Supporting Information

Riley et al. 10.1073/pnas.1400592111

SI Text

Genome Annotation

Before gene prediction, assembly scaffolds were masked using RepeatMasker (1), RepBase library (2), and most frequent (>150 times) repeats recognized by RepeatScout (3). The following combination of gene predictors was run on the masked assembly: ab initio Fgenesh (4) and GeneMark (5), homology-based Fgenesh+ (4) and Genewise (6) seeded by BLASTx (7) alignments against National Center for Biotechnology Information (NCBI) NR database, and transcriptome-based assemblies. In addition to protein-coding genes, tRNAs were predicted using tRNAscan-SE (8). All predicted proteins are functionally annotated using SignalP (9) for signal sequences, TMHMM (10) for transmembrane domains, InterProScan (11) for integrated collection of functional and structure protein domains, and protein alignments to NCBI nr, SwissProt (www.expasy.org/sprot/), KEGG (12) for metabolic pathways, and KOG (13) for eukaryotic clusters of orthologs. Interpro and SwissProt hits are used to map Gene Ontology (GO) terms (14). For each genomic locus, the best representative gene model was selected based on a combination of protein homology and EST support, which resulted in the final set of 6,903 gene models used for further analysis in this work.

Genomes Overview

Genome sizes vary over an order of magnitude in Basidiomycota (Fig. S1 and Table S1). The plant-pathogenic rusts (15) *Puccinia graminis* (88.6 Mb) and *Melampsora laricis-populina* (101.1 Mb), along with the early-diverging Agaricomycete *Auricularia delicata* (16) (74.9 Mb) feature the largest genomes, whereas the human pathogen *Malassezia globosa* (17) (9.0 Mb) and xerophilic mold *Wallemia sebi* (18) (9.8 Mb) have the smallest genomes.

Protein Clusters and Phylogeny

Gene families in 33 basidiomycetes and 30 other fungi (Table S2) with sequenced genomes were inferred by Markov chain (MCL) clustering (19) of all-vs.-all protein BLAST (7) alignments. Two clustering runs were performed. The first clustering run, used for core genes and conservation analysis, used 765,862 protein sequences resulting in 121,327 clusters and is visible at the following: http://genome.jgi-psf.org/clustering/pages/cluster/clusters.jsf?runId=2655. The second run, used for building the phylogenetic tree, used 472,010 protein sequences resulting in 73,519 clusters, and is visible at the following: http://genome.jgi-psf.org/clustering/pages/cluster/clusters.jsf?runId=2656. From this second cluster run, 183 clusters, in which each organism contributed a single protein sequence, were extracted for subsequent use in inferring the phylogeny. (Dataset S3 contains one cluster per line, with each protein in a cluster denoted by its Joint Genome Institute (JGI) protein id and JGI portal id separated by the "|" symbol).

The maximum-likelihood phylogeny (20), inferred from the protein sequences of 183 conserved gene families, along with an overview of genome size, repeat content, gene number, and gene conservation is shown in Fig. S1. The earliest-diverging nodes of the Agaricomycetes remain unclear, in particular the position of *Piriformospora indica* (Sebacinales) and the position of *Jaapia argillacea* (Jaapiales). The former species was inferred as monophyletic with *Auricularia delicata*, whereas previous multigene phylogenies (21–24) placed them in different clades along the backbone of the Agaricomycetes. The position of the Jaapiales is likewise somewhat uncertain; it forms a well-supported clade with *Gloeophyllum trabeum* and *Punctularia strigosozonata*, but that clade's position on the tree, either as a sister clade of the Polyporales, or of the clade containing the Russulales, Boletales, and Agaricales, is uncertain.

The protein clusters' KOG (13) annotations suggest that one-half of the proteins in Basidiomycota have no predicted function (Fig. S2). However, only 8% of the proteins in the "core proteome" (i.e., MCL clusters that have at least one member in all Basidiomycota) have no KOG annotation, suggesting we can predict functions for some 92% of the core proteins. In contrast, 78% of "noncore" proteins (those present in some, but not all basidiomycetes) have no KOG annotation. Protein families sporadically present in basidiomycetes are therefore mostly of unknown function and may provide clues to the unique adaptations of basidiomycete lineages.

Secondary Metabolism

Phylogenetic relationships of Basidiomycete and Ascomycete polyketide synthases (PKSs) within and among species were examined by maximum parsimony analysis of deduced amino acid sequences of 225 keto synthase (KS) domains (PF00109.17 and PF02801.13) identified in 35 Basidiomycete (the 33 used in the rest of this analysis plus *Rhodotorula graminis* and *Sporobolomyces roseus* downloaded from MycoCosm: http://jgi.doe.gov/fungi) and four Ascomycete genomes (*Aspergillus niger*, *Pichia stipitis*, *Stagonospora nodorum*, and *Trichoderma reesei*). The KS from the *Gallus gallus* fatty acid synthase (FAS) served as an outgroup. AA sequence alignments were generated with the ClustalW (25) using the Blosum multiple sequence alignment scoring matrix (26). The aligned sequences were then used to construct a gene genealogy using parsimony in PAUP* 4.0b10 (27). Statistical support for branches was generated by bootstrap analysis with 1,000 pseudoreplications.

To investigate the secondary metabolite biosynthetic potential of the Basidiomycetes, we examined the evolution of the KS domain, a key domain involved in polyketide and fatty acid synthesis, from the predicted protein sequence of 144 KS domains from 35 Basidiomycetes and 81 KS domains from four Ascomycetes. Maximum parsimony analysis of the amino acid alignment resolved the predicted peptides into five major groups (Fig. S3) corresponding to FASs, nonreducing type PKSs (NR-PKSs), and reducing type PKSs (R-PKSs) and is consistent with previous work (28–30).

