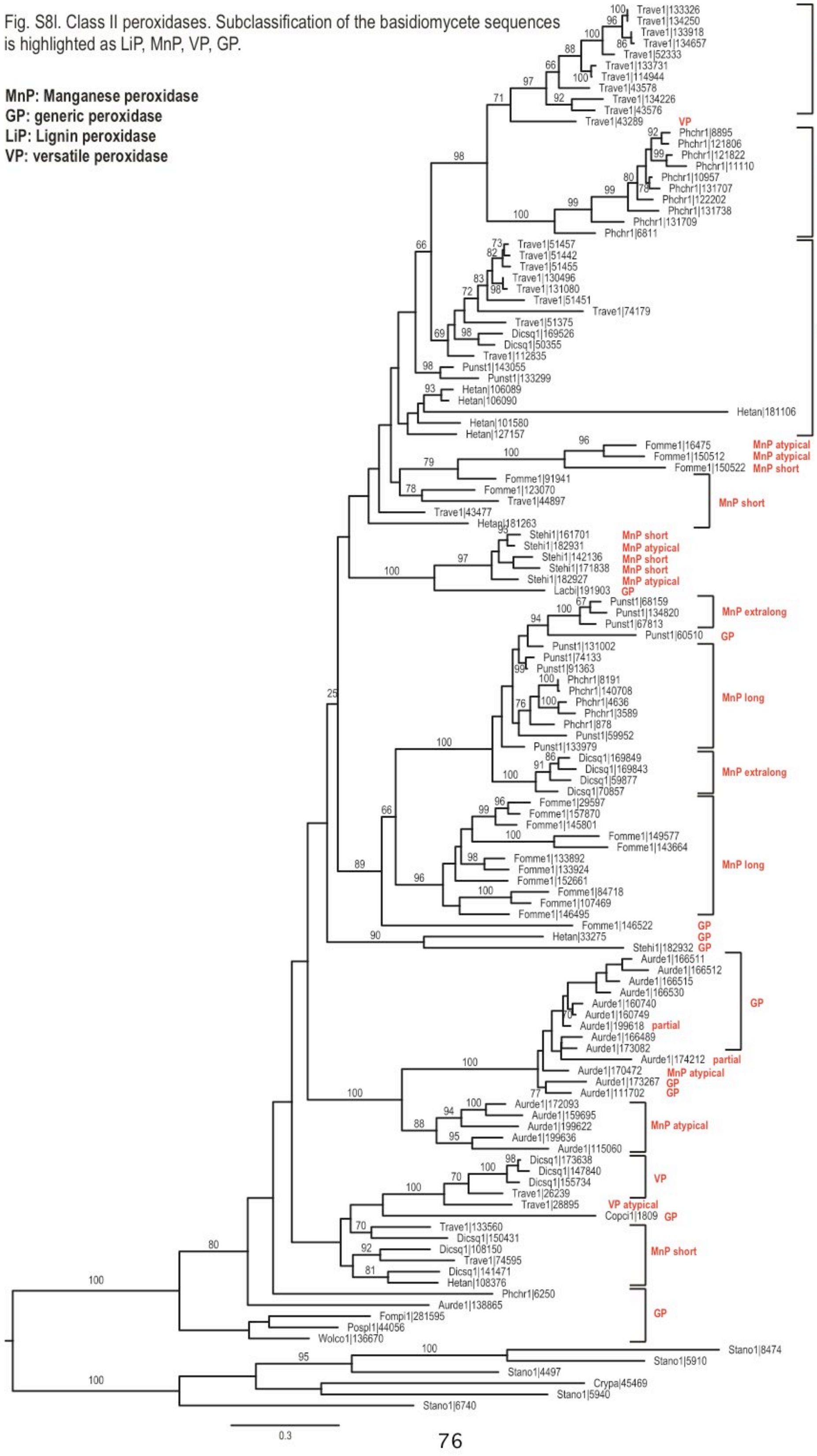
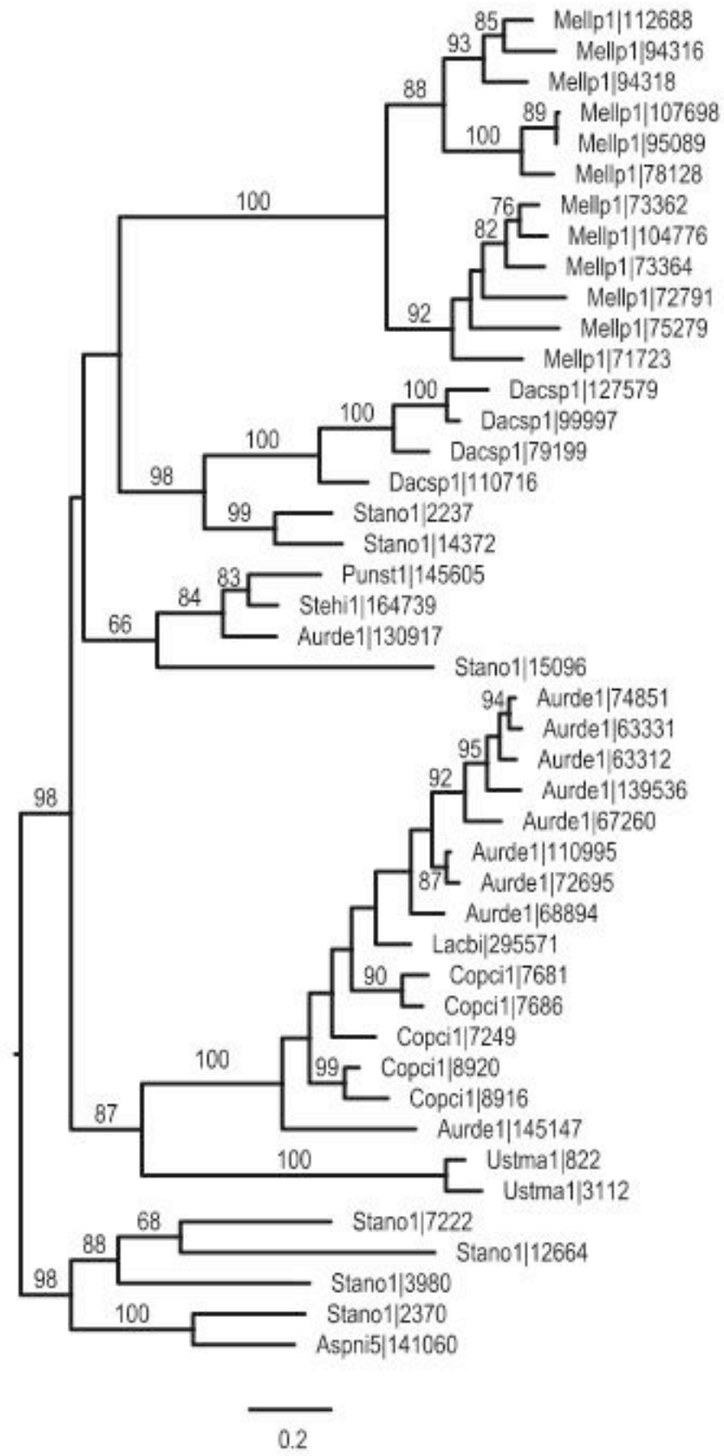
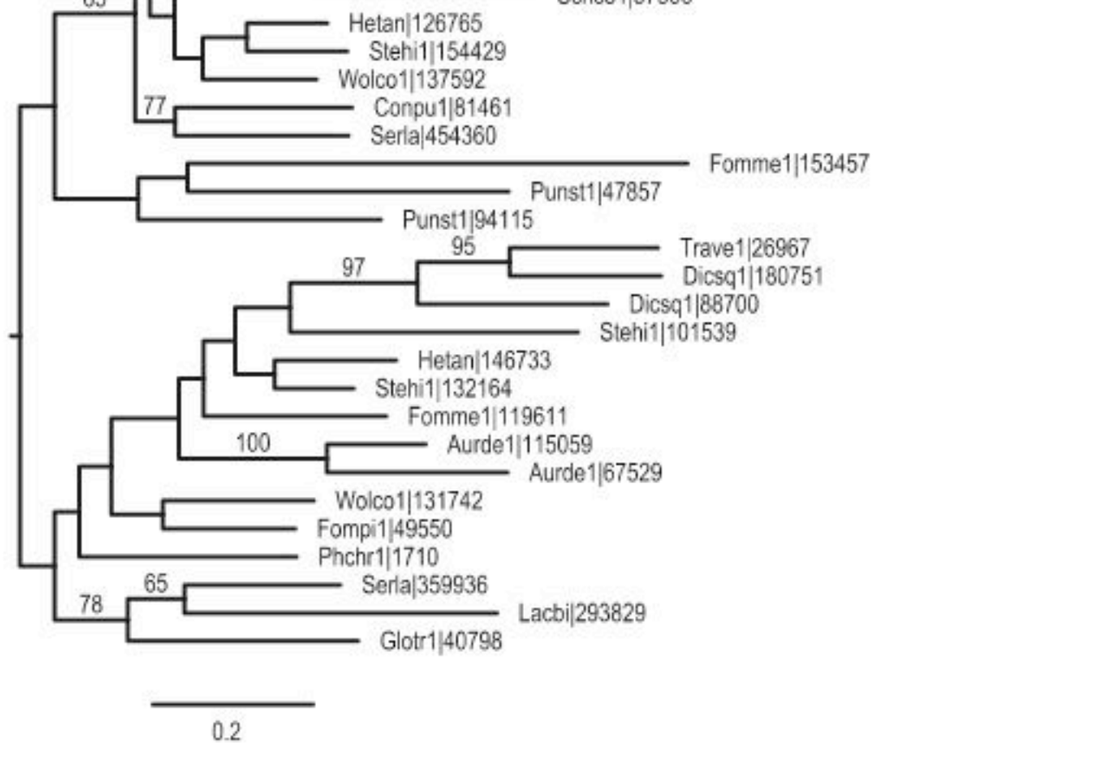
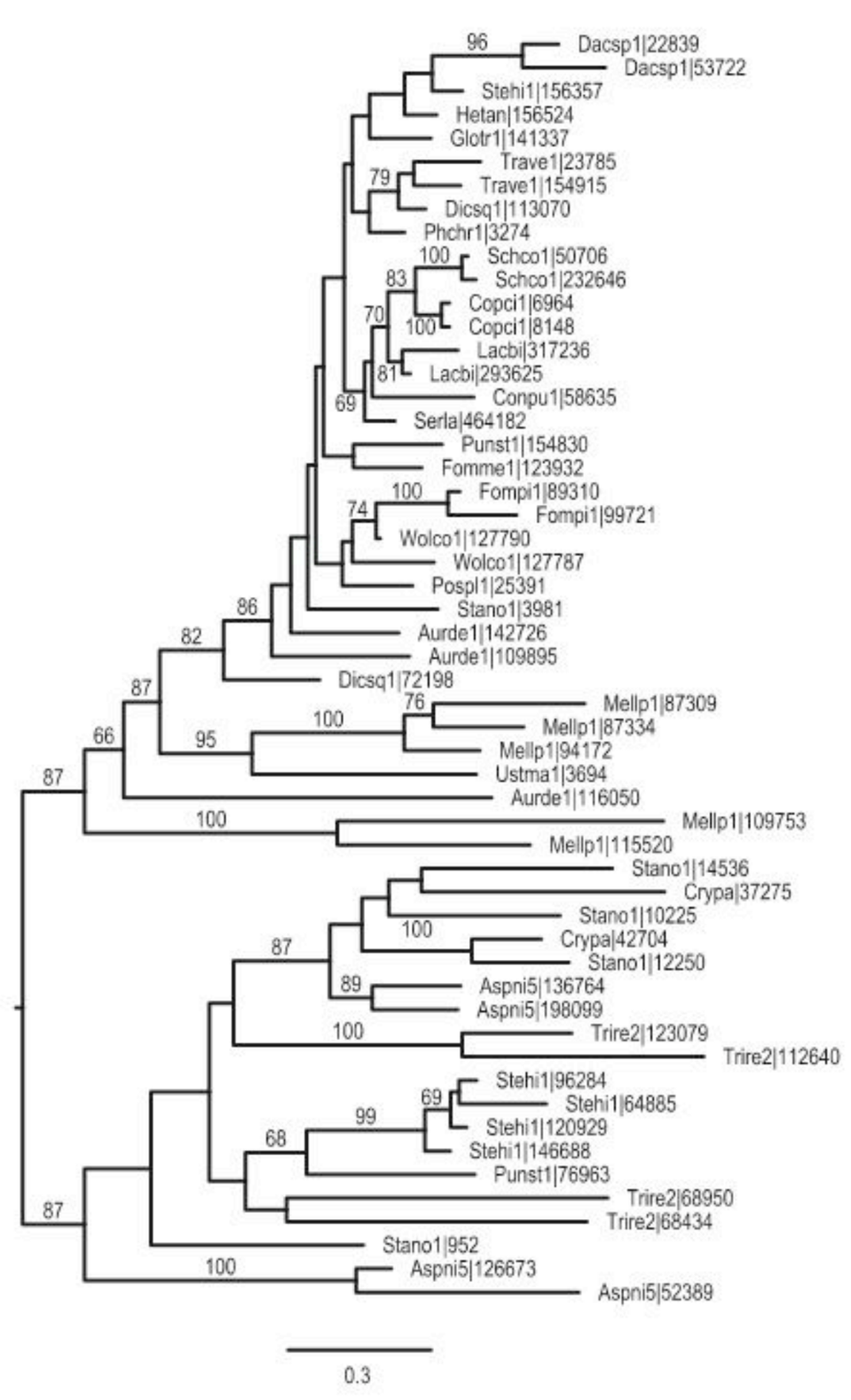


