

binding site (fig. S8A and fig. S9), Get1 must displace helix $\alpha 2^{\text{Get2}}$, which is connected to helix $\alpha 1^{\text{Get2}}$ by the flexible glycine linker. NMR analyses revealed that Get1 binding indeed causes some Get2 interactions with Get3 to disappear. Specifically, interaction with Get2 was observed in the region of L4 to A49, and upon addition of Get1, residues G24 to A49 no longer interacted with Get3 (fig. S10). This shows that helix $\alpha 2^{\text{Get2}}$ is no longer bound to Get3 in the ternary complex.

On the basis of different crystal structures of Get3, we previously proposed a model for how the Get3 ATPase regulates TA protein insertion (19). With structures of different Get3-receptor complexes as well as functional data in hand, distinct docking states can be integrated into this model (Fig. 4). Assisted by Get4/5/Sgt2, TA proteins bind to Get3-ATP-Mg²⁺ (step 1). After ATP hydrolysis, the reaction products stay trapped, and the energy gained from hydrolysis is stored in a strained conformation (19). The N terminus of Get2 tethers the Get3/TA protein complex to the ER membrane (step 2). Binding of Get1 displaces $\alpha 2^{\text{Get2}}$, and the Get3/TA protein complex is now docked to the receptor complex at the membrane (step 3). When the TA protein is released, Get3 relaxes to the closed state, and inorganic phosphate dissociates (step 4). According to the crystal structures, Get1 can stay bound to Get3 during the transition from the closed to the open state. What actually triggers opening of Get3? We favor the idea that the energy from ATP hydrolysis drives Get3 to the open state, and ADP-Mg²⁺ leaves by way of the observed tunnels. In this state, Get1 interferes with nucleotide binding and prevents closure of the dimer. Finally, binding of ATP facilitates dissociation of Get3 (step 5), which sets the stage for the next targeting cycle. As Get1-CD is rigidly linked to the TMDs, structural changes observed in the Get3/Get1 complexes can be extrapolated to the complete membrane receptor (as indicated in Fig. 4 and fig. S11). The opening of Get3 during TA protein insertion may create a force that is directly transferred to the TMDs of the receptor, which could contribute to TA protein insertion. Related structural transitions have been reported for ATP-binding cassette (ABC) transporter proteins (27, 28). In Get1, the coiled-coil domain with the tip helix may have a function similar to the coupling helix in ABC transporters and may directly communicate nucleotide-dependent changes in Get3 to the transmembrane segments as anticipated in the model above. It is now important to dissect the precise mechanism of TA protein insertion and to see whether a general concept can be derived that is shared by different membrane transport systems.

References and Notes

1. G. Blobel, B. Dobberstein, *J. Cell Biol.* **67**, 835 (1975).
2. M. Schuldiner *et al.*, *Cell* **134**, 634 (2008).
3. N. Borgese, S. Brambilla, S. Colombo, *Curr. Opin. Cell Biol.* **19**, 368 (2007).
4. B. Wattenberg, T. Lithgow, *Traffic* **2**, 66 (2001).
5. S. High, B. M. Abell, *Biochem. Soc. Trans.* **32**, 659 (2004).

6. B. C. Cross, I. Sinning, J. Lührink, S. High, *Nat. Rev. Mol. Cell Biol.* **10**, 255 (2009).
7. P. F. Egea, R. M. Stroud, P. Walter, *Curr. Opin. Struct. Biol.* **15**, 213 (2005).
8. S. O. Shan, P. Walter, *FEBS Lett.* **579**, 921 (2005).
9. A. R. Osborne, T. A. Rapoport, B. van den Berg, *Annu. Rev. Cell Dev. Biol.* **21**, 529 (2005).
10. S. Brambilla, M. Yabal, M. Makarow, N. Borgese, *J. Cell Biol.* **175**, 767 (2006).
11. B. Meineke *et al.*, *FEBS Lett.* **582**, 855 (2008).
12. B. M. Abell, C. Rabu, P. Leznicki, J. C. Young, S. High, *J. Cell Sci.* **120**, 1743 (2007).
13. S. F. Colombo, R. Longhi, N. Borgese, *J. Cell Sci.* **122**, 2383 (2009).
14. S. Stefanovic, R. S. Hegde, *Cell* **128**, 1147 (2007).
15. V. Favalaro, F. Vilardi, R. Schlecht, M. P. Mayer, B. Dobberstein, *J. Cell Sci.* **123**, 1522 (2010).
16. C. Rabu, V. Schmid, B. Schwappach, S. High, *J. Cell Sci.* **122**, 3605 (2009).
17. V. Favalaro, M. Spasic, B. Schwappach, B. Dobberstein, *J. Cell Sci.* **121**, 1832 (2008).
18. M. Schuldiner *et al.*, *Cell* **123**, 507 (2005).
19. G. Bozkurt *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21131 (2009).
20. A. Mateja *et al.*, *Nature* **461**, 361 (2009).
21. C. J. M. Suloway, J. W. Chartron, M. Zaslaver, W. M. Clemons Jr., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 14849 (2009).
22. A. Yamagata *et al.*, *Genes Cells* **15**, 29 (2010).
23. J. Hu, J. Li, X. Qian, V. Denic, B. Sha, *PLoS ONE* **4**, e8061 (2009).
24. G. E. Tusnády, I. Simon, *Bioinformatics* **17**, 849 (2001).
25. N. Borgese, E. Fasana, *Biochim. Biophys. Acta* **1808**, 937 (2011).
26. F. Vilardi, H. Lorenz, B. Dobberstein, *J. Cell Sci.* **124**, 1301 (2011).
27. K. Hollenstein, R. J. P. Dawson, K. P. Locher, *Curr. Opin. Struct. Biol.* **17**, 412 (2007).
28. J. Zaitseva *et al.*, *EMBO J.* **25**, 3432 (2006).

Acknowledgments: V.D. would like to thank M. Frech for his support of the project and acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG) (SFB 807), the Centre for Biomolecular Magnetic Resonance (BMRZ), and the Cluster of Excellence Frankfurt (Macromolecular Complexes). I.S. thanks J. Kopp and C. Siegmund from the crystallization platform of the Biochemiezentrum and the Cluster of Excellence Heidelberg (CellNetworks), the European Synchrotron Radiation Facility for access to data collection, B. Dobberstein for generous support and stimulating discussions, and acknowledges funding by the DFG (SFB 638). Coordinates and structure factors have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (PDB) with accession nos. 3SJA, 3SJB, 3SJC, and 3SJD.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1207125/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 and S2
References

18 April 2011; accepted 21 June 2011

Published online 30 June 2011;

10.1126/science.1207125

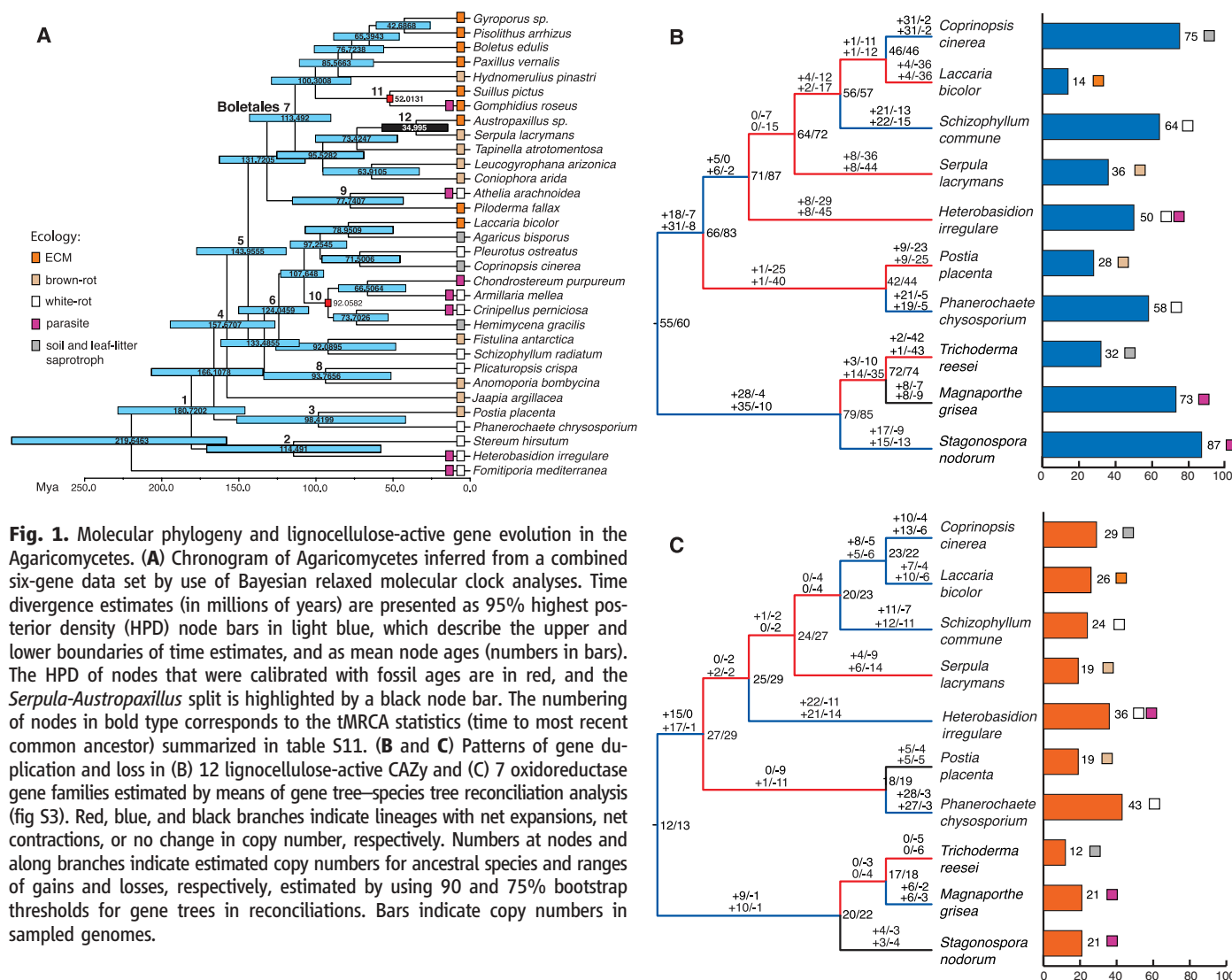
The Plant Cell Wall–Decomposing Machinery Underlies the Functional Diversity of Forest Fungi

Daniel C. Eastwood,^{1,*} Dimitrios Floudas,^{2,*} Manfred Binder,^{2,*} Andrzej Majcherczyk,^{3,*} Patrick Schneider,^{4,*} Andrea Aerts,⁵ Fred O. Asiegbu,⁶ Scott E. Baker,⁷ Kerrie Barry,⁵ Mika Bendiksby,⁸ Melanie Blumentritt,⁹ Pedro M. Coutinho,¹⁰ Dan Cullen,¹¹ Ronald P. de Vries,¹² Allen Gathman,¹³ Barry Goodell,^{9,14} Bernard Henrissat,¹⁰ Katarina Ihrmark,¹⁵ Hävard Kauserud,¹⁶ Annegret Kohler,¹⁷ Kurt LaButti,⁵ Alla Lapidus,⁵ José L. Lavin,¹⁸ Yong-Hwan Lee,¹⁹ Erika Lindquist,⁵ Walt Lilly,¹³ Susan Lucas,⁵ Emmanuelle Morin,¹⁷ Claude Murat,¹⁷ José A. Oguiza,¹⁸ Jongsun Park,¹⁹ Antonio G. Pisabarro,¹⁸ Robert Riley,⁵ Anna Rosling,¹⁵ Asaf Salamov,⁵ Olaf Schmidt,²⁰ Jeremy Schmutz,⁵ Inger Skrede,¹⁶ Jan Stenlid,¹⁵ Ad Wiebenga,¹² Xinfeng Xie,⁹ Ursula Kües,^{3,*} David S. Hibbett,^{2,*} Dirk Hoffmeister,^{4,*} Nils Högborg,^{15,*} Francis Martin,^{17,*} Igor V. Grigoriev,^{5,*} Sarah C. Watkinson^{21,*}

Brown rot decay removes cellulose and hemicellulose from wood—residual lignin contributing up to 30% of forest soil carbon—and is derived from an ancestral white rot saprotrophy in which both lignin and cellulose are decomposed. Comparative and functional genomics of the “dry rot” fungus *Serpula lacrymans*, derived from forest ancestors, demonstrated that the evolution of both ectomycorrhizal biotrophy and brown rot saprotrophy were accompanied by reductions and losses in specific protein families, suggesting adaptation to an intercellular interaction with plant tissue. Transcriptome and proteome analysis also identified differences in wood decomposition in *S. lacrymans* relative to the brown rot *Postia placenta*. Furthermore, fungal nutritional mode diversification suggests that the boreal forest biome originated via genetic coevolution of above- and below-ground biota.

Many Agaricomycete fungi have been sequenced to date (1), permitting comparative and functional genomic analyses of nutritional niche adaptation in the underground fungal networks that sustain boreal, temperate, and some subtropical forests (2). Through the se-

quencing of the brown rot wood decay fungus *Serpula lacrymans*, we conducted genome comparisons with sequenced fungi, including species representing each of a range of functional niches: brown rot and white rot wood decay, parasitism, and mutualistic ectomycorrhizal symbiosis.

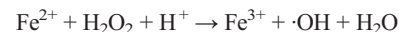


Only 6% of wood-decay species are brown rots (3), but being associated with conifer wood (4), they dominate decomposition in boreal forests. Their lignin residues contribute up to 30% of carbon in the organic soil horizons (5). Long-lived

(6) and with capacity to bind nitrogen and cations (7), these phenolic polymers condition the nutrient-poor acidic soils of northern conifer forests.

Brown rot wood decay involves an initial non-enzymic attack on the wood cell wall (8), gen-

erating hydroxyl radicals ($\cdot\text{OH}$) extracellularly via the Fenton reaction:



Hydrogen peroxide is metabolically generated by oxidase enzymes such as glyoxal oxidases and copper radical oxidases. The hydroxyl radical has a half-life of nanoseconds (8) and is the most powerful oxidizing agent of living cells. However, we do not know how it is spatially and temporally targeted to wood cell wall components. Divalent iron is scarce in aerobic environments, where the fungus is obligate and the trivalent ion is energetically favored. Phenolates synthesized by brown rot fungi, including *S. lacrymans* (9), can reduce Fe^{3+} to Fe^{2+} . Such phenolates may be modified lignin derivatives or fungal metabolites (10). After initial bond breakages in the cellulose chain, side chain hemicelluloses (arabinan and galactan) are removed, followed by main chains [xylan and mannan (11)], with subsequent hydrolysis of cellulose by synergistic glycoside hydrolases

¹College of Science, University of Swansea, Singleton Park, Swansea SA2 8PP, UK. ²Department of Biology, Clark University, Worcester, MA 01610, USA. ³Georg-August-University Göttingen, Büsgen-Institute, Büsgenweg 2, 37077 Göttingen, Germany. ⁴Friedrich-Schiller-Universität, Hans-Knöll-Institute, Beutenbergstrasse 11a, 07745 Jena, Germany. ⁵U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA. ⁶Department of Forest Sciences, Box 27, University of Helsinki, Helsinki 00014, Finland. ⁷Pacific Northwest National Laboratory, 902 Battelle Boulevard, Post Office Box 999, MSIN P8-60, Richland, WA 99352, USA. ⁸Natural History Museum, University of Oslo, Post Office Box 1172, Blindern, NO-0138, Norway. ⁹Wood Science and Technology, University of Maine, Orono, ME 04469-5755, USA. ¹⁰UMR 6098 CNRS-Universités Aix-Marseille I and II, 13288 Marseille Cedex 9, France. ¹¹Forest Products Laboratory, Madison, WI 53726, USA. ¹²Centraalbureau voor Schimmelfcultures-Royal Netherlands Academy of Arts and Sciences Fungal Biodiversity Centre, Uppsalalaan 8, 3584 CT Utrecht, Netherlands. ¹³Department of Biology, Southeast Missouri State University,

Cape Girardeau, MO 63701, USA. ¹⁴Department of Wood Science and Forest Products, 230 Cheatham Hall, Virginia Tech, Blacksburg, VA 24061, USA. ¹⁵Department of Forest Mycology and Pathology, Swedish University of Agricultural Sciences, S-750 07 Uppsala, Sweden. ¹⁶Department of Biology, University of Oslo, Post Office Box 1066 Blindern, N-0316 Oslo, Norway. ¹⁷UMR 1136, Institut National de la Recherche Agronomique (INRA)-Nancy Université, Interactions Arbres/Microorganismes, INRA-Nancy, 54280 Champenoux, France. ¹⁸Department of Agrarian Production, Public University of Navarre, 31006 Pamplona, Spain. ¹⁹Department of Agricultural Biotechnology, Seoul National University, Seoul 151-921, Korea. ²⁰Department of Wood Biology, University of Hamburg, Leuschnerstrasse 91, D-21031 Hamburg, Germany. ²¹Department of Plant Sciences, University of Oxford, Oxford OX1 3RB, UK.