BLASTP analysis against the NCBI NR database with KSs, represented by the dark blue and light blue triangles in Fig. S3, match numerous yeast and fungal FASs ($E < 1 \times 10^{-100}$), suggesting that the query proteins are involved in the synthesis of a fully reduced carbon chain typical of constitutive fatty acid biosynthesis. The largest clade (large dark blue triangle), with 79% bootstrap support, includes 38 KS domains from 34 different Basidiomycetes (Fig. S3, large blue triangle). Most Basidiomycetes possess one FAS gene; four Basidiomycetes have two FASs, whereas one, *Malassezia globosa*, lacks a FAS. *M. globosa* is associated with most skin disease in humans, including dandruff, and presumably does not need a FAS as it uses fatty acids present in sebaceous gland secretions (31). The clade with one KS from each Ascomycete examined in this study (small dark blue triangle), likely represents constitutive FASs. Three branches, represented by light blue triangles, include five KSs from two Ascomycetes; four from *A. niger* and one from *S. nodorum*. Products from these FAS-like proteins have not yet been determined. Adjacent to the FAS branches is a branch with three Ascomycete KSs (dark pink triangle), with 99% support, that by BLASTP, are similar to the keto acyl synthases or type III PKSs involved in secondary metabolite synthesis in *Aspergillus* species (32).

Domains adjacent to the KS domain present in the "PKS and PKS-like" group are typically found in nonribosomal peptide synthetetases (NRPSs). In contrast to most previously described PKSs with an NRPS module in Ascomycetes, the NRPS domain (s) here are located at the amino terminus rather than the carboxyl terminus. The first group (green) includes three Basidiomycete KS domains, which, in each case, is adjacent to a condensation (C) (PF00668.11), AMP-binding (PF00501.19), and phosphopantetheine-binding (PP-binding) (PF00550.16) domain. The next group (gray), with 83% bootstrap support, includes 14 Basidiomycete KSs present in three sister clades, each with 100% bootstrap support. The domain organization of the predicted proteins associated with each clade is significantly different but consistent within each clade. For example, the predicted proteins in the first clade include an AMP-binding domain adjacent to the KS followed by an acyl transferase (AT) (PF00698.12) and a keto reductase (KR) (PF08659.1) domain. In contrast, most of the predicted proteins in the second clade include an AMP-binding domain adjacent to the KS but lack AT and KR domains, whereas the predicted proteins in the third clade lack the AMP-binding domain but have an epimerase domain (PF01370.12) before the AT domain and lack a KR domain. The observation that five of the six fungi with KSs in this clade are white-rot fungi may suggest that the chemical product(s) generated may be important for a subset of fungi with this lifestyle, despite their distant relatedness. Adjacent to the "PKS and PKS-like group" is a branch (light pink) with two Ascomycete KSs, which BLASTP analysis suggest are homologs of 6-methylsalicylic acid synthase from *Penicillium griseofulvum* ($E = 1 \times 10^{-154}$).

The NR-PKS group consists of 48 KSs from both Basidiomycetes and Ascomycetes in a single clade with 90% bootstrap support. A majority of the Basidiomycetes KSs (26) forms a clade that is sister to a clade, with 93% support, composed of both Basidiomycete (4) and Ascomycete (10) KSs. The remaining seven Ascomycete KSs belong to a more basal clade, with 82% support. Of the 27 KSs in the Basidiomycete clade, two pairs of KSs have 100% bootstrap support, which may reflect recent gene duplication events. The 27 KSs are very well distributed with 87% of white-rot and 71% brown-rot fungi having at least one member. This group also appears to be ancient as only a few branches have significant bootstrap support.

The KSs corresponding to R-PKSs separate into two major groups corresponding to their phylum. For example, all 49 Ascomycetes KSs, representing the three previously described R-PKS clades I, II, and III (28–30), form one group, whereas the 58 Basidiomycete KSs form the other group. The limited number of orthologous gene pairs in the Ascomycete set was previously noted for a much larger KS set and it was suggested that recent gene duplication events have not contributed significantly to the expansion of this gene family (28, 29). The 58 KSs from 19 Basidiomycetes form two sister clades with 100% (42 KSs) and 64% (16 KSs) bootstrap support. Most of the KSs of this group (96%) are from wood-decaying fungi. In contrast to the Ascomycete R-PKSs noted above, nine pairs of KSs from the same fungus share between 64% and 100% bootstrap support (average, 91%), suggesting that gene duplication events contributed more significantly to PKS the expansion of this gene family. Analysis of additional, more closely related Basidiomycete genomes may suggest how long in the evolutionary past these events occurred.

Clustering of Auxiliary Redox Enzymes

The AA families of auxiliary redox enzymes are thought to break down cell wall components, including lignin (33). To better understand the role of these enzymes in wood-decaying basidiomycetes, we used the CAFE program (34) to explore their evolution according to a stochastic model of gene death and birth. Additionally, we performed hierarchical clustering of both AA families and organisms, using a table of the number of genes in each family. As done elsewhere in this study, we used a manually curated subset of the AA2 family, limited to the predicted high-oxidation potential peroxidases (PODs): lignin peroxidase (LiP), manganese peroxidase (MnP), and versatile peroxidase (VP). We refer to this subset as "POD."

The CAFE analysis revealed POD, AA3_2 (GMC oxidoreductase), and AA9 [lytic polysaccharide monooxygenase (LPMO)] as significantly departing from a random model of gene birth and death, using a family-wide significance threshold of 0.05, implying lineage-specific shifts of the duplication and loss of these genes, presumably in accordance with lifestyle. The CAFE analysis also allowed the inference of statistically significant gene family gains and losses in the various lineages. Fig. S4 shows a heat map of the number of genes in each AA family for each organism, along with the number of genes inferred to be gained or lost since that organism's split with its nearest neighbor. It is clear that the POD, AA3_2, and AA9 families have generally undergone gains in the white-rot lineages and losses in the brown-rot, mycorrhizal, and soil saprotrophic lineages, consistent with an earlier report (16). Interestingly, *Botryobasidium botryosum*, *J. argillacea* have substantial gains in the AA9 family, and *Schizophyllum commune* is also rich in these enzymes, suggesting a possible heightened importance of oxidative attack on cellulose in fungi that lack ligninolytic PODs, but that nevertheless are capable of degrading all cell wall components. *S. commune* also has undergone reductions in the AA1_1 (laccase) and AA5_1 (copper radical oxidase) families, which are abundant in white-rot lineages, and which therefore highlight *S. commune*'s break from the white-rot/brown-rot dichotomy.

