Comparative Genomics of Early-Diverging Mushroom-Forming Fungi Provides Insights into the Origins of Lignocellulose Decay Capabilities

László G. Nagy,*[1] Robert Riley,[2] Andrew Tritt,[2] Catherine Adam,[2] Chris Daum,[2] Dimitrios Floudas,[3] Hui Sun,[2] Jagjit S. Yadav,[4] Jasmyn Pangilinan,[2] Karl-Henrik Larsson,[5] Kenji Matsuura,[6] Kerrie Barry,[2] Kurt Labutti,[2] Rita Kuo,[2] Robin A. Ohm,[2,7] Sukanta S. Bhattacharya,[4] Takashi Shirouzu,[8] Yuko Yoshinaga,[2] Francis M. Martin,[9] Igor V. Grigoriev,[2] and David S. Hibbett*[10]

[1]Synthetic and Systems Biology Unit, Institute of Biochemistry, BRC-HAS, Szeged, Hungary

[2]US Department of Energy (DOE) Joint Genome Institute, Walnut Creek, CA

[3]Department of Biology, Microbial Ecology Group, Lund University, Lund, Sweden

[4]Department of Environmental Health, University of Cincinnati College of Medicine

[5]Museum of Natural History, University of Oslo, Blindern, Oslo, Norway

[6]Laboratory of Insect Ecology, Graduate School of Agriculture, Kyoto University, Kyoto, Japan

[7]Department of Microbiology, Utrecht University, Utrecht, The Netherlands

[8]National Museum of Nature and Science, Tsukuba, Ibaraki, Japan

[9]Institut National De La Recherche Agronomique, Unité Mixte De Recherche 1136 Université Henri Poincaré, Interactions Arbres/Micro-Organismes, Champenoux, France

[10]Biology Department, Clark University, Worcester, MA

***Corresponding author:** E-mail: lnagy@brc.hu; dhibbett@clarku.edu.

Associate editor: Takashi Gojobori

Abstract

Evolution of lignocellulose decomposition was one of the most ecologically important innovations in fungi. White-rot fungi in the Agaricomycetes (mushrooms and relatives) are the most effective microorganisms in degrading both cellulose and lignin components of woody plant cell walls (PCW). However, the precise evolutionary origins of lignocellulose decomposition are poorly understood, largely because certain early-diverging clades of Agaricomycetes and its sister group, the Dacrymycetes, have yet to be sampled, or have been undersampled, in comparative genomic studies. Here, we present new genome sequences of ten saprotrophic fungi, including members of the Dacrymycetes and early-diverging clades of Agaricomycetes (Cantharellales, Sebacinales, Auriculariales, and Trechisporales), which we use to refine the origins and evolutionary history of the enzymatic toolkit of lignocellulose decomposition. We reconstructed the origin of ligninolytic enzymes, focusing on class II peroxidases (AA2), as well as enzymes that attack crystalline cellulose. Despite previous reports of white rot appearing as early as the Dacrymycetes, our results suggest that white-rot fungi evolved later in the Agaricomycetes, with the first class II peroxidases reconstructed in the ancestor of the Auriculariales and residual Agaricomycetes. The exemplars of the most ancient clades of Agaricomycetes that we sampled all lack class II peroxidases, and are thus concluded to use a combination of plesiomorphic and derived PCW degrading enzymes that predate the evolution of white rot.

Introduction

Lignocellulose is a major pool of organic carbon that is composed of the recalcitrant polymer lignin along with cellulose and other carbohydrates (hemicellulose, pectin, etc.) found in plant cell walls (PCW). Among microbial decomposers, only the so-called "white-rot" species in the Agaricomycetes (mushroom-forming fungi) are able to enzymatically digest all woody PCW components, including lignin. In contrast, "brown-rot" fungi initiate degradation of cellulose through a nonenzymatic Fenton mechanism, and cause extensive loss of PCW carbohydrates, but they do not appreciably degrade lignin (Martinez et al. 2009; Lundell et al. 2014). The traditional classification of saprotrophic Agaricomycetes as either white rot or brown rot has recently been challenged by emerging genome sequences of species that produce white rot–like symptoms on wood, but lack hallmark gene families of lignin degradation, particularly the class II peroxidases (PODs), which are classified as Auxiliary Activity Family 2 (AA2) (Floudas et al. 2012, 2015; Riley et al. 2014).

In addition to AA2, lignin degradation by white-rot fungi involves several other enzyme families, such as AA1 (multicopper oxidases, MCO) and possibly dye-decolorizing peroxidases (DyP) and heme-thiolate peroxidases (HTP) (Martinez et al. 2004; Hofrichter et al. 2010; Levasseur et al.

Article

2010, 2013; Poggeler 2011; Ruiz-Duenas et al. 2011, 2013; Rytioja et al. 2014). The degradation of crystalline cellulose takes place mainly via the action of cellobiohydrolases (GH6 and GH7) and lytic polysaccharide monooxygenases (AA9 LPMOs; Vaaje-Kolstad et al. 2010; Morgenstern et al. 2014). A number of other carbohydrate-active enzyme (CAZy) families are involved in the degradation or modification of carbohydrates (Hofrichter et al. 2010; Rytioja et al. 2014). Compared with white-rot species, brown-rot fungi, as well as ectomycorrhizal (ECM) mutualists, have reduced complements of the gene-encoding enzymes that decompose lignin, PCW polysaccharides, or both (Floudas et al. 2012; Kohler et al. 2015).

A prior phylogenomic analysis by Floudas et al. (2012) suggested that ligninolytic white rot evolved about 300 Ma, based on a reconstruction of the origin of AA2 genes. Lignin is the major precursor of coal, so the finding that the origin of AA2 genes roughly coincided with the end of the Carboniferous supported the view that evolution of white rot impacted the global carbon cycle (Robinson 1990). In the analysis of Floudas et al. (2012), Auriculariales, represented by the white-rot species *Auricularia delicata*, was resolved as the sister group of the rest of the Agaricomycetes. Accordingly, the common ancestor of Agaricomycetes was reconstructed as a white-rot species. However, genome sequences were not available for several key early-diverging clades of Agaricomycetes, including some that are associated with white rot–like decay of wood, including members of the Cantharellales, Sebacinales, and Trechisporales. For instance, *Botryobasidium* spp., *Tulasnella calospora*, and *Sistotrema brinkmannii* (Cantharellales) have been collected on white-rotted logs, although they represent groups that also form mycorrhizal or endophytic associations with various plants (Kottke and Nebel 2005; Moncalvo et al. 2006; Suarez et al. 2006; Riley et al. 2014). Moreover, some species of Dacrymycetes, which is the sister group of Agaricomycetes, have been reported to produce a white rot (Seifert 1983). Specifically, *Calocera lutea* and *Calocera viscosa* were reported to cause significant lignin loss in soil-block tests (Seifert 1983), although this has been disputed (Worrall et al. 1997). If Dacrymycetes were shown to possess the capacity for lignin degradation, then it could push the evolution of white rot back in time significantly.

In this study, we considerably extend the sampling of early-diverging Agaricomycete and Dacrymycete genomes in an attempt to refine the origins of various wood-rotting mechanisms and the evolutionary roots of the white-rot enzyme complement. We present genome sequences for ten new basidiomycetes, including the first sequenced genomes for the Trechisporales, two *Calocera* species, and a second species of Auriculariales, *Exidia glandulosa*, as well as three new brown-rot species and the termite symbiont *Fibulorhizoctonia*. We reconstruct the evolution of 19 major gene families related to wood decay and suggest a hypothesis to explain the extant diversity of wood-decay strategies in the Agaricomycetes. Our results also contribute to resolving disputed phylogenetic relationships among the Agaricomycetes, in particular, the position of the Trechisporales.

Results and Discussion

New Genome Sequences

We generated new draft genome sequences for ten basidiomycetes, with a focus on early-diverging orders (table 1). These include the first representatives of the order Trechisporales (*Sistotremastrum suecicum* and *Sistotremastrum niveo-cremeum*), which has been difficult to place in prior phylogenetic studies (Binder et al. 2005). We also sampled two species of Dacrymycetes, *Calocera cornea*, and the reportedly white-rot *C. viscosa*, thereby tripling the number of genomes for the class, as well as a second species of Auriculariales, *E. glandulosa*. The other five new genomes include brown-rot species of the Gloeophyllales (*Neolentinus lepideus*) and Polyporales (*Daedalea quercina*, *Laetiporus sulphureus*), a white-rot member of the Russulales (*Peniophora* sp.), and a termite symbiont in the Atheliales (*Fibulorhizoctonia* sp.). Genome sizes for these diverse taxa range from 29 to 95 Mb, with 12,199 to 32,946 predicted protein-coding genes. The largest and most gene-rich genome was that of *Fibulorhizoctonia* sp.; its gene number of 32,946 makes it the largest Agaricomycete genome sequenced so far, ~30% larger than the genome of the ECM species *Laccaria bicolor* (Martin et al. 2008). Analyses of pairs of genomic blocks with ≥50% of genes having a best bidirectional BLAST hit suggest that high numbers of tandem and segmental duplications underlie the exceptional genome

Table 1. New Genomes Sequenced in this Study, the Nutritional Strategy, and Taxonomic Classification of the Species.

Species Name	Classification	Nutritional Strategy	No. of Predicted Genes	No. of Scaffolds	Genome Size (MB)/ Sequencing Depth
Calocera cornea	Dacrymycetes	Brown rot	13,177	545	33.24 (85×)
Calocera viscosa	Dacrymycetes	Brown rot[a]	12,378	214	29.1 (81.7×)
Daedalea quercina	Polyporales	Brown rot	12,199	357	32.74 (144×)
Exidia glandulosa	Auriculariales	White rot	26,765	1,727	78.17 (83.7×)
Fibulorhizoctonia sp.	Atheliales	White rot	32,946	1,918	95.13 (30.5×)
Laetiporus sulphureus	Polyporales	Brown rot	13,774	403	39.92 (85.2×)
Neolentinus lepideus	Gloeophyllales	Brown rot	13,164	331	35.64 (92.4×)
Peniophora sp.	Russulales	White rot	18,952	1,092	46.03 (141×)
Sistotremastrum niveo-cremeum	Trechisporales	White rot	13,080	230	35.36 (103×)
Sistotremastrum suecicum	Trechisporales	White rot	13,657	1,028	33.86 (143×)

[a]Suggested to be white rot by Seifert (1983).

size of this species. Detailed statistics for all of the new genomes are presented in table 1.

Phylogenetic Relationships

We identified 824 gene families with members in >40 species that were either single copy or had species-specific duplications but no deep paralogs (supplementary file S1, Supplementary Material online). Single-gene alignments longer than 50 residues (after excluding low-scoring alignment sites) were concatenated into a supermatrix. With the 3 different site exclusion thresholds used (0.95, 0.97, and 0.98), this resulted in 3 data sets of 201,382 (811 genes), 127,263 (731 genes), and 91,981 amino acid sites (623 genes), respectively. Our data sets show satisfying levels of completeness (each gene was present on average in 52 species), although the distribution of missing data follows a power law: The number of missing genes per species increased with the distance from crown Agaricomycetes (supplementary fig. S1, Supplementary Material online). Pucciniomycotina (rusts and relatives) and Ustilaginomycotina (smuts and relatives) are among the clades with the highest amount of missing data, which could underlie the difficulty of resolving the relationships among these clades and the Agaricomycetes.

Maximum-likelihood (ML) trees inferred using the three data sets had identical topologies, so for subsequent analyses, we focused on the data set with 623 genes (91,981 sites). Both the ML bootstrap and Bayesian Markov chain Monte Carlo (MCMC) analysis revealed strong (> 95% BS or > 0.95 Bayesian posterior probability [BPP]) support at most nodes (fig. 1). Exceptions include the nodes that unite the Amylocorticiales and Atheliales (94%), and that specify the placement of *L. sulphureus* (82%) and the relationships among Pucciniomycotina, Ustilaginomycotina, and Agaricomycotina, which have proven difficult to resolve in prior phylogenetic analyses with fewer genes, but greater taxon sampling than the present analyses (Hibbett et al. 1997; James et al. 2006; Matheny et al. 2007).

