The Paleozoic origin of enzymatic mechanisms for lignin degradation reconstructed using 31 fungal genomes

Dimitrios Floudas[1], Manfred Binder[1], Robert Riley[2], Kerrie Barry[2], Robert A. Blanchette[3], Bernard Henrissat[4], Angel T. Martínez[5], Robert Otillar[2], Joseph W. Spatafora[6], Jagjit S. Yadav[7], Andrea Aerts[2], Isabelle Benoit[8, 9], Alex Boyd[6], Alexis Carlson[1], Alex Copeland[2], Pedro M. Coutinho[4], Ronald P. de Vries[8, 9], Patricia Ferreira[10], Keisha Findley[11], Brian Foster[2], Jill Gaskell[12], Dylan Glotzer[1], Paweł Górecki[13], Joseph Heitman[11], Cedar Hesse[6], Chiaki Hori[14], Kiyohiko Igarashi[14], Joel A. Jurgens[3], Nathan Kallen[1], Phil Kersten[12], Annegret Kohler[15], Ursula Kües[16], T. K. Arun Kumar[17], Alan Kuo[2], Kurt LaButti[2], Luis F. Larrondo[18], Erika Lindquist[2], Albee Ling[1], Vincent Lombard[4], Susan Lucas[2], Taina Lundell[19], Rachael Martin[1], David J. McLaughlin[17], Ingo Morgenstern[20], Emanuelle Morin[15], Claude Murat[15], Laszlo G. Nagy[1], Matt Nolan[2], Robin A. Ohm[2], Aleksandrina Patyshakuliyeva[9], Antonis Rokas[21], Francisco J. Ruiz-Dueñas[5], Grzegorz Sabat[22], Asaf Salamov[2], Masahiro Samejima[14], Jeremy Schmutz[23], Jason C. Slot[21], Franz St. John[12], Jan Stenlid[24], Hui Sun[2], Sheng Sun[11], Khajamohiddin Syed[7], Adrian Tsang[20], Ad Wiebenga[9], Darcy Young[1], Antonio Pisabarro[25], Daniel C. Eastwood[26], Francis Martin[15], Dan Cullen[12], Igor V. Grigoriev[2]*, and David S. Hibbett[1]*
*To whom correspondence should be addressed

1. 1Biology Department, Clark University, Worcester, Massachusetts 01610, USA
2. US Department of Energy Joint Genome Institute, Walnut Creek, California, 94598, USA
3. Department of Plant Pathology, University of Minnesota, St. Paul, Minnesota 55108, USA
4. Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, CNRS UMR 6098, 13288 Marseille Cedex 9, France
5. Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, E-28040 Madrid, Spain
6. Department of Botany & Plant Pathology, Oregon State University, Corvallis, Oregon 97331, USA
7. Environmental Genetics and Molecular Toxicology Division, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267, USA
8. Microbiology & Kluyver Centre for Genomics of Industrial Fermentation, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands
9. CBS-KNAW Fungal Biodiversity Centre, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands
10. Department of Biochemistry and Molecular and Cellular Biology and Institute of Biocomputation and Physics of Complex Systems, Zaragoza University, 50009, Zaragoza, Spain
11. Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA
12. USDA Forest Products Laboratory, Madison, Wisconsin 53726, USA
13. Institute of Informatics, Warsaw University, Warsaw, 02-097, Poland

14. Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan
15. INRA, UMR 1136, INRA-Nancy Université, Interactions Arbres/Microorganismes, 54280 Champenoux, France
16. Molecular Wood Biotechnology and Technical Mycology, Büsgen-Institute, Georg-August-University Göttingen, Büsgenweg 2, D-37077 Göttingen, Germany
17. Department of Plant Biology, University of Minnesota, St. Paul, Minnesota, 55108, USA
18. Departamento de Genética Molecular y Microbiología, Facultad de Ciencias Biológicas, Pontifica Universidad Católica de Chile, Casilla 114-D, Santiago 833-1010, Chile
19. Department of Applied Chemistry and Microbiology, Viikki Biocenter, P.O. Box 56, Biocenter 1, University of Helsinki, FIN-00014 Helsinki, Finland
20. Centre for Structural and Functional Genomics, Concordia University, 7141 Sherbrooke Street West, Montreal H4B 1R6 Quebec, Canada
21. Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, 37235, USA
22. University of Wisconsin Biotechnology Center, Wisconsin 53726, USA
23. HudsonAlpha Institute for Biotechnology, Huntsville, Alabama 35806, USA
24. Department of Forest Mycology and Pathology, Swedish University of Agricultural Sciences, Box 7026, Ulls v 26A, 750 07, Uppsala, Sweden
25. Genetics and Microbiology Research Group, Public University of Navarre, 31006, Pamplona, Spain
26. College of Science, University of Swansea, Singelton Park, Swansea SA2 8PP, UK

June 2012

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

The Paleozoic origin of enzymatic mechanisms for lignin degradation reconstructed using 31 fungal genomes