*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: d.c.eastwood@swansea.ac.uk

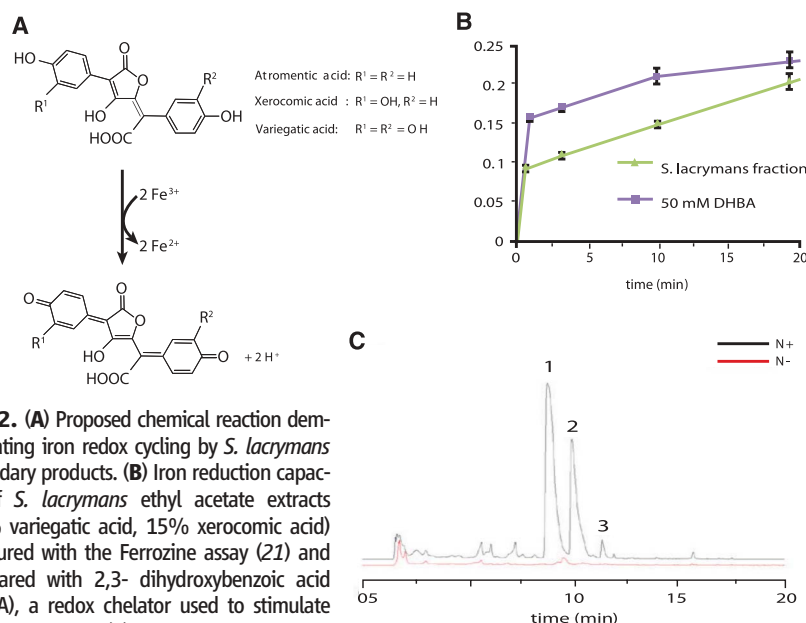


Fig. 2. (A) Proposed chemical reaction demonstrating iron redox cycling by *S. lacrymans* secondary products. (B) Iron reduction capacity of *S. lacrymans* ethyl acetate extracts (60% variiegatic acid, 15% xerocomic acid) measured with the Ferrozine assay (21) and compared with 2,3-dihydroxybenzoic acid (DHBA), a redox chelator used to stimulate Fenton systems. (C) Comparison of HPLC chromatograms of *S. lacrymans* ethyl acetate extracts as a function of nitrogen supply. Red trace, nitrogen rich medium (+N); black trace, nitrogen-depleted minimal medium (-N). The identity of the compounds was confirmed with mass spectrometry and by their ultraviolet-visual spectrum (1, variiegatic acid; 2, xerocomic acid; 3, atramentic acid).

(GHs). Residual lignin is demethylated. In contrast, white rot fungi decompose both cellulose and lignin, with free radical attack theorized to break a variety of bonds in the lignin phenylpropanoid heteropolymer.

S. lacrymans is in the Boletales, along with several ectomycorrhizal lineages (Fig. 1A) (12). *S. lacrymans* is thus phylogenetically distant from brown rot *Postia placenta* (Polyporales) (13), as well as other sequenced ectomycorrhizal fungi (14, 15), parasites, and white rot wood decomposers (16). We estimated divergence dates in fungal phylogeny using the data set of Binder *et al.* [supporting online material (SOM), molecular clock analyses] (17), with two well-characterized fungal fossils that were used to calibrate the minimum ages of the marasmioid (Fig. 1A, node 10) and suilloid clades (Fig. 1A, node 11). The estimated age of the split between *Serpula* and its ectomycorrhizal sister-group *Austropaxillus* (53.1 to 15 million years ago) (Fig. 1A and table S11) suggests that transition from brown rot saprotrophy to mutualistic symbiosis occurred after rosid (Eurosids I) became widespread (Fig. 1A) (18). Diversification in fungal nutritional modes occurred alongside diversification of angiosperms and gymnosperms, as these fungi are currently associated with members of both gymnosperms (Pinaceae) and angiosperms (18).

S. lacrymans comprises two subgroups that diverged in historic time (19), *S. lacrymans* var. *shastensis*, which is found in montane conifer forest, and *S. lacrymans* var. *lacrymans*, which is a cause of building dry rot. Two *S. lacrymans* var. *lacrymans* complementary monokaryons (haploids of strain S7), S7.9 (A2B2) and S7.3 (A1B1) (20),

were sequenced via Sanger and 454 pyrosequencing, respectively. The genome of *S. lacrymans* S7.9 was 42.8 megabase pairs (Mbp), containing 12,917 gene predictions (21).

We analyzed 19 gene families of enzymes for lignocellulose breakdown: carbohydrate active enzymes (CAZys; www.cazy.org) (22) (GHs and carbohydrate esterases) and oxidoreductases (table S9). Losses and expansions in these families were compared across 10 fungi, including Agaricomycetes, with a range of nutritional modes (Fig. 1, B and C, and table S9). Convergent changes in enzyme complement were found in the two independently evolved brown rot species, with parallels in the ectomycorrhizal *Laccaria bicolor* (fig. S3 and table S9). The inferred most recent common ancestor of the Agaricales, Boletales, and Polyporales is predicted to be a white rot with 66 to 83 hydrolytic CAZy genes and 27 to 29 oxidoreductases (Fig. 1, B and C). Brown rot and ectomycorrhizal fungi have the fewest hydrolytic CAZy genes. Brown rot fungi have the fewest oxidoreductases, not because of gene losses but because of gene duplications in white rot species.

Both brown rot and ectomycorrhizal fungi lacked class II peroxidases, which are used by white rot fungi in depolymerizing the lignin matrix of wood and unmasking usable cellulose embedded within it. This family was expanded in the white rots *Coprinopsis cinerea*, *Phanerochaete chrysosporium*, and *Schizophyllum commune*, with 29, 43, and 24 genes, respectively, with only 19 each in *S. lacrymans* and *P. placenta*. Oxidoreductases conserved in brown rot fungi included iron and quinone reductases and multicopper oxidases (fig. S3 and table S8). Absence of ligninolysis in

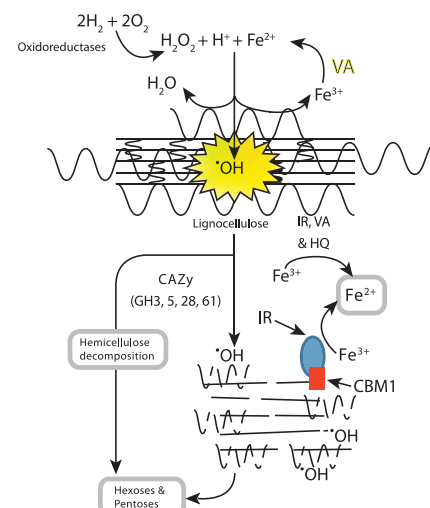


Fig. 3. Schematic overview of the proposed mechanism of wood decay by *S. lacrymans*. Scavenging mycelium colonizes a new food source, inducing VA production and expression of oxidoreductase enzymes, which drive hydroxyl radical attack on the lignocellulose composite. CAZy gain access to the weakened composite structure and break down accessible carbohydrates. Cellulose-binding iron reductase targets $\cdot\text{OH}$ -generating Fenton's reaction on cellulose chains, releasing chain ends for hydrolysis and assimilation. IR, iron reductase; HQ, hydroxyquinones; CBM, cellulose-binding module.

brown rots raises the question of how they achieve pervasive cellulolysis in wood with the lignin matrix intact.

GH gene families had parallel patterns of losses and expansion in both brown rots and ectomycorrhizas. CAZy families GH5 (endoglucanases, hydrolyzing cellulose) and GH28 (pectinases, hydrolyzing intercellular cohesive polysaccharides in plant tissues) were expanded in both brown rot species, in which they might facilitate intercellular enzyme diffusion, and retained in *L. bicolor*, in which they might facilitate intercellular penetration of living roots. Both brown rot species lacked GH7 (endoglucanase/cellobiohydrolase CBHI), and GH61 genes—with unknown function but recently implicated in oxidative attack on polysaccharides (23)—were reduced. GH6 (cellobiohydrolase CBHII) and cellulose-binding modules (CBM1), which were absent from *P. placenta* (13), were present in *S. lacrymans*. One CBM was associated with an iron reductase in a gene (*S. lacrymans* S7.9 database protein ID, 452187) originally derived from a cellobiose dehydrogenase (fig. S5).

The general utility of the conserved suite of GH genes in wood decay by *S. lacrymans* was supported through transcriptomic and proteomic analysis. Carbohydrate-active enzymes accounted for 50% of proteins identified (table S14), and 33.9% of transcripts regulated greater than 20-fold by *S. lacrymans* growing on pine wood as compared with glucose medium (fig. S4).

Cellulose-, pectin-, and hemicellulose-degrading enzymes (GH families 5, 61, 3, and 28) were prominent, and GH5 endoglucanase (*S. lacrymans* S7.9 database protein ID, 433209) and GH74 endoglucanase/xyloglucanase (*S. lacrymans* S7.9 database protein ID, 453342) were up-regulated greater than 100-fold.

We conclude that brown rot fungi have cast off the energetically expensive apparatus of ligninolysis and acquired alternative mechanisms of initial attack. Wood decomposition by *S. lacrymans* may involve metabolically driven nonenzymatic disruption of lignocellulose with internal breakage of cellulose chains by highly localized $\cdot\text{OH}$ radical action. Mycelia in split plates mimicking realistic nutrient heterogeneity (fig. S1) produced variegatic acid (VA), an iron-reducing phenolate (Fig. 2, A to C), via the Boletales atromentin pathway, which was recruited in *S. lacrymans* for the Fenton's reaction. The genome was rich in secondary metabolism genes (table S15), including a putative atromentin locus (24). Mycelium imports amino acids to sites of wood colonization (25), which is consistent with observed up-regulation of oligopeptide transporters on wood (table S12). Localizing variegatic acid production to well-resourced parts of the mycelium could enhance Fenton's chemistry in contact with wood.

Wood colonization is presumably followed by coordinated induction of the decay machinery revealed in the wood-induced transcriptome (Fig. 3 and fig. S4). GHs and oxidoreductases accounted for 20.7% of transcripts, accumulating greater than fourfold on wood relative to glucose medium (fig. S4 and table S12). Iron reduction mechanisms included an enzyme harboring a C terminal cellulose-binding module (*S. lacrymans* S7.9 database protein ID, 452187) (fig. S5) that is up-regulated 122-fold on wood substrate (fig. S4 and table S12). This enzyme, which is present in *Ph. chrysosporium* but absent from *P. placenta* (26), is a potential docking mechanism for localizing iron reductase activity, and hence $\cdot\text{OH}$ generation, on the surface of microcrystalline cellulose. Cellulose-targeted iron reduction, combined with substrate induction of variegatic acid biosynthesis, might explain the particular ability of brown rot fungi in Boletales to degrade unassociated microcrystalline cellulose without the presence of lignin (27).

Thus, comparative genomics helps us understand the molecular processes of forest soil fungi that drive the element cycles of forest biomes (28). Sequenced forest Agaricomycetes revealed shared patterns of gene family contractions and expansions associated with emergences of both brown rot saprotrophy and ectomycorrhizal symbiosis. In Boletales, loss of aggressive ligninolysis might have permitted brown rot transitions to biotrophic ectomycorrhiza, which is promoted in soils impoverished in nitrogen by brown rot residues, and by the nutritional advantage conferred by the connection to a mycorrhizal network. *S. lacrymans* and other fungi cultured with conifer roots (29) ensheath *Pinus sylvestris* roots

with a mantle-like layer (fig. S6), suggesting nutrient exchange.

The chronology of divergences in extant fungal nutritional mode (Fig. 1A) matches the predicted major diversification in conifers (18), suggesting that the boreal forest biome may have originated via genetic coevolution of above- and below-ground biota.

References and Notes

1. F. Martin *et al.*, *New Phytol.* **190**, 818 (2011).
2. F. Martin, in *Biology of the Fungal Cell*, R. J. Howard, N. A. R. Gow, Eds. (Springer Berlin, Heidelberg, 2007), vol. 8, pp. 291–308.
3. R. L. Gilbertson, *Mycologia* **72**, 1 (1980).
4. D. S. Hibbett, M. J. Donoghue, *Syst. Biol.* **50**, 215 (2001).
5. K.-E. Eriksson, R. A. Blanchette, P. Ander, *Microbial and enzymatic degradation of wood and wood components* (Springer-Verlag, Berlin, New York, 1990).
6. B. D. Lindahl *et al.*, *New Phytol.* **173**, 611 (2007).
7. R. R. Northup, Z. Yu, R. A. Dahlgren, K. A. Vogt, *Nature* **377**, 227 (1995).
8. B. Goodell *et al.*, *J. Biotechnol.* **53**, 133 (1997).
9. T. Shimokawa, M. Nakamura, N. Hayashi, M. Ishihara, *Holzforchung* **58**, 305 (2005).
10. V. Arantes, A. M. Milagres, T. R. Filley, B. Goodell, *J. Ind. Microbiol. Biotechnol.* **38**, 541 (2011).
11. S. F. Curling, C. A. Clausen, J. E. Winandy, *Int. Biodeterior. Biodegradation* **49**, 13 (2002).
12. M. Binder, D. S. Hibbett, *Mycologia* **98**, 971 (2006).
13. D. Martinez *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 1954 (2009).
14. F. Martin *et al.*, *Nature* **464**, 1033 (2010).
15. F. Martin *et al.*, *Nature* **452**, 88 (2008).
16. D. Martinez *et al.*, *Nat. Biotechnol.* **22**, 695 (2004).
17. M. Binder, K. H. Larsson, P. B. Matheny, D. S. Hibbett, *Mycologia* **102**, 865 (2010).
18. A. J. Eckert, B. D. Hall, *Mol. Phylogenet. Evol.* **40**, 166 (2006).
19. H. Kauserud *et al.*, *Mol. Ecol.* **16**, 3350 (2007).
20. O. Schmidt, *Holzforchung* **54**, 221 (2000).
21. Materials and methods are available as supporting material on Science Online.
22. B. L. Cantarel *et al.*, *Nucleic Acids Res.* **37**, D233 (2009).
23. G. Vaaje-Kolstad *et al.*, *Science* **330**, 219 (2010).
24. P. Schneider, S. Bouhired, D. Hoffmeister, *Fungal Genet. Biol.* **45**, 1487 (2008).
25. M. Tialka, M. Fricker, S. Watkinson, *Appl. Environ. Microbiol.* **74**, 2700 (2008).
26. A. Vanden Wymelenberg *et al.*, *Appl. Environ. Microbiol.* **76**, 3599 (2010).
27. T. Nilsson, J. Ginns, *Mycologia* **71**, 170 (1979).
28. B. O. Lindahl, A. F. S. Taylor, R. D. Finlay, *Plant Soil* **242**, 123 (2002).
29. R. Vasiliauskas, A. Menkis, R. D. Finlay, J. Stenlid, *New Phytol.* **174**, 441 (2007).

Acknowledgments: J. Schilling, University of Minnesota, and D. Barbara, University of Warwick, critically reviewed the manuscript; T. Marks designed graphics; and B. Wackler and M. Zomorodi gave technical assistance. Assembly and annotations of *S. lacrymans* genomes are available at www.jgi.doe.gov/Serpula and DNA Data Bank of Japan/European Molecular Biology Laboratory/GenBank, accessions nos. AECQB000000000 and AEQC000000000. The complete microarray expression data set is available at the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo/) accession no. GSE27839. The work was conducted by the U.S. Department of Energy Joint Genome Institute and supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231. Further financial support is acknowledged in the supporting online material on Science Online.

Supporting Online Material

www.sciencemag.org/cgi/content/full/science.1205411/DC1
Materials and Methods
SOM Text
Figs. S1 to S6
Tables S1 to S15
References (30–89)

10 March 2011; accepted 20 June 2011
Published online 14 July 2011;
10.1126/science.1205411

The Leukemogenicity of AML1-ETO Is Dependent on Site-Specific Lysine Acetylation

Lan Wang,¹ Alexander Gural,¹ Xiao-Jian Sun,² Xinyang Zhao,¹ Fabiana Perna,¹ Gang Huang,¹ Megan A. Hatlen,¹ Ly Vu,¹ Fan Liu,¹ Haiming Xu,¹ Takashi Asai,¹ Hao Xu,¹ Tony Deblasio,¹ Silvia Menendez,¹ Francesca Voza,¹ Yanwen Jiang,³ Philip A. Cole,⁴ Jinsong Zhang,⁵ Ari Melnick,³ Robert G. Roeder,² Stephen D. Nimer^{1*}

The chromosomal translocations found in acute myelogenous leukemia (AML) generate oncogenic fusion transcription factors with aberrant transcriptional regulatory properties. Although therapeutic targeting of most leukemia fusion proteins remains elusive, the posttranslational modifications that control their function could be targetable. We found that AML1-ETO, the fusion protein generated by the t(8;21) translocation, is acetylated by the transcriptional coactivator p300 in leukemia cells isolated from t(8;21) AML patients, and that this acetylation is essential for its self-renewal-promoting effects in human cord blood CD34⁺ cells and its leukemogenicity in mouse models. Inhibition of p300 abrogates the acetylation of AML1-ETO and impairs its ability to promote leukemic transformation. Thus, lysine acetyltransferases represent a potential therapeutic target in AML.