The results of the double-hierarchical clustering based on AA families shown in Fig. S5 is in full agreement with previous results based on the analysis of a different set of fungal genomes (33). Thus, families AA1_1, AA3_2, AA3_3, POD, and AA5_1 cluster together in Fig. S5 in accord with the known cooperation of these enzymes (33). A difference between the results of Levasseur et al. (33) and the analysis reported here is that our clustering also includes LPMOs, whereas Levasseur et al. focused exclusively on ligninolytic enzymes. Interestingly, family AA9 LPMOs clustered with families AA1_1, AA3_2, AA3_3, POD, and AA5_1, suggesting a certain degree of cooperation of the set of enzyme families during the breakdown of plant cell walls.

In an analysis based exclusively on the ligninolytic machinery, Levasseur et al. unexpectedly found *S. commune* in the BR group. Here, we show that the addition of family AA9 LPMOs to the oxidoreductase gene sets used for the clustering is sufficient to take *S. commune* out of the BR group and to place this fungus near *B. botryosum*. *J. argillacea* is most similar to white-rot fungus *A. delicata*.

Assays for Oxidoreductases Related to Degradation of Lignin

Three media were evaluated. Carbon-limited B3 and nitrogen-limited B3 medium (35, 36) were incubated statically for 5, 9, and 11 d. Cultures were flushed with O_2 after 2 d. Nutrient-starved B3 media have been widely used for production of LiP, MnP, and glyoxal oxidase. A more complex medium contained 0.5% (wt/vol) Wiley-mill ground *Populus grandidentata* (aspen) as sole carbon source in Highley's basal salts (37). These cultures were incubated on a rotary shaker at 150 rpm. In all cases, 20 mL of medium was inoculated with mycelium-covered agar plugs. *Phanerochaete chrysosporium* was incubated at 37 °C, and all other cultures remained at room temperature (~20 °C). Based on appearance and total extracellular protein, white-rot fungi *P. chrysosporium*, *Ceriporiopsis subvermispora* and brown-rot fungi *Wolfiporia cocos* and *Postia placenta* grew at the expected rates and were therefore harvested at day 5. The sequenced *Jaapia* *argillacea* MUCL-33604 SS and *Botryobasidium botryosum* FD172 SS-1 monokaryons exhibited relatively slower growth, so additional cultures were harvested after 9- and 11-d incubation.

Cultures were harvested by filtration through Miracloth (Calbiochem). The B3 culture filtrates were then passed through a 0.45-μm filter before concentration 6- to 10-fold with a 10-kDa cutoff Microsep spin concentrator (Pall). Filtrates from cultures containing ground *Populus* required a low-speed centrifugation before the 0.45-μm filter and Microsep concentration.

Protein concentration was determined by the Bradford assay (Sigma-Aldrich) according manufacturer's instructions. Measurement of MnP activity was based on the oxidative dimerization of 2,6-dimethoxyphenol (2,6-DMP) (38). The reaction mixture contained 100 μM 2,6-DMP, 100 μM $MnSO_4$, 50 mM sodium tartrate (pH 4.5), 50 μM H_2O_2, and culture filtrate in 1,000 μL. LiP activity was determined by conversion of veratryl alcohol (Sigma-Aldrich) to veratraldehyde in the presence of H_2O_2 (39). Fifty millimolar sodium tartrate (pH 3.0) served as buffer and the culture filtrate was mixed with 2 mM veratryl alcohol and 0.4 mM H_2O_2 in 1,000 μL at room temperature. Laccase (Lac) activity was assayed with 2,2′-azonodi-3-ethylbenzothiazoline-6-sulfuric acid (ABTS) (Boehringer) as a substrate in 30 mM glycine/HCl buffer (pH 3.0) at room temperature (40). The reaction contained 14 μM ABTS and culture filtrate in 1,000 μL. Copper radical oxidase activity was measured as optimized for glyoxal oxidase (GLOX) using methylglyoxal as substrate (41).

As expected, we observed MnP and LiP activity after 5 d in B3 cultures inoculated with the white-rot fungi *P. chrysosporium* and *C. subvermispora* (Table S4). Also consistent with many earlier studies, methylglyoxal oxidation was highest in *P. chrysosporium* cultures, likely due to GLOX. ABTS oxidation indicated laccase activity in *C. subvermispora* in B3 media. Consistent with the absence of PODs, veratryl alcohol and DMP oxidation could not be detected in the *J. argillacea* or *B. botryosum* culture filtrates. ABTS oxidation, typical of laccase activity, was also absent despite the presence of a laccase-encoding gene in the *B. botryosum* genome.