The Trechisporales (*Sistotremastrum* spp.) was placed as the sister group of the Phallomycetidae (represented by *Sphaerobolus*) in ML analyses, with strong support (100% BS), but Bayesian analyses placed *Sphaerobolus* as the sister group to the Trechisporales plus more derived Agaricomycetes (0.96 BPP; supplementary fig. S2, Supplementary Material online). The Russulales was placed as the sister group of a large clade containing the Polyporales, Gloeophyllales, Corticiales, Jaapiales, and Agaricomycetidae in the ML analyses (100% BS), whereas in the Bayesian tree it formed a clade together with the Polyporales, although without statistical support (BPP 0.59). Both conformations contradict previous multilocus- and genome-based placements of the Russulales (Hibbett et al. 1997; Binder and Hibbett 2002; Matheny et al. 2007).

Genomic Innovations in the Agaricomycetes

A considerable proportion of gene families in the genomes of Agaricomycetes are specific to the class (fig. 1). Enrichment analyses of PFAM domains revealed 37 PFAM domains that are exclusive to and a further 270 that are significantly overrepresented in the Agaricomycetes relative to other fungi ($P \leq 0.05$, Fisher's exact test; supplementary file S2, Supplementary Material online). Among carbohydrate-active protein families, we found members of the GH52, GH76, CBM19, polysaccharide lyases, and a family of putative cellulases (PF12876) to be significantly overrepresented in the Agaricomycetes ($P \leq 0.05$, Fisher's exact test). Of the 62 species examined, GH52 is specific to the Auriculariales, where it occurs both in *Auricularia* and *Exidia*. This family displays beta-xylosidase activity as its only known function and has been implicated in lignocellulosic plant biomass decomposition by various bacteria and the ascomycete *Trichoderma reesei* (Adav et al. 2012), but to our knowledge it has not been reported in the basidiomycetes. Similarly, a family of putative cellulases (PF12876, $P = 0.007$, Fisher's exact test) was found specific to the Auriculariales within the Agaricomycetes (but present in rusts and other fungi). GH76 are alpha-mannanases and have been found significantly upregulated in *Trichoderma* cultures grown on softwood as the major carbon source, mainly consisting of glucomannans (Adav et al. 2012). Polysaccharide lyases, enzymes involved in pectin degradation (van den Brink and de Vries 2011), were found in all Agaricomycete species with a considerable expansion in the Auriculariales. We also detected significant overrepresentation ($P = 0.000008$, Fisher's exact test) of fungal carbohydrate-binding module (CBM) 19 family members, which are chitin-binding proteins found in fungal chitinases.

Evolution of Gene Families Related to Wood Decay

We reconstructed the evolution of 18 gene families, belonging to 3 functional groups: Oxidoreductases related to the degradation of lignin or lignin-like compounds, CAZys active on polysaccharide main chains, and other CAZys related to wood decay, which show various catalytic activities on carbohydrates (table 2). On average, the 62 genomes contained 133 genes in the 18 gene families, of which 33, 56, and 34 are oxidative CAZys, CAZys active on polysaccharide main chains, and other CAZys related to wood decay, respectively. Of the newly sequenced early-diverging species, *C. cornea* and *C. viscosa* have low numbers of CAZys (49 and 44, respectively), whereas *E. glandulosa* has 216 and *Si. niveo-cremeum* and *Si. suecicum* have 133 and 135 genes, respectively (table 3). *Fibulorhizoctonia* sp. has a high number of genes (239), comparable with white-rot species, but no AA2s.

Evolution of Decay-Related Oxidoreductases

Reconstructed ancestral copy numbers at internal nodes of the phylogeny suggest that a burst of diversification (from 42 to 85 copies) occurred after the divergence of the Auriculariales and remaining Agaricomycetes (supplementary fig. S3, Supplementary Material online), which correlates with the elaboration of the lignin decomposition capacity of early Agaricomycete ancestors. Analysis of a random set of gene families not related to lignocellulose decomposition (supplementary fig. S4, Supplementary Material online) suggests that

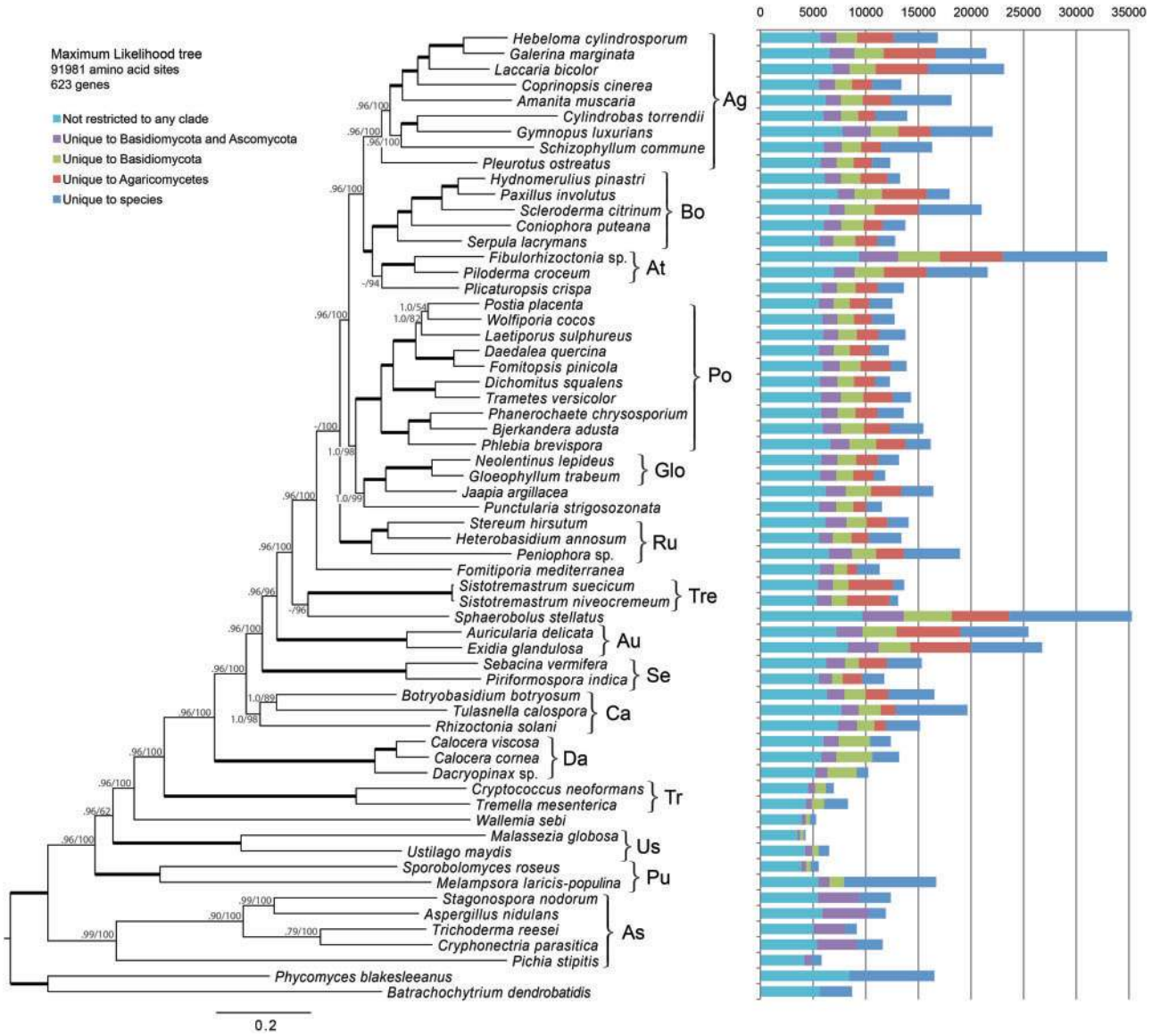


Fig. 1. Species phylogeny and genome dynamics across 62 fungal genomes with an emphasis on the Agaricomycetes (left panel). Numbers above branches of the tree represent Bayesian posterior probabilities and ML bootstrap percentages, respectively. Thickened branches received maximal support from both analyses. Major high-level clades are marked from top to bottom as Agaricales (Ag), Boletales (Bo), Atheliales (At), Polyporales (Po), Gloeophyllales (Glo), Russulales (Ru), Trechisporales (Tre), Auriculariales (Au), Sebacinales (Se), Cantharellales (Ca), Dacrymycetes (Da), Tremellomycetes (Tr), Ustilaginomycotina (Us), Pucciniomycotina (Pu), and Ascomycota (As). Bars correspond to the number of genes in the genome colored by their phylogenetic distribution.

this pattern deviates from the general evolutionary dynamics of Agaricomycete genome size. Other significant oxidoreductase expansions were inferred in the lineages leading to *Sphaerobolus* and *Fibulorhizoctonia*. However, these expansions correlate well with the results of control analyses of a random set of genes, suggesting that they may reflect genome-wide processes of gene family expansion, rather than adaptive expansion of decay-related oxidoreductase gene families.

Within decay-related oxidoreductases, AA2s show a similar early expansion (fig. 2) that is robust to the choice of gene tree–species tree reconciliation algorithm (supplementary fig. S4, Supplementary Material online). The early expansion was observed in both the TreeFix and Notung analyses, although Notung yielded considerably lower copy number estimates along the backbone of the tree and consequently fewer gene losses (fig. 2 and supplementary fig. S4, Supplementary Material online). Of the early-diverging lineages, the sampled species of Cantharellales, Sebacinales, and Dacrymycetes all lack AA2s, whereas the most recent common ancestor (mrca) of the Auriculariales and other Agaricomycetes has 3 reconstructed genes, followed by 21 genes in the mrca of the Trechisporales plus later diverging orders and 16 copies in the mrca of the Auriculariales. Looking more closely at these expansions, we found that the expansion in the Auriculariales agrees well with general genome-size evolution (supplementary fig. S3, Supplementary Material online), suggesting that the AA2 expansion in the Auriculariales might also be related to general genome-size dynamics. However, the expansion in the mrca of the Trechisporales plus later

Table 2. Gene Families Sampled in this Study and their Proposed Functions in Wood Degradation.

Functional Gene Family Category	Abbreviation	Activity Related to Wood Degradation
Oxidoreductases	AA2	Lignin degradation
	DyP	Acting on lignin or lignin derivatives
	HTP	Possible action on lignin derivatives
	AA1_1 and AA1_2[a]	Lignin degradation
CAZys active on polysaccharide main chains	GH5-5[b]	Endoglucanase
	GH5-7	Endomannanase
	GH6	Cellobiohydrolase
	GH7	Cellobiohydrolase
	AA9	Cellulose cleaving oxidoreductase
	GH10	Endoxylanase
	GH12	Endoglucanase
	GH28	Pectinase activity
	GH45	Endoglucanase
	GH74	Xyloglucanase
Other CAZys related to wood decay	GH3	β-Glucosidase/β-xylosidase
	GH43	α-Arabinofuranosidase/β-xylosidase
	CE1	Acetyl-xylan-esterase, ferruloyl esterase, cinnamoyl esterase
	CE16	Acetyl-xylan-esterase, ferruloyl esterase, cinnamoyl esterase

[a]Including both Basidiomycete Laccases and the Laccase-like gene family.
[b]Subfamily classification follows Aspeborg et al. (2012).

diverging orders is very striking (from 3 genes to 21 gene copies) and has likely provided the raw material for the evolution of extant suites of AA2s in white-rot fungi.

We detected 7,209 genes encoding cytochrome p450s in the 62 genomes, which are distributed across 267 different subfamilies following Nelson (2006). Genomes of Agaricomycetes and Dacrymycetes are enriched in p450s compared with other basidiomycetes and ascomycetes (supplementary fig. S9, Supplementary Material online). Expansions of p450s appear to be correlated with the diversification patterns observed for many CAZys, but not decay-related oxidoreductases. No significant contractions of p450s were inferred in brown rot or mycorrhizal lineages. Certain cytochrome p450s are thought to play roles in detoxifying diverse compounds, including secondary products of lignin-olysis and xenobiotic compounds. Cytochrome p450s are expanded in the Dacrymycete–Agaricomycete clade, but their role in the origin and maintenance of white rot is unclear.