Fig. S8m. Heme Thiolate Peroxidases group 1 (upper left), group 2 (upper right), group 3 (lower left) and Dye Decolorizing Peroxidases (lower right).







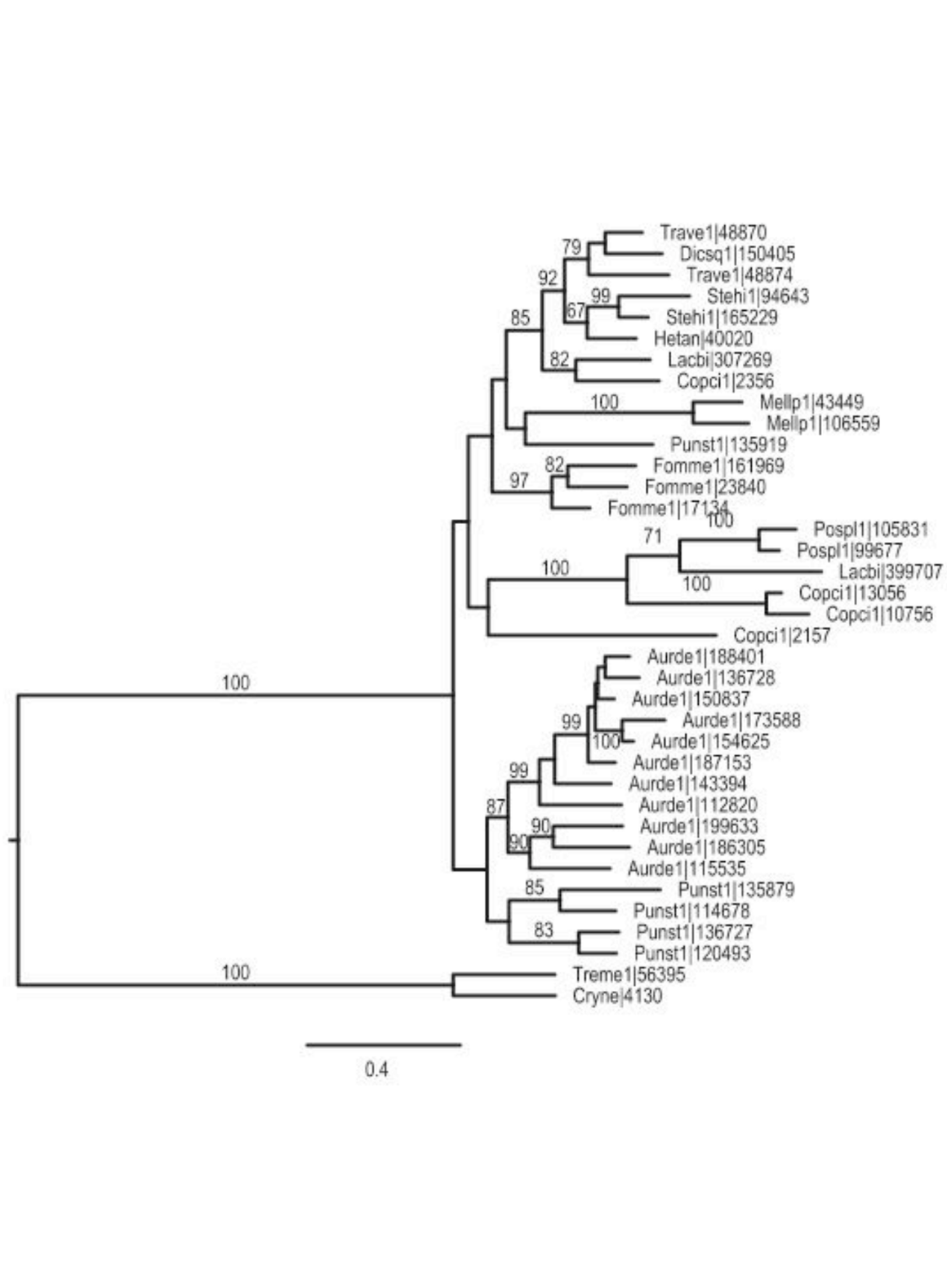


Fig. S8n. Copper radical oxidases 1, 2, 6 and glyoxal oxidases (major tree) and copper radical oxidases 3-5 (tree in frame). Protein models are indicated with green or brown if they are present in white rot species or brown rot species respectively. The glyoxal oxidases are absent from all the brown species sampled and present in *Auricularia delicata* which indicates the presence of the gene in the common ancestor of Agaricomycetes.

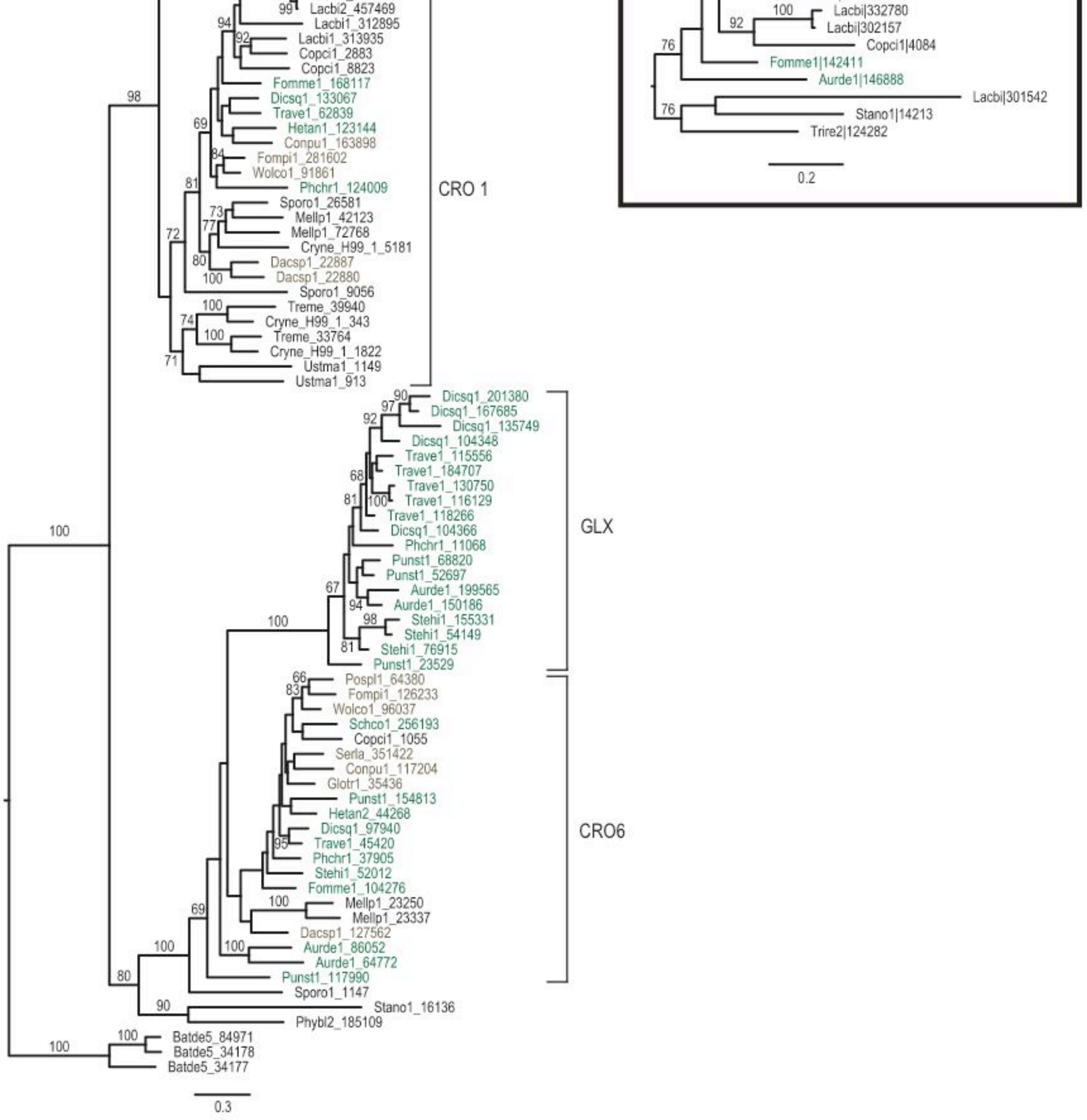


Fig. S8o. Basidiomycete and ascomycete laccases ‘sensu stricto’