Histone-modifying enzymes can regulate the binding of specific chromatin-binding proteins to histone marks and can change the affinity of the histones for DNA (1, 2). These

enzymes also affect nonhistone proteins, and posttranslational modifications of transcription factors such as p53 or AML1 (which is required for definitive hematopoietic development) can



www.sciencemag.org/cgi/content/full/science.1205411/DC1

Supporting Online Material for

The Plant Cell Wall–Decomposing Machinery Underlies the Functional Diversity of Forest Fungi

Daniel C. Eastwood,* Dimitrios Floudas, Manfred Binder, Andrzej Majcherczyk, Patrick Schneider, Andrea Aerts, Fred O. Asiegbu, Scott E. Baker, Kerrie Barry, Mika Bendiksby, Melanie Blumentritt, Pedro M. Coutinho, Dan Cullen, Ronald P. de Vries, Allen Gathman, Barry Goodell, Bernard Henrissat, Katarina Ihrmark, Hävard Kauserud, Annegret Kohler, Kurt LaButti, Alla Lapidus, José L. Lavin, Yong-Hwan Lee, Erika Lindquist, Walt Lilly, Susan Lucas, Emmanuelle Morin, Claude Murat, José A. Oguiza, Jongsun Park, Antonio G. Pisabarro, Robert Riley, Anna Rosling, Asaf Salamov, Olaf Schmidt, Jeremy Schmutz, Inger Skrede, Jan Stenlid, Ad Wiebenga, Xinfeng Xie, Ursula Kües, David S. Hibbett, Dirk Hoffmeister, Nils Högberg, Francis Martin, Igor V. Grigoriev, Sarah C. Watkinson

*To whom correspondence should be addressed. E-mail: d.c.eastwood@swansea.ac.uk

Published 14 July 2011 on *Science Express*
DOI: 10.1126/science.1205411

This PDF file includes:

Materials and Methods

SOM Text

Figs. S1 to S6

Tables S1 to S15

References (30–89)

SUPPORTING ONLINE MATERIAL

TABLE OF CONTENTS

| | | |
|------|--|----|
| 1 | MATERIALS AND METHODS..... | 3 |
| 1.1 | Nucleic acid extraction for genomic study..... | 3 |
| 1.2 | Sequencing and assembly..... | 3 |
| 1.3 | EST clustering and assembly..... | 4 |
| 1.4 | Genome annotation..... | 5 |
| 1.5 | Molecular clock analysis..... | 6 |
| 1.6 | Comparative study of the evolution of carbohydrate active and oxidoreductase enzyme gene repertoires..... | 6 |
| 1.7 | Comparative transcriptomic analysis of <i>Serpula lacrymans</i> growth on defined glucose media and on <i>Pinus sylvestris</i> sapwood..... | 7 |
| 1.8 | Proteomic analysis of <i>Serpula lacrymans</i> grown on wood..... | 9 |
| 1.9 | Genome analysis of natural product genes and secondary metabolite analysis under nutrient asymmetry..... | 10 |
| 1.10 | <i>S. lacrymans</i> interaction with <i>Picea sylvestris</i> seedlings..... | 11 |
| 2 | ADDITIONAL RESULTS AND DISCUSSION..... | 11 |
| 2.1 | Genome assembly, annotation and analysis..... | 11 |
| 2.2 | Molecular clock analysis..... | 12 |
| 2.3 | Comparative study of the evolution of carbohydrate active and oxidoreductase enzyme gene repertoires..... | 12 |
| 2.4 | Comparison between transcriptomic and proteomic data..... | 15 |
| 2.5 | Natural product genes in the <i>Serpula lacrymans</i> genome and description of the atromentin locus involved in variegatic acid production..... | 17 |
| 2.6 | <i>S. lacrymans</i> interaction with <i>Picea sylvestris</i> seedlings..... | 17 |
| 3 | FURTHER FUNDING SOURCES..... | 18 |
| 4 | SUPPLEMENTARY FIGURES..... | 19 |
| | Fig S1. LCMS/MS analysis of <i>S. lacrymans</i> S7.9 split plate extract and purified Fractions..... | 19 |
| | Fig S2. CAFE anaylis of the total number of protein families in each species node..... | 20 |
| | Fig S3. Reconciliation analysis of lignocellulose-active enzymes from sequenced fungal genomes..... | 21 |
| | Fig S4. Functional characterisation of <i>S. lacrymans</i> transcripts with significant differentiation when grown on wood..... | 39 |
| | Fig S5. <i>S. lacrymans</i> S7.9 protein models with similar iron reductase domain..... | 40 |
| | Fig S6. Interaction between <i>S. lacrymans</i> and roots of <i>Picea sylvestris</i> | 41 |
| 5 | SUPPLEMENTARY TABLES..... | 42 |
| | Table S1. Genomic libraries included in the <i>Serpula lacrymans</i> genome assembly...42 | |
| | Table S2. Summary statistics of the output of the <i>S. lacrymans</i> S7.9 whole genome shot gun assembly..... | 42 |
| | Table S3. <i>S. lacrymans</i> final assembly statistics..... | 42 |
| | Table S4. Predicted gene models and supporting lines of evidence..... | 43 |
| | Table S5. Characteristics of predicted gene models..... | 43 |
| | Table S6. Functional annotation of proteins..... | 43 |
| | Table S7. Comparison of the putative CAZy enzymes from genome sequenced fungi with differing nutritional modes..... | 44 |
| | Table S8. Comparison of the putative oxidoreductase enzymes from genome sequenced fungi with differing nutritional modes..... | 46 |
| | Table S9. Gene families and organisms used for reconciliation analysis..... | 48 |

| | |
|--|----|
| Table S10. Number of orthologs between <i>Serpula</i> strains and with other Agaricomycetes..... | 49 |
| Table S11. Posterior probability distribution for divergence times for major lineages in the Agaricomycetes..... | 49 |
| Table S12. <i>S. lacrymans</i> transcriptomic comparison from cultures grown on either glucose or wood wafers..... | 50 |
| Table S13. Extracellular proteins from <i>S. lacrymans</i> S7 grown in solid state wood Culture..... | 55 |
| Table S14. Comparison of proteins identified as expressed on solid state wood culture with the transcriptomic study..... | 56 |
| Table S15. Comparison of secondary metabolic genes in sequenced basidiomycetes | 57 |
| 6 REFERENCES..... | 58 |

1. MATERIALS AND METHODS

1.1 Nucleic acid extraction for genomic study

DNA extraction. *Serpula lacrymans* monokaryons S7.9 and S7.3, prepared from *S. lacrymans* S7 dikaryon by Dr O. Schmit, University of Hamburg, were inoculated in Czapek dox (30 gL⁻¹ sucrose, 1 gL⁻¹ monosodium glutamate, 1 gL⁻¹ KH₂PO₄, 0.5 gL⁻¹ MgSO₄ x7H₂O, 0.01 gL⁻¹ FeSO₄ x7H₂O) liquid medium at 20°C for 21 days. Cultures were harvested and filtered between muslin before freeze drying. Harvested mycelia were ground under liquid nitrogen and DNA extracted following the CTAB protocol described on the JGI website (<http://my.jgi.doe.gov/general/index.html>). DNA quantity and quality was verified by agarose gel electrophoresis and Nanodrop spectrophotometry (Thermo Scientific, Hertfordshire, UK).

RNA extraction. *S. lacrymans* S7.9 was cultured under a range of conditions to maximise the number of genes expressed. The fungus was cultured on 1) Czapek dox medium (described above), 2) Czapek dox where either sucrose and monosodium glutamate was reduced to 3 gL⁻¹ or 0.1 gL⁻¹ respectively, 3) Czapek dox where the sucrose was replaced by carboxymethyl cellulose, 4) complex medium consisting of 20 gL⁻¹ malt extract (BD Difco, Oxford, UK) and 5 gL⁻¹ yeast extract, 6) aged culture, grown in Czapek dox for 10 weeks, and 7) heat and cold shock, cultures grown for 21 days in Czapek dox at 20 °C were placed at either 25 °C or 4 °C for four hours prior to harvesting. Mycelium was harvested and freeze dried as above. Total RNA was extracted using the phenol:chloroform protocol described previously (30). RNA quantity and quality was verified by formamide agarose gel electrophoresis and Nanodrop spectrophotometry (Thermo Scientific).

1.2 Sequencing and assembly

Serpula lacrymans S7.9 was sequenced using Sanger sequencing on ABI 3730XL capillary machines (Life Technologies, California, USA). Three different sized libraries were used as templates for the plasmid subclone sequencing process and both ends were sequenced. 234,528 reads from the 2.5 kb sized library, 291,744 reads from the 6.4 kb sized library, and 29,952 reads from a 39.8 kb fosmid library were sequenced (table S1). A total of 564,672 reads were assembled using a modified version of Arachne (31) v.20071016 with parameters maxcliq1=100, correct1_passes=0 and BINGE_AND_PURGE=True (see table S2 for scaffold and

contigs totals). This produced 68 scaffold sequences, with L50 of 2.9 Mb, 22 scaffolds larger than 100 kb, and total scaffold size of 43.0 Mb. Each scaffold was screened against bacterial proteins, organelle sequences and GenBank using megablast against Genbank NR and blastp against a set of known microbial proteins. No scaffolds were identified as contamination. We classified additional scaffolds as unanchored rDNA (18 scaffolds), mitochondrion (1 scaffold), and small repetitive (3 scaffolds). Additional scaffolds were removed if they consisted of greater than 95% 24mers that occurred 4 other times in the scaffolds larger than 50kb or if the scaffold contained only unanchored rDNA sequences. The final assembly contains 46 scaffolds that cover 42.4 Mb of the genome with a contig L50 of 228.0 kb and a scaffold L50 of 2.9 Mb (table S3)

Serpula lacrymans S7.3 genome was sequenced using 454 (Roche, Connecticut, USA) pyrosequencing platforms, which included 6 half runs of unpaired 454 Titanium data (table S1). All general aspects of library construction and sequencing can be found at the JGI website (<http://www.jgi.doe.gov/>). After filtering for low quality reads and contaminants, the resulting 454 reads were assembled with the Newbler assembler version 2.3-PreRelease-9/14/2009 to the final estimated assembled coverage of 24X (table S1) and resulted in 6088 contigs with an N/L50 of 247/45.5Kb. To improve the assembly it was scaffolded against *Serpula lacrymans* S7.9, using Nucmer show-tilings (32). The final assembly included 2128 scaffolds with an N/L50 of 7/2.9Mb (table S3).

1.3 EST clustering and assembly

Serpula lacrymans S7.9 total RNA was used to extract PolyA RNA using oligo (dT) magnetic beads (Absolutely mRNA™ Purification kit, Stragene, Agilent Technologies, California, USA). PolyA RNA was reversed transcribed using Superscript III (Invitrogen) using a dT₁₅VN₂ primer. Second strand cDNA was synthesized by nick translation with *E. coli* DNA ligase, *E. coli* DNA polymerase I, and RNase H and blunt end repaired using T4 polymerase (Invitrogen, Life Technologies, California, USA). The dscDNA was fragmented and 300-800 base pair fragments were gel purified using a 2% agarose gel. The purified fragments were then used to create the 454 single stranded cDNA library as described below (454 library preparation kit, Roche).

The fragment ends were polished using T4 ligase and T4 polynucleotide kinase (Roche). Adaptors containing primer sequences and a biotin tag were ligated to the fragment ends (Roche). The fragments with properly ligated adaptors were immobilized onto magnetic streptavidin coated beads (Roche). Nicks or gaps between the adaptors and the dscDNA fragments were repaired using the fill-in polymerase (Roche). The non-biotinylated strands of the immobilized dscDNA fragments were melted off to generate the single stranded cDNA library for 454 sequencing.

The ESTs were evaluated for the presence of polyA tails (which if present were removed) then evaluated for length, removing ESTs with fewer than 50 bases remaining. Additionally, ESTs consisting of more than 50% low complexity sequence were removed from the final set of ESTs. The resulting set of ~751k ESTs were used for clustering.

For clustering, ESTs were evaluated with malign, a kmer based alignment tool (Chapman, unpublished), which clusters ESTs based on sequence overlap (kmer = 16, seed length requirement = 32 alignment ID >= 98%). EST clusters were then each

assembled using CAP3 (33) to form consensus sequences. Clustering and assembly of all ESTs resulted in 107,971 consensus sequences and 69,437 singlets.

1.4 Genome annotation

Both genomes were annotated using the JGI annotation pipeline, which takes multiple inputs (scaffolds, ESTs, and known genes) and runs several analytical tools for gene prediction and annotation, and deposits the results in the JGI Genome Portal (<http://www.jgi.doe.gov/Serpula>) for further analysis and manual curation.

Genomic assembly scaffolds were masked using RepeatMasker (34) and the RepBase library of 234 fungal repeats (35). tRNAs were predicted using tRNAscan-SE (36). Using the repeat-masked assembly, several gene prediction programs falling into three general categories were used: 1) *ab initio* - FGENESH (37); GeneMark (38) trained on full-length *Serpula lacrymans* genes, 2) *homology-based* - FGENESH+; Genewise (39) seeded by BLASTx alignments against GenBank's database of non-redundant proteins (NR: <http://www.ncbi.nlm.nih.gov/BLAST/>), and 3) *EST-based* - EST_map (<http://www.softberry.com/>) seeded by *Serpula lacrymans* EST contigs. Genewise models were extended where possible using scaffold data to find start and stop codons. EST BLAT alignments (40) were used to extend, verify, and complete the predicted gene models. The resulting set of models was then filtered for the best models, based on EST and homology support, to produce a non-redundant representative set. This representative set (12,917 and 14,495 genes in S7.9 and S7.3, respectively – tables S4-S5) was subject to further analysis and manual curation. Measures of model quality include proportions of the models complete with start and stop codons (>82% of models), consistent with ESTs (>63% of models covered over ≥75% of exon length), supported by similarity with proteins from the NCBI NR database (>68% of models). Quality metrics for gene models are summarized in tables S4-S5.

All predicted gene models were functionally annotated using SignalP (41), TMHMM (42), InterProScan (43), BLASTp (44) against nr, and hardware-accelerated double-affine Smith-Waterman alignments (deCypherSW; http://www.timelogic.com/decypher_sw.html) against SwissProt (<http://www.expasy.org/sprot/>), KEGG (45), and KOG (46). KEGG hits were used to assign EC numbers (<http://www.expasy.org/enzyme/>), and Interpro and SwissProt hits were used to map GO terms (<http://www.geneontology.org/>). Functional annotations are summarized in Table S6. Manual curation of the automated annotations was performed by using the web-based interactive editing tools of the JGI Genome Portal to assess predicted gene structures, assign gene functions, and report supporting evidence.

Multigene families were assembled from 158,123 predicted proteins found in *S. lacrymans*, representatives from Agaricomycotina, Pucciniomycotina and Ustilagomycotina phyla using the MCL algorithm (47) with inflation parameter set to 3.0. Multigene families were then analyzed for evolutionary changes in protein family size using the CAFE program (48). The latter program uses a random birth and death process to model gene gain and loss across a user specified tree structure. The distribution of family sizes generated under the random model provides a basis for assessing the significance of the observed family size differences among taxa (p -value 0.001). CAFE estimates for each branch in the tree whether a protein family has not changed, has expanded or contracted. A phylogenetic tree was constructed using 203 highly conserved, core gene representatives of *S. lacrymans* and other basidiomycetes (see the FunyBASE at <http://genome.jouy.inra.fr/funybase/>).

1.5 Molecular clock analysis

A subset of the 191 taxa multigene dataset from Binder et al. (2010) (17) was supplemented with sequences of *Postia placenta*, *Phanerochaete chrysosporium*, *Heterobasidion irregulare* (previously *H. annosum*), and *Fomitiporia mediterranea*. Alignments for three nuclear ribosomal genes (nuc-ssu = 1783 bp, nuc-lsu = 1399 bp, 5.8S = 159 bp) and three protein coding genes (*rpb1* = 427 bp, *rpb2* = 2087 bp, *tef1* = 1016 bp) were made separately using MacClade 4.08 (49). The concatenated alignment including 32 species and a total of 6871 sites was built and edited using TextWrangler (Bare Bones Software, Inc.).