Evolution of CAZys Active on Polysaccharide Main Chains and Other CAZy Families Related to Wood Decay

Like the decay-related oxidoreductases, CAZys active on polysaccharide main chains (GH5, GH6, GH7, AA9 LPMO, GH10, GH12, GH28, GH45, GH74) and other CAZys related to wood decay (GH3, GH43, GH51, CE1, CE16) show expansions early in Agaricomycete evolution, although less pronounced than in oxidative CAZys, due to high ancestral copy numbers along the backbone of the tree (supplementary figs. S6 and S7, Supplementary Material online). As in the case of decay-related oxidoreductases, brown-rot Dacrymycetes have very low copy numbers: 16–17 copies in CAZys families active on polysaccharide main chains and 16–21 copies in other CAZy families related to wood decay, whereas the Cantharellales (in particular *Rhizoctonia* and *Tulasnella*) and Sebacinales have moderate to high numbers of main chain CAY-s (66–133 genes) and other CAZys related to wood decay (16–43 genes).

Table 3. The Number of CAZy Enzymes in the Different Functional Categories in Newly Sequenced Early-Diverging Agaricomycete Genomes.

CAZy Class	*Average*	*Calco*	*Calvi*	*Exigl*	*Sissu*	*Sisni*	*Fibsp*
Oxidative CAZy	34	10	12	81	66	63	75
Class II Peroxidases	8	0	0	37	17	15	0
CAZys active on poly-saccharide main chains	56	17	16	81	52	53	102
Other CAZys related to wood decay	33	17	21	74	17	17	62
Total	131	44	49	272	152	148	241

Within CAZys active on polysaccharide main chains, of particular interest are AA9 LPMOs, a large family involved in cellulose decomposition. This gene family shows a significant expansion in Agaricomycetes (from 44 to 62 gene copies; fig. 2 and supplementary fig. S8, Supplementary Material online) prior to the expansion of decay-related oxidoreductases. Our analyses reconstructed high ancestral copy numbers already in the mrca of the Dikarya, suggesting that some components of lignocellulose attack are shared between the Ascomycetes and the Agaricomycetes.

Evolutionary Origins of White Rot

Our analyses refine previous hypotheses on the origin of white rot (Floudas et al. 2012). We extended the sampling of early-diverging Agaricomycete genomes significantly, including suspect white-rot species. We find that none of the lineages arising prior to the split of the Auriculariales and the rest of the Agaricomycetes possess AA2s. We found that AA2s and in a wider sense decay-related oxidoreductases began to diversify after Sebacinales diverged from the backbone of the tree. For the mrca of the Auriculariales and later-diverging Agaricomycetes, we inferred three AA2 gene copies, which are consistent with the finding of Floudas et al. (2012, using complementary methods). Previous CAFÉ results suggested a significant expansion of PODs on the branch leading to the node uniting Auriculariales and residual Agaricomycetes (Node A in Floudas et al. 2012). Using a much denser sampling of taxa, our analyses push this expansion to the next more recent node, that is, the mrca of Trechisporales, *Sphaerobolus*, and residual Agaricomycetes. This suggests that the majority of oxidative enzyme diversity used by white-rot fungi has emerged after the origins of the first white-rot species.

Gene Contents Suggest Functional Similarities between Decay Modes of Early-Diverging and Certain Derived Agaricomycetes and Soft-Rot Ascomycetes

Our analyses suggest that the mrcas of the Dikarya and Basidiomycota already had diverse suites of CAZys active

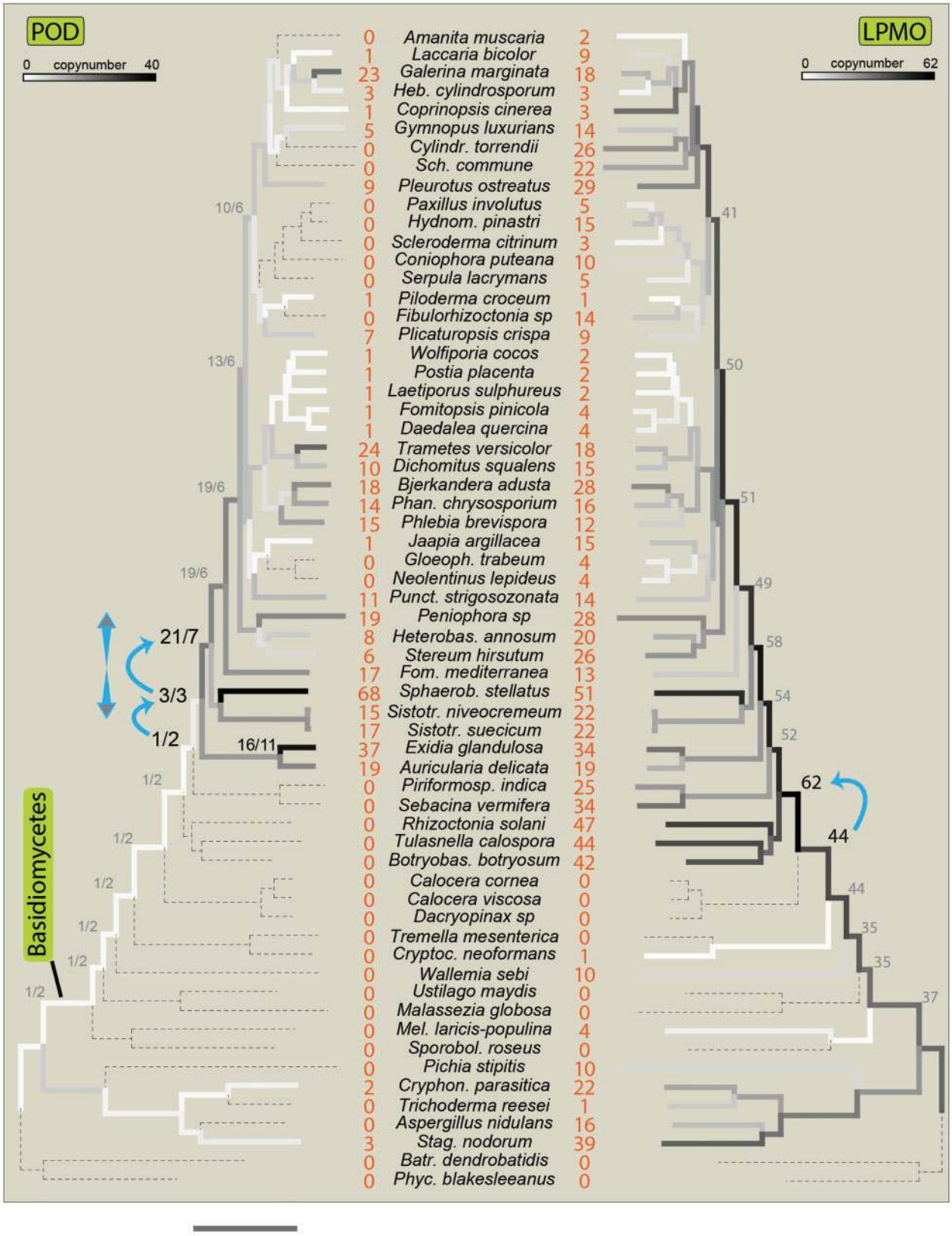


FIG. 2. A comparison of reconstructed copy numbers through the evolution of PODs and LPMOs in the Agaricomycetes. Branches are colored according to reconstructed ancestral copy numbers. Dashed line marks nodes with no genes in the family. Numerical values are shown for key nodes. Copy numbers in extant species are shown in red. For PODs, first and second numbers correspond to copy numbers inferred using TreeFix and Notung, respectively. For visualization, the copy number in *Sphaerobolus* (68) has been truncated at 40.

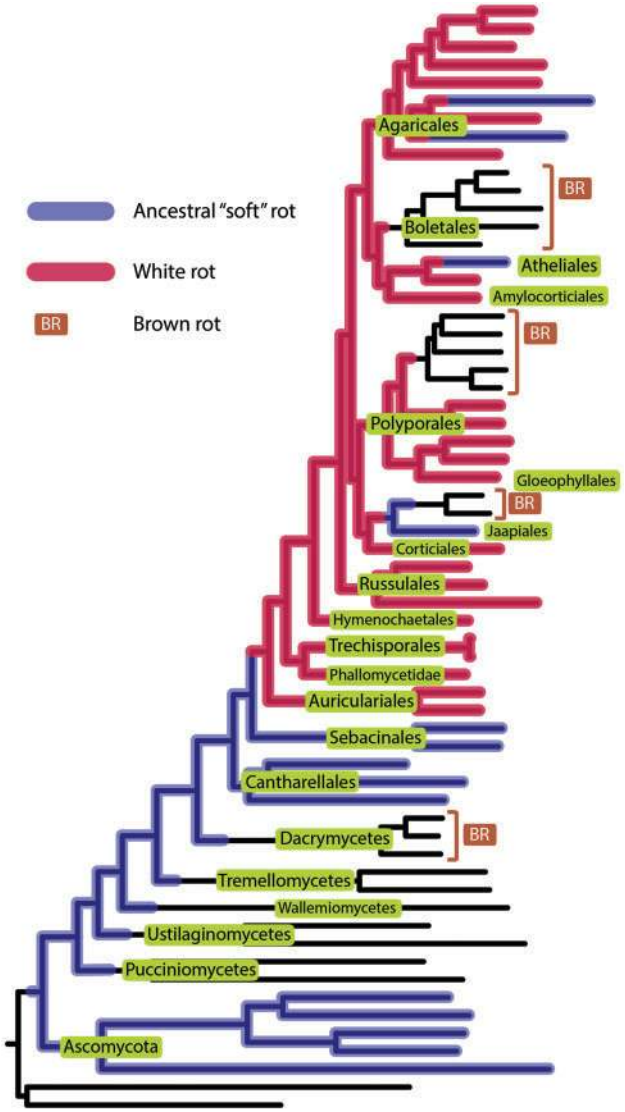


Fig. 3. A hypothesis for the evolution of decay types in the Agaricomycetes. Blue marks the presence of an "ancestral" wood-decay enzyme complement shared between Asco- and Basidiomycetes. Purple layer over the ancestral rot shows the evolution of white rot marked by the expansion of oxidative gene families involved in lignin decomposition (PODs, MCOs, HTPs, DyPs). In brown rotters and mycorrhizal mutualists, both enzyme complements were lost (black lines), whereas in the Atheliales, *Schizophyllum*, and *Cylindrobasidium* (Agaricales) only oxidative enzymes were lost, leading to a decomposition strategy similar to the ancestral rot-type.

on polysaccharide main chains and other CAZys related to wood decay, in particular AA9 LPMOs, but primitively lacked AA2s and other ligninolytic oxidoreductases. Taxa such as *Botryobasidium* (Cantharellales) may manifest a decay mode that is plesiomorphic for the Agaricomycetes as a whole. In contrast, decay-related oxidoreductase families mostly originated later in the evolution of Agaricomycetes, around the divergence of Auriculariales and more derived clades. Subsequently, the white-rot apparatus has been modified by differential loss or retention of various classes of decay enzymes, including the following: 1) Losses of both decay-related oxidoreductases and AA9 LMPOs (and CAZys active on polysaccharide main chains in general) and 2) loss of decay-related oxidoreductases, but retention of CAZys active on polysaccharide main chains. The first pattern is observed in brown-rot fungi, which are efficient decayers of massive woody substrates that have evolved a nonenzymatic Fenton mechanism for carbohydrate extraction, and ECM fungi, which no longer depend on enzymatic digestion of lignocellulose to provide carbon nutrition (Floudas et al. 2012; Kohler et al. 2015). The second pattern restores an ancestral gene complement similar to that of *Botryobasidium* and ascomycetes and it is observed in species such as *Schizophyllum commune*, *Jaapia argillacea*, and *Cylindrobasidium torrendii*. These taxa, and *Botryobasidium*, have been generally considered to be saprotrophs, but they have limited ability to degrade wood. It is possible that in nature they scavenge carbohydrates in communities that include ligninolytic white-rot fungi, or possibly function as mycorrhizal symbionts, mycoparasites, or plant parasites (Riley et al. 2014; Floudas et al. 2015). The decay apparatus of all these taxa, which are neither white rot or brown rot, resembles soft rot, which is associated with ascomycetes.