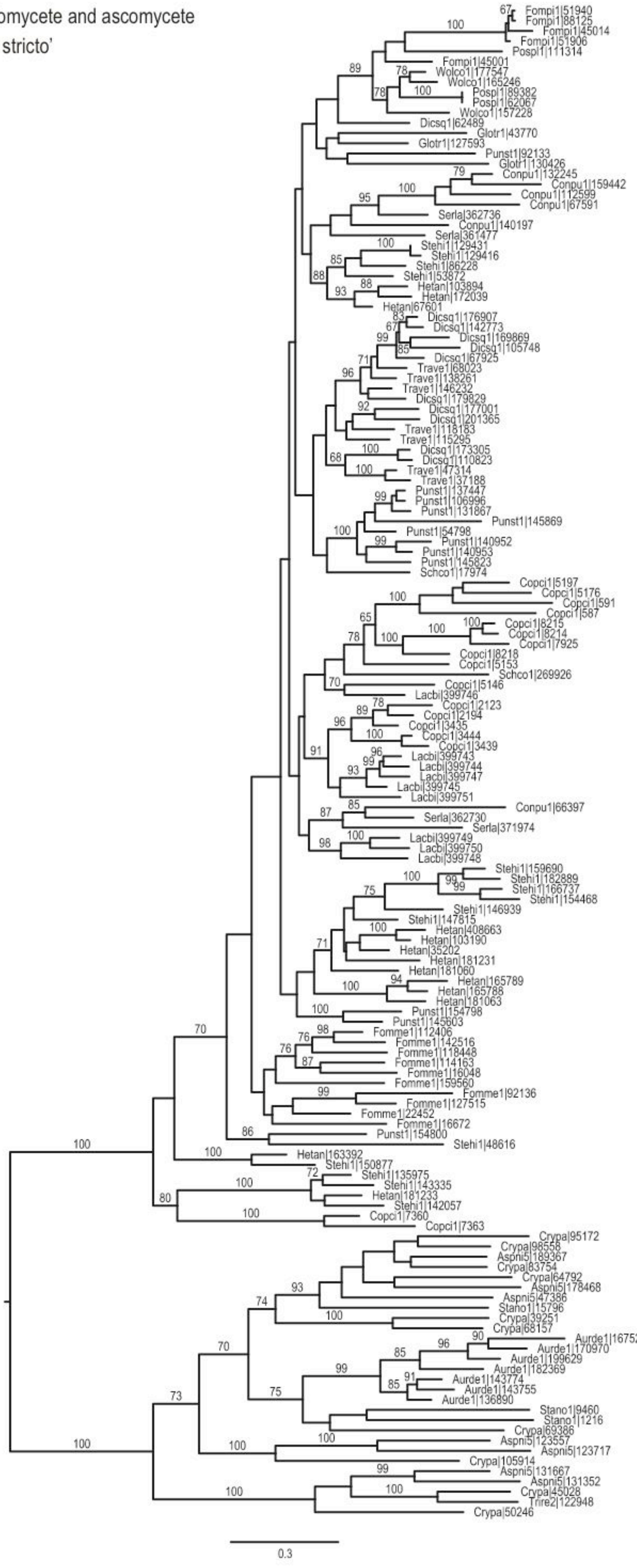


Fig. S8p. potential Fet 3 ferroxidases and ferroxidases/laccases
(L-ascorbate oxidases and fungal pigments MCOs are not analyzed here)

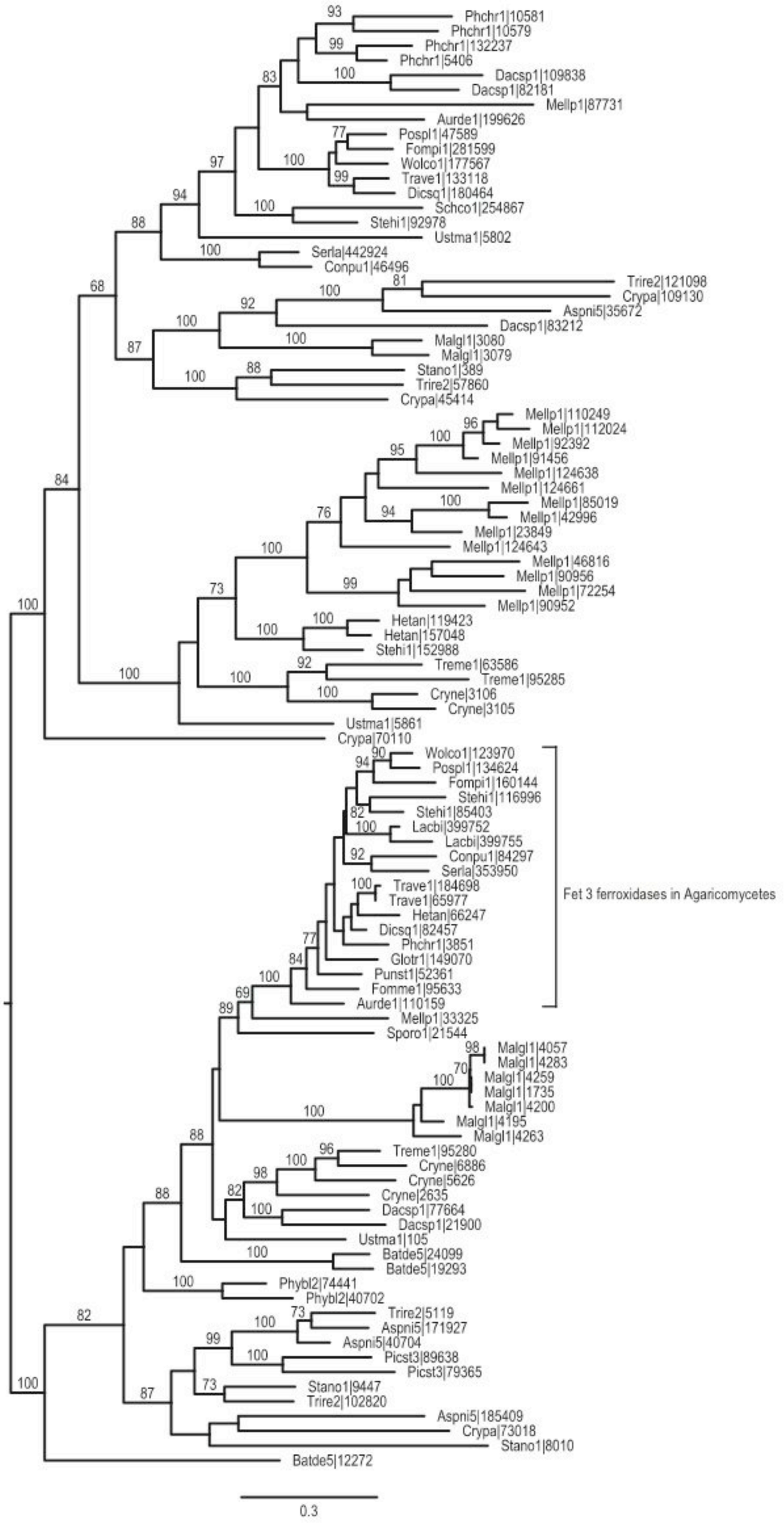


Fig S8q. Potential iron reducing glycopeptides (upper tree)
and potential cellobiose dehyrogenases (lower tree)

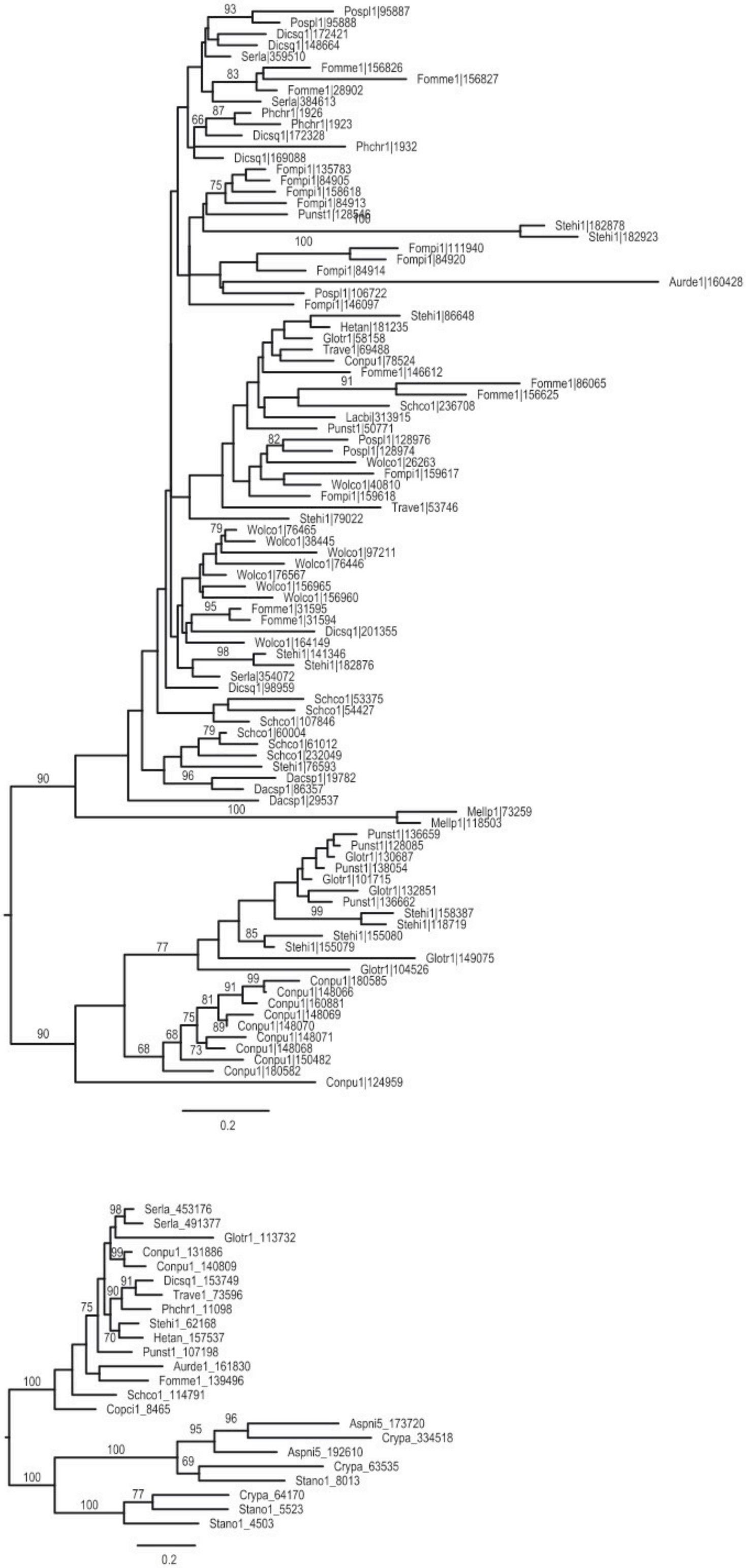


Fig. S9. Net changes in copy numbers in gene families estimated with Notung at EWT = 75%. Blue = net gains; Red = net losses. Families shown have significant family-wide p-values in CAFE analyses, except GH6 and CDH. Changes in terminal branches shown to right of taxon labels.