Dimitrios Floudas[1], Manfred Binder[1], Robert Riley[2], Kerrie Barry[2], Robert A. Blanchette[3], Bernard Henrissat[4], Angel T. Martínez[5], Robert Otillar[2], Joseph W. Spatafora[6], Jagjit S. Yadav[7], Andrea Aerts[2], Isabelle Benoit[8,9], Alex Boyd[6], Alexis Carlson[1], Alex Copeland[2], Pedro M. Coutinho[4], Ronald P. de Vries[8,9], Patricia Ferreira[10], Keisha Findley[11], Brian Foster[2], Jill Gaskell[12], Dylan Glotzer[1], Paweł Górecki[13], Joseph Heitman[11], Cedar Hesse[6], Chiaki Hori[14], Kiyohiko Igarashi[14], Joel A. Jurgens[3], Nathan Kallen[1], Phil Kersten[12], Annegret Kohler[15], Ursula Kües[16], T. K. Arun Kumar[17], Alan Kuo[2], Kurt LaButti[2], Luis F. Larrondo[18], Erika Lindquist[2], Albee Ling[1], Vincent Lombard[4], Susan Lucas[2], Taina Lundell[19], Rachael Martin[1], David J. McLaughlin[17], Ingo Morgenstern[20], Emanuelle Morin[15], Claude Murat[15], Laszlo G. Nagy[1], Matt Nolan[2], Robin A. Ohm[2], Aleksandrina Patyshakuliyeva[9], Antonis Rokas[21], Francisco J. Ruiz-Dueñas[5], Grzegorz Sabat[22], Asaf Salamov[2], Masahiro Samejima[14], Jeremy Schmutz[23], Jason C. Slot[21], Franz St. John[12], Jan Stenlid[24], Hui Sun[2], Sheng Sun[11], Khajamohiddin Syed[7], Adrian Tsang[20], Ad Wiebenga[9], Darcy Young[1], Antonio Pisabarro[25], Daniel C. Eastwood[26], Francis Martin[15], Dan Cullen[12], Igor V. Grigoriev[2]*, and David S. Hibbett[1]*

*To whom correspondence should be addressed

[1]Biology Department, Clark University, Worcester, Massachusetts 01610, USA

[2]US Department of Energy Joint Genome Institute, Walnut Creek, California, 94598, USA

[3]Department of Plant Pathology, University of Minnesota, St. Paul, Minnesota 55108, USA

[4]Architecture et Fonction des Macromolécules Biologiques, Aix-Marseille Université, CNRS UMR 6098, 13288 Marseille Cedex 9, France

[5]Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, E-28040 Madrid, Spain

[6]Department of Botany & Plant Pathology, Oregon State University, Corvallis, Oregon 97331, USA

[7]Environmental Genetics and Molecular Toxicology Division, Department of Environmental Health, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267, USA

[8]Microbiology & Kluyver Centre for Genomics of Industrial Fermentation, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[9]CBS-KNAW Fungal Biodiversity Centre, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

[10]Department of Biochemistry and Molecular and Cellular Biology and Institute of Biocomputation and Physics of Complex Systems, Zaragoza University, 50009, Zaragoza, Spain

[11]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA

[12]USDA Forest Products Laboratory, Madison, Wisconsin 53726, USA

[13]Institute of Informatics, Warsaw University, Warsaw, 02-097, Poland

[14]Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan

[15]INRA, UMR 1136, INRA-Nancy Université, Interactions Arbres/Microorganismes, 54280 Champenoux, France

[16]Molecular Wood Biotechnology and Technical Mycology, Büsgen-Institute, Georg-August-University Göttingen, Büsgenweg 2, D-37077 Göttingen, Germany

[17]Department of Plant Biology, University of Minnesota, St. Paul, Minnesota, 55108, USA

[18]Departamento de Genética Molecular y Microbiología, Facultad de Ciencias Biológicas, Pontifica Universidad Católica de Chile, Casilla 114-D, Santiago 833-1010, Chile

[19]Department of Applied Chemistry and Microbiology, Viikki Biocenter, P.O. Box 56, Biocenter 1, University of Helsinki, FIN-00014 Helsinki, Finland

[20]Centre for Structural and Functional Genomics, Concordia University, 7141 Sherbrooke Street West, Montreal H4B 1R6 Quebec, Canada

[21]Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, 37235, USA

[22]University of Wisconsin Biotechnology Center, Wisconsin 53726, USA

[23]HudsonAlpha Institute for Biotechnology, Huntsville, Alabama 35806, USA

[24]Department of Forest Mycology and Pathology, Swedish University of Agricultural Sciences, Box 7026, Ulls v 26A, 750 07, Uppsala, Sweden

[25]Genetics and Microbiology Research Group, Public University of Navarre, 31006 Pamplona, Spain

[26]College of Science, University of Swansea, Singleton Park, Swansea SA2 8PP, UK

Wood is a major pool of organic carbon that is highly resistant to decay, owing largely to the presence of lignin. The only organisms capable of significant lignin decay are white rot Fungi in the Agaricomycetes, which also contains non-lignin-degrading brown rot and ectomycorrhizal species. Comparative analyses of 31 fungal genomes (12 newly generated) suggest that lignin-degrading peroxidases expanded in the lineage leading to the ancestor of the Agaricomycetes, which is reconstructed as a white rot species, and then contracted in parallel lineages leading to brown rot and mycorrhizal species. Molecular clock analyses suggest that the origin of lignin degradation roughly coincided with the sharp decrease in the rate of organic carbon burial around the end of the Carboniferous period.