Gene Content Suggests that Calocera cornea *Is a Brown-Rot Species*

Among the analyzed basidiomycete lineages, Dacrymycetes have one of the lowest numbers of lignocellulose decomposing CAZys, similar in their gene content to Tremellomycetes. They possess 16–17, 16–21, and 9–12 genes of CAZys active on polysaccharide main chains, other CAZys related to wood decay, and oxidative CAZys, respectively (table 3 and figs. S2 and S4–S7, Supplementary Material online). Notably, they lack AA2s, the hallmark gene family of white-rot species as well as AA9 LPMOs. Thus, although *C. viscosa* has been implicated in causing white rot in soil-block tests (Seifert 1983), based on its gene content it resembles typical brown-rot taxa. However, they are unique among brown-rot lineages in that they are derived not from white-rot ancestors, but apparently from saprotrophs with an ancestral soft rot-like decay mode.

Fibulorhizoctonia *Is a Termite Symbiont that Possesses Diverse White-Rot Enzymes*

Fibulorhizoctonia sp. is an anamorphic genus (teleomorph: *Athelia* spp.) first reported as "termite balls," fungal sclerotia that mimic real eggs in termite nests (Matsuura et al. 2000; Matsuura 2006). It has also been reported as a carrot-spoilage species from cold storage facilities (de Vries et al. 2008). Of decay-related gene families, *Fibulorhizoctonia* lacks AA2s, but contains several other oxidative enzymes such as DyPs, HTPs, and AA1 (supplementary fig. S3, Supplementary Material online), suggesting ligninolytic activities. *Fibulorhizoctonia* also possesses a rich complement of AA9 LPMOs, in contrast to its ECM sister taxon, *Piloderma croceum*, which has a highly reduced decay apparatus, as is typical of ECM symbionts (Kohler et al. 2015). The nature of the association between *Fibulorhizoctonia* and its termite hosts remains obscure (Matsuura et al. 2000; Matsuura 2006), but gene contents suggest that the fungus possesses decay capabilities that could potentially play a role in the symbiosis.

Sistotremastrum *spp. (Trechisporales) Represent an Early-Diverging Lineage of White-Rot Fungi*

The phylogenetic position and nutritional modes in the Trechisporales have been a matter of debate since the description of the order (Binder and Hibbett 2002; Larsson et al. 2004; Larsson 2007; Matheny et al. 2007). Our ML analyses uniformly recovered it as a sister group of the Phallomycetidae (BS 96%; fig. 1), whereas Bayesian MCMC placed it as sister to Phallomycetidae plus other Agaricomycetes (BPP 0.96). *Sistotremastrum* spp. possess multiple copies of AA2s (15–17 copies), HTPs (35), and MCOs (14-15) and are thus equipped for enzymatic decomposition of lignin. This is consistent with the occurrence of *Sistotremastrum* spp. on decaying wood (without close contact with soil) and observations that they can be grown on microbiological media. In comparison, other members of the Trechisporales (mostly belonging to the Hydnodontaceae, e.g., *Trechispora* spp.) occur on soil or wood in close contact with soil and have been found in ECM root tips (Dunham et al. 2007), suggesting that they form mycorrhizal interactions with plants. However, genomes for the putatively non-saprotrophic Trechisporales are not yet available.

Conclusions

Saprotrophic mushroom-forming fungi have evolved a wide range of strategies for decomposing lignocellulose, including white rot, brown rot, and hitherto unclassified soft rot-like wood-decay mechanisms (Floudas et al. 2012; Kohler et al. 2015). These decay modes have drawn upon both ancestral and novel or clade-specific enzyme complements. Previous studies suggested that oxidative enzyme families involved in lignin degradation underwent a burst of diversification within the Agaricomycetes, most likely in the common ancestor of the Auriculariales and other mushroom-forming fungi (Floudas et al. 2012), but taxon sampling was too limited to resolve the precise origins of lignocellulolytic mechanisms within the early Agaricomycetes and its sister group, the Dacrymycetes. We refined understanding of the origins of white rot and other decay modes by analyzing genomes of early-diverging Basidiomycetes from the Cantharellales (*Botryobasidum*, *Tulasnella*), Sebacinales (*Sebacina*, *Piriformospora*), Trechisporales (*Sistotremastrum* spp.), and reportedly white-rot Dacrymycetes (*C. viscosa*). Our results indicate that basidiomycete groups branching off before the Auriculariales lack key enzymes of lignin decomposition; Trechisporales have a robust white-rot apparatus, and we are now able to resolve that this order diverged after the origin of Auriculariales. However, the most ancient clades of Agaricomycetes already had diverse repertoires of gene families involved in cellulose decomposition. CAZys active on polysaccharide main chains, in particular AA9 LPMOs, are expanded in the Cantharellales and to some extent in the Sebacinales. Our analysis of an expanded sample of early-diverging Agaricomycete genomes is consistent with prior suggestions that the first white-rot species may have been the mrca of Auriculariales and more derived clades (Floudas et al. 2012). However, this inference should be tested further through genome sequencing of putatively saprotrophic members of early-diverging clades, such as *Craterocolla cerasi* (Sebacinales) and *S. brinkmannii* (Cantharellales).

Methods

Strains, Growth Conditions, and Nucleic Acid Extraction

For DNA and RNA extraction, *L. sulphureus* (93-53, deposited in Forest Products Laboratory, FPL), *D. quercina* (L15889 ss-12, FPL), *N. lepideus* (HHB-14362, FPL), *C. viscosa* (*TUFC12733*), and *C. cornea* (HHB-12733-sp, FPL) were grown in 1-l flasks with liquid media of malt extract (20 g/l) and yeast extract (0.5 g/l) at 28 °C in darkness until sufficient amount of mycelium was available. *Fibulorhizoctonia* sp. (CBS 109695, CBS), *Si. suecicum* (HHB-10207 ss-3, FPL), *Si. niveo-cremeum* (HHB-9708 ss-1, FPL), *E. glandulosa* (HHB-12029-sp, FPL), and *Peniophora* sp. (unidentified strain) were grown on vitamin medium (malt extract 1 g/l, glucose 4 g/l, ammonium tartrate 1 g/l, KH_2PO_4 0.2 g/l, $MgSO_4$ 0.1 g/l, NaCl 0.02 g/l, $CaCl_2$ 0.026 g/l, $ZnSO_4$ 0.88 mg/l, $MnSO_4$ 0.8 mg/l, $FeCl_3$ 0.8 mg/l, and BME 100× vitamin solution/sigma/1 ml/l) at 28 °C in darkness. Genomic DNA and RNA were extracted using the Qiagen Blood & Cell Culture DNA Kit and Qiagen RNeasy Midi Kit with in-column DNAse digestion, respectively. Decay characteristics for the new sequences species, where known, are based on reports from in vitro cultural studies and field observations of decayed substrates (Nakasone 1990; Ginns and Lefebvre 1993; Ryvarden and Gilbertson 1993; Worrall et al. 1997).

Sequencing, Assembly, and Annotation

All genomes and transcriptomes in this study were sequenced using whole genome sequencing on Illumina platform. For fragment Illumina libraries, ~1 µg of genomic DNA was sheared using the Covaris E210 (Covaris) and size selected using Agencourt Ampure or SPRI Beads (Beckman Coulter). The DNA fragments were treated with end repair, A-tailing, and adapter ligation using the TruSeq DNA Sample Prep Kit (Illumina) or the KAPA-Illumina library creation kit (KAPA biosystems). LFPE (ligation-free paired end) mate pair fragments were generated using the 5500 SOLiD Mate-Paired Library Construction Kit (SOLiD). A total of 15 µg of genomic DNA was sheared using the Covaris g-TUBETM (Covaris) and gel size selected for 4 kb. For *Fibulorhizoctonia* sp. CBS 109695, two additional 8 and 20 kb libraries were produced from 25 µg each using a similar protocol. The sheared DNA was end repaired, ligated with biotinylated internal linkers, circularized, and treated with plasmid safe to remove noncircularized products. The circularized DNA was nick translated and treated with T7 exonuclease and S1 nuclease to generate fragments containing internal linkers with genomic tags on each end. The mate pair fragments were A-tailed and purified using Strepavidin bead selection (Invitrogen). The purified fragments were ligated with Illumina adaptors and amplified with Illumina primers (Illumina) to generate the final library.mRNA was purified from 5 µg of total RNA using the Dynabeads mRNA Purification Kit (Invitrogen) and

chemically fragmented to 200–250 bp (Ambion). mRNA was reverse transcribed with SuperScript II using random hexamers. Second strand cDNA was synthesized using dNTP/dUTP mix (Thermo Scientific), *Escherichia coli* DNA ligase, *E. coli* DNA polymerase I, and *E. coli* RnaseH (Invitrogen). The fragments were treated with end-repair, A-tailing, and ligation of Illumina compatible adapters (IDT, Inc) using the KAPA-Illumina library creation kit (KAPA biosystems). Second strand cDNA was removed by AmpErase UNG (Applied Biosystems) to generate strandedness and enriched with 10 cycles of PCR.

Libraries were prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v3, and Illumina's cBot instrument to generate a clustered flowcell for sequencing. Sequencing of the flowcell was performed on the Illumina HiSeq2000 sequencer using a TruSeq SBS sequencing kit, v3, following a 2 × 150 bp (for Illumina fragments and transcriptomes) and 2 × 100 bp (for LFPE) indexed run recipe.

Illumina reads of stranded RNA-seq data were de novo assembled into consensus sequences using Rnnotator (v. 2.5.6 or later) (Martin et al. 2010). Genome assemblies were produced from a single Illumina fragment and a single long mate-pair libraries using AllPathsLG (Gnerre et al. 2011). For *Si. suecicum*, a 3-kb long mate-pair library was constructed in silico from the initial assembly with Velvet (Zerbino and Birney 2008) and then reassembled together with the initial data set using AllPathsLG. For *Fibulorhizoctonia* sp, data from additional 8 and 20 kb Illumina libraries were included in the AllPathsLG assembly, which then was also improved using pbJelly (English et al. 2012) employing PacBio reads, error corrected using Illumina data. The PacBio reads were obtained from two unamplified 10 and 20 kb libraries, which were generated using Pacific Biosciences standard template preparation protocol from 5 and 50 μg of genomic DNA sheared using Covaris g-Tubes into > 10 and > 20 kb fragments, respectively. The sheared DNA fragments were then prepared using Pacific Biosciences SMRTbell template preparation kit, where the fragments were treated with DNA damage repair, had their ends repaired so that they were blunt ended, and 5′ phosphorylated. Pacific Biosciences hairpin adapters were then ligated to the fragments to create the SMRTbell template for sequencing. The SMRTbell templates were then purified using exonuclease treatments and size selected using AMPure PB beads (for 10 kb libraries) and Sage Sciences Blue Pippin Size-Selection system (for 20 kb). Sequencing primer was then annealed to the SMRTbell templates and sequencing polymerase version XL (for 10 kb) and P4 (for 20 kb) was bound to them. The prepared SMRTbell template libraries were then sequenced on a Pacific Biosciences RSII sequencer using Version C2 chemistry and 2-h sequencing movie run times.

All genomes were annotated using the JGI Annotation Pipeline and made available via the JGI fungal portal MycoCosm (Grigoriev et al. 2014). Genome assemblies and annotation were also deposited at DDBJ/EMBL/GenBank under the following accessions: LKHQ00000000, LKHT00000000, LKHU00000000, LOAW00000000, LNCK00000000, LOAX00000000, LNCL00000000, LOAU00000000, LOAV00000000, and LOAY00000000.