Twelve tMRCA's (time to most recent common ancestors) were specified using BEAUTi v.1.4.8 (50) to define the monophyletic groups resolved in previous analyses (17, 51, 52). Fossil calibrations were applied to the marasmiod clade in the Agaricales and to the Suillineae in the Boletales, reflecting the minimum age estimates for both groups. *Archaeomarasmium legetti* from mid-Cretaceous amber was found in a well-characterized layer of clay in New Jersey that dates to 94-90 Mya (53) and we calibrated the marasmiod clade (fig 1A) including *Armillaria mellea*, *Chondrostereum purpureum*, *Crinipellis perniciosus*, and *Hemimycena gracilis* with this estimate. The permineralized suilloid ectomycorrhiza fossil associated with pine roots in the middle Eocene Princeton chert of British Columbia (54) was used to calibrate the Suillineae (represented by *Gomphidius roseus* and *Suillus pictus*) with 50 myr as a minimum age estimate. We estimated the absolute timing for the split between the saprotrophic *Serpula* lineage and ectomycorrhizal *Austropaxillus* lineage under the GTR model using an uncorrelated relaxed clock with normal rate distribution, invoking the Yule birth-death process of speciation as the tree prior. Three Bayesian relaxed clock analyses were run for 10 million generations using BEAST v1.4.8 (50), saving trees every 1000th generation. The resulting log file was inspected with Tracer v1.4 (55) to confirm that the estimated sample sizes for statistics represent appropriate values of posterior distributions. The runs converged to stable likelihood values after one million generations and 9000 ultrametric trees from each run were analyzed in TreeAnnotator v1.4.8 (50) to estimate the 95% credible node intervals referred to as highest posterior densities (HPD), which mark the lower and upper boundaries of the time estimates.

1.6 Comparative study of the evolution of lignocellulolytic carbohydrate active and oxidoreductase enzyme gene repertoires

Genomes and focal gene families. Carbohydrate active enzyme (CAZy) and oxidoreductase families involved in lignocellulose decomposition in *S. lacrymans* were compared with other genome sequenced fungi exhibiting a range of nutritional modes (tables S7-8). A subset of these fungi and enzymes were selected for further analysis to investigate lignocellulolytic gene repertoires (table S9). The analyses included data from 7 Basidiomycota and 3 Ascomycota genomes, basidiomycetes *Serpula lacrymans* S7.9 v1.0 *Phanerochaete chrysosporium* v2.0, *Postia placenta* MAD-698, *Heterobasidion irregulare* v1.0 (previously identified as *H. annosum*), *Schizophyllum commune* v1.0, *Laccaria bicolor*, *Coprinopsis cinerea* okayama 7#130, and the ascomycetes *Trichoderma reesei* v2.0 (anamorph of *Hypochrea jecorina*), *Magnaporthe grisea* 70-15 and *Stagonospora nodorum*. The data for the first six basidiomycetes and *T. reesei* were retrieved using the databases of Joint Genome Institute (www.jgi.doe.gov), while the data for *C. cinerea*, *M. grisea* and *S.*

nodorum were retrieved using the databases of the Broad Institute (www.broadinstitute.org).

Nineteen gene families with various roles in wood degradation were targeted (table S9), separated into two broad functional classes of gene families, including genes encoding Carbohydrate-Active enzymes (CAZY's, e.g., cellobiohydrolases, endoglucanases, esterases) and various oxidoreductases implicated in lignocellulolysis (e.g. Class II peroxidases, multicopper oxidases). In families containing enzymes with diverse roles, e.g. glycoside hydrolases (GH) 5, 28 and 61 and carbohydrate esterases (CE) 1 and 16, only enzymes with direct lignocellulolytic activity were considered in the analysis, e.g. endoglucanases and mannanases in GH5 and cinnamoyl esterases in CE16. In addition highly truncated sequences were also omitted from the analysis. This was done to ensure duplications and losses of enzymes related to wood decomposition would be identified and not masked by fluctuations in functional groups not involved in decomposition. The query sequences used in blast searches were obtained from CAZY base (www.cazy.org), UniProt (www.uniprot.org), or relevant studies.

Blast searches, alignment, and phylogenetic analysis. Blastp searches were conducted with the e-value set to -5, using the 'best models' database for each genome. The recovered protein models were checked for intron-exon structure, functional domains and putative function as displayed on the genome browser or by using Pfam (<http://pfam.sanger.ac.uk/>) and InterProScan (<http://www.ebi.ac.uk/Tools/InterProScan/>). Alignments were performed using MAFFT version 6 (<http://mafft.cbrc.jp/alignment/software>, (56)) and adjusted manually in MacClade 4.08 OS X (49), with alternative models being used in some cases, including reprediction of truncated models using FGENESH (<http://linux1.softberry.com/berry.phtml>). The latter models carry the suffix 'mod'. Poorly aligned areas were excluded. The appropriate protein evolution model was selected for each dataset using ProtTest1.4.mac (57) and an unrooted maximum likelihood analysis was performed in RAxML Blackbox with 1000 rapid bootstrap replicates with the final ML search under the GAMMA+P-Invar model (<http://phylobench.vital-it.ch/raxml-bb/>, (58)).

Gene tree/species tree reconciliation analysis. Gene family phylogenies were reconciled with an organismal phylogeny reflecting current understanding of relationships among the ten genomes (59) using Notung (60, 61). The default settings for costs of duplications and losses were used (1.5 and 1.0 respectively), and two analyses were performed for each gene family, using edge weight thresholds (EWT) set to 90 or 75 (bootstrap frequencies). Gene trees were rooted to minimize the cost of duplications and losses, and topological rearrangements were performed according to the EWT setting. The implied duplications and losses for each gene family under both EWT settings were mapped onto the species tree, and the total duplications and losses across all CAZY gene families and all oxidoreductase gene families were compiled and mapped separately on the species tree.

*1.7 Comparative transcriptomic analysis of *Serpula lacrymans* growth on defined glucose media and on *Pinus sylvestris* sapwood*

Sample preparation. *S. lacrymans* S7 was grown on shavings of *Pinus sylvestris* (L.) sapwood placed on moist soil separated by a nylon net for a period of ten days at 20°C and 80% relative humidity. The nylon net covered with *S. lacrymans* mycelium was removed and RNA was extracted following the CTAB protocol outlined below. Control cultures were prepared using MMN agar medium (5 gL⁻¹

glucose, 0.5 gL⁻¹ NH₄NO₃, 0.5 gL⁻¹ KH₂PO₄, 0.5 gL⁻¹ MgSO₄·7H₂O, 0.167 gL⁻¹ thamine).

RNA extraction. The RNA was extracted from the culture isolates using a 2% cetyltrimethyl ammonium bromide (CTAB) protocol. Briefly, mycelia were ground under liquid nitrogen and extraction buffer added (CTAB, polyvinyl pyrrolidone and β-mercaptoethanol) with acid phenol. Chloroform:isoamyl alcohol extraction and lithium chloride precipitation was carried out according to established protocols (30). All samples were treated with RQ1 DNASE (Promega, Stockholm, Sweden) according to manufacturer's instructions. The quantity of the RNA was analyzed with NanoDrop and Qubit™ fluorometer (Invitrogen). The quality of the extracted RNA was determined using 2100 Bioanalyzer (Agilent Technologies, Edinburgh, UK). To determine the integrity of the extracted RNA, the samples were divided into two, where one sample was incubated in 37°C for 2 hours and reanalyzed on the Agilent 2100 Bioanalyzer to establish whether the RNA had been degraded.

cDNA synthesis and amplification. Total RNA (50-100 ng) was used to synthesize first strand cDNA. First and second strand cDNA synthesis and amplification was performed using the SMARTer™ pico PCR cDNA synthesis kit and Advantage® 2 PCR kit (Clontech, Saint-Germain, France), according to the protocol from the manufacturer. The quantity of the cDNA was analyzed with NanoDrop and Qubit™ fluorometer (Invitrogen). The quality was determined using the 2100 Bioanalyzer.

Array design and analysis. The *Serpula lacrymans* S7.9 custom-exon expression array (4 x 72K) manufactured by Roche NimbleGen Systems Limited (Madison, WI) (<http://www.nimblegen.com/products/exp/index.html>) contained five independent, nonidentical, 60-mer probes per gene model coding sequence. For 12,797 of the 12,917 annotated protein-coding gene models probes could be designed. For 8 gene models no probes could be generated and 112 gene models shared all five probes with other gene models. Included in the array were 2130 random 60-mer control probes and labelling controls. For 1200 randomly chosen gene models, technical duplicates were included on the array.

Single dye labeling of samples, hybridization procedures, data acquisition, background correction and normalization were performed at the NimbleGen facilities (NimbleGen Systems, Reykjavik, Iceland) following their standard protocol. Microarray probe intensities were quantile normalized across all chips. Average expression levels were calculated for each gene from the independent probes on the array and were used for further analysis. Raw array data were filtered for non-specific probes (a probe was considered as non-specific if it shared more than 90% homology with a gene model other than the gene model it was made for) and renormalized using the ARRAYSTAR software (DNASTAR, Inc. Madison, WI, USA). For 994 gene models no reliable probe was left. A transcript was deemed expressed when its signal intensity was three-fold higher than the mean signal-to-noise threshold (cut-off value) of the random oligonucleotide probes present on the array (50 to 100 arbitrary units). Gene models with an expression value higher than three-fold the cut-off level were considered as transcribed. The maximum signal intensity values were ~65000 arbitrary units. A Student *t*-test with false discovery rate (FDR) (62) multiple testing correction was applied to the data using the ARRAYSTAR software (DNASTAR). Transcripts with a significant *p*-value (<0.05) were considered as differentially expressed. The complete expression dataset is available as series (accession number GSE27839) at the Gene Expression Omnibus at NCBI (<http://www.ncbi.nlm.nih.gov/geo/>).

1.8 Proteomic analysis of *Serpula lacrymans* grown on wood

Fungal cultures. *Serpula lacrymans* S7 (dikaryon) culture was maintained on malt agar medium (2% malt extract and 1.5% agar) at 20°C in the dark. Millet culture was prepared from 100 g millet (soaked in water overnight and autoclaved) inoculated with three 10 mm diameter pieces of 7 days old agar culture and incubated for 3 weeks until substrate was completely overgrown by the fungus. *Picea abies* wood was chopped to particles of approximately 1-5x0.5x1 mm and dried for three days at 80°C (moisture content of 4.4% and water content of 4.2%). Wood cultures of *S. lacrymans* with calcium silicate were prepared by autoclaving 100 g wood mixed with 2 g meta calcium silicate (CaSiO₃, reagent grade, Alfa Aesar, Karlsruhe, Germany) and 240 ml water in 1 L preserving jar and inoculated with about 2 g of the millet culture. Samples without calcium were prepared in the same way omitting calcium silicate. The initial moisture content and water content of the wood cultures was 280% and 74%, respectively. Cultures in several replications were incubated for 30 days at 20°C in dark.

Protein extraction. Extracellular proteins were extracted from each culture using 500 ml of 500 mM Tris buffer pH 7.0 containing 1% (v/v) Tween 80 and 1 mM PMSF (phenylmethanesulfonylfluoride). Samples were evacuated to remove air from the material, sonicated twice for 5 min and incubated for about 40 min. Separated liquids from parallel samples were combined and stored at -20°C.

After thawing, samples were centrifuged at 25.000 g for 60 min (4°C) and collected supernatants kept in an ice-bath. Sodium chloride and sodium deoxycholate were added to the samples to final concentrations of 1 M and 0.05%, respectively (63). Proteins were precipitated by addition of TCA from 100% stock solution to 10% final concentration; 100% trichloroacetic acid (TCA) stock solution contained 100 g TCA in 45.4 ml water. After mixing, samples were allowed to stand on ice for 30 min. Thereafter, phosphotungstic acid was added from 10% stock solution to the final concentration of 0.5% (64). Samples were mixed well and incubated on ice overnight. Precipitated proteins were collected by centrifugation at 25.000 g for 30 min and further processed as previously described (65). Total protein in this step was determined by the Bradford Reagent (Pierce, Germany). Protein aliquots for 2D-electrophoresis were freeze-dried and samples for shot gun analysis suspended in 100 mM ammonium bicarbonate.

Shotgun protein identification. Protein digestion and peptide fractionation was performed as previously described (66, 67). Briefly, aliquots of approximately 500 µg total protein from each experiment (three replicates) suspended in 100 mM ammonium bicarbonate were digested with sequencing grade trypsin (Promega, Germany) in an enzyme:substrate ratio of 1:40 (w/w) at 37°C for 16 h. Thereafter, samples were reduced with dithiothreitol (DTT) and alkylated with iodoacetic acid. After addition of a new amount of trypsin samples were digested again for 30 min at 58°C (68). The digested peptides were de-salted with a C18 Sep-Pak (Waters, Milford MA) and dried in vacuum centrifuge. Samples were dissolved in 8 M urea and total peptide amount determined by BCA Protein Reagent (Pierce, Germany) calibrated with a tryptic bovine serum albumen (BSA) digest. Samples containing 300 µg total peptides were brought up with 8 M urea and IPG buffer 3-10 (Amersham) to 350 µl and applied to a 18 cm IPG dry strip (Amersham, Munich, Germany). Using an IPGphor II (Amersham), the following focusing protocol was applied: active rehydration for 6 h at 20 V, current limit of 50 µA per strip at 20 °C, 1 h 300 V, 1 h gradient to 1000 V, 3 h gradient to 4000 V, 3 h gradient to 8000 V, 8000 V upto 50

kVh. Each gel strip was cut into 26 sections and peptides extracted with three sequential solutions containing 0%, 50%, and 100% acetonitrile in water with 0.1% trifluoroacetic acid. Supernatants were combined, dried and redissolved in 5% formic acid. Samples were cleaned of salts and residual oil using STAGE tips (69), dried and stored at -20 °C before analysis.

LC-MS/MS analysis. Peptide analysis was performed using 1100 LC (Agilent, Böblingen, Germany) interfaced to an Esquire3000 ion trap mass spectrometer (Bruker-Daltonics, Germany) via an electrospray ionization (ESI) unit. Each peptide fraction was dissolved in 5 µl of 5% formic acid and 4 µl samples were loaded on a 180 µm i.d. capillary column packed with 3 µm Reprosil-Pur C18-AQ (Dr. Maisch GmbH, Ammerbuch, Germany), conditioned in 98% of solvent A (0.1% formic acid in water) and 2% of solvent B (0.1% formic acid in 90% acetonitrile). After 20 min isocratic elution at 2 µl/min peptides were eluted by gradients of solvent B: 15% in 5 min, 40% in 90 min, 50% in 5 min, and 90% in 5 min. Mass spectrometer was setup to take four averages of MS-spectra (200 to 1500 m/z) and four averages of MS/MS-spectra (200 to 3000 m/z) of two most abundant precursor ions. The Dynamic Exclusion was set to non-single charged precursor ions and an exclusion time of 5 min. The MS/MS spectra were extracted by DataAnalysis (V. 3.0, Bruker Daltonics) and peptides identified using Mascot (V. 2.2, Matrix Science, UK). Target database was constructed from annotated genomes of *S. lacrymans* S7.9 and S7.3 and the SwissProt database. Searches against a decoy database, created by randomizing the target database, indicated the false discovery rate (FDR) of 0.25%. All searches were run as a tryptic digest with one missing cleavage allowed, fixed carbamidomethylation of cysteine and variable oxidation of methionine. Mass tolerances were set to 1.4 Da and 0.4 Da for the MS and MS/MS spectra, respectively. Mascot results were extracted from raw DAT-files using a VB-script and transferred to an SQL-database (Microsoft SQL Server 2005). SQL queries were used to extract proteins with at least two peptides with scores higher than the corresponding identity score. Analogous data processing was performed on false/positive searches of the decoy database and additional proteins were identified using an average peptide scoring (APS) method (70, 71) with restriction to proteins matched by at least two confidently identified peptides.

1.9 Genome analysis of natural product genes and secondary metabolite analysis under nutrient asymmetry

Genome analysis. The *S. lacrymans* genome was analysed for natural product genes putatively involved in secondary metabolite production.

Culture preparation. *S. lacrymans* S7 was grown on solid standard 'Serpula Czapek Dox' medium (per liter: 30g sucrose, 1g monosodium glutamate, 1g KH₂PO₄, 0.5g MgSO₄ · 7H₂O, 0.01g FeSO₄ · x7H₂O, 20g agar). After 3-4 weeks, mycelia including the solid medium of six plates (10 cm in diameter) were shredded and exhaustively extracted with ethyl acetate. After solvent evaporation under reduced pressure the crude extract was solved in methanol.