Lignin is a heterogeneous polymer that provides strength and rigidity to wood, protects cellulose and hemicellulose from microbial attack (*1*), and is the major precursor of coal (*2*). Genomic studies of wood decay organisms have focused on model systems for white rot (in which all plant cell wall components are degraded), such as *Phanerochaete chrysosporium (3),* and brown rot (in which lignin is modified but not appreciably degraded), such as *Postia placenta (4)* and *Serpula lacrymans (5)*. However, these species represent just two of the 18 recognized orders of Agaricomycetes, of which five contain brown rot taxa. To reconstruct the evolution of lignin decay mechanisms, we

analyzed 31 diverse fungal genomes, including twelve newly-sequenced species of Agaricomycotina (Table 1). The new genomes comprise six white rot species, five brown rot species, and one mycoparasite, representing nine orders (Figs. 1, S1-5; see Supporting Online Material for information on strains, culture conditions, genome size and assembly).

To estimate phylogenetic relationships, we constructed datasets using 71 or 26 single-copy genes, with varying alignment criteria and treatments for fast-evolving sites, yielding matrices of 10002 to 34257 amino acids, which we analyzed with maximum likelihood (ML) and Bayesian methods (*6, 7*) (see Supporting Online Material for details). All but six nodes receive maximal support values in all analyses and the rest are strongly supported (bootstrap ≥99% or posterior probability ≥0.99) in at least three analyses. The tree topology is consistent with prior analyses (*8*) and resolves four independent brown rot lineages (Figs. 1a, S6).

We next searched all 31 genomes for 27 gene families encoding oxidoreductases and carbohydrate-active enzymes (CAZymes) that have been implicated in wood decay (*9, 10*) (Table 1). CAZymes, particularly those acting on crystalline cellulose, are abundant in white rot genomes, which have 61-148 (avg. 87) copies of genes encoding CAZymes, representing 14-17 gene families, whereas brown rot genomes have 32-68 copies (avg. 46) from 9-12 families. The ectomycorrhizal *Laccaria bicolor* resembles brown rot species in this regard, possessing 28 CAZyme genes in eight families (Table 1). Notably, glycoside hydrolase (GH) families GH6 and GH7, which include cellobiohydrolases that are involved in the attack of crystalline cellulose (*9*), are present in all white rot lineages, but they are absent in brown rot

lineages (except Boletales) and *L. bicolor*. A similar pattern of enrichment in white rot genomes is shown by genes encoding GH61 enzymes, which have a copper-dependent oxidative mechanism for disrupting crystalline cellulose (*11*), and cellulose binding modules (CBM1), which effectively increase the concentration of the enzymes on the surface of crystalline cellulose (*12*) (Table 1).

Here, we focus on fungal class II peroxidases (PODs), which have been shown to be important in lignin degradation in *P. chrysosporium* and others (*13-15*). (The remaining gene families are discussed in Supporting Online Material; Figs. S7-19.) We classified PODs in four major groups, including three ligninolytic forms, lignin peroxidase (LiP), manganese peroxidase (MnP) and versatile peroxidase (VP), and a fourth POD type, defined here as “generic peroxidase” (GP), which is expected to include non-ligninolytic low redox-potential peroxidases with catalytic properties similar to those of the peroxidase of *Coprinopsis cinerea* or the product of the *nop*A gene in *P. chrysosporium* (*14*) (further subclassification of PODs is described in Supporting Online Material).

Consistent with a central role for PODs in lignin degradation, white rot species have 5 to 26 copies (avg. 14) of genes encoding ligninolytic PODs, but all brown rot species lack these enzymes, as do the ectomycorrhizal (ECM) *Laccaria bicolor,* the soil saprotroph *C. cinerea,* and *Schizophyllum commune,* which has been regarded as a white rot fungus, but has a limited capacity to degrade lignin (*15*). Moreover, analyses of gene diversification using binary state speciation analysis (*19*) confirmed that the rate of duplication of POD genes is elevated in white rot lineages vs. non-white rot lineages (see Supporting Online Material for details).

LiPs possess a tryptophan residue on the surface of the enzyme corresponding to Trp171 in *P. chrysosporium* LiP-H8 that enables direct oxidation of lignin compounds via long-range electron transfer, and MnPs possess two or three residues corresponding to Glu35, Glu39 and Asp175 of *P. chrysosporium* MnP1 that function in binding Mn (*16, 17*). VPs possess both the Trp171 homologue and Mn-binding residues, while all are lacking in GPs. To reconstruct functional evolution of PODs, we performed Bayesian and ML analyses (*6, 18*) using the GPs of Ascomycota as outgroups and reconstructed the evolution of the key residues using BayesTraits (*20*). The ancestor of all PODs lacked the Mn-binding and Trp171 residues, suggesting that it was non-ligninolytic (Fig. 1b). The most recent common ancestor of all ligninolytic Agaricomycete PODs is reconstructed as an MnP, with a single origin of LiP (gain of Trp171 and loss of Mn-binding residues), leading to parallel expansions in *P. chrysosporium* and *T. versicolor* (each with 10 LiP copies; Figs. 1b, S7, S17). Two origins of VP are resolved in the Polyporales, leading to *T. versicolor* and *D. squalens* (each with three VP copies; Figs. 1b, S7). VPs are also produced in the “oyster mushroom” *Pleurotus ostreatus* (Agaricales) (*21*), indicating further convergent evolution of this class of enzymes.