To assess gene family enrichment in Agaricomycetes, we used Markov clustering (MCL) (11917018) with an inflation parameter of 2.0 to identify protein families in the full comparative set of organisms. This allowed the identification of Agaricomycete-specific gene families. Protein functions were then assigned by assigning the protein sequences to protein domains in the Pfam v22 database (24288371) using a TimeLogic Decypher machine. Then, for each Pfam, we identified four protein sets: *a* = Agaricomycete-specific proteins with a given Pfam; *b* = other proteins with the Pfam; *c* = Agaricomycete-specific proteins without the Pfam; *d* = other proteins without the Pfam. We then used Fisher's exact test as implemented in the fisher_exact function of the scipy.stats Python package to assign odds ratios and two-sided *P* values (supplemental file S2, Supplementary Material online).

Species Phylogeny

Taxon Sampling, Data Collection

Sixty-two species have been included in this study with an emphasis on early diverging Agaricomycetes (fig 1). Representative Asco- and Zygomycete species have been used as outgroups. To identify gene families that can be used in phylogeny reconstruction, we first clustered predicted protein sequences from all 62 genomes using the MCL algorithm with an inflation parameter of 2.0 (van Dongen 2000) followed by a screen for clusters of single-copy gene families to use in phylogeny reconstruction. However, this proved to be an overly restrictive strategy, as we found that the resulting suite of clusters had very low taxon occupancy (<15 species per cluster). Therefore, we further screened the clusters allowing the occurrence of inparalogs (paralogs restricted to a single species), but discarding gene families containing deep paralogs. To do this we inferred multiple alignments for each cluster by using PRANK v. 111130 (+F turned off) (Loytynoja and Goldman 2008) and estimated ML gene trees using RAxML 7.2.6 (under the PROTGAMMAWAG model, otherwise with default parameters) (Stamatakis 2006).

Next, we screened the gene trees for deep paralogs, discarding the gene family if gene duplications were observed deep in the gene tree (i.e., not on a terminal branch). If only inparalogs were found in a gene family, we retained the paralog closest to the root as judged by branch lengths on the ML gene trees. Gene families identified using the above strategy were added to the array of strictly single-copy gene families. We removed low-scoring regions of the protein alignments using the method we developed previously (Kohler et al. 2015), which takes PRANK posterior probabilities for each alignment site as a criterion for site exclusion. Three different posterior probability thresholds were used (0.95, 0.97, and 0.98). In addition to low-scoring alignment sites, we removed parsimony uninformative gapped sites. These sites would not contribute phylogenetic signal to the analyses; however, their removal reduced the memory requirements of phylogenetic analyses. After the removal of low-scoring alignment sites, we concatenated single-protein alignments that contained

sequence data for more than 40 species and were longer than 50 amino acid residues into a supermatrix.

Phylogenomic Analyses

We inferred organismal phylogenies using ML and Bayesian inference. For ML inference, we used the PTHREADS version of RAxML 7.2.8 (Stamatakis 2006) and a partitioned model where each gene was treated as a separate partition with parameters of the PROTGAMMAWAG model estimated separately for each partition. Initial analyses of the three data set with thresholds 0.95, 0.97, and 0.98 yielded identical tree topologies, so more thorough analyses were subsequently performed only on the data set with an exclusion threshold of 0.98. We ran 1000 bootstrap replicates, using the rapid hill climbing algorithm and a partitioned model where each protein family was considered as a separate partition. Bootstrap support values were mapped to the ML tree using the SumTrees script of the Dendropy package (Sukumaran and Holder 2010). Bayesian MCMC inference was performed in PhyloBayes 3.3 (Lartillot et al. 2013) with 3 replicates of one chain each for 10,000 cycles. The CAT mixture model of protein evolution was used and the concatenated alignment was treated as one partition. Burn-in values were established by checking convergence in likelihood values and clade posterior probabilities. Trees from three independent runs (excluding the burn-in) were summarized in a 50% majority rule consensus tree by using SumTrees.

Analyses of Gene Family Evolution in Decay-Related Gene Families

We included 18 families of PODs and CAZys in the analyses following Floudas et al. (2015). We divided the gene families into three major groups (table 2), oxidative enzymes that are putatively involved in lignin decomposition (AA2, DyP, HTP, AA1_1, AA1_3), CAZys active on polysaccharide main chains (GH5, GH6, GH7, GH10, GH11, GH12, GH28, GH45, and AA9 LPMOs), as well as other CAZys related to wood decay (GH3, GH43, CE1, and CE16). We assembled data sets for 18 gene families involved in PCW degradation. Each data set was assembled using InterPro and PFAM domains and the JGI cluster run http://genome.jgi-psf.org/clm/run/agaricomycotina.34;JsZNg?organismsGroup=agaricomycotina (last accessed July 6, 2015). In order to identify the clusters that contain the copies for POD, GH5-5, GH5-7, GH74, GH51, CE1, and CE16, we used annotated proteins belonging in those families (table 2; Floudas et al. 2012; Kohler et al. 2015). We also identified the CBM1 copies found in each genome by searching for the corresponding PFAM domain PF00734

Gene Phylogenies and Gene Tree–Species Tree Reconciliation

To reconstruct gene family expansions and contractions in the selected families, we first aligned each gene family using the Probabilistic Alignment Kit (PRANK) v. 111130 (Loytynoja and Goldman 2008) using default parameters. Overly divergent sequences, gene fragments, low-scoring alignment regions, as well as leading and trailing gaps were removed manually from the alignments. Subsequently, we inferred ML gene trees using RAxML (see above) and performed gene tree species tree reconciliations using TreeFix (Wu et al. 2013). TreeFix was launched with the manually trimmed alignments, the estimated ML gene tree, and the ML organismal phylogenetic tree as input. We used the default reconciliation model (duplication/loss cost) and RAxML as the estimator of site-wise likelihoods. We allowed rerooting of the input gene trees during the analyses if it decreased the duplication/loss score. *P* values were obtained by the Shimodaira–Hasegawa test (Goldman et al. 2000) and a 0.05 significance cutoff was applied. We used the reconciled gene trees to map gene duplications and losses onto the organismal phylogeny using the corresponding scripts of the COMPARE pipeline (Nagy et al. 2014).

To control for the effects of general genome dynamics through evolution and potential sequencing-related copy number artifacts, we have also obtained a duplication/loss mapping of a random sample of gene trees from the 62 genomes. For this, we randomly chose 9,000 gene families, inferred multiple alignments and ML gene trees (as above) for each and reconciled them by using TreeFix with the parameters described above. Subsequently, we mapped duplications and losses onto the organismal phylogeny using Dollo-parsimony and reconstructed the number of duplications and losses for each branch of the species tree.

For class II PODs, we have also performed gene tree–species tree reconciliation using Notung 2.6 (Chen et al. 2000). For the Notung analyses, we ran a ML bootstrap analysis first with 100 bootstrap replicates on the manually trimmed alignment in RAxML 7.2.8. This bootstrapped gene tree was then reconciled with the species tree in Notung using an edge weight threshold of 90% and the inferred duplications and losses mapped onto the species tree as described above.

Supplementary Material

Supplementary files S1 and S2 and figures S1–S9 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We are grateful to Liam Revell for his help with the PhyTools package. L.G.N. was supported by the Lendulet Programme of the Hungarian Academy of Sciences under contract no. LP2014/12-2014. This research was supported in part by US National Science Foundation awards DEB-0933081 and DEB-1208719 to D.S.H. The work conducted by the US. Department of Energy Joint Genome Institute is supported by the Office of Science of the US. Department of Energy under contract no. DE-AC02-05CH11231.

References

Adav SS, Chao LT, Sze SK. 2012. Quantitative secretomic analysis of *Trichoderma reesei* strains reveals enzymatic composition for lignocellulosic biomass degradation. *Mol Cell Proteomics*. 11:M111.012419.

Aspeborg H, Coutinho PM, Wang Y, Brumer H, Henrissat B 2012. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). *BMC Evol Biol.* 12:186.

Binder M, Hibbett DS. 2002. Higher-level phylogenetic relationships of Homobasidiomycetes (mushroom-forming fungi) inferred from four rDNA regions. *Mol Phylogenet Evol*. 22:76–90.

Binder M, Hibbett DS, Larsson K-H, Larsson E, Langer E. 2005. The phylogenetic distribution of resupinate forms in the homobasidiomycetes. *Systematics and Biodiversity*. 3:113–157.

Chen K, Durand D, Farach-Colton M. 2000. NOTUNG: a program for dating gene duplications and optimizing gene family trees. *J Comput Biol*. 7:429–447.

de Vries RP, de Lange ES, Wosten HA, Stalpers JA. 2008. Control and possible applications of a novel carrot-spoilage basidiomycete, *Fibulorhizoctonia psychrophila*. *Antonie Van Leeuwenhoek* 93: 407–413.

Dunham SM, Larsson KH, Spatafora JW. 2007. Species richness and community composition of mat-forming ectomycorrhizal fungi in old- and second-growth Douglas-fir forests of the HJ Andrews Experimental Forest, Oregon, USA. *Mycorrhiza* 17:633–645.

English AC, Richards S, Han Y, Wang M, Vee V, Qu J, Qin X, Muzny DM, Reid JG, Worley KC, et al. 2012. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS One* 7:e47768.

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martinez AT, Otillar R, Spatafora JW, Yadav JS, et al. 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336:1715–1719.

Floudas D, Held BW, Riley R, Nagy LG, Koehler G, Ransdell AS, Younus H, Chow J, Chiniquy J, Lipzen A, et al. 2015. Evolution of novel wood decay mechanisms in Agaricales revealed by the genome sequences of *Fistulina hepatica* and *Cylindrobasidium torrendii*. *Fungal Genet Biol*. 76:78–92.

Ginns J, Lefebvre MNL. 1993. Lignicolous corticioid fungi of North America. *Mycol Mem*. 19:1–247.

Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, et al. 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A*. 108:1513–1518.

Goldman N, Anderson JP, Rodrigo AG. 2000. Likelihood-based tests of topologies in phylogenetics. *Syst Biol*. 49:652–670.

Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, et al. 2014. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res*. 42: D699–D704.

Hibbett DS, Pine EM, Langer E, Langer G, Donoghue MJ. 1997. Evolution of gilled mushrooms and puffballs inferred from ribosomal DNA sequences. *Proc Natl Acad Sci U S A*. 94:12002–12006.

Hofrichter M, Ullrich R, Pecyna MJ, Liers C, Lundell T. 2010. New and classic families of secreted fungal heme peroxidases. *Appl Microbiol Biotechnol*. 87:871–897.

James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, et al. 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443:818–822.

Kohler A, Kuo A, Nagy LG, Morin E, Barry KW, Buscot F, Canback B, Choi C, Cichocki N, Clum A, et al. 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nat Genet*. 47:410–415.

Kottke I, Nebel M. 2005. The evolution of mycorrhiza-like associations in liverworts: an update. *New Phytol*. 167:330–334.

Larsson KH. 2007. Re-thinking the classification of corticioid fungi. *Mycol Res*. 111:1040–1063.

Larsson KH, Larsson E, Koljalg U. 2004. High phylogenetic diversity among corticioid homobasidiomycetes. *Mycol Res*. 108:983–1002.

Lartillot N, Rodrigue N, Stubbs D, Richer J. 2013. PhyloBayes MPI: phylogenetic reconstruction with infinite mixtures of profiles in a parallel environment. *Syst Biol*. 62:611–615.

Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels*. 6:41.

Levasseur A, Saloheimo M, Navarro D, Andberg M, Pontarotti P, Kruus K, Record E. 2010. Exploring laccase-like multicopper oxidase genes from the ascomycete *Trichoderma reesei*: a functional, phylogenetic and evolutionary study. *BMC Biochem*. 11:32.

Loytynoja A, Goldman N. 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* 320:1632–1635.

Lundell T, Makela MR, de Vries RP, Hilden K. 2014. Genomics, lifestyles and future prospects of wood-decay and litter-decomposing basidiomycota. In: Martin FM, editor. Advances in Botanical Research, Vol. 70. Fungi. London: Academic. p. 329–370.