1. Smit AFA, Hubley R, Green P (1996–2010) RepeatMasker Open-3.0. Available at www.repeatmasker.org. Accessed November 2002.
2. Jurka J, et al. (2005) Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res* 110(1-4):462–467.
3. Price AL, Jones NC, Pevzner PA (2005) De novo identification of repeat families in large genomes. *Bioinformatics* 21(Suppl 1):i351–i358.
4. Salamov AA, Solovyev VV (2000) *Ab initio* gene finding in *Drosophila* genomic DNA. *Genome Res* 10(4):516–522.
5. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M (2008) Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res* 18(12):1979–1990.
6. Birney E, Clamp M, Durbin R (2004) GeneWise and Genomewise. *Genome Res* 14(5):988–995.
7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410.
8. Lowe TM, Eddy SR (1997) tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25(5):955–964.
9. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* 10(1):1–6.
10. Melén K, Krogh A, von Heijne G (2003) Reliability measures for membrane protein topology prediction algorithms. *J Mol Biol* 327(3):735–744.
11. Quevillon E, et al. (2005) InterProScan: Protein domains identifier. *Nucleic Acids Res* 33(Web Server issue):W116–W120.
12. Kanehisa M, et al. (2006) From genomics to chemical genomics: New developments in KEGG. *Nucleic Acids Res* 34(Database issue):D354–D357.
13. Koonin EV, et al. (2004) A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol* 5(2):R7.
14. Ashburner M, et al.; The Gene Ontology Consortium (2000) Gene ontology: Tool for the unification of biology. *Nat Genet* 25(1):25–29.
15. Duplessis S, et al. (2011) Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc Natl Acad Sci USA* 108(22):9166–9171.
16. Floudas D, et al. (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336(6089):1715–1719.
17. Xu J, et al. (2007) Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. *Proc Natl Acad Sci USA* 104(47):18730–18735.
18. Padamsee M, et al. (2012) The genome of the xerotolerant mold *Wallemia sebi* reveals adaptations to osmotic stress and suggests cryptic sexual reproduction. *Fungal Genet Biol* 49(3):217–226.
19. Enright AJ, Van Dongen S, Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30(7):1575–1584.
20. Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688–2690.
21. Hibbett DS (2006) A phylogenetic overview of the Agaricomycotina. *Mycologia* 98(6):917–925.
22. Matheny PB, Gossmann JA, Zalar P, Kumar TKA, Hibbett DS (2006) Resolving the phylogenetic position of the Wallemiomycetes: An enigmatic major lineage of Basidiomycota. *Can J Bot* 84(12):1794–1805.
23. Matheny PB, et al. (2007) Contributions of rpb2 and tef1 to the phylogeny of mushrooms and allies (Basidiomycota, Fungi). *Mol Phylogenet Evol* 43(2):430–451.
24. Hibbett DS, et al. (2007) A higher-level phylogenetic classification of the Fungi. *Mycol Res* 111(Pt 5):509–547.
25. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948.
26. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci USA* 89(22):10915–10919.
27. Swofford DL (2003) *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4* (Sinauer Associates, Sunderland, MA).
28. Kroken S, Glass NL, Taylor JW, Yoder OC, Turgeon BG (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc Natl Acad Sci USA* 100(26):15670–15675.
29. Baker SE, et al. (2006) Two polyketide synthase-encoding genes are required for biosynthesis of the polyketide virulence factor, T-toxin, by *Cochliobolus heterostrophus*. *Mol Plant Microbe Interact* 19(2):139–149.
30. Brown DW, Butchko RA, Baker SE, Proctor RH (2012) Phylogenomic and functional domain analysis of polyketide synthases in *Fusarium*. *Fungal Biol* 116(2):318–331.
31. Dawson TL, Jr. (2007) *Malassezia globosa* and *restricta*: Breakthrough understanding of the etiology and treatment of dandruff and seborrheic dermatitis through whole-genome analysis. *J Investig Dermatol Symp Proc* 12(2):15–19.
32. Seshime Y, Juvvadi PR, Kitamoto K, Ebizuka Y, Fujii I (2010) Identification of csypyrone B1 as the novel product of *Aspergillus oryzae* type III polyketide synthase CsyB. *Bioorg Med Chem* 18(12):4542–4546.
33. 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(1):41.
34. De Bie T, Cristianini N, Demuth JP, Hahn MW (2006) CAFE: A computational tool for the study of gene family evolution. *Bioinformatics* 22(10):1269–1271.
35. Brown A, Sims PFG, Raeder U, Broda P (1988) Multiple ligninase-related genes from *Phanerochaete chrysosporium*. *Gene* 73(1):77–85.
36. Kirk TK, Schultz E, Conners WJ, Lorentz LF, Zeikus JG (1978) Influence of culture parameters on lignin metabolism by *Phanerochaete chrysosporium*. *Arch Microbiol* 117:277–285.
37. Hori C, et al. (2013) Genomewide analysis of polysaccharides degrading enzymes in 11 white- and brown-rot Polyporales provides insight into mechanisms of wood decay. *Mycologia* 105(6):1412–1427.
38. Wariishi H, Valli K, Gold MH (1992) Manganese(II) oxidation by manganese peroxidase from the basidiomycete *Phanerochaete chrysosporium*. Kinetic mechanism and role of chelators. *J Biol Chem* 267(33):23688–23695.
39. Tien M, Kirk TK (1984) Lignin-degrading enzyme from *Phanerochaete chrysosporium*: Purification, characterization, and catalytic properties of a unique H_2O_2-requiring oxygenase. *Proc Natl Acad Sci USA* 81(8):2280–2284.
40. Niku-Paavola ML, Karhunen E, Salola P, Raunio V (1988) Ligninolytic enzymes of the white-rot fungus *Phlebia radiata*. *Biochem J* 254(3):877–883.
41. Kersten PJ, Kirk TK (1987) Involvement of a new enzyme, glyoxal oxidase, in extracellular H_2O_2 production by *Phanerochaete chrysosporium*. *J Bacteriol* 169(5):2195–2201.