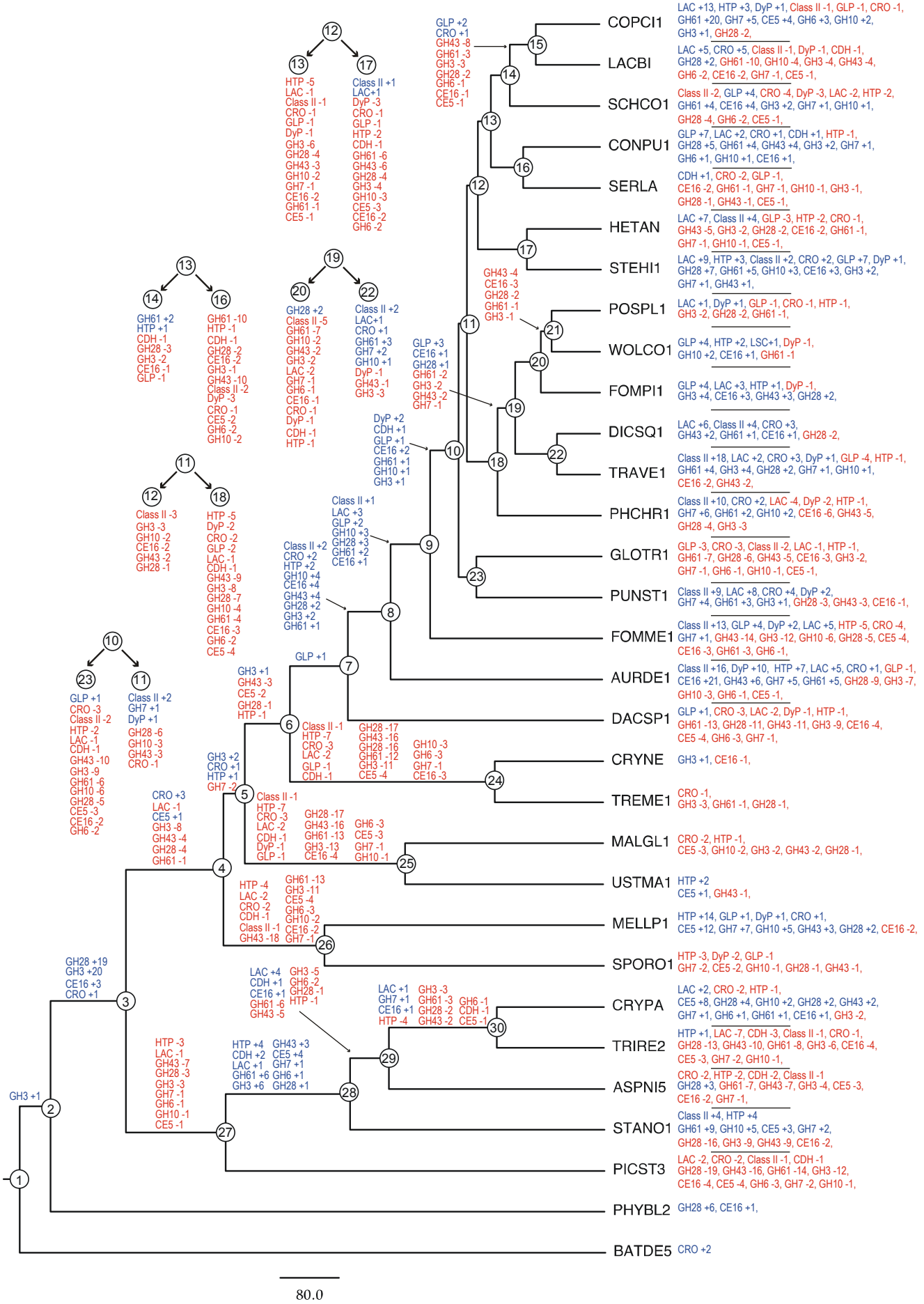
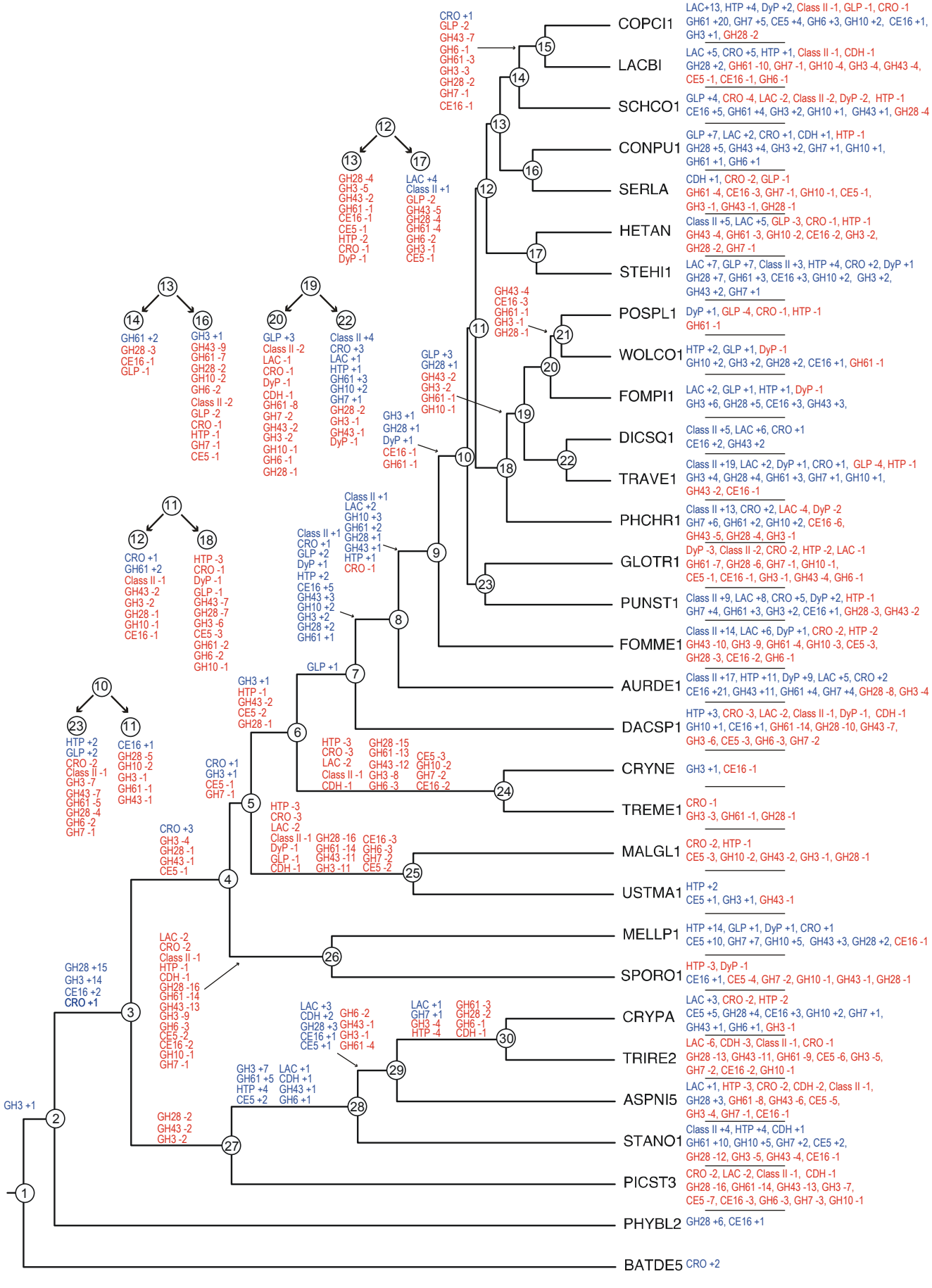


Fig. S10. Net changes in copy numbers in gene families estimated with Notung at EWT = 90%. Blue = net gains; Red = net losses. Families shown have significant family-wide p-values in CAFE analyses, except GH6 and CDH. Changes in terminal branches shown to right of taxon labels.



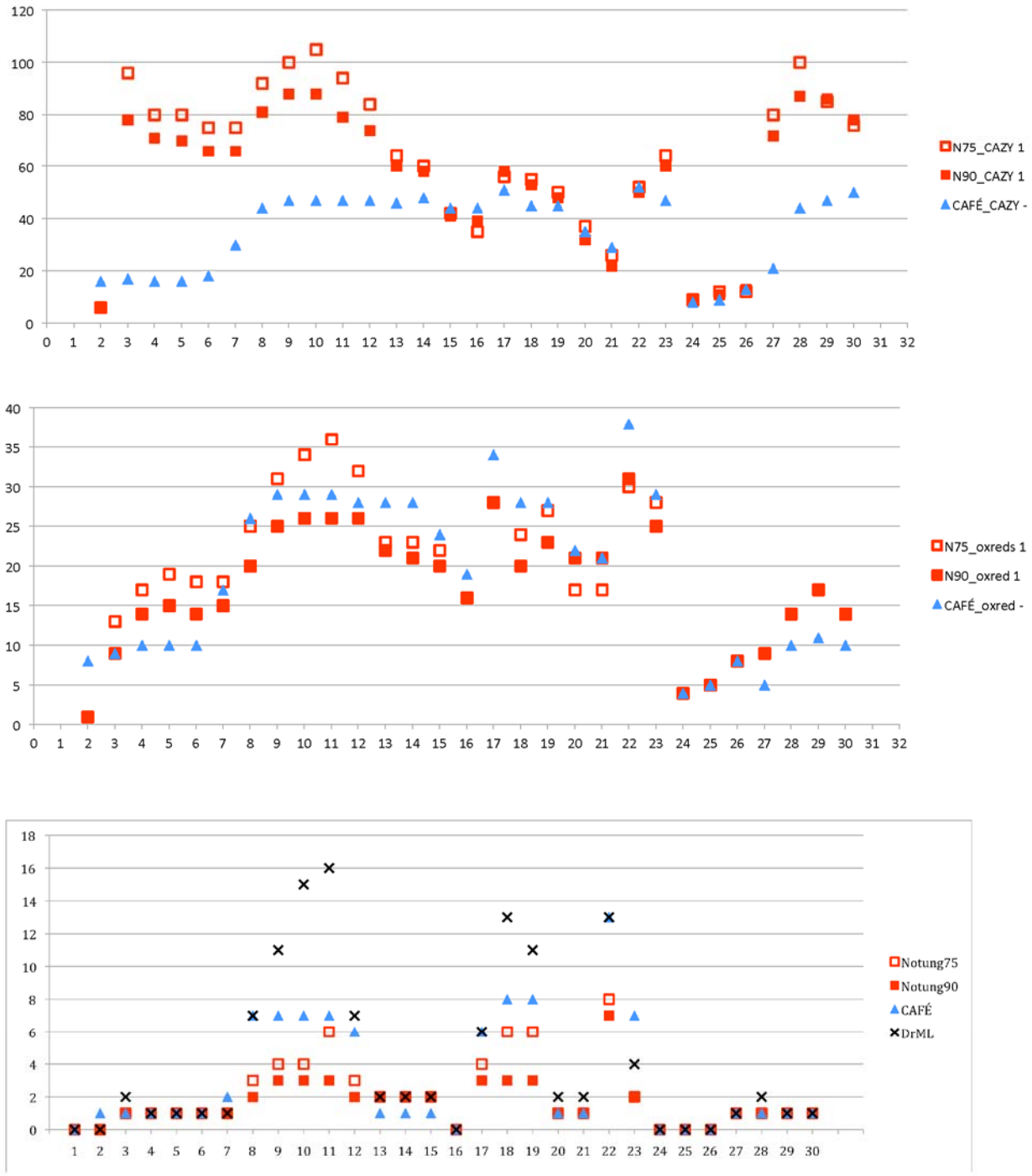


Fig. S11. Estimates of gene counts at internal nodes in the organismal phylogeny obtained with CAFE, Notung, and DrMLX-axis: internal node numbers, with numbering corresponding to Fig. 1 (main text): node 3 is the ancestor of Dikarya and node 8 is the common ancestor of Agaricomycetes. Y-axis: estimated gene copy number. Top: sums of gene counts for CAZyme gene families (CAFE and Notung only). Middle: sums of counts for oxidoreductase gene families (CAFE and Notung only). Bottom: counts for class II fungal peroxidases (including DrML).

Fig. S12. Reconciliation of class II fungal peroxidase phylogeny and species phylogeny with Notung at EWT = 75%. Subclassification of class II peroxidases is indicated.