To localize the diversification of PODs in the organismal phylogeny, we performed gene tree/species tree reconciliation analyses using CAFÉ (*22*), Notung (*23*), and DrML (*24*). All methods suggest that a single POD gene copy was present in the common ancestor of Basidiomycota, with parallel losses in lineages leading to the Pucciniomycotina, Ustilaginomycotina, Tremellomycetes, and *Dacryopinax sp.* (Fig. 1a). Diversification of PODs began in the lineage leading to the most recent common

ancestor of the Agaricomycetes (node “A” in Fig. 1a), which is reconstructed as having two to seven POD gene copies in the various analyses. The “backbone” nodes in the Agaricomycete phylogeny (nodes labeled “B” in Fig. 1a) are reconstructed as having 3-16 POD gene copies, suggesting that duplicated PODs were retained throughout the early evolution of Agaricomycetes. Subsequently, all analyses suggest that there were parallel expansions of POD genes in the terminal lineages leading to at least seven of the eight white rot species, as well as parallel contractions within lineages leading to the brown rot species *Dacryopinax sp.*, *Gloeophyllum trabeum*, the Boletales, and the brown rot Polyporales (Fig. 1).

In addition to multiple PODs, reconciliation analyses suggest that the ancestor of the Agaricomycetes possessed one or two genes encoding dye-decolorizing peroxidases (DyP), which are heme peroxidases that have been shown to degrade lignin model compounds (*25, 26*), as well 5-8 genes encoding oxidases (including glyoxal oxidase) involved in peroxide generation (*27*) (details in Supporting Online Material). Collectively, these results suggest that the ancestor of Agaricomycetes was a white rot species that possessed a ligninolytic system with PODs, DyPs, and multiple pathways for H_2O_2 production.

To place the origin of lignin degradation in the context of geologic time, we performed Bayesian relaxed molecular clock analyses using BEAST (*18*) and PhyloBayes (*7*), with fossil-based calibrations at three nodes, including the ancestors of the Boletales (*28*), Agaricales (*29*), and Ascomycota (*30*) (see Supporting Online Material for details of molecular clock analyses). The mean age of the Agaricomycetes is ca. 290 mya in both BEAST and PhyloBayes analyses, with the mean age of the

Agaricomycotina being ca. 430-470 mya. These age estimates are consistent with basidiomycete fossils that were not used as calibration points, including hyphae with clamp connections from the Mississippian (*31*) (ca. 330 mya). BEAST analyses of the POD genes, calibrated with the split between Ascomycota and Basidiomycota according to the organismal phylogeny, suggest that the first ligninolytic MnP arose ca. 295 mya (Fig. 1), which is slightly earlier than (and therefore consistent with) the oldest definitive white rot fossils from the Permian (ca. 260 mya) and Triassic (ca. 230 mya) (*32*).

Organic carbon derived from lignin accumulated at an exceptionally high rate during the Carboniferous and Permian, resulting in the formation of vast coal deposits (*33*). A frequently cited explanation for this phenomenon is that decay was inhibited in the anoxic sediments of the widespread Permo-Carboniferous swamp forests. Our results are consistent with a complementary hypothesis (*2*), which posits that the sharp decline in the rate of organic carbon burial at the end of the Permo-Carboniferous was caused, at least in part, by the evolution of lignin decay capabilities in white rot Agaricomycetes.

Figure 1. **A.** Organismal phylogeny (chronogram) produced with BEAST. Light blue bars are 95% highest posterior density intervals for node ages; mean ages of selected nodes (millions of years) are in parentheses. Blue and red branches indicate significant expansion and contraction, respectively, of PODs inferred using CAFÉ. Numbers in red following following taxon names are POD gene counts. Numbers in red at nodes separated by commas, are numbers of POD gene copies estimated with CAFÉ, Notung (with two different edge weight threshold settings), and DrML, respectively. Node labled

A is the ancestor of Agaricomycetes, while nodes labeled B are "backbone" nodes in Agaricomycetes (see text). Asterisks indicate nodes that do not receive maximal support in all analyses (support values are in Fig. S6). Full species names are listed in a footnote to Table 1. **B.** POD gene phylogeny estimated in BEAST with ancestral state reconstructions for manganese-binding site (colored pies) and Trp171 residues (black and white pies) estimated with BayesTraits. Bars to right of gene IDs indicate presence of functional residues (*17*). Mean ages for selected nodes in parentheses are followed by 95% highest posterior density ranges.