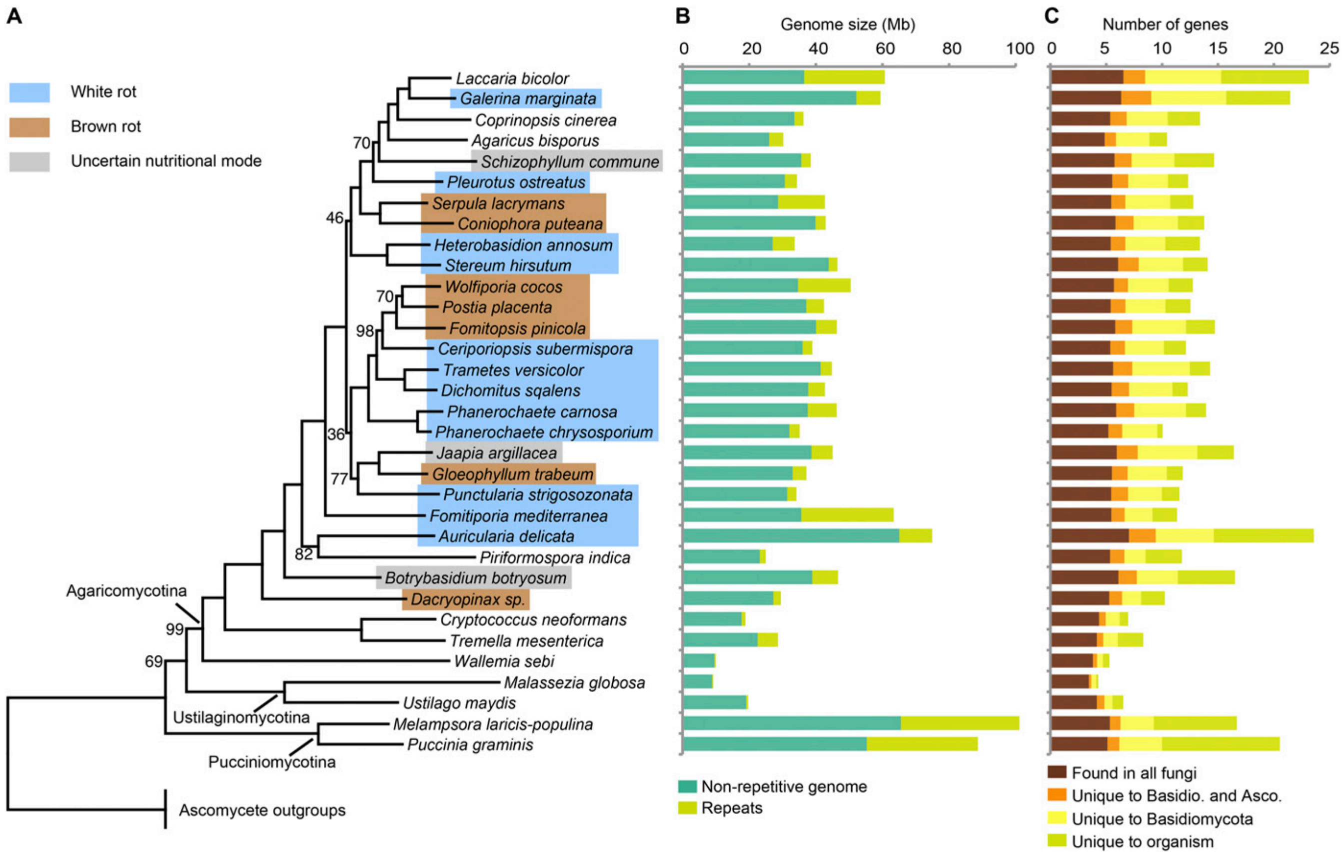

Fig. S1. Phylogeny, genome size and repeats, and gene conservation among basidiomycetes. (*A*) Maximum-likelihood tree of 33 basidiomycetes based on concatenated alignments of 183 widely conserved genes. Ascomycete outgroups are omitted from the figure. Bootstrap values of branches are 100% except where indicated. (*B*) Repeat content in basidiomycetes is highly variant, ranging from 1% to 44%. (*C*) Conservation of genes in basidiomycetes.

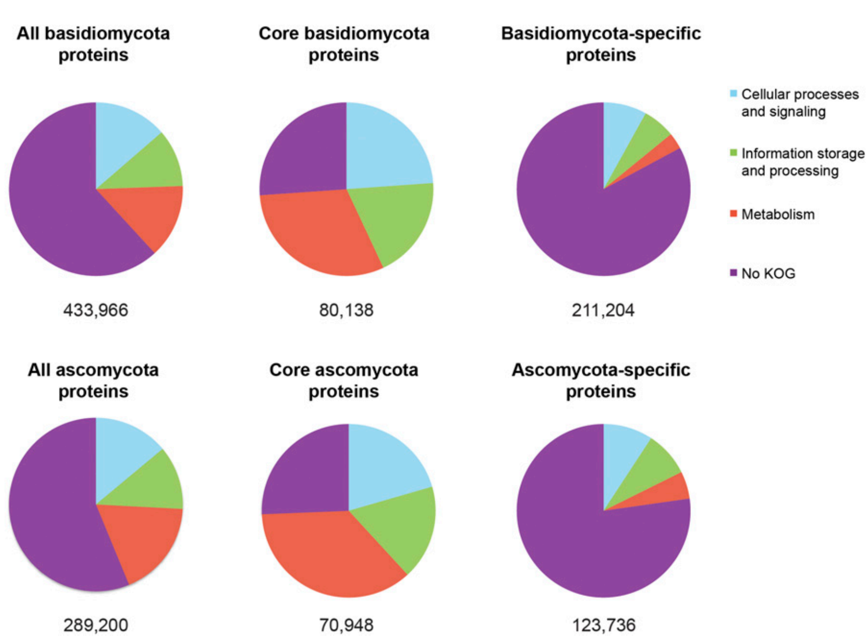

Fig. S2. Core genes of Basidiomycota. Notice that one-half of basidiomycete proteins have no KOG annotation (function unknown); 92% of core basidiomycete proteins have a KOG annotation (putative function predicted); and 78% of noncore basidiomycete proteins have no KOG annotation (function unknown).