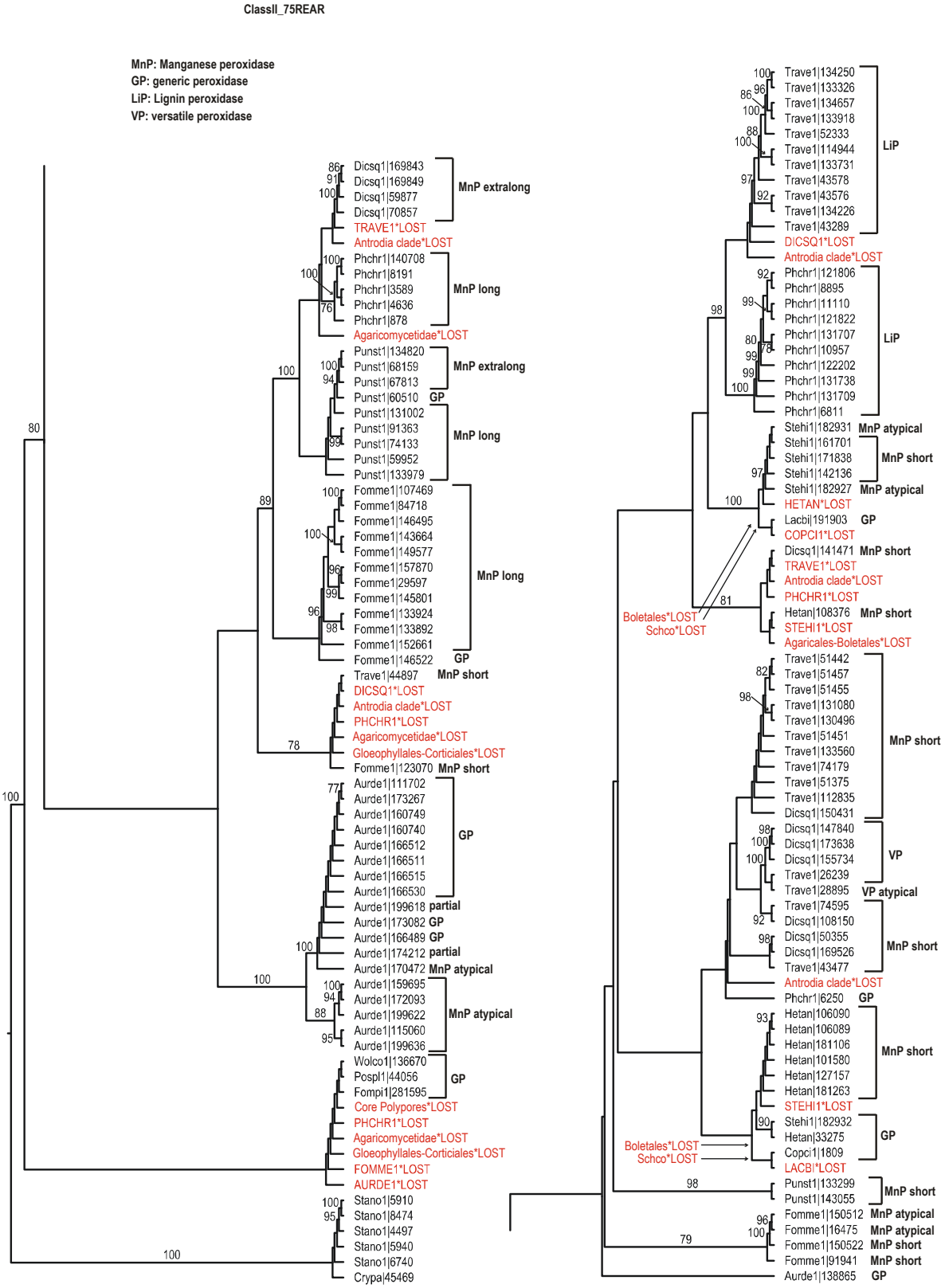


Fig. S13. Reconciliation of class II fungal peroxidase gene phylogeny (using bootstrap consensus tree) and species phylogeny inferred with DrML.

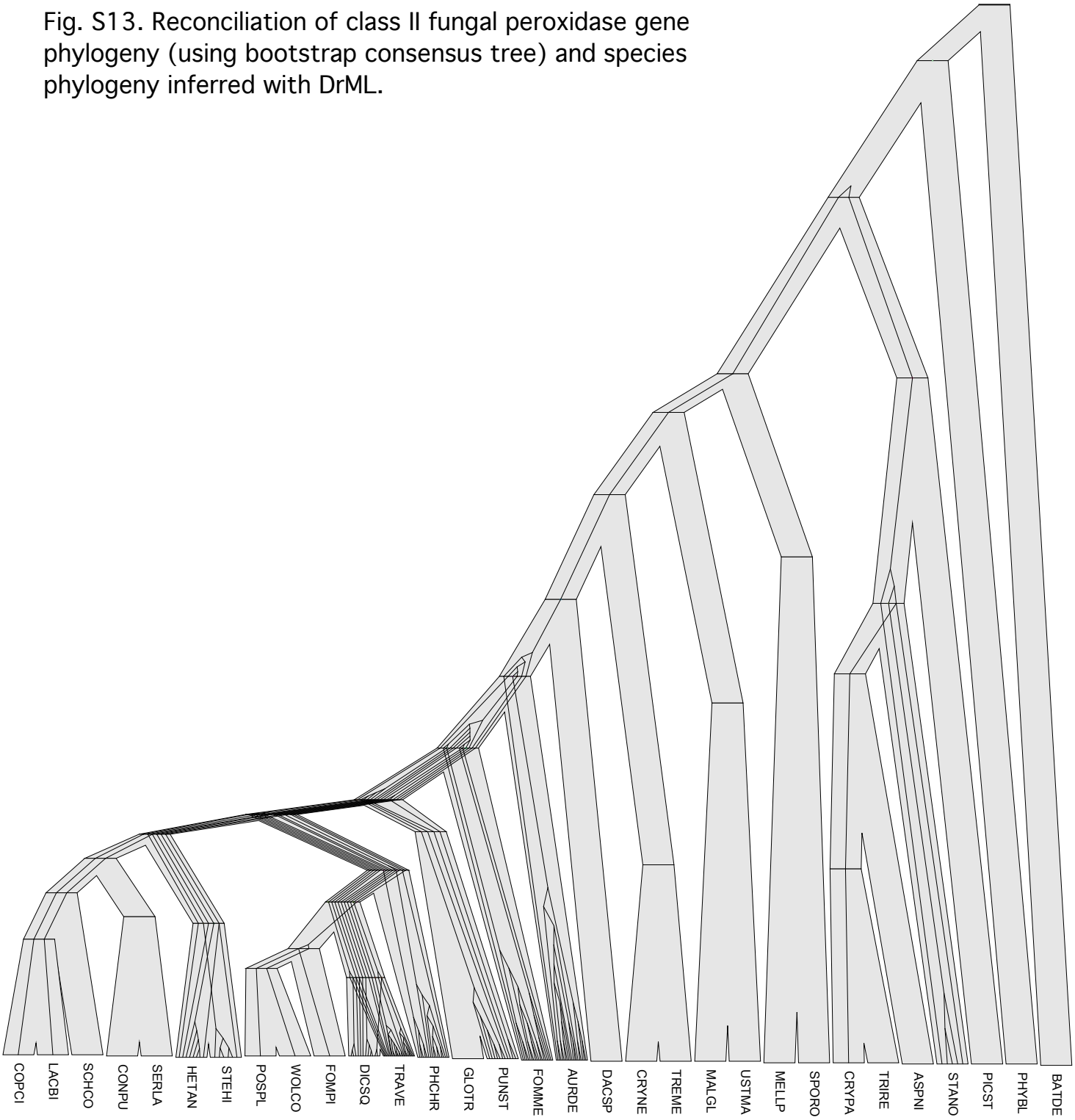


Fig. S14. Distributions of copy numbers of genes encoding class II fungal peroxidases inferred with DrML using 100 bootstrapped gene trees. Modal values are indicated.

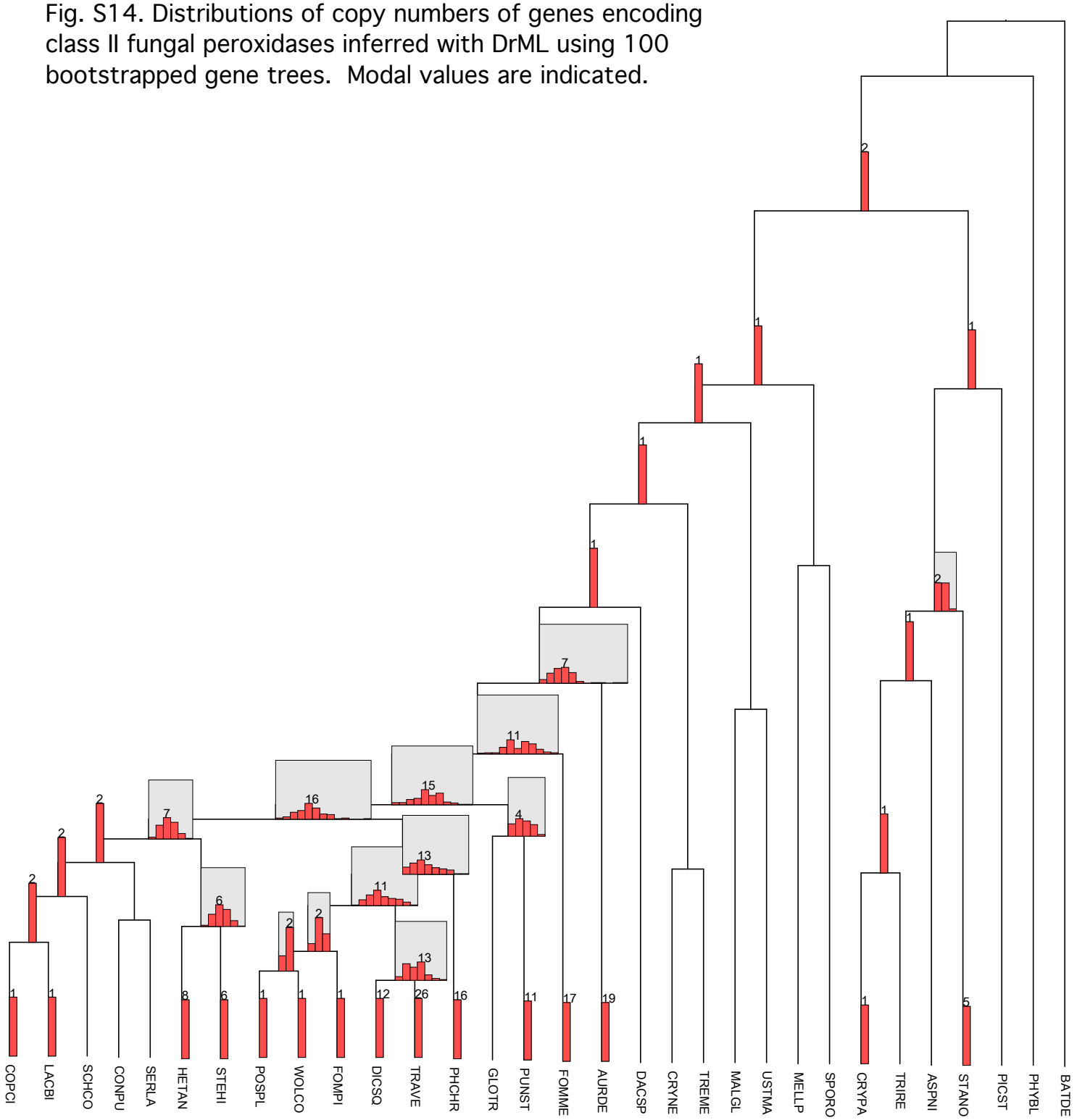


Fig. S15. Distributions of gene duplications (green histograms) and losses (yellow histograms) in class II fungal peroxidases inferred with DrML using 100 bootstrapped gene trees. Modal values are indicated.