References and Notes

1. K. E. Eriksson, R. Blanchette, P. Ander, *Microbial and enzymatic degredation of wood and wood components*. (Springer-Verlag, Berlin, 1990).
2. J. M. Robinson, Lignin, land plants, and fungi: Biological evolution affecting Phanerozoic oxygen balance. *Geology* **18**, 607 (1990).
3. D. Martinez *et al.*, Genome sequence of the lignocellulose degrading fungus Phanerochaete chrysosporium strain RP78. *Nat Biotechnol* **22**, 695 (Jun, 2004).
4. D. Martinez *et al.*, Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion. *Proceedings of the National Academy of Sciences* **106**, 1954 (2009).
5. D. C. Eastwood *et al.*, The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science* **333**, 762 (2011).
6. A. Stamatakis, RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688 (2006).
7. N. Lartillot, T. Lepage, S. Blanquard, PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* **25**, 2286 (2009).
8. P. B. Matheny *et al.*, Contributions of rpb2 and tef1 to the phylogeny of mushrooms and allies (Basidiomycota, Fungi). *Mol Phylogenet Evol* **43**, 430 (May, 2007).
9. P. Baldrian, V. Valášková, Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiology Reviews* **32**, 501 (2008).
10. B. L. Cantarel *et al.*, The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* **37**, D233 (2009).
11. R. J. Quinlan *et al.*, Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences* **108**, 15079 (2011).
12. D. Guillén, S. Sánchez, R. Rodríguez-Sanoja, Carbohydrate-binding domains: multiplicity of biological roles. *Applied Microbiology and Biotechnology* **85**, 1241 (2010).

13. A. T. Martínez, F. J. Ruiz-Dueñas, M. J. Martínez, J. C. d. Río, A. Gutiérrez, Enzymatic delignification of plant cell wall: from nature to mill. *Current Opinion in Biotechnology* **20**, 348 (2009).
14. L. Larrondo, A. González, A. T. Pérez, D. Cullen, R. Vicuña, The nop gene from *Phanerochaete chrysosporium* encodes a peroxidase with novel structural features. *Biophysical Chemistry* **116**, 167 (2005).
15. O. Schmidt, W. Liese, Variability of wood degrading enzymes of *Schizophyllum commune*. *Holzforschung* **34**, 67 (1980).
16. K. Hammel, D. Cullen, Role of fungal peroxidases in biological ligninolysis. *Curr Opin Plant Biol* **11**, 349 (2008).
17. F. J. Ruiz-Duenas *et al.*, Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. *Journal of Experimental Botany* **60**, 441 (2009).
18. A. J. Drummond, A. Rambaut, BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**, 214 (2007).
19. R. Fitzjohn, W. P. Maddison, S. P. Otto, Estimating trait-dependent speciation and extinction rates from incompletely resolved phylogenies. *Systematic Biology* **58**, 595 (2009).
20. D. Barker, A. Meade, M. Pagel, Constrained models of evolution lead to improved prediction of functional linkage from correlated gain and loss of genes. *Bioinformatics* **23**, 14 (January 1, 2007, 2007).
21. F. J. Ruiz-Dueñas, E. Fernández-Fueyo, M. J. Martínez, A. T. Martínez, *Pleurotus ostreatus* heme peroxidases: An in silico analysis from the genome sequence to the enzyme molecular structure. *Comptes Rendus Biologies*, (2011).
22. T. De Bie, CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269 (2006).
23. D. Durand, H. B. V., B. Vernot, A hybrid micro-macroevolutionary approach to gene tree reconstruction. *Journal of Computational Biology* **13**, 320 (2006).
24. P. Górecki, G. J. Burleigh, O. Eulenstein, Maximum likelihood models and algorithms for gene tree evolution with duplications and losses. *BMC Bioinformatics* **12**, S15 (2011).
25. C. Liers, C. Bobeth, M. Pecyna, R. Ullrich, M. Hofrichter, DyP-like peroxidases of the jelly fungus *Auricularia auricula-judae* oxidize nonphenolic lignin model compounds and high-redox potential dyes. *Applied Microbiology and Biotechnology* **85**, 1869 (2010).
26. M. Hofrichter, R. Ullrich, M. Pecyna, C. Liers, T. Lundell, New and classic families of secreted fungal heme peroxidases. *Applied Microbiology and Biotechnology* **87**, 871 (2010).
27. A. Vanden Wymelenberg *et al.*, Structure, organization, and transcriptional regulation of a family of copper radical oxidase genes in the lignin-degrading basidiomycete Phanerochaete chrysosporium. *Applied and environmental microbiology* **72**, 4871 (Jul, 2006).
28. B. A. LePage, R. S. Currah, R. A. Stockey, G. W. Rothwell, Fossil ectomycorrhizae from the Middle Eocene. *American Journal of Botany* **84**, 470 (1997).
29. D. S. Hibbett, D. Grimaldi, M. J. Donoghue, Fossil mushrooms from Cretaceous and Miocene ambers and the evolution of homobasidiomycetes. *American Journal of Botany* **84**, 981 (1997).
30. T. N. Taylor, H. Hass, H. Kerp, M. Krings, R. T. Hanlin, Perithecial ascomycetes from the 400 million year old Rhynie chert: an example of ancestral polymorphism. *Mycologia* **97**, 269 (2005).

31. M. Krings, N. Dotzler, G. Galtier, T. N. Taylor, Oldest fossil basidiomycete clamp connections. *Mycoscience* **52**, 18 (2011).
32. S. P. Stubblefield, T. N. Taylor, Wood decay in silicified gymnosperms from Antarctica. *Botanical Gazzette* **147**, 116 (1986).
33. R. A. Berner, *The Phanerozoic carbon cycle: CO2 and O2*. (Oxford University press, Oxford, 2004), pp. 150.