Martin F, Aerts A, Ahren D, Brun A, Danchin EG, Duchaussoy F, Gibon J, Kohler A, Lindquist E, Pereda V, et al. 2008. The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452:88–92.

Martin J, Bruno VM, Fang Z, Meng X, Blow M, Zhang T, Sherlock G, Snyder M, Wang Z. 2010. Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. *BMC Genomics* 11:663.

Martinez D, Challacombe J, Morgenstern I, Hibbett D, Schmoll M, Kubicek CP, Ferreira P, Ruiz-Duenas FJ, Martinez AT, Kersten P, et al. 2009. Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion. *Proc Natl Acad Sci U S A*. 106:1954–1959.

Martinez D, Larrondo LF, Putnam N, Gelpke MD, Huang K, Chapman J, Helfenbein KG, Ramaiya P, Detter JC, Larimer F, et al. 2004. Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nat Biotechnol*. 22:695–700.

Matheny PB, Wang Z, Binder M, Curtis JM, Lim YW, Nilsson RH, Hughes KW, Hofstetter V, Ammirati JF, Schoch CL, et al. 2007. Contributions of rpb2 and tef1 to the phylogeny of mushrooms and allies (Basidiomycota, Fungi). *Mol Phylogenet Evol*. 43:430–451.

Matsuura K. 2006. Termite-egg mimicry by a sclerotium-forming fungus. *Proc Biol Sci*. 273:1203–1209.

Matsuura K, Tanaka C, Nishida T. 2000. Symbiosis of a termite and a sclerotium-forming fungus: sclerotia mimic termite eggs. *Ecol Res*. 15:405–414.

Moncalvo JM, Nilsson RH, Koster B, Dunham SM, Bernauer T, Matheny PB, Porter TM, Margaritescu S, Weiss M, Garnica S, et al. 2006. The cantharelloid clade: dealing with incongruent gene trees and phylogenetic reconstruction methods. *Mycologia* 98:937–948.

Morgenstern I, Powlowski J, Tsang A. 2014. Fungal cellulose degradation by oxidative enzymes: from dysfunctional GH61 family to powerful lytic polysaccharide monooxygenase family. *Brief Funct Genomics*. 13:471–481.

Nagy LG, Ohm RA, Kovacs GM, Floudas D, Riley R, Gacser A, Sipiczki M, Davis JM, Doty SL, de Hoog GS, et al. 2014. Latent homology and convergent regulatory evolution underlies the repeated emergence of yeasts. *Nat Commun*. 5:4471.

Nakasone K. 1990. Cultural studies and identification of wood-inhabiting Corticiaceae and selected Hymenomycetes from North America. *Mycol Mem*. 15:1–412.

Nelson DR. 2006. Cytochrome P450 nomenclature. *Mycol Mem*. 320:1–10.

Poggeler S. 2011. Evolution of multicopper oxidase genes in coprophilous and non-coprophilous members of the order sordariales. *Curr Genomics*. 12:95–103.

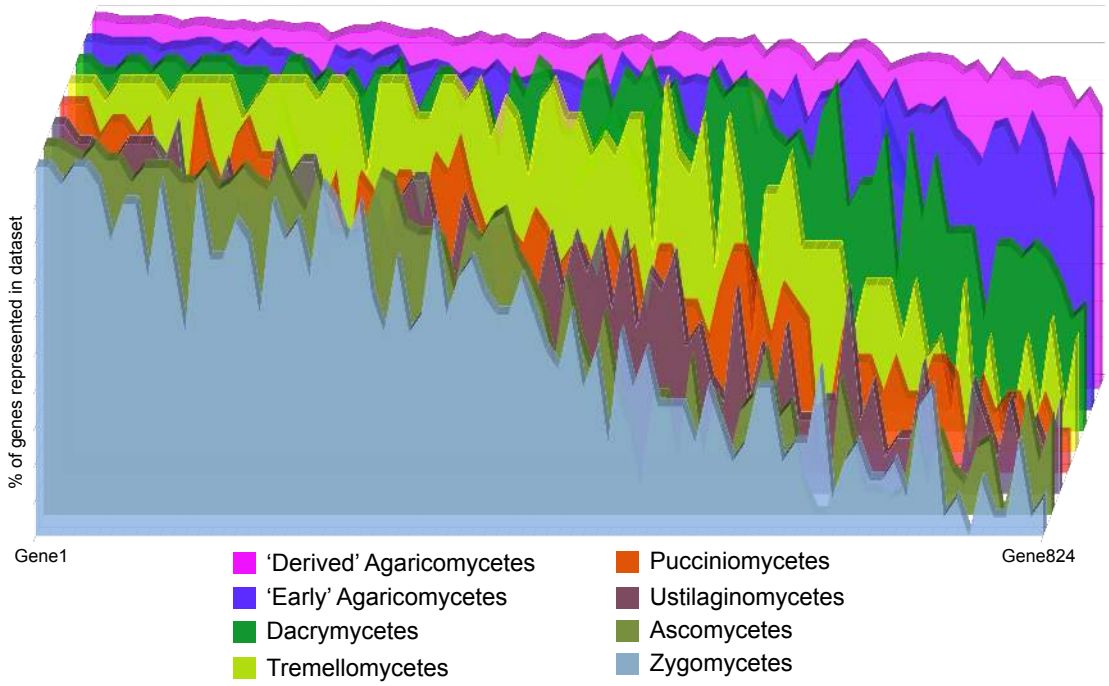
Riley R, Salamov AA, Brown DW, Nagy LG, Floudas D, Held BW, Levasseur A, Lombard V, Morin E, Otillar R, et al. 2014. Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. *Proc Natl Acad Sci U S A*. 111:9923–9928.

Robinson J. 1990. Lignin, land plants, and fungi: biological evolution affecting Phanerozoic oxygen balance. *Geology* 18:607–610.

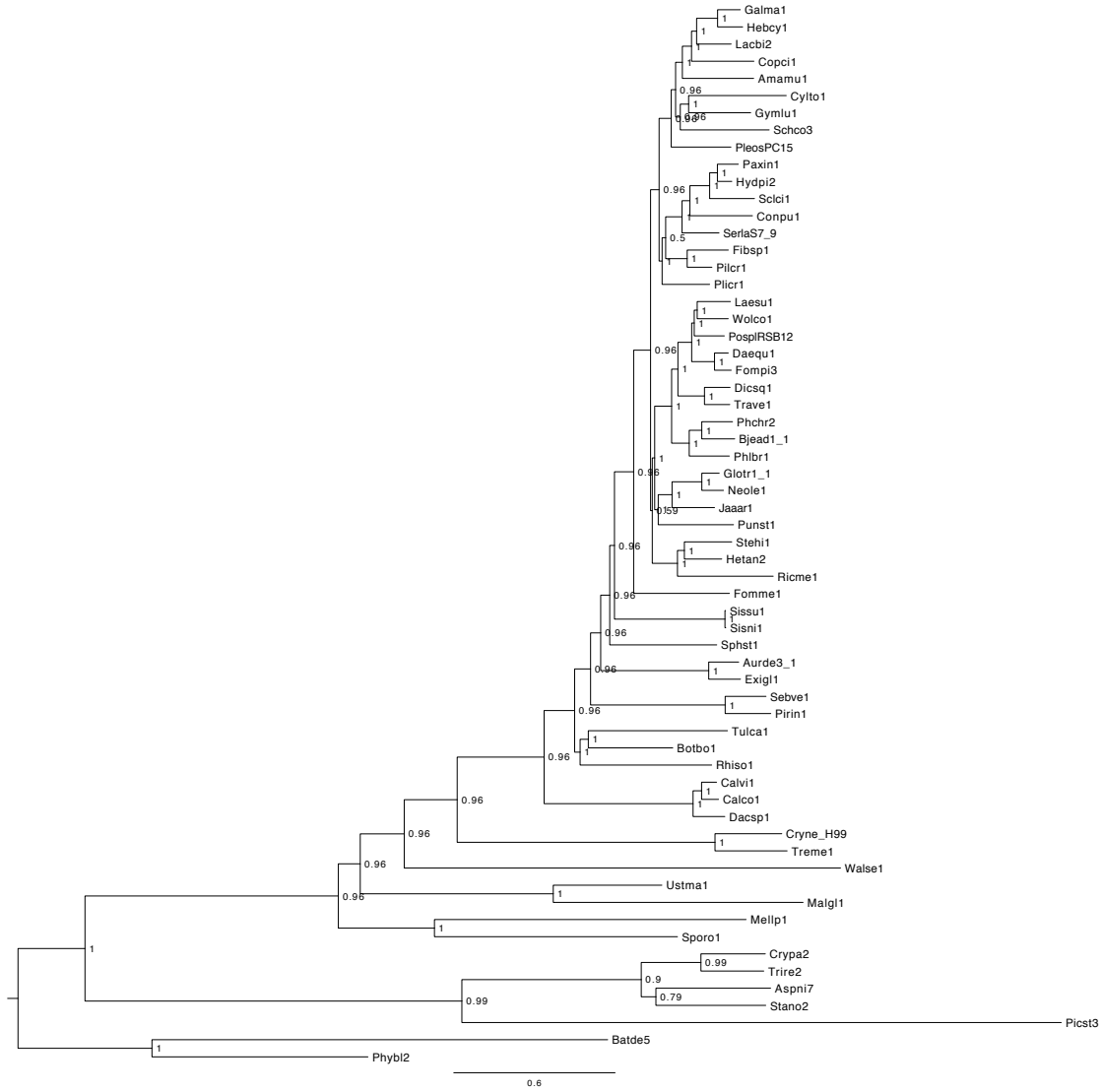
Ruiz-Duenas FJ, Fernandez E, Martinez MJ, Martinez AT. 2011. *Pleurotus ostreatus* heme peroxidases: an in silico analysis from the genome sequence to the enzyme molecular structure. *C R Biol*. 334:795–805.

Ruiz-Duenas FJ, Lundell T, Floudas D, Nagy LG, Barrasa JM, Hibbett DS, Martinez AT. 2013. Lignin-degrading peroxidases in Polyporales: an