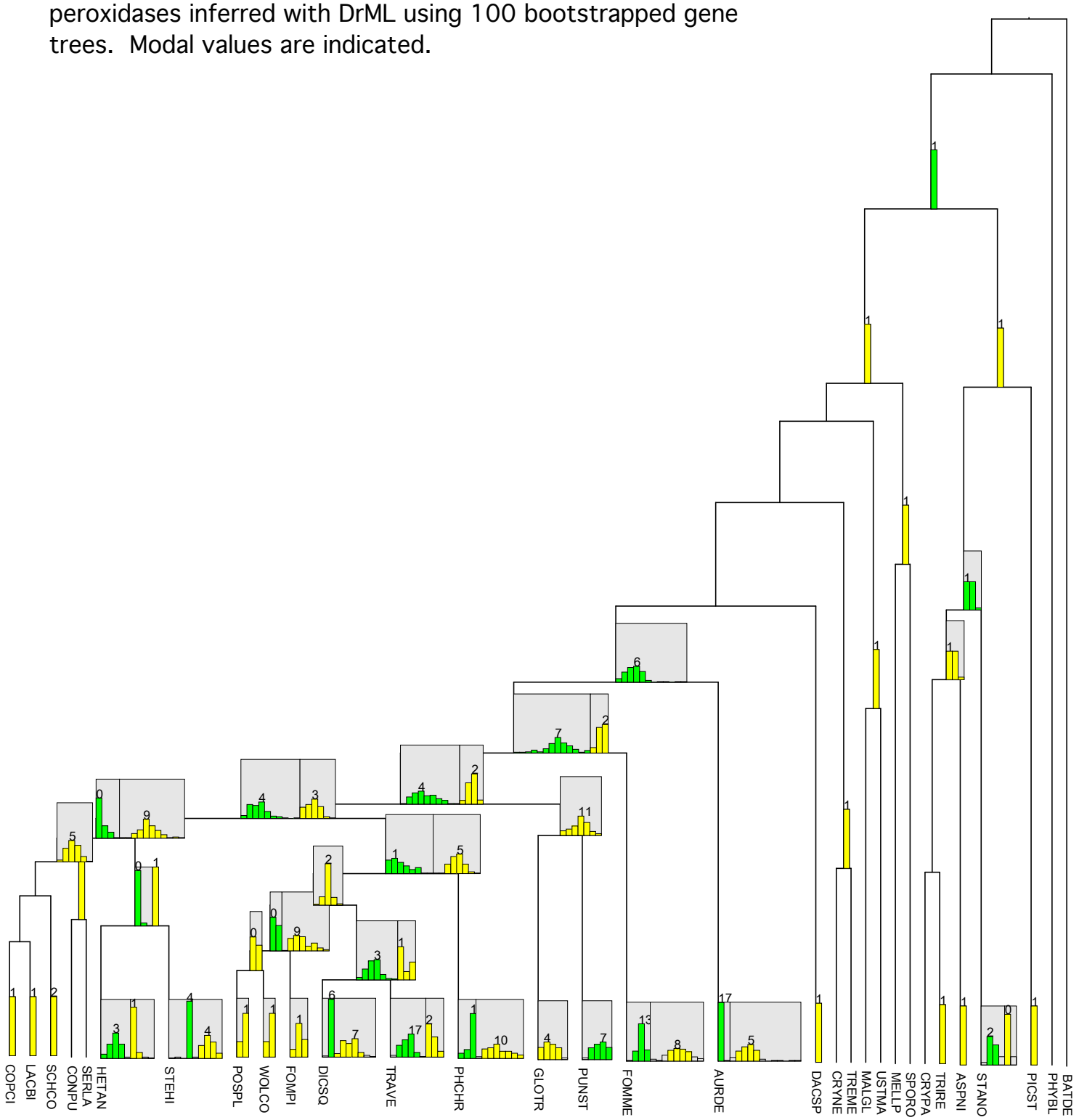


Fig. S16. Distributions of rate parameters for class II peroxidase evolution estimated with binary state speciation analysis. Top: rates of gene duplication (=speciation) in white rot (blue) vs. brown rot (red) lineages. Middle: rates of gene loss (=extinction). Bottom: rates of transformation from white rot to brown rot (red) and vice versa (blue).

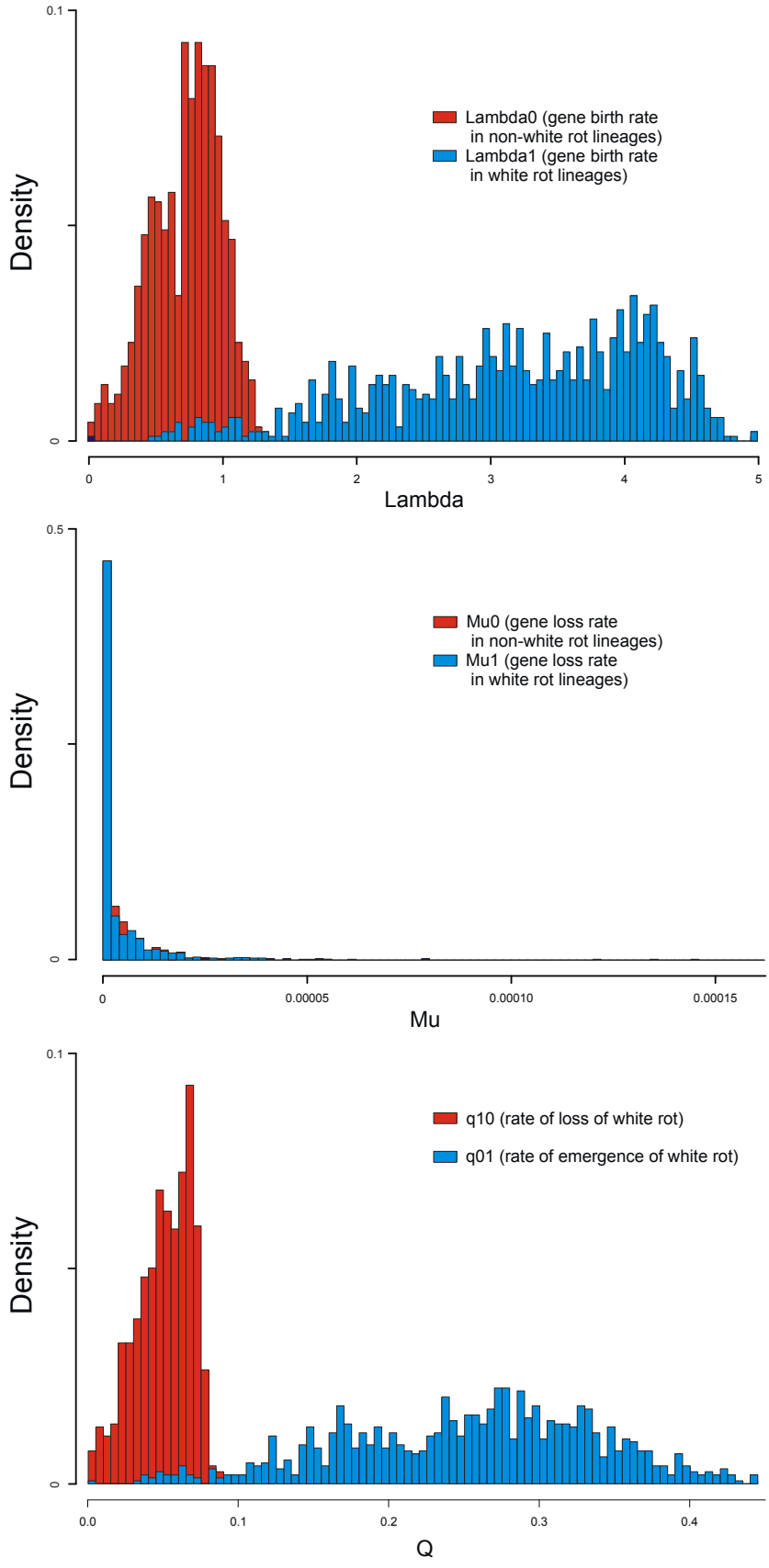


Fig. S17. Number and proportion of selected CAZyme genes including cellulose binding module (CBM1). Left graph: total number of genes in families GH6, GH7, GH10, GH5, GH61, GH74, CE1, and CE15 (blue) and numbers of genes containing CBM1. Right graph: proportion of genes in each family containing CBM1.

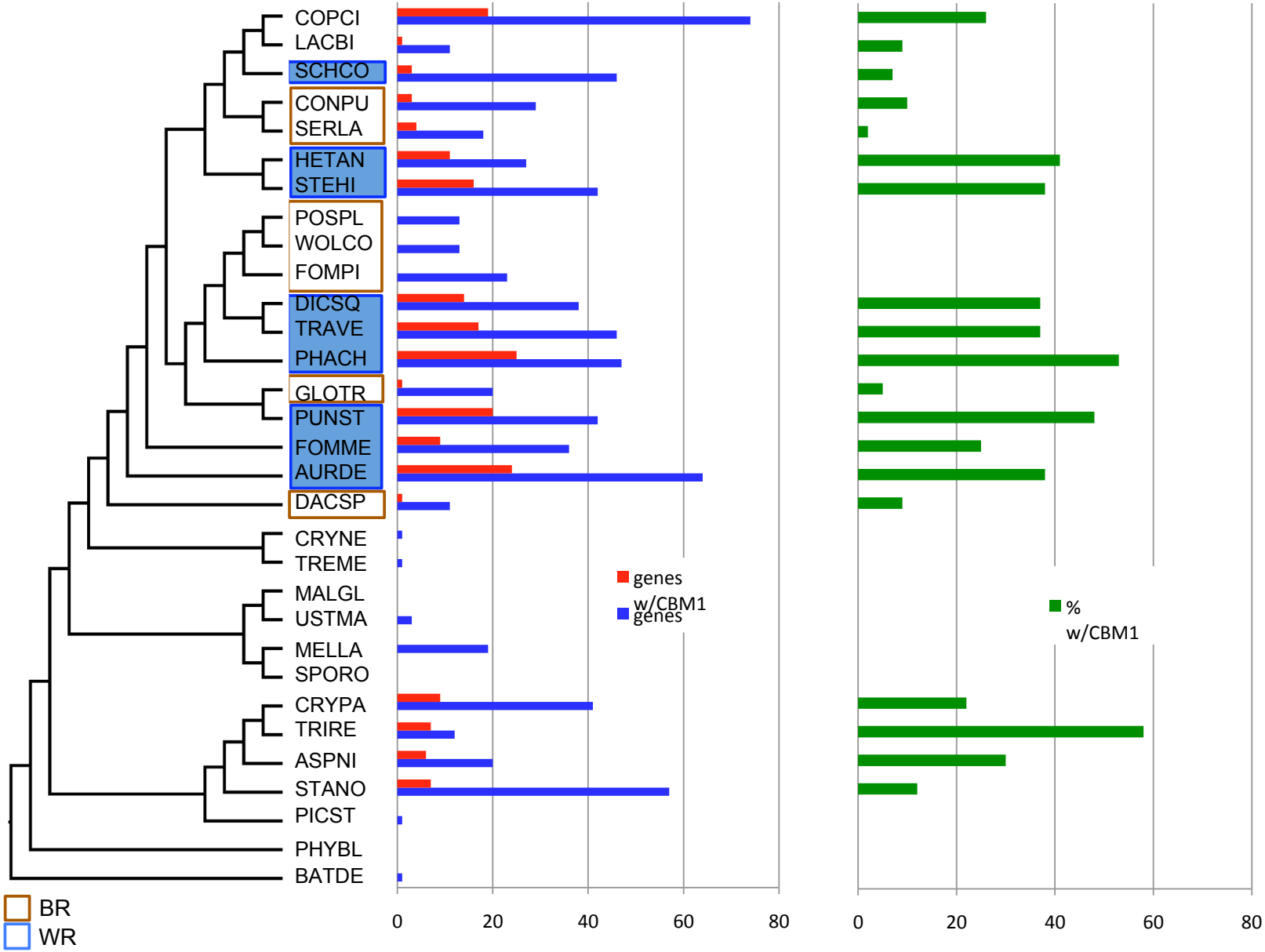
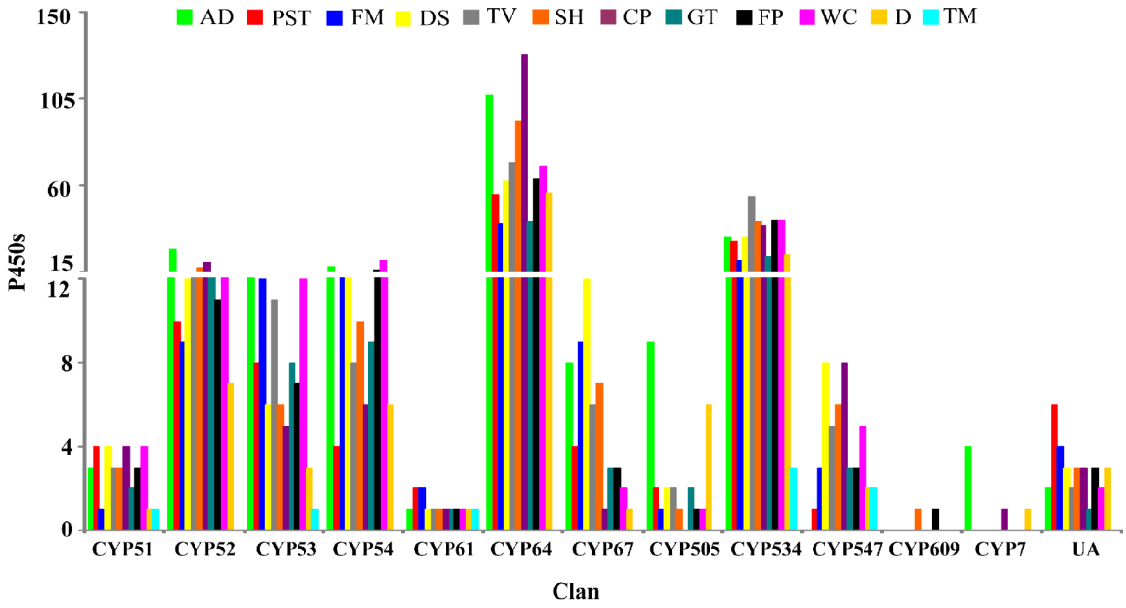