Acknowledgements: The work conducted by the U.S. Department of Energy Joint Genome Institute was supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231. This research was also supported by the Assembling the Fungal Tree of Life (AFTOL) project under U.S. National Science Foundation awards DEB-0732968 to DSH, DEB-0732993 to JWS, and DEB-0732550 to DJM. The authors thank Ronald H. Petersen for the strain of *A. delicata*. Assemblies and annotations of the twelve genomes reported here are available from the JGI fungal portal MycoCosm (http://jgi.doe.gov/fungi) and from DDBJ/EMBL/GenBank under the following accessions: AFVO00000000; AEIT00000000; AEUS00000000; AEID00000000; AEJJ00000000; AEHC00000000; AFVP00000000; AEGM00000000; AEGX00000000; AEJI00000000; AFVY00000000; AEHD00000000. Aligned sequence data for organismal and gene family phylogenies and molecular clock analyses, and secretome results are available at DRYAD (http://datadryad.org/) accession no. XXXXX.

Supporting Online Material

Supplementary Methods, Results and Discussion

Tables S1 to S16

Figs. S1 to S22

References

Table 1. Gene contents in 17 CAZyme and 11 oxidoreductase families in the genomes of 20 Agaricomycotina and eleven other Fungi.

Taxonomy[1]		Basidiomycota																								Ascomycota					Ch	Mu
		Agaricomycotina																				Ust		Pucc		Pez				Sc		
		Agaricomycetes																	Dac	Trem												
		Ha	Sh	Pu	Fm	Ad	Tv	Ds	Pc	Pp	Wc	Fp	Gt	Sl	Cp	Lb	Sc	Cc	Da	Tm	Cn	Mg	Um	Ml	Sr	An	Cr	Tr	Sn	Ps	Bd	Pb
Ecology[2]		WR								BR						ECM	WR	S	BR	MP	AP	AP	PP	PP	Y	AP	PP	S	PP	Y	AP	S
CAZymes																																
Genes[3]	*p*[4]																															
GH3[7]	0.000	11	15	12	7	12	11	7	9	5	7	12	9	9	12	2	11	7	8	3	7	1	3	3	3	16	15	11	16	7	1	2
GH5[8]	0.305	7	6	6	6	8	5	5	5	5	4	5	5	8	8	3	3	6	5	0	0	0	1	7	2	4	7	3	5	0	1	1
GH6	0.916	1	1	1	2	2	1	1	1	0	0	0	0	1	2	0	1	5	0	0	0	0	0	0	0	2	2	1	4	0	0	0
GH7	0.000	1	3	5	2	6	4	4	8	0	0	0	0	0	2	0	2	6	0	0	0	0	0	9	0	2	5	2	5	0	0	0
GH10	0.004	2	6	5	4	4	6	5	6	4	4	2	3	1	3	0	5	6	3	0	0	0	2	6	0	2	4	1	7	1	0	0
GH11	0.687	0	1	1	0	3	0	0	1	0	0	0	0	0	0	0	1	6	0	0	0	0	1	0	0	3	4	3	7	0	0	0
GH12	0.483	4	5	2	3	1	5	3	2	2	2	2	2	1	4	3	1	1	1	0	0	0	0	5	0	3	5	2	4	0	0	0
GH28	0.000	8	17	13	16	10	11	7	4	7	9	13[9]	10	7	13	7	3	3	6	0	1	0	1	3	0	22	21	4	4	0	0	9
"GH"61	0.000	10	16	14	13	19	18	15	15	2	2	4	4	5	10	5	22	35	0	0	1	0	0	0	0	7	12	3	29	0	0	0
GH74	0.923	1	2	2	4	1	1	1	4	0	0	0	1	1	0	0	1	1	0	0	0	0	0	0	0	1	2	1	0	0	0	0
GH43	0.000	4	10	7	6	26	3	7	4	1[10]	1[10]	7	5	1	6	0	12	4	5	0	0	1	2	4	0	7	14	2	10	0	0	0
CE1	0.129	1	1	2	0	3	3	0	4	0	0	0	1	0	0	0	4	3	0	0	0	0	1	0	0	1	2	0	6	0	0	0
CE16	0.000	5	10	8	6	29	7	10	2	5	6	11	6	3	6	3	10	5	4	1	0	0	0	0	2	3	7	2	2	0	0	2
CE5	0.002	0	1	1	0	3	0	0	0	0	0	0	0	0	1	1	2	6	0	0	0	0	4	14	0	5	15	4	11	0	0	0
CE8	0.347	3	4	6	3	3	2	3	2	1	1	2	2	2	2	4	2	0	3	0	0	0	1	5	0	3	4	0	2	0	0	4
CE12	0.796	2	3	0	2	1	0	2	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	2	0	1	0	0	0
CE15	0.488	1	1	2	1	6	2	2	2	1	1	1	1	0	0	0	2	8	1	0	0	0	0	0	0	0	2	1	1	0	0	0
Oxidoreductases																																
POD	0.000	8	6	11	17	19	26	12	16	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	5	0	0	0