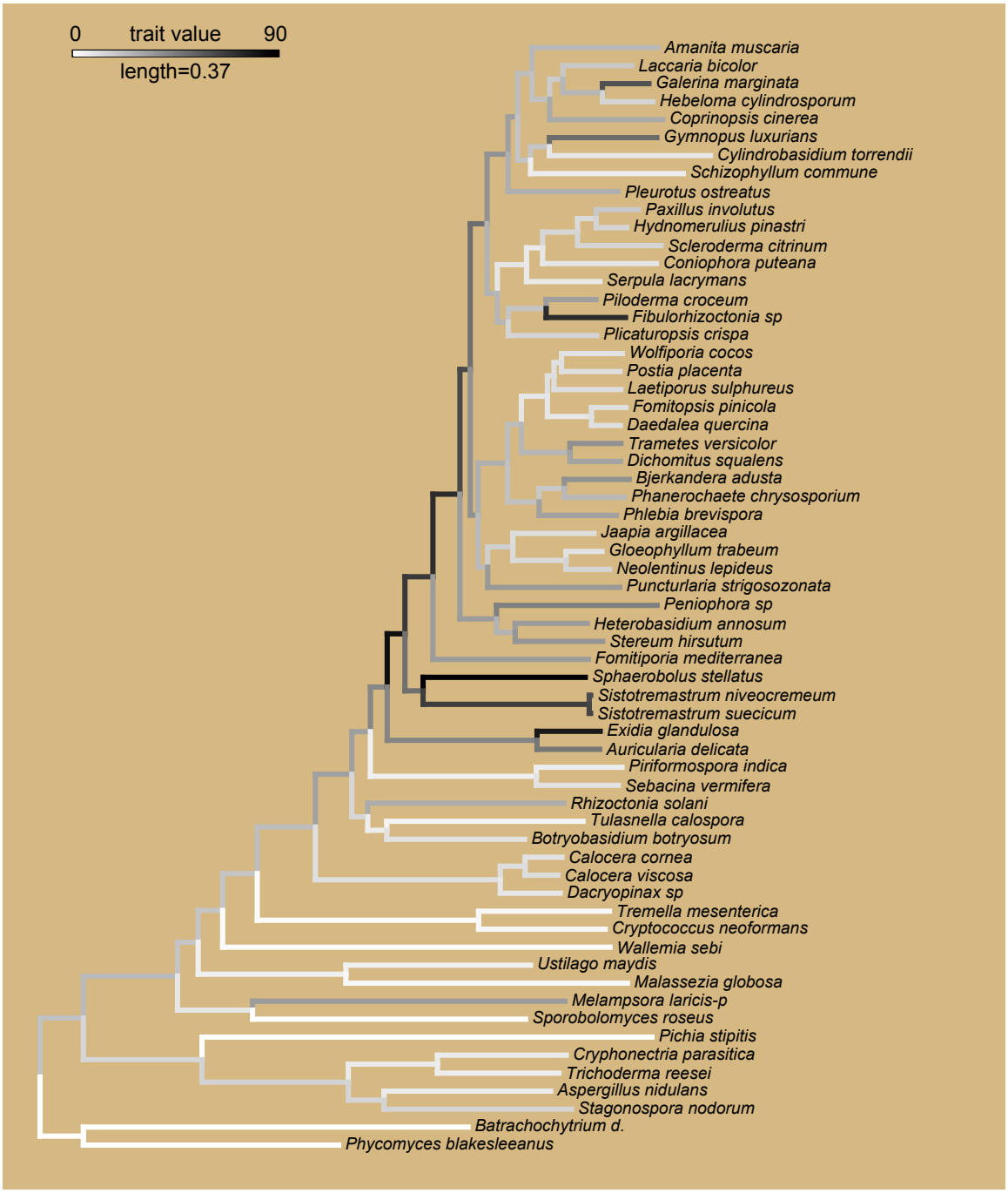
evolutionary survey based on 10 sequenced genomes. *Mycologia* 105:1428–1444.
Rytioja J, Hilden K, Yuzon J, Hatakka A, de Vries RP, Makela MR. 2014. Plant-polysaccharide-degrading enzymes from Basidiomycetes. *Microbiol Mol Biol Rev*. 78:614–649.
Ryvarden L, Gilbertson RL. 1993. European polypores. Vol. 1–2. Oslo (Norway): Fungiflora.
Seifert K. 1983. Decay of wood by the Dacrymycetales. *Mycologia* 75:1011–1018.
Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22:2688–2690.
Suarez JP, Weiss M, Abele A, Garnica S, Oberwinkler F, Kottke I. 2006. Diverse tulasnelloid fungi form mycorrhizas with epiphytic orchids in an Andean cloud forest. *Mycol Res*. 110:1257–1270.
Sukumaran J, Holder MT. 2010. DendroPy: a Python library for phylogenetic computing. *Bioinformatics* 26:1569–1571.
Vaaje-Kolstad G, Westereng B, Horn SJ, Liu ZL, Zhai H, Sorlie M, Eijsink VGH. 2010. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* 330: 219–222.
van den Brink J, de Vries RP. 2011. Fungal enzyme sets for plant polysaccharide degradation. *Appl Microbiol Biotechnol*. 91: 1477–1492.
van Dongen S. 2000. Graph clustering by flow simulation. PhD thesis. Utrecht (The Netherlands): University of Utrecht.
Worrall J, Anagnost S, Zabel RA. 1997. Comparison of wood decay among diverse lignicolous fungi. *Mycologia* 89:199–219.
Wu YC, Rasmussen MD, Bansal MS, Kellis M. 2013. TreeFix: statistically informed gene tree error correction using species trees. *Syst Biol*. 62:110–120.
Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res*. 18: 821–829.



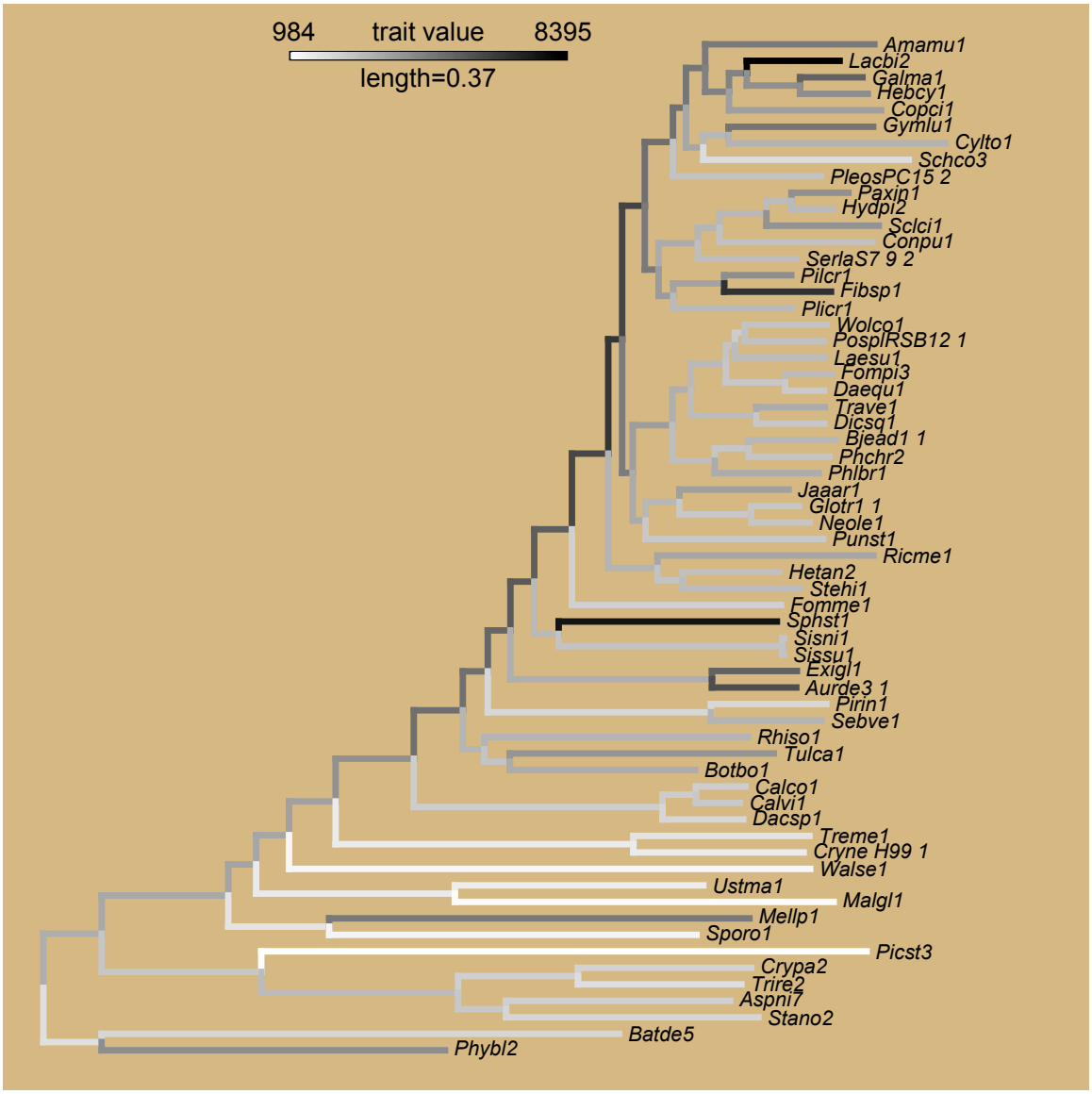
FigS1. Taxon occupancy across 824 gene families that were single-copy across the 62 genomes or contained inparalogs but no deep paralogs.



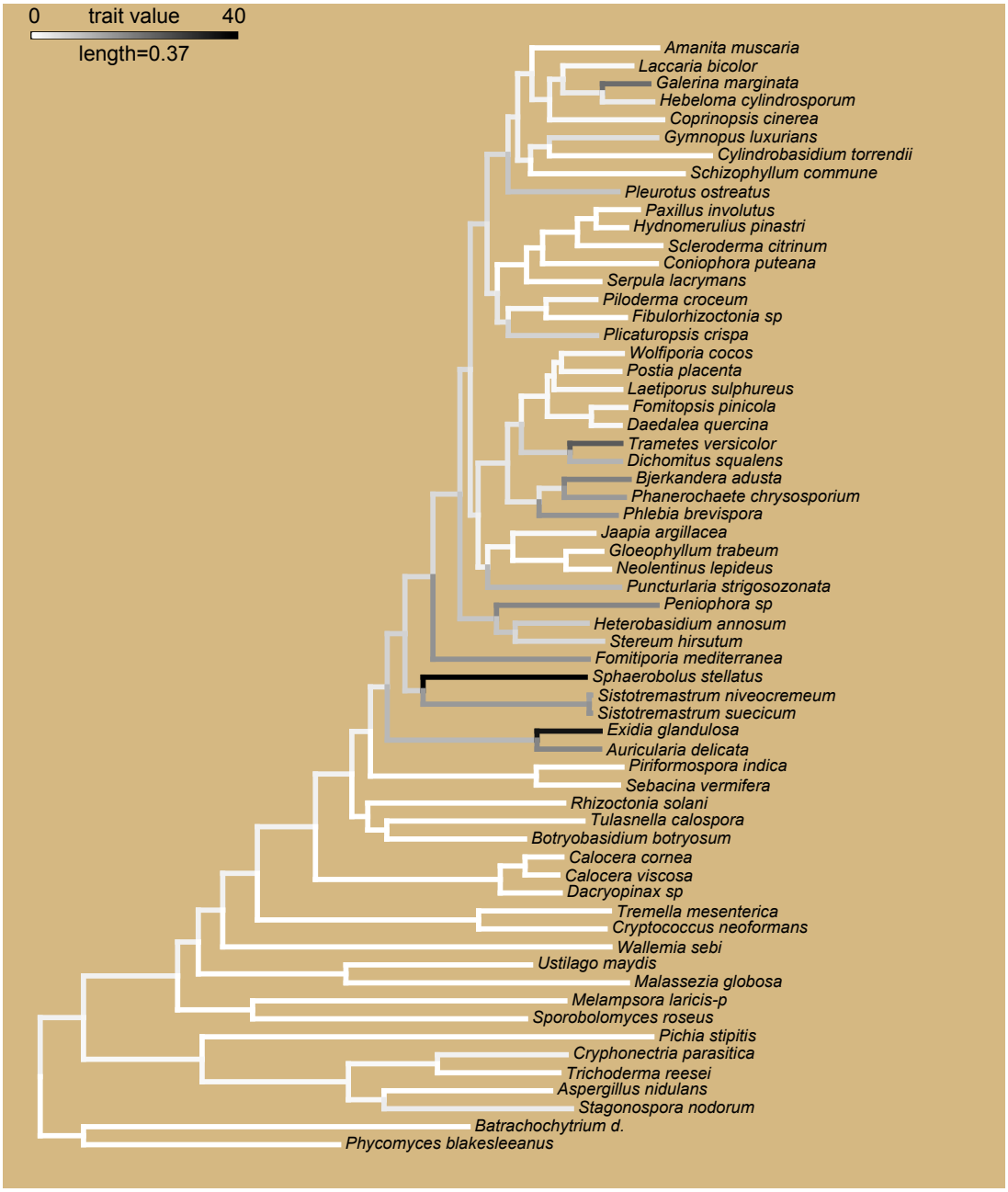
FigS2. Bayesian 50% Majority Rule consensus tree inferred using PhyloBayes.