Fig. S18. P450ome clan-level classification in twelve Agaricomycotina genomes. X-axis represents the individual P450 clans. Y-axis represents member P450s in each clan. UA stands for Unassigned clan. AD, *Auricularia delicata*; PST, *Punctularia strigoso-zonata*; FM, *Fomitiporia mediterranea*; DS, *Dichomitus squalens*; TV, *Trametes versicolor*; SH, *Stereum hirsutum*; CP, *Coniophora puteana*; GT, *Gloeophyllum trabeum*; FP, *Fomitopsis pinicola*; WC, *Wolfiporia cocos*; D, *Dacryopinax sp*; and TM, *T. mesenterica*.



(next page)

Fig. S19. Chromosomal clustering of genes encoding class II fungal peroxidases. Maximum Likelihood phylogeny of class II peroxidase genes in 31 SAP fungi, showing the gene clusters with which they are associated. Gene families indicated in the legend correspond to protein alignments described elsewhere (see section 5. Analyses of decay-related gene families). Scale bar in the legend refers to nucleotide distances in kilobase pairs in gene cluster figures. Lines connect protein names to the corresponding clustered genes. Colors of sequence names and lines were arbitrarily chosen to distinguish between clusters from different genomes.

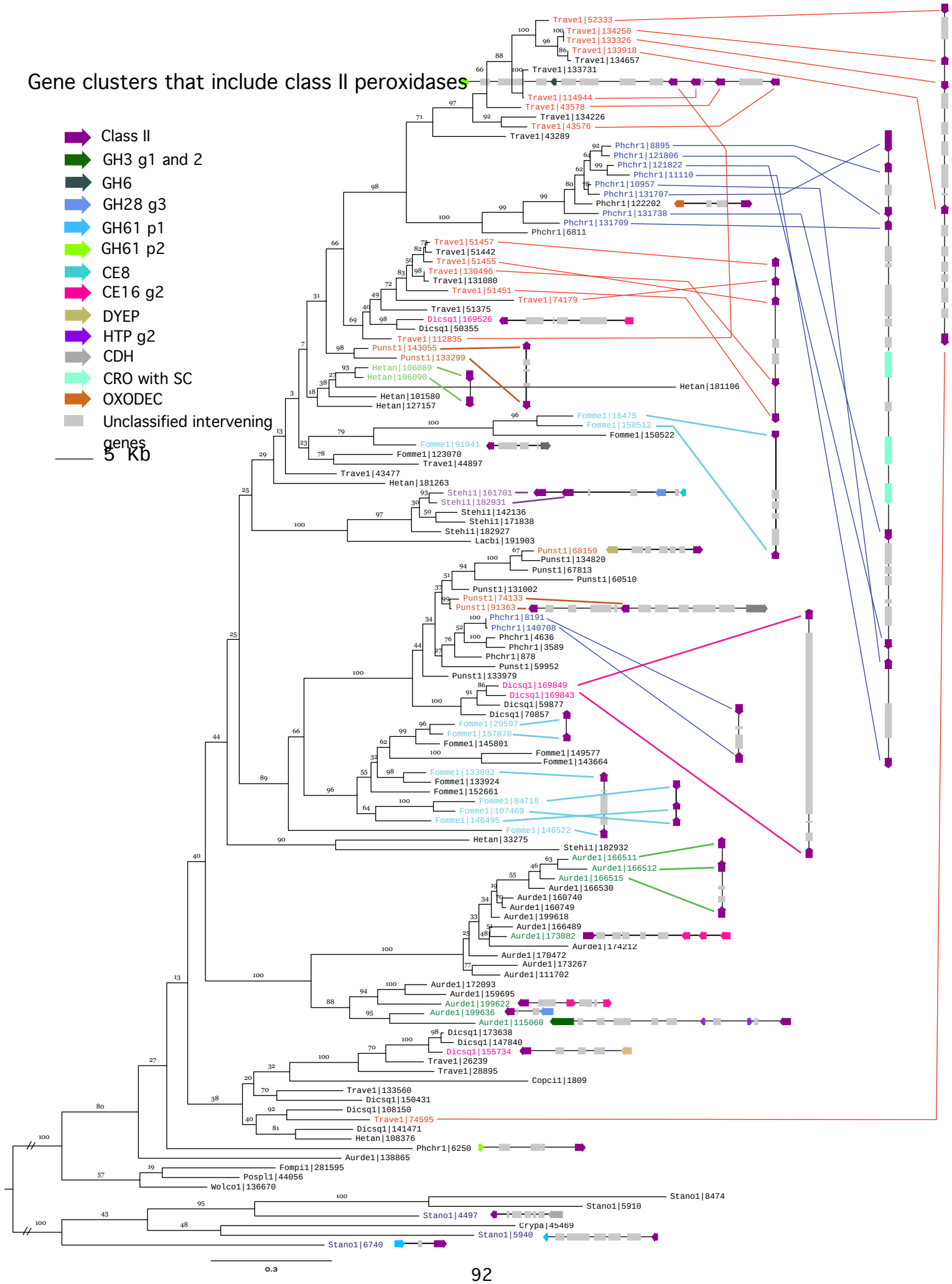

Gene clusters that include class II peroxidases
Class II
GH3 g1 and 2
GH6
GH28 g3
GH61 p1
GH61 p2
CE8
CE16 g2
DYEP
HTP g2
CDH
CRO with SC
OXODEC
Unclassified intervening genes
5 Kb
Trave1|52333
Trave1|134250
Trave1|133326
Trave1|133918
Trave1|134657
Trave1|133731
Trave1|114944
Trave1|43578
Trave1|134226
Trave1|43576
Trave1|43289
Phchr1|8895
Phchr1|121806
Phchr1|121822
Phchr1|11110
Phchr1|10957
Phchr1|131707
Phchr1|122202
Phchr1|131738
Phchr1|131709
Phchr1|6811
Trave1|51457
Trave1|51442
Trave1|51455
Trave1|130496
Trave1|131080
Trave1|51451
Trave1|74179
Trave1|51375
Dicsq1|169526
Dicsq1|50355
Trave1|112835
Punst1|143055
Punst1|133299
Hetan|106089
Hetan|106090
Hetan|181106
Hetan|101580
Hetan|127157
Fomme1|16475
Fomme1|150512
Fomme1|150522
Fomme1|91941
Fomme1|123070
Trave1|44897
Trave1|43477
Hetan|181263
Stehi1|161701
Stehi1|182931
Stehi1|142136
Stehi1|171838
Stehi1|182927
Lacbi|191903
Punst1|68159
Punst1|134820
Punst1|67813
Punst1|60510
Punst1|131002
Punst1|74133
Punst1|91363
Phchr1|8191
Phchr1|140708
Phchr1|4636
Phchr1|3589
Phchr1|878
Punst1|59952
Punst1|133979
Dicsq1|169849
Dicsq1|169843
Dicsq1|59877
Dicsq1|70857
Fomme1|29597
Fomme1|157870
Fomme1|145801
Fomme1|149577
Fomme1|143664
Fomme1|133892
Fomme1|133924
Fomme1|152661
Fomme1|84718
Fomme1|107469
Fomme1|146495
Fomme1|146522
Hetan|33275
Stehi1|182932
Aurde1|166511
Aurde1|166512
Aurde1|166515
Aurde1|166530
Aurde1|160740
Aurde1|160749
Aurde1|199618
Aurde1|166489
Aurde1|173082
Aurde1|174212
Aurde1|170472
Aurde1|173267
Aurde1|111702
Aurde1|172093
Aurde1|159695
Aurde1|199622
Aurde1|199636
Aurde1|115060
Dicsq1|173638
Dicsq1|147840
Dicsq1|155734
Trave1|26239
Trave1|28895
Copci1|1809
Trave1|133560
Dicsq1|150431
Dicsq1|108150
Trave1|74595
Dicsq1|141471
Hetan|108376
Phchr1|6250
Aurde1|138865
Fompi1|281595
Pospl1|44056
Wolco1|136670
Stano1|8474
Stano1|5910
Stano1|4497
Crypa|45469
Stano1|5940
Stano1|6740
0.3

Fig. S20. Micromorphological onservations of wood decay. All sections are transverse sections. The dotted bar on lower right is the magnification and microns are listed on the lower extreme right. a-h are white rot species; i-l are brown rot species. The decay caused by each species is illustrated on aspen (*Populus* sp.) or pine (*Pinus* sp.).

Fig. S20a. *Auricularia delicata,* pine. Cell wall erosion of the tracheids consisting of a thinning of the cell wall from the cell lumen toward the middle lamella was obserced in some cells and a complete degradation of all cell wall layers in others. Attack of all cell wall components (lignin, cellulose and hemicellulose) is typical of a white rot pattern of decay.

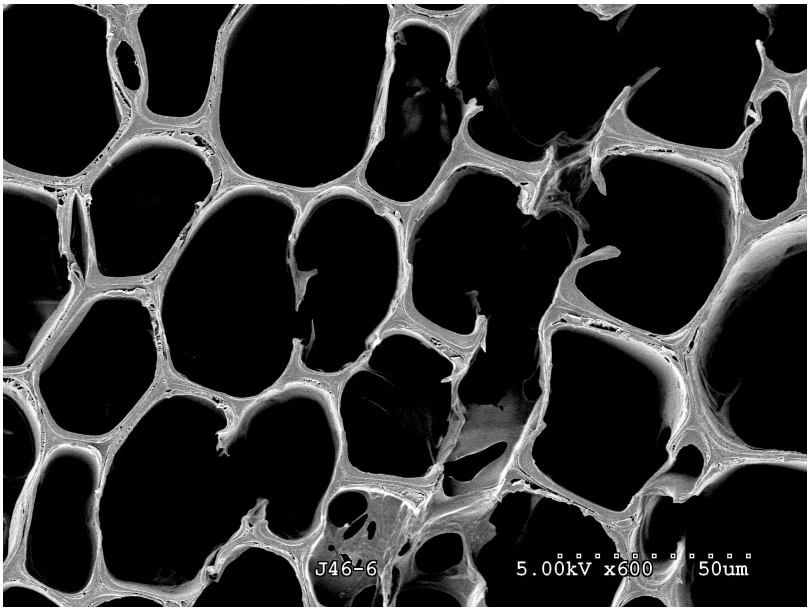


Fig. S20b. *Fomitiporia mediterranea*, aspen. A localized white rot with groups of cells extensively degraded. Cell walls that were eroded and thinned have separated and collapsed creating voids in the wood.