MCO	**0.000**	17	20	13	11	10	10	13	5	5	5	7	4	6	8	11	6	17	5	4	5	9	6	20	1	14	16	8	9	2	3	5
CRO[5]	**0.000**	5	8	9	4	9[4]	9	9	7	3	4	4	2	3	6	11	2	6	3	2	3	1	3	4	3	0	0	1	2	0	3	1
CDH[6]	**0.575**	1	1	1	1	1	1	1	1	0	0	0	1	2	2	0	1	1	0	0	0	0	0	0	0	2	3	0	3	0	0	0
Cytb562	**0.177**	1	1	0	0	0	1	1	1	0	0	0	0	2	3	0	2	0	0	0	0	0	0	0	0	0	1	0	2	0	0	0
OXO	**0.026**	3	3	2	3	3	5	5	7	5	4	5	4	3	2	1	5	1	2	0	0	0	0	1	0	2	2	3	3	0	0	0
GLP	**0.000**	1	11	6	8	1	2	6	3	5	10	10	6	3	11	1	7	0	3	0	0	0	0	2	0	0	0	0	0	0	0	0
QRD	**0.496**	2	1	3	3	4	1	1	4	1	1	1	3	2	2	2	4	3	1	1	2	0	1	1	1	1	1	1	1	4	1	4
DyP	**0.000**	1	2	5	3	11	2	1	0	2	0	0	0	0	0	2	0	4	0	1	1	0	0	2	0	0	0	0	0	0	0	0
HTP	**0.000**	5	10	8	4	16	3	4	3	5	5	4	6	3	2	5	3	8	6	0	0	0	3	17	0	5	2	4	13	0	0	0
P450	**0.000**	144	215	144	130	249	190	187	149	250	206	190	130	164	238	101	115	139	126	9	13	7	17	28	7	156	125	71	125	10	9	52

[1]Species: New genomes: Ad = *A. delicata*, Cp = *C. puteana*, Da = *Dacryopinax sp*., Ds = *D. squalens*, Fm = *F. mediterranea*, Fp = *F. pinicola*, Gt = *G. trabeum*, Pu = *P. strigosozonata*, Sh = *S. hirsutum*, Tm = *T. mesenterica*, Tv = *T. versicolor*, Wc = *W. cocos.* Others: An = *A. niger*, Bd = *B. dendrobatidis*, Cc = *C. cinerea*, Cn = *C. neoformans*, Cr = *C. parasitica*, Ha = *H. annosum* (has been reclassified as *H. irregulare*), Lb = *L. bicolor*, Mg = *M. globosa*, Ml = *M. laricis-populina*, Pb = *P. blakesleeanus*., Pc = *P. chrysosporium*, Pp = *P. placenta*, Ps = *P. stipitis*, Sc = *S. commune*, Sl = *S. lacrymans*, Sn = *S. nodorum*, Sr = *S. roseus*, Tr = *T. reesei*, Um = *U. maydis*. Higher taxa: Dac = Dacrymycetes, Trem = Tremellomycetes, Ust = Ustilaginomycotina, Pucc = Pucciniomycotina, Pez = Pezizomycotina, Sc = Saccharomycotina, Ch = Chytridiomycota, Mu = Mucoromycotina.

[2]WR = white rot, BR = brown rot, ECM = mycorrhiza, S = non-wood decay saprotroph, MP = mycoparasite, AP = animal pathogen/parasite, PP = plant pathogen, Y = yeast.

[3]POD= Clas-II peroxidases, MCO= multicopper oxidases, CRO= copper-radical oxidases, CDH= cellobiose dehydrogenase, Cytb562= cytochrome b562, OXO= oxalate oxidase/decarboxylases, GLP= Fe(III)-reducing glycopeptides, QRD= quinone reductases, DyP= dye-decolorizing peroxidases, HTP= heme-thiolate peroxidases, P450= cytochromes P450.

[4]*p*-values indicate strength of rejection of model of random diversification in CAFÉ analyses.

[5]One of four CRO2 genes in *A. delicata* is a potential pseudogene.