Pigment identification. The crude extract was analyzed on an Agilent 1200 HPLC instrument equipped with a Zorbax Eclipse XDB C-18 column (150 x 4.6 mm, 3.5 µm particle size) and a guard column (Agilent Technologies). The following gradient was applied (solvent A: water, solvent B: acetonitrile): initial hold for 2 min at 5% B, then linear increase to 95% B within 20 min, at a flow rate of 0.5 ml/min. Chromatograms were recorded at λ = 254 nm. Variegatic acid and its precursors, atromentic acid and xerocomic acid, were identified by their masses during

electrospray mass spectrometry (positive and negative mode) and by their UV spectra, which were compatible to those published (72).

Sample preparation. The crude extract was adsorbed into a 3 ml Bakerbond silica gel solid phase extraction cartridge, followed by sequential elution with cyclohexane, ethyl acetate, and methanol. Variegatic acid and derivatives desorbed into two ethyl acetate fractions, one of which (4 mg) was composed of 60% variegatic acid and 15% xerocomic acid and was tested in the ferrozine assay, along with crude extract. HPLC analysis of the crude and two ethyl acetate fractions were recorded, data showed that the sample of fraction 3 was the most purified, while the raw extract the least (fig S1).

Ferrozine iron reduction assay. Three *Serpula* extract samples were analyzed for their iron reducing capabilities using a ferrozine assay. Ferrozine [3-(2-pyridyl)-5,6-bis(4-phenylsulfonic acid)-1,2,4-triazine] reacts with divalent iron and forms a stable magenta colored complex that can easily be determined and measured photometrical at a fixed wave length of 562nm (73). Three *S. lacrymans* extracts prepared as outlined above (raw extract and 2 ethyl acetate purified fractions) before being assessed for iron reducing capability. The *Serpula* extract samples were diluted in 5.0 ml 50% (v/v) ethanol solution and care was taken to ensure they were well mixed into solution and stored tightly covered at 4 °C. Samples (0.5 µM) were assayed in a reaction consisting of 50 µM FeCl₃ and 10 mM ferrozine in 1 M acetic acid and 1M sodium acetate buffer (pH 4.5), in 10 ml total volume. Control samples containing 50 µM 2,3-dihydroxybenzoic acid (DHBA) as iron reducing agent were included as a reference. After 5, 20, 60, and 120 minutes incubation, the ferrozine was added and measurements were recorded after a further 2 minutes incubation at 562 nm using a spectrophotometer. Exposure to air was minimized and sample vials were purged with nitrogen to prevent oxidation.

1.10 *S. lacrymans* interaction with *Picea sylvestris* seedlings

Experimental microcosm systems were constructed using the established mycorrhizal synthesis system (29). Two-week-old seedlings were aseptically inoculated with agar plugs from of *S. lacrymans* S7 in Petri dishes with growth substrate of sterilized fine sphagnum peat : vermiculite : 1/10 strength liquid MMN mixture in the ratio 1 : 4 : 2. Five replicate microcosms were constructed for each tree species/fungus combination (two tree species × three fungi × five replicates, or 30 microcosms in total). Inoculated microcosms were incubated in a growth chamber at 20 ° C, with a 16 h light : 8 h dark photoperiod.

2. ADDITIONAL RESULTS AND DISCUSSION

2.1 Assembly, annotation and analysis

The 42.4 MB genome of *Serpula lacrymans* S7.9 was sequenced to 8x read depth coverage using Sanger platform and assembled into 46 scaffolds using Arachne (37). The second strain S7.3 was sequenced with 454 (Roche) pyrosequencing and assembled using assembled S7.9 as a template for scaffolding (table S3). 12,917 and 14,495 genes were predicted in S7.9 and S7.3, respectively, using a combination of gene predictors and validated with ESTs (tables S4-S5). On average orthologs between two strains show 98.5% amino acid identity and arranged into large scaffold-long syntenic blocks while only a half of this number (5,000-5600) are orthologous to other Agaricomycetes with 57-61% identity (table S10) and still significant syntenic blocks. Interestingly, S7.3 shows more genes than S7.9, that may affect differences in functional profiles (tables S6, 10). In *S. lacrymans*, 581 gene families showed a

significant expansion, 3,571 showed no change and 906 families had undergone contraction by comparison to a putative common ancestor Basidiomycota having 5,058 gene families (fig S2).

2.2 Molecular clock analysis

All Bayesian searches converged independently on a tree topology that is consistent with previously published phylogenetic inferences (17, 52). The relaxed molecular clock analyses with normal fossil calibration priors produced broadly spaced estimates for the HDP of the non-calibrated nodes (table S11, fig 1A). This result is not unexpected since the approach taken was aimed at estimating the minimum node ages conservatively. Our main objective was to provide a minimum age estimate for the split between *Serpula lacrymans* and *Austropaxillus* spp. within a comprehensive phylogenetic framework. The posterior time estimate for the Agaricomycetidae (Agaricales, Amylocorticiales, Atheliales, Boletales, and Jaapiales) dates the most recent common ancestor at 166.1 Mya (95% HPD 189.8 – 126.5) and this estimate largely overlaps with the fossil record of Pinaceae as potential hosts dating back to the lower Jurassic (74–76). The minimum timing for the origin of the Boletales is estimated at 113.4 mya (HPD 140.5 – 87.3), which is consistent with the findings of Hibbett and Matheny (2009) (77) suggesting that the Boletales are young enough to have been associated with Pinaceae or rosids (Eurosids I).

Collectively, previous studies (12, 19, 78) suggest that *S. lacrymans* has an Asian origin and it became cosmopolitan by gradually expanding its natural range over the northern hemisphere in association with pinaceous hosts. Contemporary *Austropaxillus* spp. on the other hand are bound to *Nothofagus* and to some extent to *Eucalyptus* and have a southern hemisphere distribution extending over Australia, Tasmania, New Zealand, Papua New Guinea, and parts of South America. In addition, the ancestral nutritional mode of the most exclusive clade containing *Serpula* and *Austropaxillus* was estimated as brown-rot saprotrophic with a single transitional event to ECM leading to *Austropaxillus* (12). The split between *Serpula* and *Austropaxillus* estimated at 34.9 mya (HPD 53.1 – 15.0) coincides with the separation of South America and Australia from Antarctica about 31 mya ago (79, 80). This finding is consistent with a vicariant distribution of *Austropaxillus* and its main host, *Nothofagus*, although dispersal may have played a major role in the current biogeography of the tree hosts (81). Thus, it appears plausible that the transition from a saprotrophic ecology to ECM in *Austropaxillus* occurred after a host switch from Pinaceae to Nothofagaceae before the complete break-up of Pangaea. Alternatively, extinct or unsampled *Serpula*-like fungi have acquired the ability to function as ECM partners and morphological elaboration from resupinate to stipitate-pileate forms became manifested with the occurrence of *Austropaxillus*. A comparative genomic study between *Serpula* and *Austropaxillus* would ultimately help to clarify whether *Serpula* is genetically predisposed to enter ectomycorrhizal associations.

2.3 Comparative study of the evolution of lignocellulolytic carbohydrate active and oxidoreducases enzyme gene repertoires

CAFE analysis identified gene families that had undergone significant expansion and contraction (fig S2). A more detailed analysis of genes involved in lignocellulose decomposition revealed a link between brown rot evolution and gene repertoires. 783 protein models were retrieved and used in the phylogenetic and reconciliation analyses. 767 protein models were included in the final mapping of

gene losses and duplications, including 521 from basidiomycete genomes and 246 from ascomycete genomes.

Lignocellulolytic Carbohydrate Active enzymes (CAZY's). The lowest number of CAZY gene families selected for study in basidiomycetes was found in *L. bicolor* (14 genes), while in ascomycetes. *T. reesei* (32 genes) had the fewest CAZY genes. The highest number of selected CAZY gene families in basidiomycetes and ascomycetes were identified in *Coprinopsis cinerea* (75 genes) and *St. nodorum* (87 genes) respectively (fig 1B).

The distribution of gene copies was not homogeneous among the genomes and among the gene families. *C. cinerea*, *Sc. commune* and *Ph. chrysosporium* had representative copies from all the CAZY families included in the study, while *L. bicolor* lacked copies in half of the families, *S. lacrymans* and *P. placenta* lacked copies for some of the families and *H. irregulare* (previously classified as *H. annosum*) only for the Glycoside Hydrolases family 11 (tables S7 and 9). The Glycoside Hydrolases family 74 included the lowest number of gene copies (8 copies) while the Glycoside Hydrolases families 61 and 3 had the most prominent representation in the dataset with 139 and 104 copies respectively (table S9). The lineages leading to *C. cinerea*, *Sc. commune*, *Ph. chrysosporium* and *St. nodorum* have undergone net expansions in CAZY gene families, while the lineages leading to *L. bicolor*, *S. lacrymans*, *P. placenta*, *H. irregulare* and *T. reesei* have undergone net contractions (fig 1B).

The common ancestor of the basidiomycetes was estimated to have 66 to 83 total gene copies for the selected CAZY gene families (fig 1B), depending on the edge weight threshold (EWT). The genomes of *C. cinerea*, *Sc. commune* and *Ph. chrysosporium* had similar gene numbers to the common basidiomycete ancestor, although the similarity did not coincide in gene diversity, as several gene families had undergone expansions and contractions. The rest of the basidiomycete genomes had a decreased number of genes compared to the common ancestor (fig 1B).

A more detailed examination of the results for *S. lacrymans*, *P. placenta* and *L. bicolor*, which have undergone parallel reductions in CAZY gene families, suggested that: 1) a number of gene families are dispensable and lost in all or two out of the three genomes (table S9), 2) the three species bear a low number of lineage specific duplications (figs 1B & S3A-M) much lower than the lineage specific duplications for the other species, 3) despite the shared similarities among the three genomes, *L. bicolor* bears a lower number of copies than the two brown rot species for most of the gene families that all three species include gene copies (table S9).

H. irregulare also had an increased number of lineage specific gene losses (fig 1B) but the gene distribution pattern observed was different from the three species referred above. Thus in spite of the increased losses, *H. irregulare* had at least one gene copy in all but one gene family (table S9).

Oxidoreductases. The highest number of copies for oxidoreductase gene families was present in *Ph. chrysosporium* (43 copies) followed by *H. irregulare* (36 copies), while the lowest number of copies were identified in *S. lacrymans* and *P. placenta* (19 copies each) and *T. reesei* (fig 1C & table S8). The differences among the genomes in total gene copies were less prominent in the dataset compared to the CAZY dataset, with the exception of *Ph. chrysosporium*, *H. irregulare* and *T. reesei*.

As with the CAZs, the distribution of oxidoreductase gene copies was not equally dispersed among the gene families and genomes. The genomes of *Ph. chrysosporium* and *H. irregulare* include at least one copy for each one of the gene families studied, while the rest of the genomes lack copies for one or more families

(table S8). The cellobiose dehydrogenases had the lowest number of gene copies (12 copies), while the highest number of gene copies was in the multicopper oxidases (92 copies).

The summarized reconciliation results suggest that the lineages leading to *C. cinerea*, *L. bicolor*, *Sc. commune*, *H. irregulare* and *Ph. chrysosporium* have undergone net expansion in oxidoreductases while the lineage leading to *S. lacrymans* has undergone net gene contraction, and the lineage leading to *P. placenta* is almost constant regarding gene copies numbers (fig 1C).

Gene families encoding Class II peroxidases, which are implicated in ligninolysis, had expanded independently in the lineages of *Ph. chrysosporium* and *H. irregulare*, while the rest of basidiomycetes either have had no expansions or bore a low number of losses (fig S3N). Genes for this family were absent from the genomes of *S. lacrymans* and *Sc. commune*. Furthermore, the multicopper oxidases have undergone extensive gene duplication in the genomes of *C. cinerea*, *H. irregulare* and *L. bicolor*, and thus gene expansions are associated not only with white rot species (fig S3O). Regarding the four gene families involved in oxidative degradation of cellulose in relation to the Fenton reaction (cellobiose dehydrogenase, oxalate oxidases/decarboxylases, iron reducing glycopeptides and quinone reductases), all the wood degrading species, with the exception of *H. irregulare*, had an increased number of gene copies as compared to *C. cinerea* and *L. bicolor*. *Serpula lacrymans* and *P. placenta* had lineage specific duplications for some of these families (fig S3P-S), and the former was the only species among the basidiomycetes that has duplications in cellobiose dehydrogenases (fig S3T).

Discussion. Reconciliation analyses suggest that the 12 CAZY gene families in *S. lacrymans*, *P. placenta*, *L. bicolor* and *T. reesei* have undergone net contraction. A reduction in CAZY gene content was noted previously for the brown rot species *Postia placenta* (Polyporales), including losses of genes encoding exocellobiohydrolases and cellulolytic enzymes with cellulose binding modules (CBM1) (13). A similar pattern of reduction in CAZYS is evident in *S. lacrymans* (Boletales), which represents an independent origin of brown rot, as well as in the ectomycorrhizal *L. bicolor* (Agaricales) (15). Surprisingly, *T. reesei*, which is a very efficient cellulolytic fungal species, also has reduced gene content for several CAZY families (table S9). The reconciliation results support those findings especially in this case where *T. reesei* was compared with the aggressive plant pathogens *M. grisea* and *St. nodorum*. In contrast, the white rot basidiomycete *Sc. commune* contains a great variety of CAZY enzymes (82), which reconciliation analysis shows to have arisen through multiple recent duplications in the lineage leading to *Sc. commune*.

Brown rot has evolved independently in multiple lineages of basidiomycetes (4). Gene content and reconciliation analyses suggest that these transitions have been characterized by extensive gene losses in the CAZY families, and that the common ancestor of Agaricomycetes (Basidiomycota) was able to utilize cellulose and hemicellulose in a similar fashion as contemporary white rot species. The brown rot mechanisms of *S. lacrymans* and *P. placenta* demonstrate remarkable convergence in their cellulose degradation biochemistry, but they are distinct as *S. lacrymans* has a cellobiohydrolase of the Glycoside Hydrolase family 6 as well as genes that carry a CBM1 domain for the families 6 and 5. Another brown rot species of the Boletales, *Coniophora puteana*, also produces cellobiohydrolases (83) and genes coding for cellobiohydrolases of both the Glycoside Hydrolase families 6 and 7 have been cloned from this species (84). The differences in gene content between *S. lacrymans* and *P. placenta* reflect the independent origin of the two brown rot lineages.

Evolution of a mycorrhizal lifestyle in *L. bicolor* also appears to involve reductions in CAZY gene families. Similar patterns of gene loss have been suggested in the evolution of other mycorrhizal taxa, including the basidiomycetes *Amanita bisporigera* (85) and the ascomycete *Tuber melanosporum* (14).

The convergent evolutionary events in gene losses and duplications among gene families related to wood degradation for *L. bicolor*, *S. lacrymans* and *P. placenta* as they were highlighted above especially for the CAZY gene families, indicate a possible transition from a brown rot lifestyle towards a mycorrhizal one. This is further supported by a study (17) that placed several brown rot genera as basal in the Boletales. The most prominent finding in that study was the placement of *Austropaxillus*, a south hemisphere mycorrhizal genus, as sister group of *S. lacrymans* and *Serpula himantioides*.

2.4 Comparison between *S. lacrymans* transcriptomic and proteomic data

S. lacrymans gene transcripts on wood and glucose medium were analysed using NimbleGen microarrays containing 11,804 probes produced from the *S. lacrymans* S7.9 genome database. A total of 517 gene features were significantly ($P < 0.01$) regulated 4-fold or greater between treatments, 300 on wood substrate and 217 on glucose. Secreted proteins were extracted from wood-based cultures, separated by 2D-gel electrophoresis and peptides identified by LC-MS/MS. Twenty-eight putatively secreted proteins were identified, 22 (78%) of which were in 10% of the most highly expressed transcripts and 17 (68%) were differentially regulated on wood.

A comparison of transcriptomes of white rot *Ph. chrysosporium* and the closely related brown rot *P. placenta* showed that *Ph. chrysosporium* expressed a conventional system of synergistically acting endo- and exo- hydrolase enzymes and relatively few oxidoreductases (26). *P. placenta* produced a larger proportion of hemicellulose degrading and oxidoreductase enzymes, fewer cellulases, and appeared to lack exocellobiohydrolase activity (26). There was evidence that a hydroquinone-driven Fenton's system for non-enzymic hydroxyl radical cleavage of cellulose, typical of brown rot decay fungi takes the place of the enzymic hydrolysis of α -cellulose in *P. placenta* (13). The *S. lacrymans* wood-induced transcriptome showed similarities to both *P. placenta* and *Ph. chrysosporium*, with glycoside hydrolase (GH) and oxidoreductase enzymes accounting for 35.7% of genes upregulated on wood, while lipase and sugar transporter transcripts were also increased (table S12). GH enzymes accounted for 50% proteins identified (table S13) and a large proportion of transcripts were regulated more than 20 fold on wood (33.9%, fig S4), with GH families 5, 61, 3 and 28 represented more than once. Two endoglucanases (Prot id: 433209 and 453342) were expressed more than 100 fold more on wood, and hemicellulose-degrading enzymes (1,3- β -galactosidase, mannanases, endoxylanases, galactouronases, glucuronidases) were prominent. The *S. lacrymans* GH6 cellobiosehydrolase gene (absent in *P. placenta*) was not highly expressed, suggesting that, as in *P. placenta*, the depolymerisation of cellulose by *S. lacrymans* occurs in the absence of known exo-acting cellobiohydrolases.