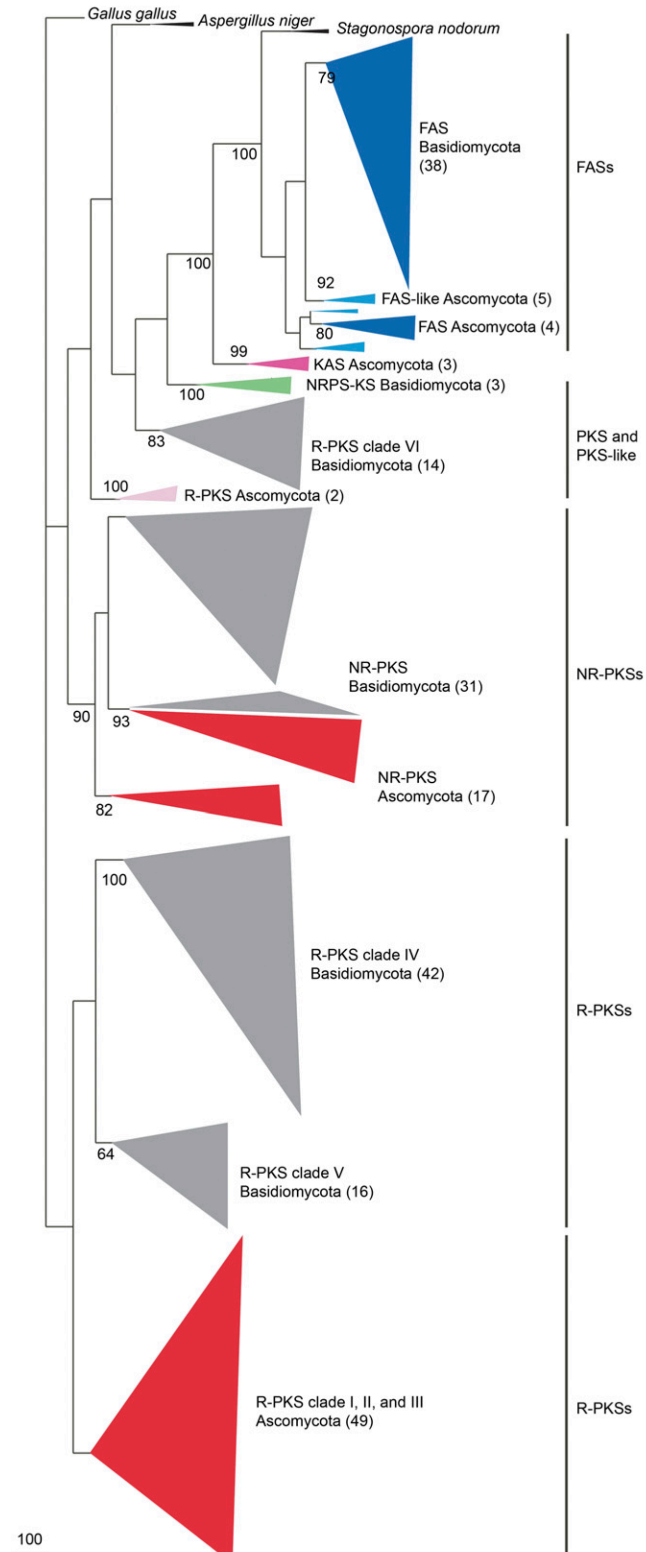

Fig. S3. Phylogenetic tree of secondary metabolism enzymes. Proteins with a ketoacyl synthase (KS) domain (PF00109 and PF02801) were divided into three categories: fatty acid synthase (FAS), nonreducing polyketide synthase (NR-PKS), and reducing PKS (R-PKS) based on their phylogenetic relationships.

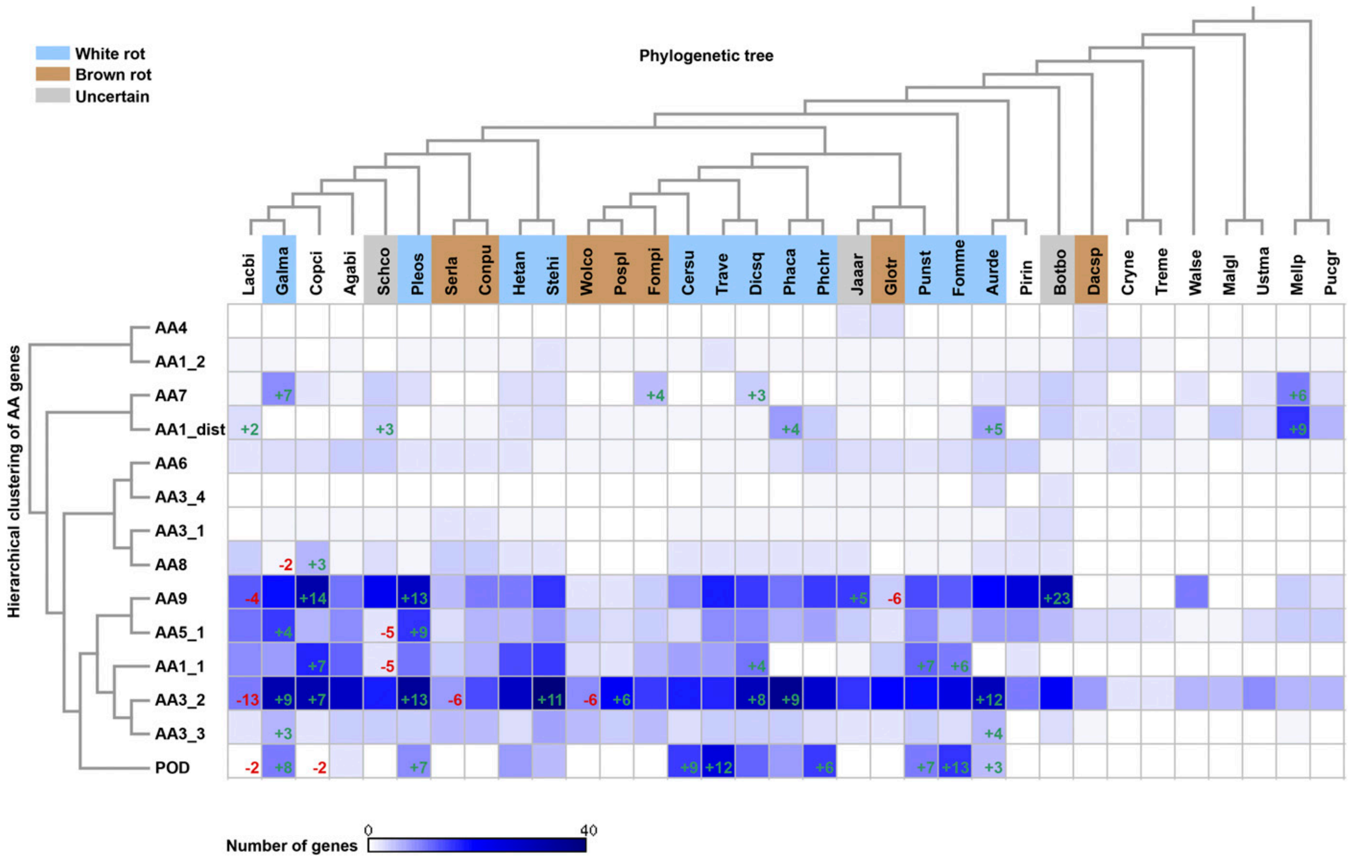

Fig. S4. Heat map of the number of genes in each auxiliary redox enzyme (AA) family for each organism, along with the number of genes inferred, using the CAFE program, to be gained or lost since that organism's divergence with its nearest neighbor. The gene families are hierarchically clustered.