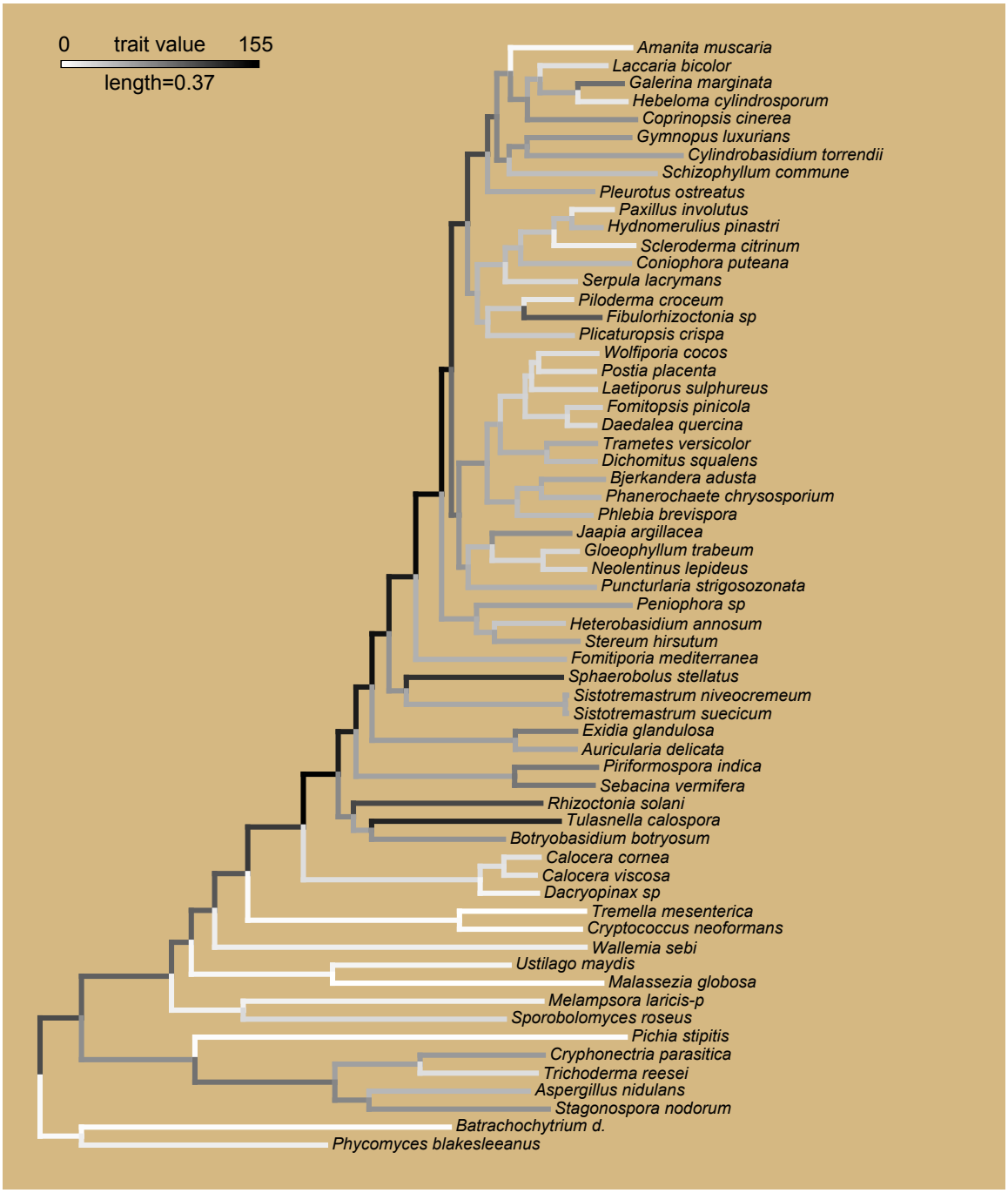
FigS3. Oxidative CAZYs. Note: For visualization, the copy number in *Sphaerobolus* (325) has been truncated at 90 for better visibility of shades.



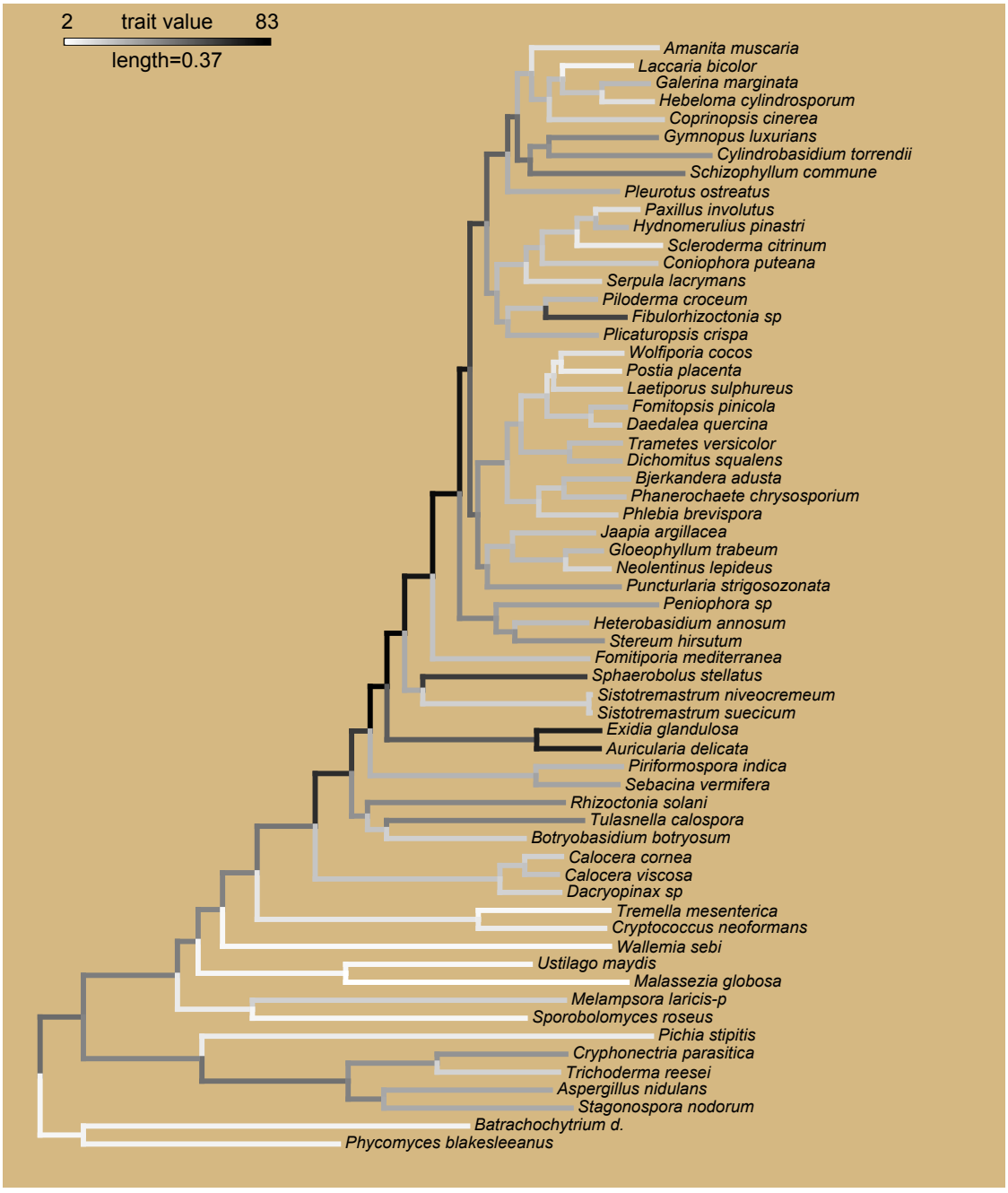
FigS4. Evolutionary dynamics of gene content in Agaricomycete genomes. To investigate whether the early expansion we inferred for oxidative CAZy-s represents a departure from the background genome-dynamics, we reconstructed copy number evolution in a large random sample of gene families not relevant to lignocellulose decomposition (supplementary data x for the list of gene families). This reconstruction shows a markedly different pattern with a gradual increase in mean gene family size along the backbone of the Agaricomycetes, with the highest diversity of orthogroups (6863 orthogroups) inferred for the mrca of the Polyporales, Boletales and Agaricales.



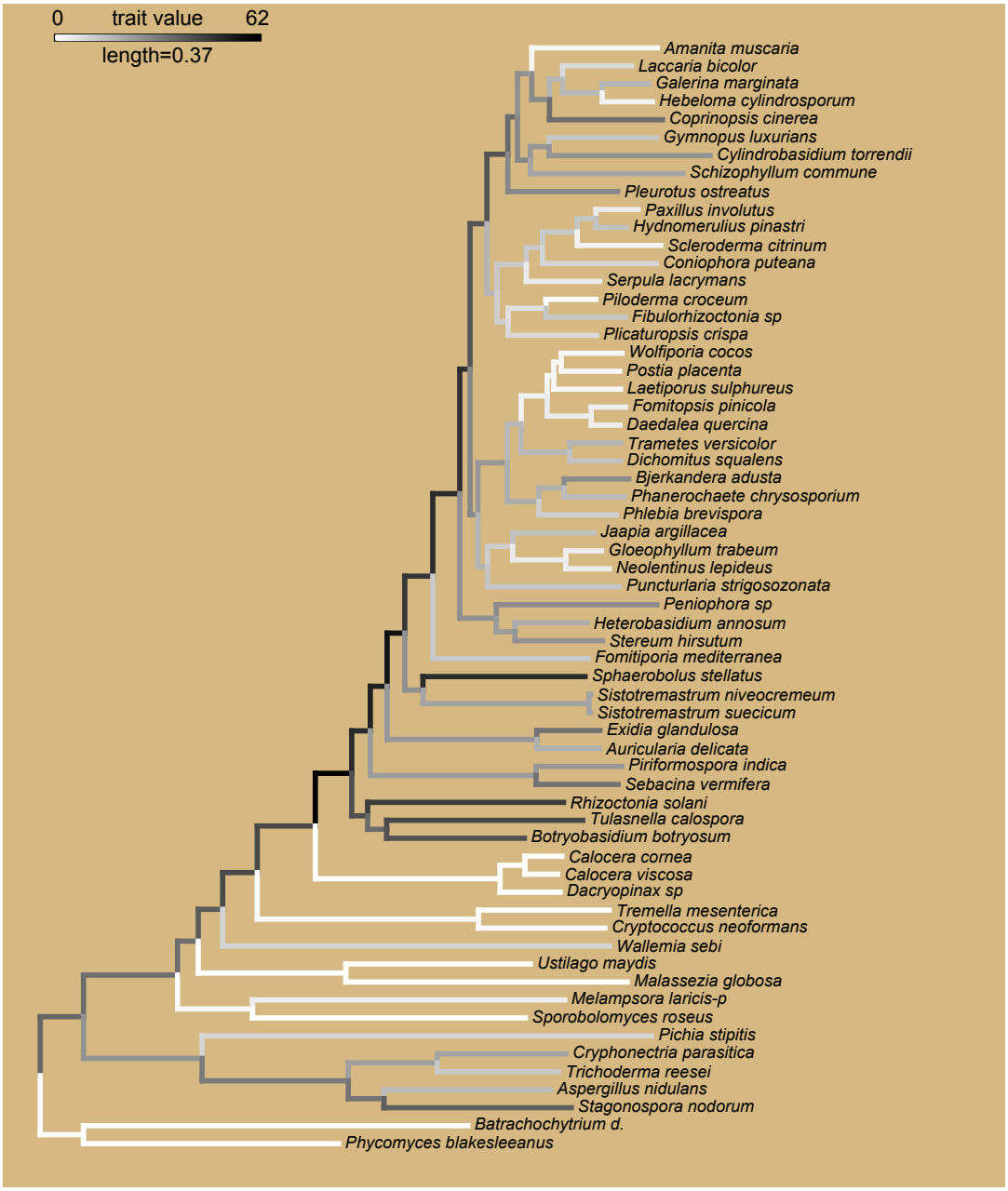
FigS5. Copy number evolution of Class-II-peroxidases as reconstructed by Notung. For visualization, copy number in *Sphaerobolus* (68) has been truncated to 40 for better visibility of shades.



FigS6. Copy number evolution of CAZY families active on polysaccharide main chains.



FigS7. Copy number evolution of Other CAZY families related to wood-decay.



FigS8. Copy number evolution of Lytic Polysaccharide Monooxygenases.

FigS9. Copy number evolution of cytochrome p450-s.