Also highly represented amongst the *S. lacrymans* transcripts from wood cultures were cytochrome P450 (18 genes), aldo-keto reductases (11 genes) and short chain dehydrogenases (8 genes), while FAD-dependent pyridine nucleotide disulphide oxidoreductase (451883) and glucose-methanol-choline oxidoreductase (439506) were upregulated 109 and 38 fold respectively. Aromatic ring hydroxylases, alcohol dehydrogenases, multicopper oxidase and aldehyde dehydrogenases were also

identified. No evidence of upregulated hydroquinone synthesis was discovered; 1,4-benzoquinone reductase and hydroquinol-1,2-dioxygenases described in *P. placenta* were not differentially regulated or highly expressed in *S. lacrymans*, although a ferric reductase-like protein (478753) possibly involved in iron redox cycling was expressed (fig S5). Unexpectedly, neither *S. lacrymans* cellobiose dehydrogenase CDH gene (453176 and 453175) was highly or differentially expressed.

Transcriptomic data was obtained from *S. lacrymans* S7 (dikaryon) cultures grown on wood shavings on soil for 10 and 30 days and compared with mycelium grown on a minimal medium with glucose as the main carbon source. For ease of presentation 10 day wood cultures were compared with glucose medium. Genes were identified as significantly differentially regulated (up or down) if transcript levels 4 fold or more between treatments. A total of 517 microarray features were identified as being differentially regulated, 300 had greater expression in the wood samples compared with glucose medium and 217 lower levels (fig S4; table S12 for gene list and annotation of 50 most differentially transcribed on wood).

Proteomic data was obtained from *S. lacrymans* S7 cultures growing on wood particles (*Picea abies* about 1-5x0.5x1 mm) for 30 days either in the presence or absence of calcium ions (meta calcium silicate, 2g/100g wood). Low quantities of extracellular proteins were obtained, leading to the identification of 39 proteins, 29 of which having an extracellular signal peptide motif. Twenty four proteins were characterised from the *S. lacrymans* S7.9 monokaryon and 15 from the *S. lacrymans* S7.3 monokaryon (table S13), proteins identified were compared with transcriptomic-derived data (table S14). The microarray was designed using gene models of the monokaryon S7.9, therefore, proteins identified from *S. lacrymans* S7.3 were compared against the S7.9 database by BLAST and the ortholog identified was compared with the transcriptomic data. In some instances (e.g. closely related genes families such as glycoside hydrolase 3) more than one S7.9 gene model gave the same Blast similarity score, in which case the transcript level of each S7.9 ortholog was examined.

The effect of calcium on protein expression was examined for the proteomic study, 11 proteins appeared to be associated with cultures in the presence of calcium ions only; 6 proteins were detected on wood without calcium and 22 were associated with the fungus growing on wood both with and without calcium. As calcium was not used in the transcriptomic study, the comparison between the two was conducted using the 28 proteins identified as being present in wood when calcium was not added.

Analysis revealed that 3 proteins were not represented by a probe on the microarray (protein number 1, 6 and 29), presumably because a suitable probe could not be designed. Therefore comparison is based on 25 proteins rather than 28. Seventeen (68%) proteins identified in the proteomic study were highlighted as differentially regulated in the transcriptomic study. Of the remaining 8 proteins, 6 had transcript levels in the top 10% signal intensity in wood cultures, but their expression on glucose was also high and therefore these genes were not identified as strongly up-regulated in the transcriptomic study. Of the proteins identified in the proteomic study 22 (88%) were in the top 10% of genes with greatest transcript levels from the transcriptomic study and 9 (36%) were in the top 1%. Glycoside hydrolase and carboxylesterase enzymes were well represented in both studies.

2.5 Natural product genes in the *Serpula lacrymans* genome and description of the atromentin locus involved in variegatic acid production

Serpula species are known as producers of small molecule natural products, and three distinct classes of metabolites were reported in the literature: i) members of the pulvinic acid family of pigments (86) which are common among members of the Boletales, ii) the himanimides (87), and iii) polyine acids (88). Genes for five polyketide synthases (PKSs), 15 nonribosomal peptide synthases (NRPSs) or NRPS-like enzymes, and two PKS/NRPS-hybrids, were identified. For example, none of the known secondary *S. lacrymans* metabolites requires PKS activity during biosynthesis. Some products of NRPS genes may participate in the fungus' primary metabolism, as they resemble α -amino adipate reductases (NPS5, NPS9, NPS10, NPS11) or siderophore-synthesizing enzymes (NPS4). Two genes (*nps6* and *nps8*) whose products show an identical domain organization encode putative PKS/NRPS hybrid enzymes. Also, putative genes for 10 terpene cyclases were found, some of which may catalyze secondary product formation.

Putative atromentin locus and pigment formation. Atromentin is the central intermediate *en route* to pulvinic acid-derived pigments, such as xerocomic, variegatic, and atromentic acid. A cluster of genes resembling the atromentin biosynthesis locus in *Tapinella panuoides* (24) was found on scaffold 9, thus making these very likely candidates to govern pigmentation of *S. lacrymans* fruiting bodies and undifferentiated mycelia. We expect the central enzymes are encoded by i) the gene *nps3* (74% identical amino acids to the *T. panuoides* quinone synthetase AtrA, an NRPS-like enzyme) and *amt1* (59% identical amino acids to the *T. panuoides* L-tyrosine:2-oxoglutarate aminotransferase AtrD). As in *T. panuoides*, a third reading frame in opposite transcriptional direction putatively encoding an alcohol dehydrogenase was identified between the aminotransferase and quinone synthetase genes.

Tailoring enzymes. Numerous reading frames which may encode tailoring enzymes were identified. Most of these are located in the vicinity of NRPS and PKS genes thus suggesting the biosynthesis gene cluster paradigm holds up for the Boletales the same way as it does for aspergilli and numerous other filamentous fungi (89).

Prenyl transferase activity is required during himanimide assembly. Consistent with this, genes for two putative metal independent prenyl transferases - *ppt1* on scaffold 12 (clustered with *nps1*) and *ppt2* on scaffold 17 - were detected. Halogenated natural products have not yet been described from *S. lacrymans*. However, two genes (*hal1* on scaffold 18, clustered with a PKS gene, and *hal2* on scaffold 5) which very likely code for flavin-dependent halogenases indicate the capacity to synthesize as yet unknown halogenated secondary products. Although some pathways may be tightly regulated we expect, based on the genomic data (table S15), the *S. lacrymans* secondary metabolome to be much more diverse than evident from current chemical data.

2.6 *S. lacrymans* interaction with *Picea sylvestris* seedlings

In co-culture with *Picea sylvestris* seedlings, *S. lacrymans* S7 was observed to grow towards and around the roots of the plant (fig S6). In addition, the roots were observed to form short lateral branches commonly associated with mycorrhizal associations (fig S6B). However, while microscopic investigation revealed a close association of the fungus and root cells, no evidence of Hartig net or true mantle necessary for mycorrhizal formation were observed (fig S6C). The observations

suggest a transient interaction between fungus and plant in which *P. sylvestris* seedling in attempting to form symbiotic association interact with a fungus foraging for nutrients. The interaction does not proceed to form structures expected for mycorrhizal associations and it is not clear from this experiment to what extent this is due to the plant not recognising a compatible fungal partner or *S. lacrymans* lacking the necessary machinery to develop the association further. The data do suggest a close relationship which can form between tree roots and fungi and point to a scenario in which saprotrophic fungi could evolve into partners in a mycorrhizal mutualistic association.

3. Further funding sources

Funding is acknowledged from the "Wealth Out of Waste" programme funded by a EPSRC grant to the Warwick Innovative Manufacturing Centre, INRA, the Region Lorraine Council, the European Commission within the Project ENERGYPOPLAR (FP7-211917), the Network of Excellence EVOLTREE (FP6-016322), and the US Department of Energy – Oak Ridge National Laboratory Scientific Focus Area for Genomics Foundational Sciences (Project Plant–Microbe Interfaces), FORMAS, Carl Tryggers Foundation, KSLA, Ångpanneföreningen, Ministry of Science and Culture of Lower Saxony (VW-Vorab 11-76251-99-9/04 ZN 2043/ZN 2128) and the Deutsche Bundesstiftung Umwelt, ANR (grant ANR-07-BIOE-006), the Natural Environment Research Council, UK (grant UK NER/A/S/2002/882) and Helsinki University Research Fund and Academy of Finland. The CAZy database is funded in part by GIS-IBiSA.

4. Supplementary Figures

Figure S1. LCMS/MS analysis of *S. lacrymans* S7.9 split plate extract and purified fractions used to test for iron reducing activity in the Ferrozine assay

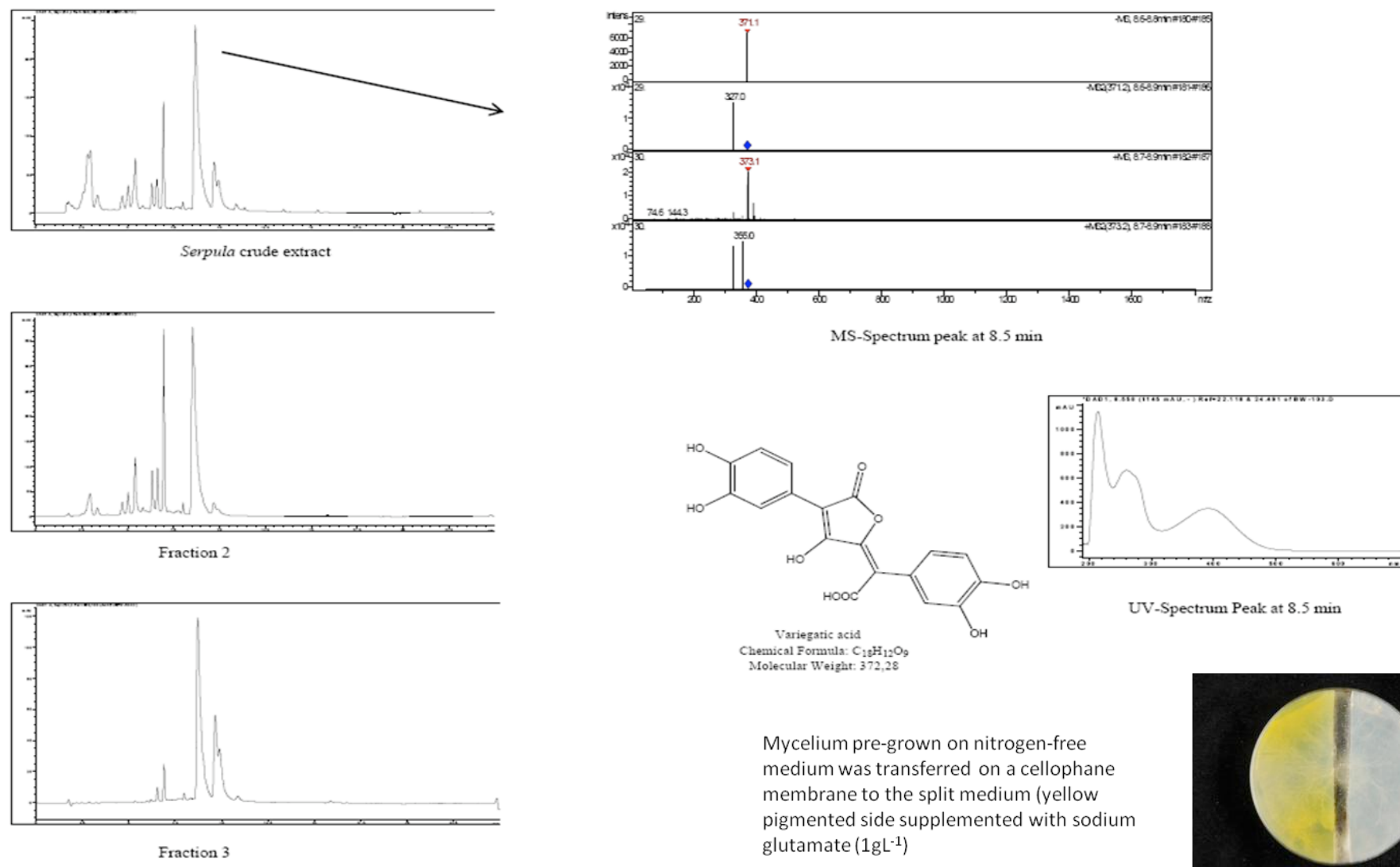
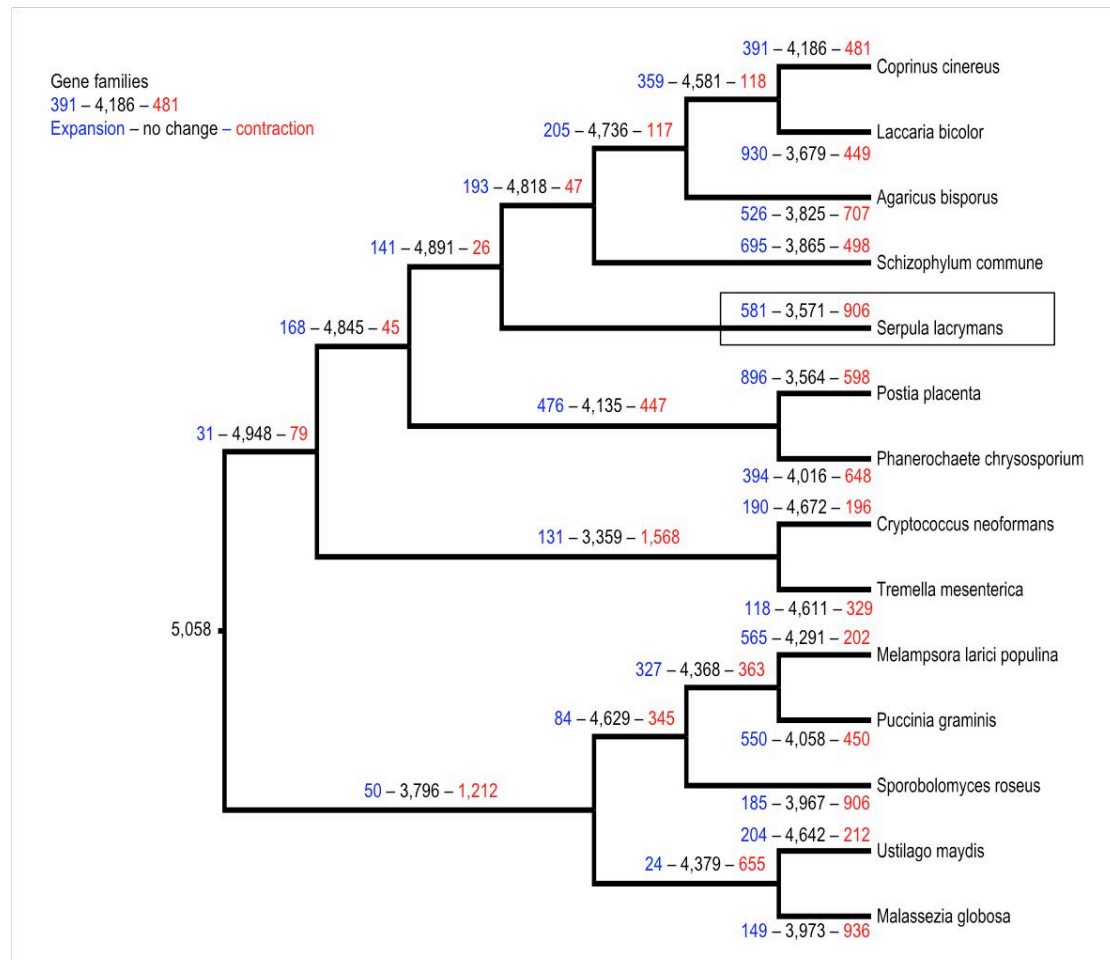