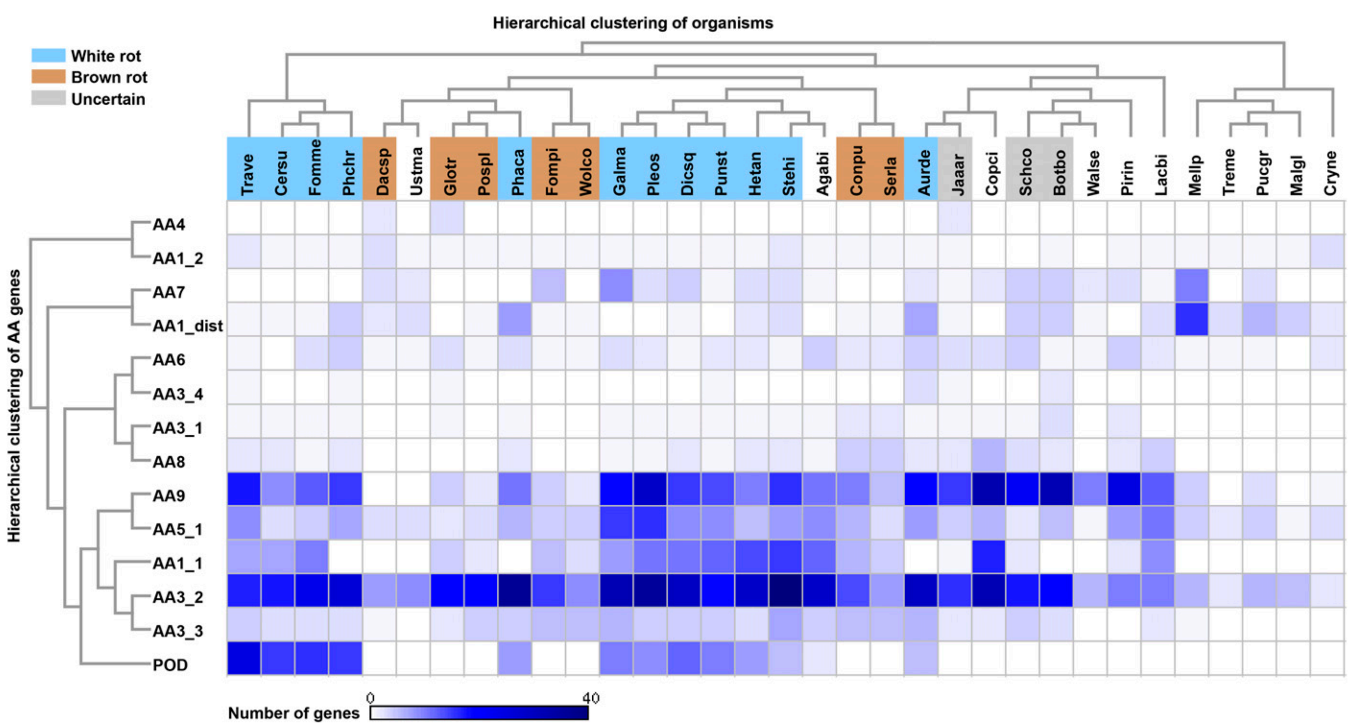

Fig. S5. Double-hierarchical clustering of auxiliary redox enzyme (AA) families. Gene families and organisms are hierarchically clustered.

Table S1. Statistics on genomes and gene annotations for basidiomycete fungi

Subphylum	Order	Organism	Ref.	JGI portal id	Genome size, Mb	Gene no.	Median gene length	Median intron length	% GC genome	Median exons per gene	Gene density, genes/Mb
Agaricomycotina	Agaricales	*Agaricus bisporus*	1	Agabi_varbisH97_2	30.2	10,438	1,497	55	46	5	345
		Coprinopsis cinerea	2	Copci1	36.3	13,393	1,477	57	52	5	369
		Galerina marginata		Galma1	59.4	21,461	1,349	58	48	4	361
		Laccaria bicolor	3	Lacbi2	64.9	19,036	1,209	56	47	4	293
		Pleurotus ostreatus		PleosPC15_2	34.3	12,330	1,394	56	51	5	359
		Schizophyllum commune	4	Schco2	38.5	13,210	1,475	54	57	5	343
	Auriculariales	*Auricularia delicata*	5	Aurde1	74.9	23,577	1,309	55	59	4	315
	Boletales	*Serpula lacrymans*	6	SerlaS7_9_2	42.8	12,917	1,317	58	46	4	302
		Coniophora puteana	5	Conpu1	43.0	13,761	1,636	58	53	5	320
	Cantharellales	*Botryobasidion botryosum*		Botbo1	46.7	16,526	1,407	64	52	4	354
	Corticiales	*Punctularia strigosozonata*	5	Punst1	34.2	11,538	1,542	57	55	5	338
	Dacrymycetales	*Dacryopinax* sp.	5	Dacsp1	29.5	10,242	1,379	53	52	4	347
	Gloeophyllales	*Gloeophyllum trabeum*	5	Glotr1_1	37.2	11,846	1,501	56	54	5	319
	Hymenochaetales	*Fomitiporia mediterranea*	5	Fomme1	63.4	11,333	1,540	60	46	5	179
	Jaapiales	*Jaapia argillacea*		Jaaar1	45.1	16,419	1,402	59	50	4	364
	Polyporales	*Dichomitus squalens*	5	Dicsq1	42.7	12,290	1,562	59	56	5	287
		Ceriporiopsis subvermispora	7	Cersu1	39.0	12,125	1,678	56	54	5	311
		Fomitopsis pinicola	5	Fompi1	46.3	14,724	1,372	56	56	4	318
		Phanerochaete carnosa	8	Phaca1	46.3	13,937	1,448	57	53	4	301
		Phanerochaete chrysosporium	9	Phchr1	35.1	10,048	1,363	54	57	5	286
		Postia placenta	10	PosplRSB12_1	42.5	12,541	1,492	58	50	5	295
		Trametes versicolor	5	Trave1	44.8	14,296	1,489	58	58	4	319
		Wolfiporia cocos	5	Wolco1	50.5	12,746	1,633	56	53	5	252
	Russulales	*Heterobasidion annosum*	11	Hetan2	33.6	13,405	1,283	59	52	4	398
		Stereum hirsutum	5	Stehi1	46.5	14,072	1,718	65	51	5	303
	Sebacinales	*Piriformospora indica*	12	Pirin1	25.0	11,767	1,283	50	51	4	471
	Tremellales	*Cryptococcus neoformans*	13	Cryne_H99_1	18.9	6,967	1,685	56	48	5	369
		Tremella mesenterica	5	Treme1	28.6	8,313	1,741	77	47	5	290
	Wallemiales	*Wallemia sebi*	14	Walse1	9.8	5,284	1,503	47	40	3	538
Pucciniomycotina	Pucciniales	*Puccinia graminis*	15	Pucgr1	88.6	20,534	1,329	91	43	4	232
		Melampsora laricis-populina	15	Mellp1	101.1	16,831	1,376	80	41	4	166
Ustilaginomycotina	Ustilaginales	*Ustilago maydis*	16	Ustma1	19.7	6,522	1,578	97	54	1	331
	Malasseziales	*Malassezia globosa*	17	Malgl1	9.0	4,286	1,218	50	52	1	478

Gene models for all genomes (except *C. cinerea*, *C. neoformans*, *M. globosa*, *P. indicus*, *P. graminis*, and *U. maydis*; see references) were predicted using the JGI Annotation Pipeline, which employs a variety of gene-modeling algorithms and quality filtering to select the best models.