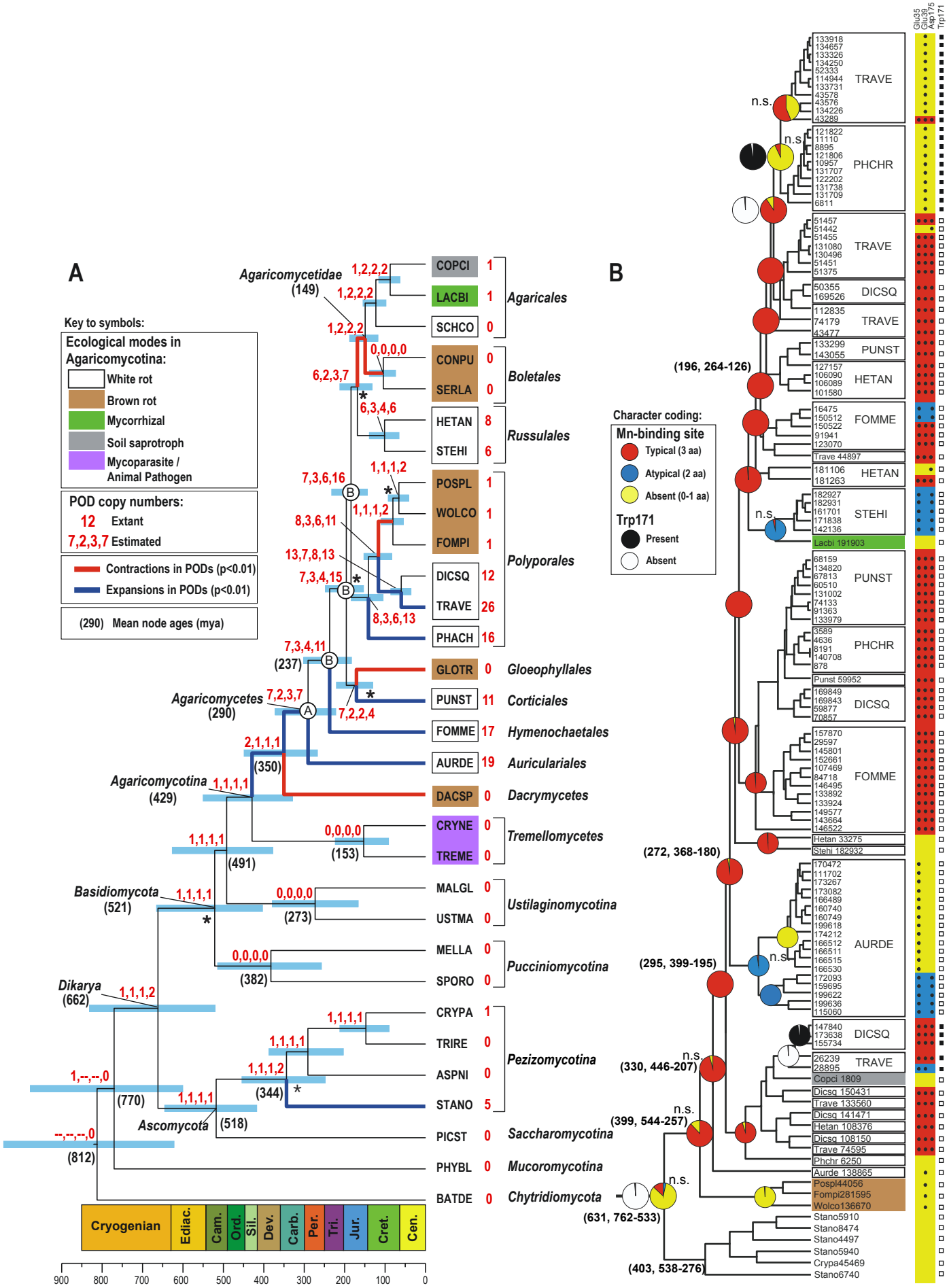
[6]One CDH gene in *C. puteana* lacks a cyt domain and may not be functional.

[7]GH3 does not include β-N-acetylhexosaminidase genes

[8]GH5 includes only models with similarity to endo-1,4-β-D-glucanases and mannan endo-β-1,4-mannosidases.

[9]One model (Fompi1 162677) is a potential pseudogene.

[10]The two models of Pospl1 and Wolco1 were missing from the original dataset as the clusters failed to recover them.


A
Key to symbols:
Ecological modes in Agaricomycotina:
White rot
Brown rot
Mycorrhizal
Soil saprotroph
Mycoparasite / Animal Pathogen
POD copy numbers:
12 Extant
7,2,3,7 Estimated
Contractions in PODs (p<0.01)
Expansions in PODs (p<0.01)
(290) Mean node ages (mya)
Agaricomycetidae (149)
Agaricomycetes (290)
Agaricomycotina (429)
Basidiomycota (521)
Dikarya (662)
Ascomycota (518)
COPCI 1
LACBI 1
SCHCO 0
Agaricales
CONPU 0
SERLA 0
Boletales
HETAN 8
STEHI 6
Russulales
POSPL 1
WOLCO 1
FOMPI 1
DICSQ 12
TRAVE 26
PHACH 16
Polyporales
GLOTR 0
Gloeophyllales
PUNST 11
Corticiales
FOMME 17
Hymenochaetales
AURDE 19
Auriculariales
DACSP 0
Dacrymycetes
CRYNE 0
TREME 0
Tremellomycetes
MALGL 0
USTMA 0
Ustilaginomycotina
MELLA 0
SPORO 0
Pucciniomycotina
CRYPA 1
TRIRE 0
ASPNI 0
STANO 5
Pezizomycotina
PICST 0
Saccharomycotina
PHYBL 0
Mucoromycotina
BATDE 0
Chytridiomycota
Cryogenian
Ediac.
Cam.
Ord.
Sil.
Dev.
Carb.
Per.
Tri.
Jur.
Cret.
Cen.
900 800 700 600 500 400 300 200 100 0
B
Character coding:
Mn-binding site
Typical (3 aa)
Atypical (2 aa)
Absent (0-1 aa)
Trp171
Present
Absent
Glu35
Glu39
Asp175
Trp171
TRAVE
PHCHR
DICSQ
PUNST
HETAN
FOMME
STEHI
AURDE
n.s.
(196, 264-126)
(272, 368-180)
(295, 399-195)
(330, 446-207)
(399, 544-257)
(631, 762-533)
(403, 538-276)
Copci 1809
Lacbi 191903
Pospi44056
Fompi281595
Wolco136670
Stano5910
Stano8474
Stano4497
Stano5940
Crypa45469
Stano6740