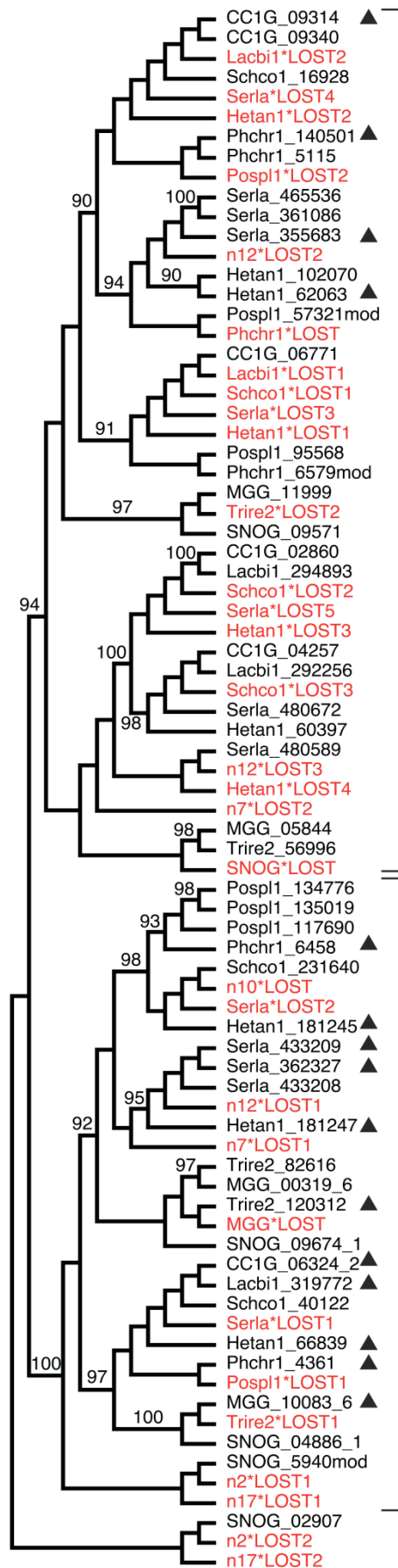
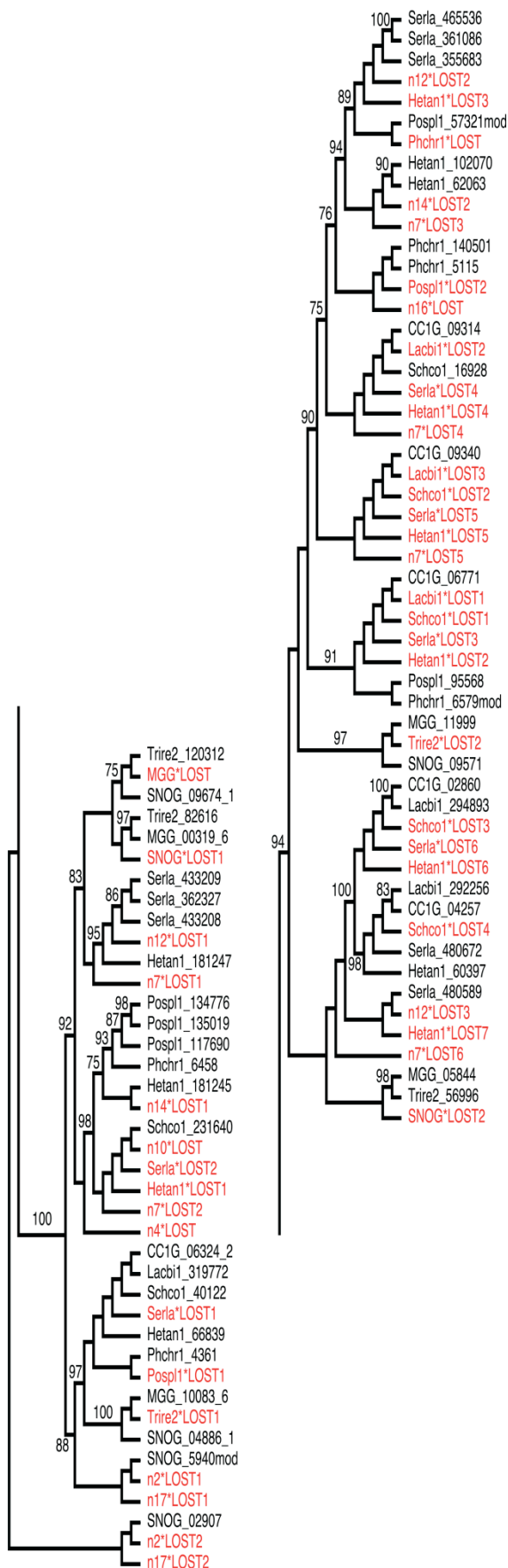
Figure S2. CAFE analysis of the total number of protein families in each species node. The figure represents the total number of protein families in each species or node. We take into consideration families with at least 10 members and 2 species. The numerals on branches show numbers of expanded (left, blue), unchanged (middle, black) or contracted (right, red) protein families along lineages by comparison to the putative pan-proteome.



3B.

Glycoside hydrolase family 5, EWT 75

Glycoside hydrolase family 5, EWT 90

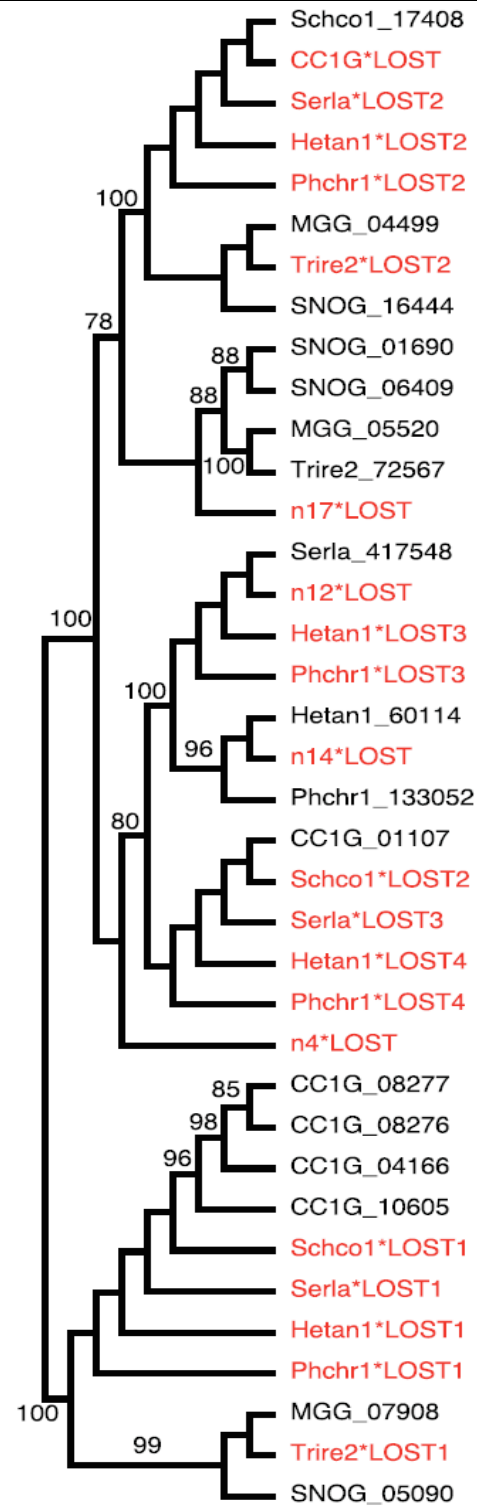


putative
mannanases

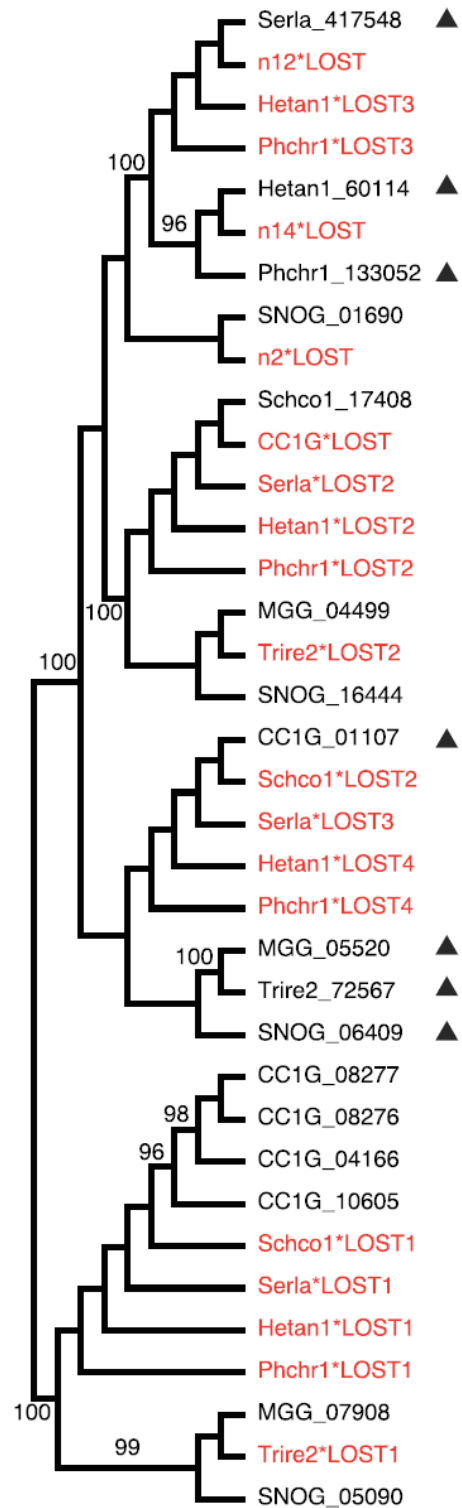
putative
endoglucanases

3C.

Glycoside hydrolase family 6, EWT 75

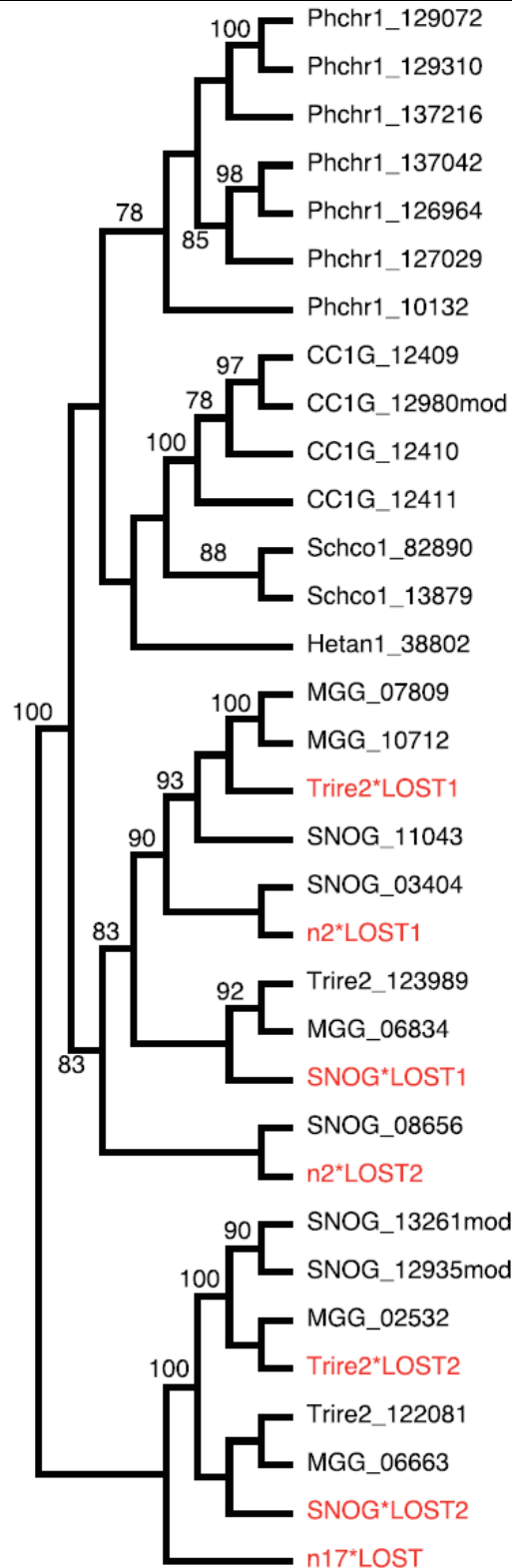


Glycoside hydrolase family 6, EWT 90

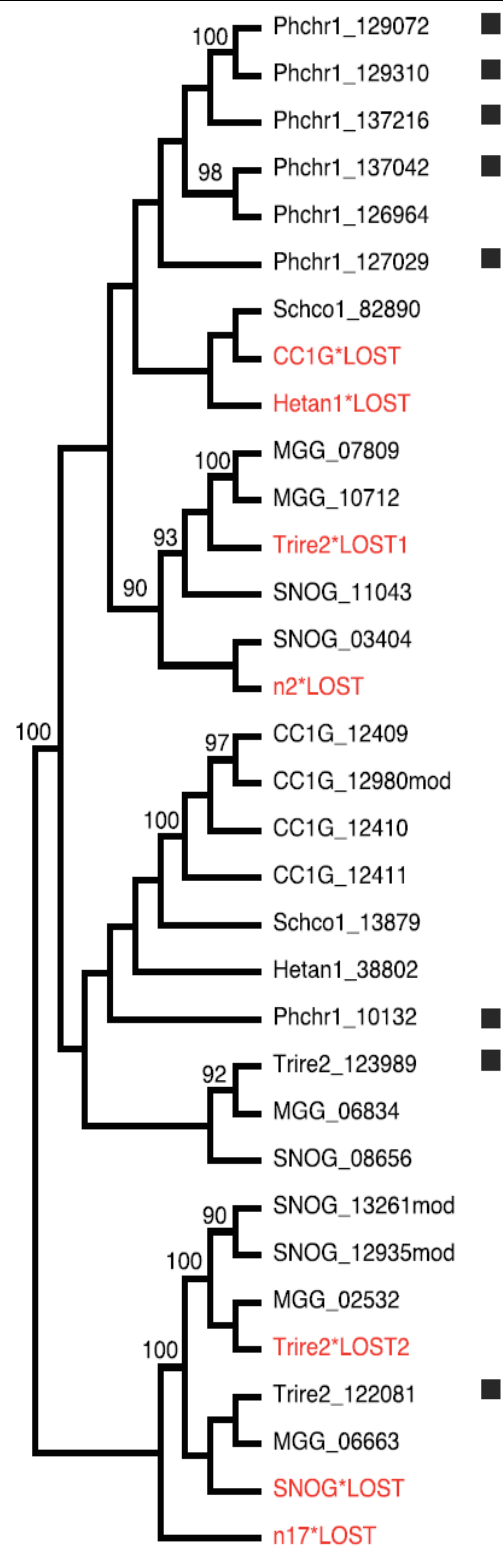


3D.