1. Morin E, et al. (2012) Genome sequence of the button mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche. *Proc Natl Acad Sci USA* 109(43):17501–17506.
2. Stajich JE, et al. (2010) Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom *Coprinopsis cinerea* (*Coprinus cinereus*). *Proc Natl Acad Sci USA* 107(26):11889–11894.
3. Martin F, et al. (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452(7183):88–92.
4. Ohm RA, et al. (2010) Genome sequence of the model mushroom *Schizophyllum commune*. *Nat Biotechnol* 28(9):957–963.
5. Floudas D, et al. (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336(6089):1715–1719.
6. Eastwood DC, et al. (2011) The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science* 333(6043):762–765.
7. Fernandez-Fueyo E, et al. (2012) Comparative genomics of *Ceriporiopsis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. *Proc Natl Acad Sci USA* 109(14):5458–5463, and correction (2012) 109(21):8352.
8. Suzuki H, et al. (2012) Comparative genomics of the white-rot fungi, *Phanerochaete carnosa* and *P. chrysosporium*, to elucidate the genetic basis of the distinct wood types they colonize. *BMC Genomics* 13:444.
9. Martinez D, et al. (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nat Biotechnol* 22(6):695–700.
10. Martinez D, et al. (2009) Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. *Proc Natl Acad Sci USA* 106(6):1954–1959.
11. Olson A, et al. (2012) Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. *New Phytol* 194(4):1001–1013.
12. Zuccaro A, et al. (2011) Endophytic life strategies decoded by genome and transcriptome analyses of the mutualistic root symbiont *Piriformospora indica*. *PLoS Pathog* 7(10):e1002290.
13. Loftus BJ, et al. (2005) The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science* 307(5713):1321–1324.
14. Padamsee M, et al. (2012) The genome of the xerotolerant mold *Wallemia sebi* reveals adaptations to osmotic stress and suggests cryptic sexual reproduction. *Fungal Genet Biol* 49(3):217–226.
15. Duplessis S, et al. (2011) Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc Natl Acad Sci USA* 108(22):9166–9171.
16. Kämper J, et al. (2006) Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* 444(7115):97–101.
17. Xu J, et al. (2007) Dandruff-associated *Malassezia* genomes reveal convergent and divergent virulence traits shared with plant and human fungal pathogens. *Proc Natl Acad Sci USA* 104(47):18730–18735.

Table S2. Nonbasidiomycete fungi used for comparative purposes

Subkingdom	Phylum	Subphylum	Organism	JGI portal ID
Dikarya	Ascomycota	Pezizomycotina	*Stagonospora nodorum SN15*	Stano2
			Aspergillus niger ATCC 1015	Aspni5
			Aspergillus nidulans	Aspnid1
			Botrytis cinerea	Botci1
			Fusarium graminearum	Fusgr1
			Fusarium oxysporum	Fusox1
			Leptosphaeria maculans	Lepmu1
			Magnaporthe grisea	Maggr1
			Mycosphaerella graminicola	Mycgr3
			Nectria hematococca	Necha2
			Neurospora crassa OR74A	Neucr1
			Neurospora tetrasperma FGSC 2508 mat A	Neute_matA2
			Neurospora tetrasperma FGSC 2509 mat a	Neute_mat_a1
			Pyrenophora tritici-repentis	Pyrtr1
			Pyrenophora teres f. teres	Pyrtt1
			Sclerotinia sclerotiorum	Sclsc1
			Sporotrichum thermophile	Spoth2
			Thielavia terrestris	Thite2
			Trichoderma atroviride	Triat2
			Trichoderma reesei	Trire2
			Tuber melanosporum	Tubme1
			Verticillium dahliae	Verda1
		Saccharomycotina	*Candida tenuis* NRRL Y-1498	Cante1
			Dekkera bruxellensis CBS 2499	Dekbr2
			Pichia stipitis	Picst3
			Saccharomyces cerevisiae S288C	Sacce1
			Spathaspora passalidarum NRRL Y-27907	Spapa3
Fungi incertae sedis	Chytridiomycota	N/A	*Batrachochytrium dendrobatidis* JAM81	Batde5
	Mucoromycotina	Mucoromycotina	*Phycomyces blakesleeanus* NRRL1555	Phybl2
			Rhizopus oryzae 99–880	Rhior3

Table S3. CAZymes grouped by three major substrates

Substrate preference	CAZymes and auxiliary redox enzymes
Cellulose	GH6
	GH7
	GH12
	GH44
	GH45
	AA9 (formerly GH61)
	AA10 (formerly CBM33)
	CBM1
Hemicelluloses	GH10
	GH11
	GH26
	GH30
	GH51
	GH74
	GH43
Pectin	GH43
	GH28
	GH53
	GH78
	GH88
	GH105
	PL1
	PL3
	PL4
	CE8
	CE12

Table S4. Enzyme activities after 5 d

	B3 C-limited				B3 N-limited				Wiley-milled aspen			
Organism	LiP	MnP	GLOX	Lac	LiP	MnP	GLOX	Lac	LiP	MnP	GLOX	Lac
Phchr	ND	9.44	0.087	0.74	6.74	11.8	9.19	ND	5.57	ND	0.708	ND
Cersu	ND	4.05	0.082*	7.95	ND	3.83	0.01*	6.15	ND	ND	ND	ND
Pospl	ND	ND	0.587	0.056	ND	ND	0.713	0.056	ND	ND	0.694	0.341
Wolco	ND	ND	0.634	ND	ND	ND	0.796	ND	ND	ND	0.541	1.00
Jaaar	ND	ND	0.027*	ND	ND	ND	0.056*	ND	ND	ND	0.014*	ND
Botbo	ND	ND	ND	ND	ND	ND	0.056*	ND	ND	ND	0.014*	ND

In nanomoles per milliliter per minute. ND, not detected.
*Trace levels below reliable measurement. Assessment of *J. argillacea* and *B. botryosum* cultures after days 9 and 11 d were essentially identical, i.e., no reliable activity.

Other Supporting Information Files

Dataset S1 (XLSX)
Dataset S2 (XLSX)
Dataset S3 (TXT)