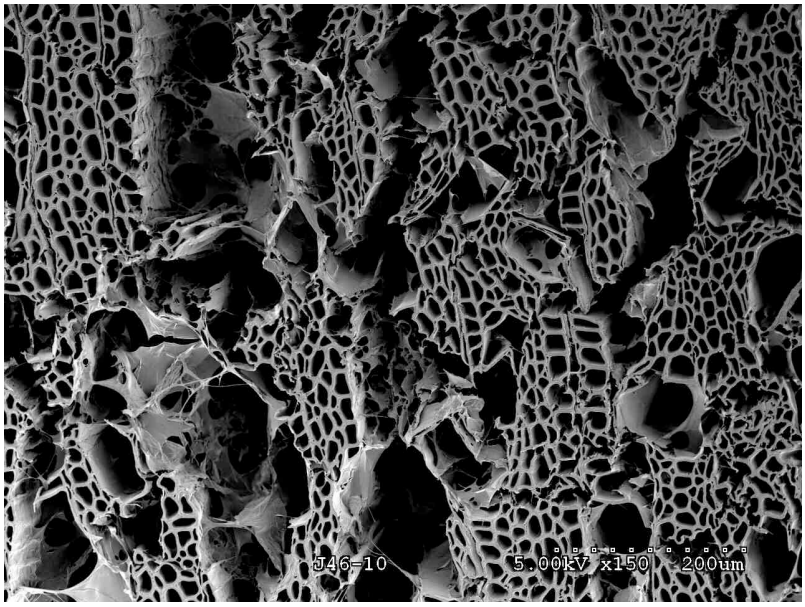


Fig. S20c. *Punctularia strigoso-zonata*, aspen. This fungus caused an aggressive white rot with complete degradation of all cells in some areas of the wood. Large voids were produced where all fibers, parnechyma and vessel elements were degraded.

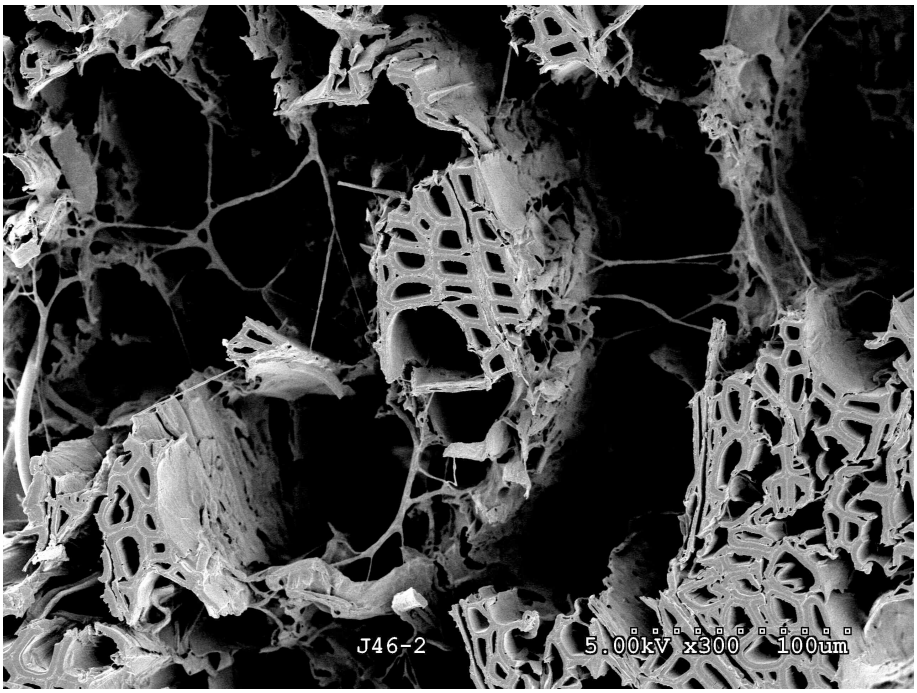


Fig. S20d. *Stereum hisrutum*, pine. Advanced stages of white rot with thinning of tracheid walls and some completely eroded cells.

Fig. S20e. *Dacryopinax* sp., aspen. Brown rot degradation in all wood cells. A diffuse attack causing a depolymerization of the cellulose occurred throughout the wood. Weakened cells do not maintain strength and appear to have a wavy appearance or are collapsed.

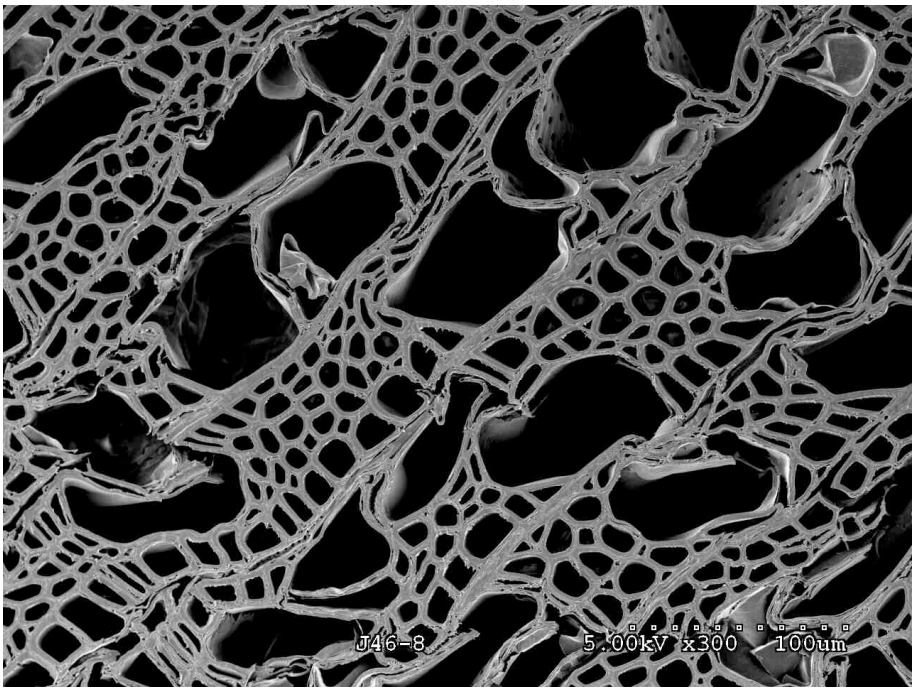


Fig. S20f. *Gloeophyllum trabeum*, pine. Extensive brown rot in tracheids resulting in a diffuse attack of all cell wall layers. In contrast to white rot, cells are not eroded or completely degraded but depolymerization and degradation of the cellulose causes a weakening of the trachieds and they become distorted. The altered cell walls appear somewhat swollen with indistinct cell wall layers and cells bend and collapse.

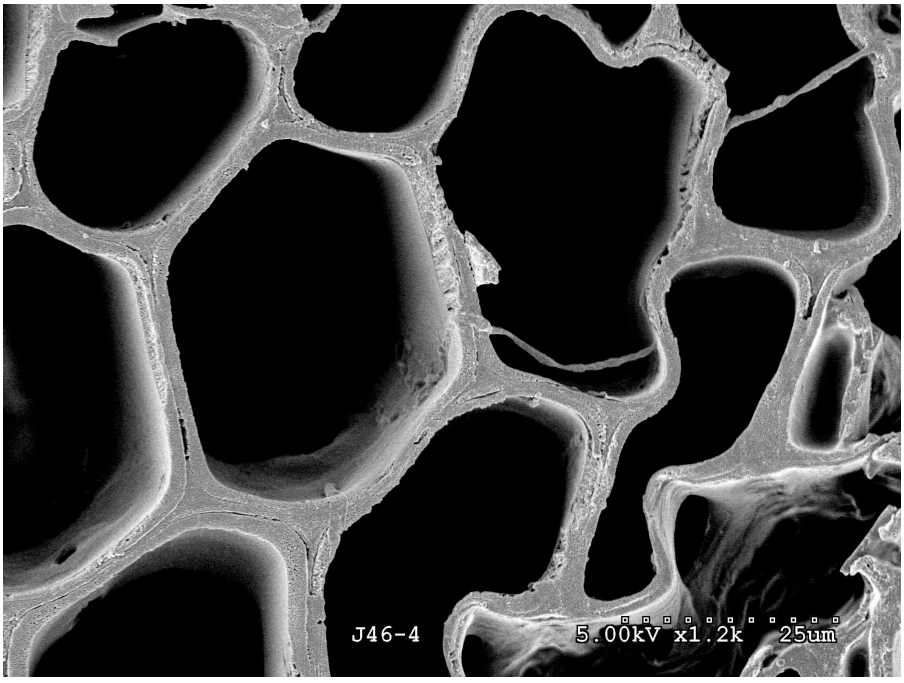


Fig. S21. Symbiosis genes up-regulated in *Laccaria bicolor* and their homologs in saprotophic and pathogenic Fungi.

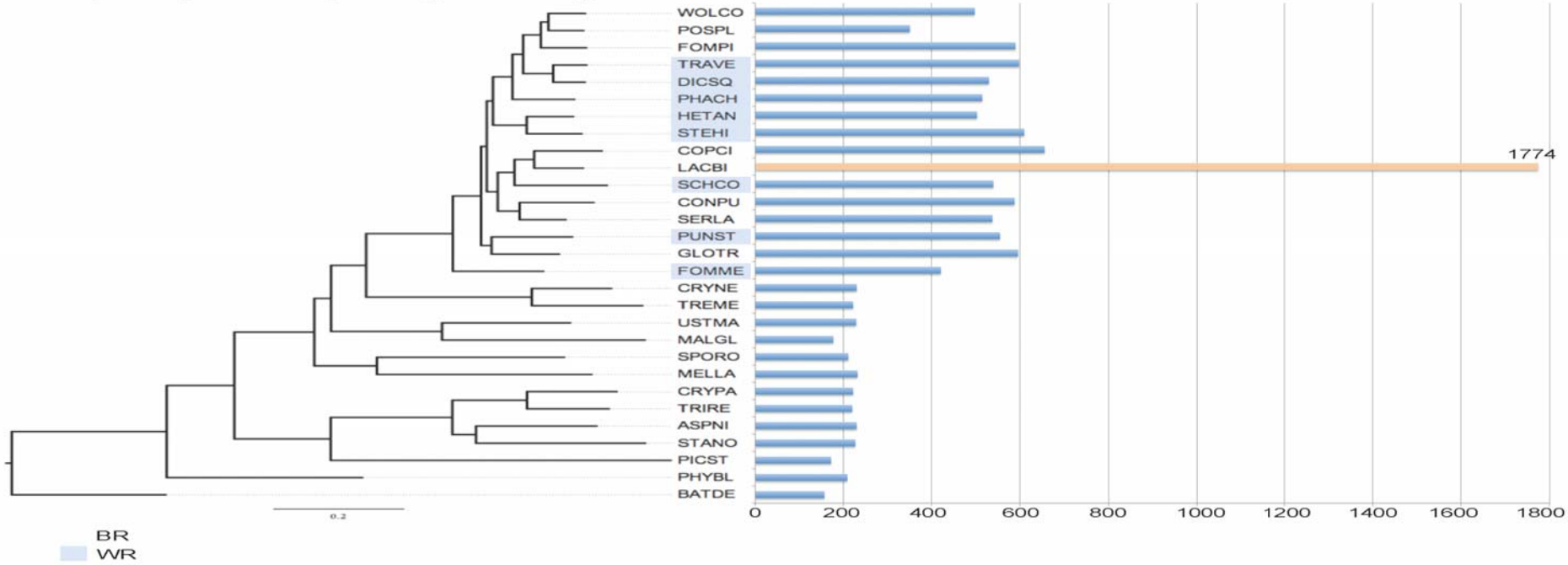


Fig. S22. Symbiosis genes down-regulated in *Laccaria bicolor* and their homologs in saprotophic and pathogenic Fungi.