Glycoside hydrolase family 7, EWT 75

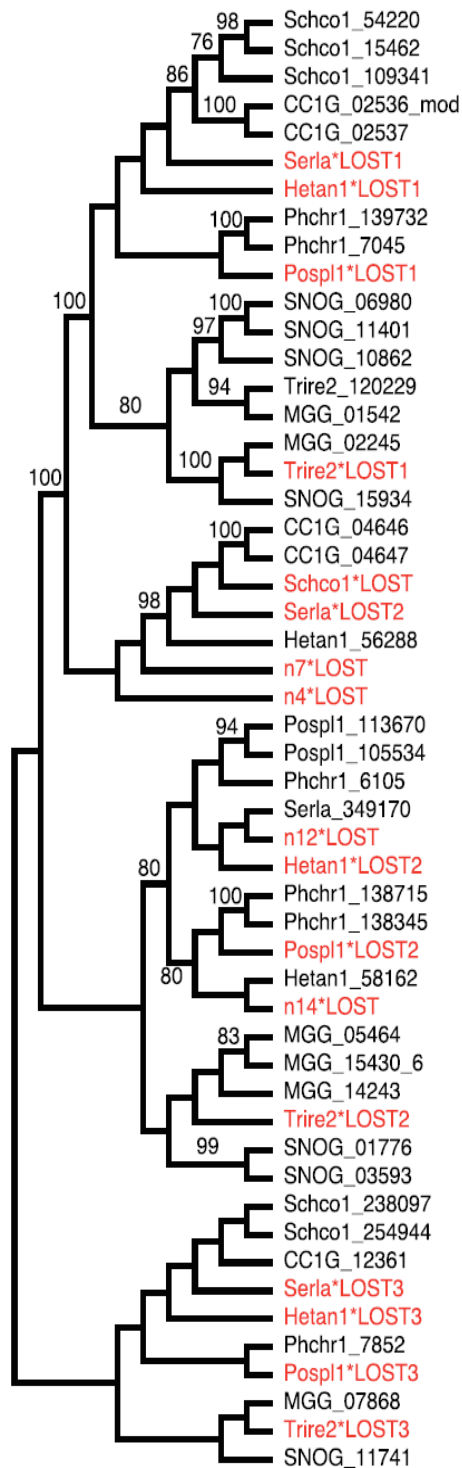


Glycoside hydrolase family 7, EWT 90

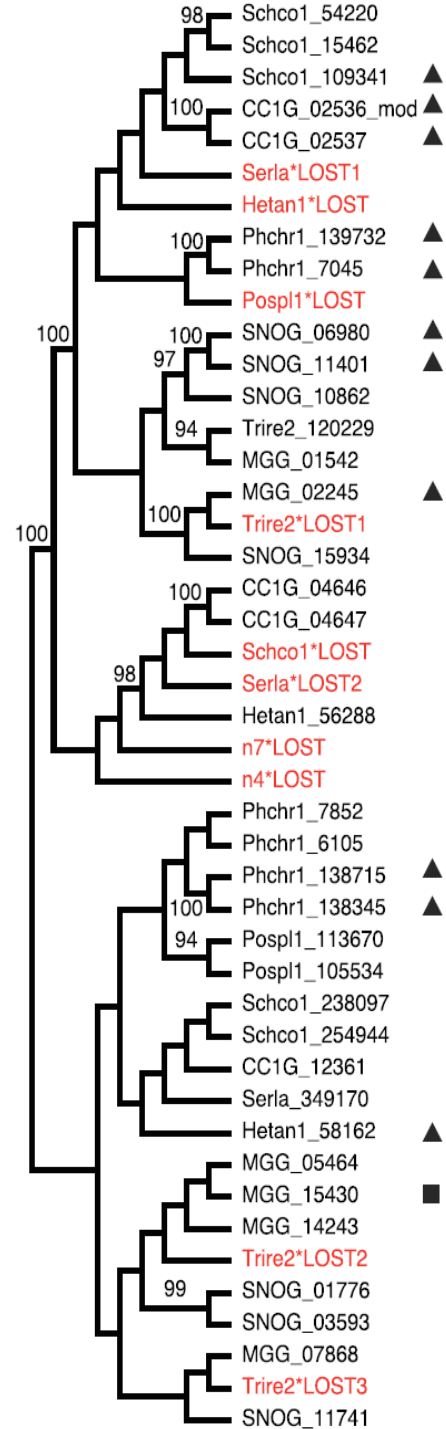


3E.

Glycoside hydrolase family 10, EWT 75

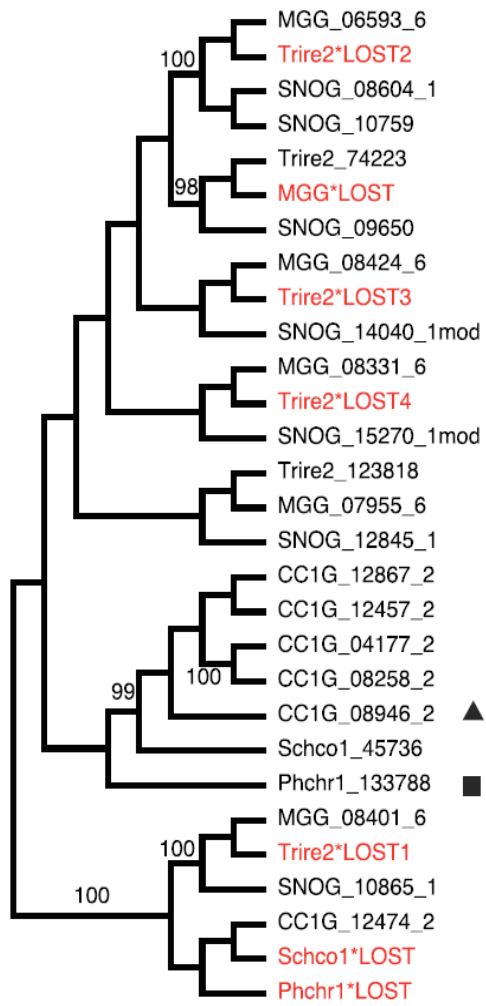


Glycoside hydrolase family 10, EWT 90



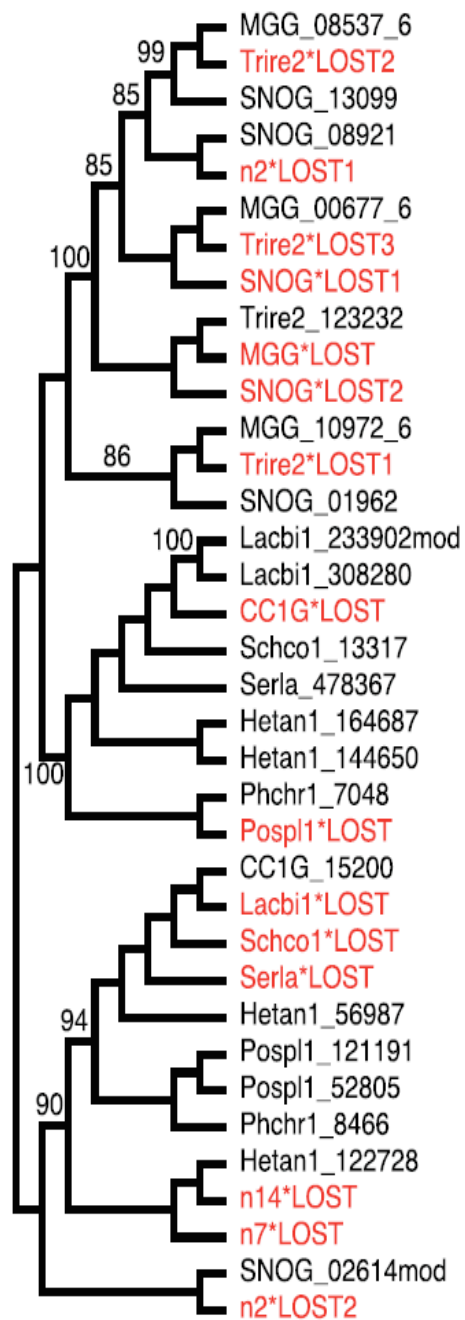
3F.

Glycoside hydrolase family 11, EWT 75 & 90

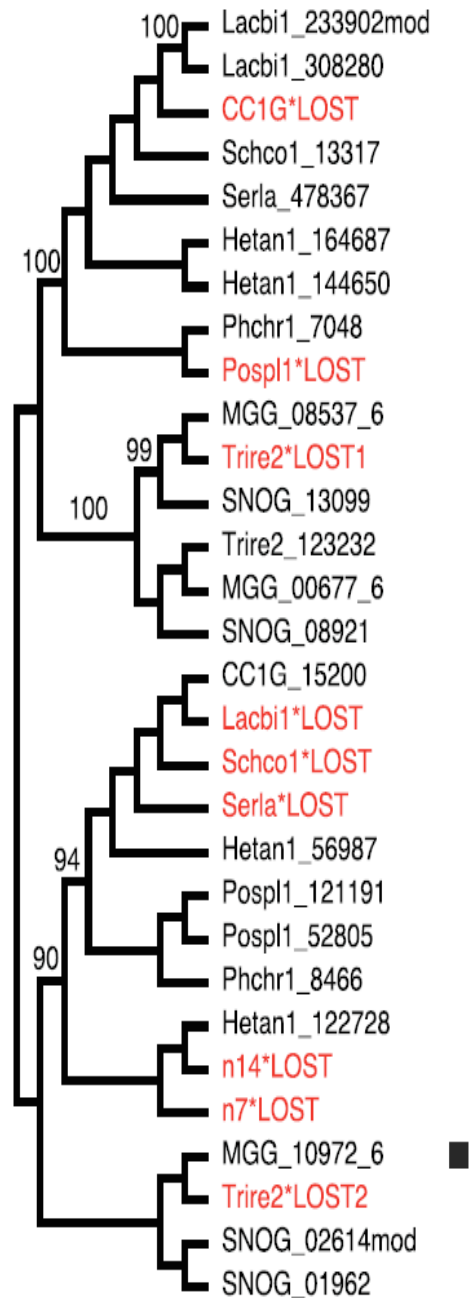


3G.

Glycoside hydrolase family 12, EWT 75

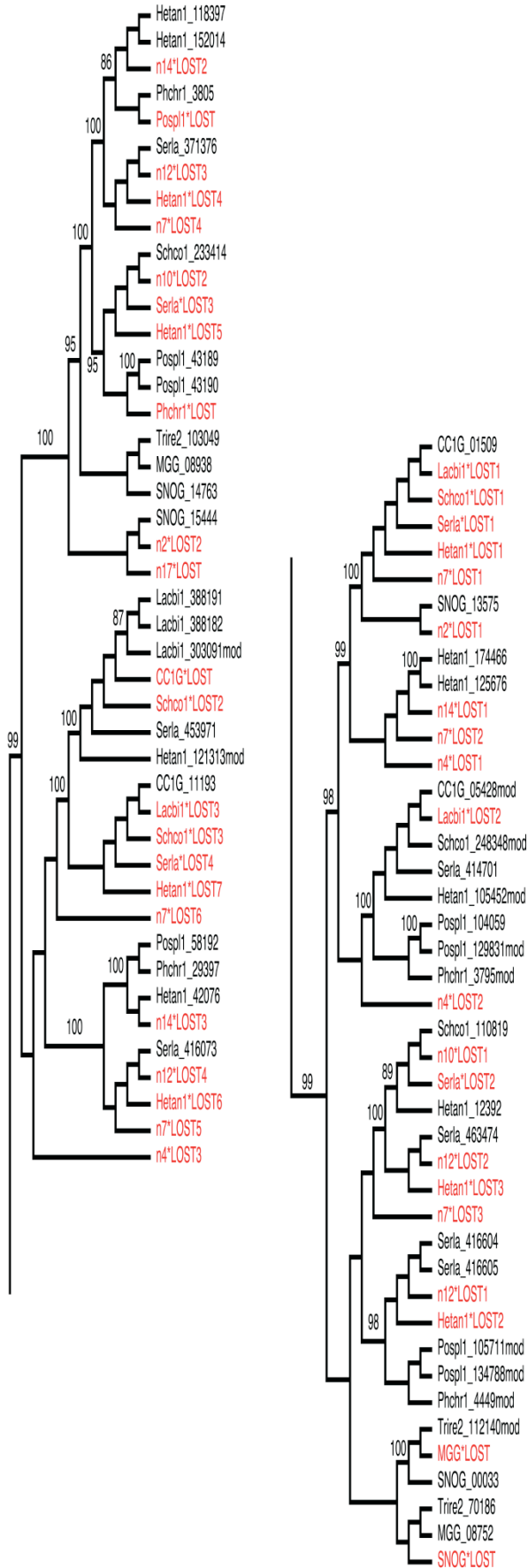


Glycoside hydrolase family 12, EWT 90

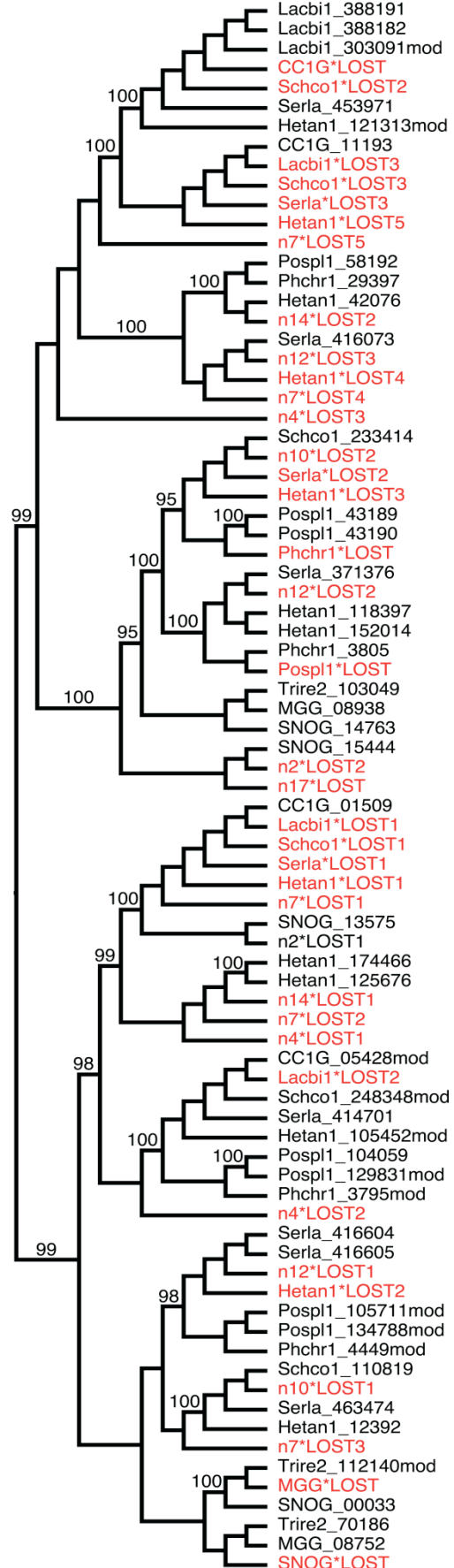


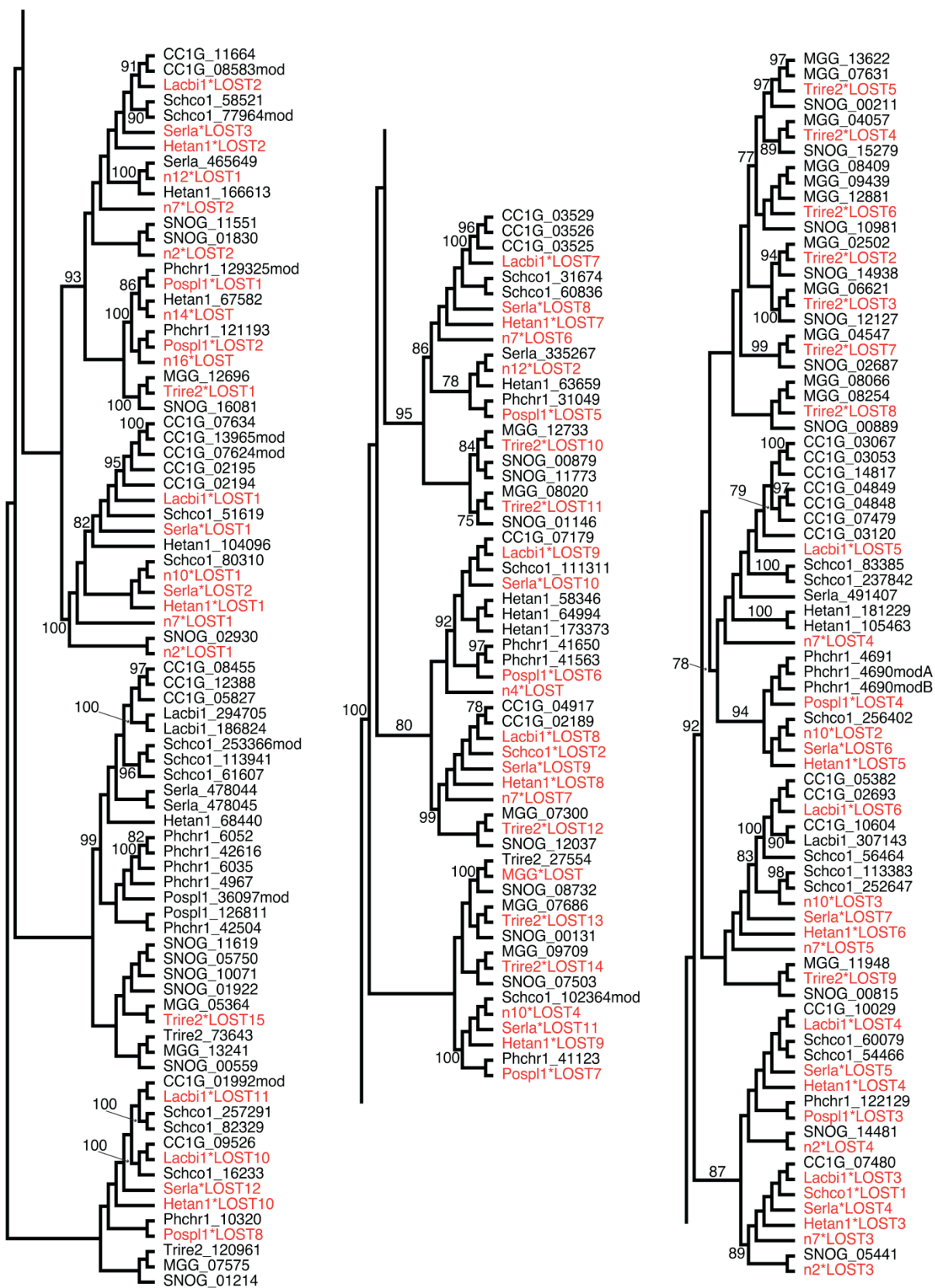
3H.

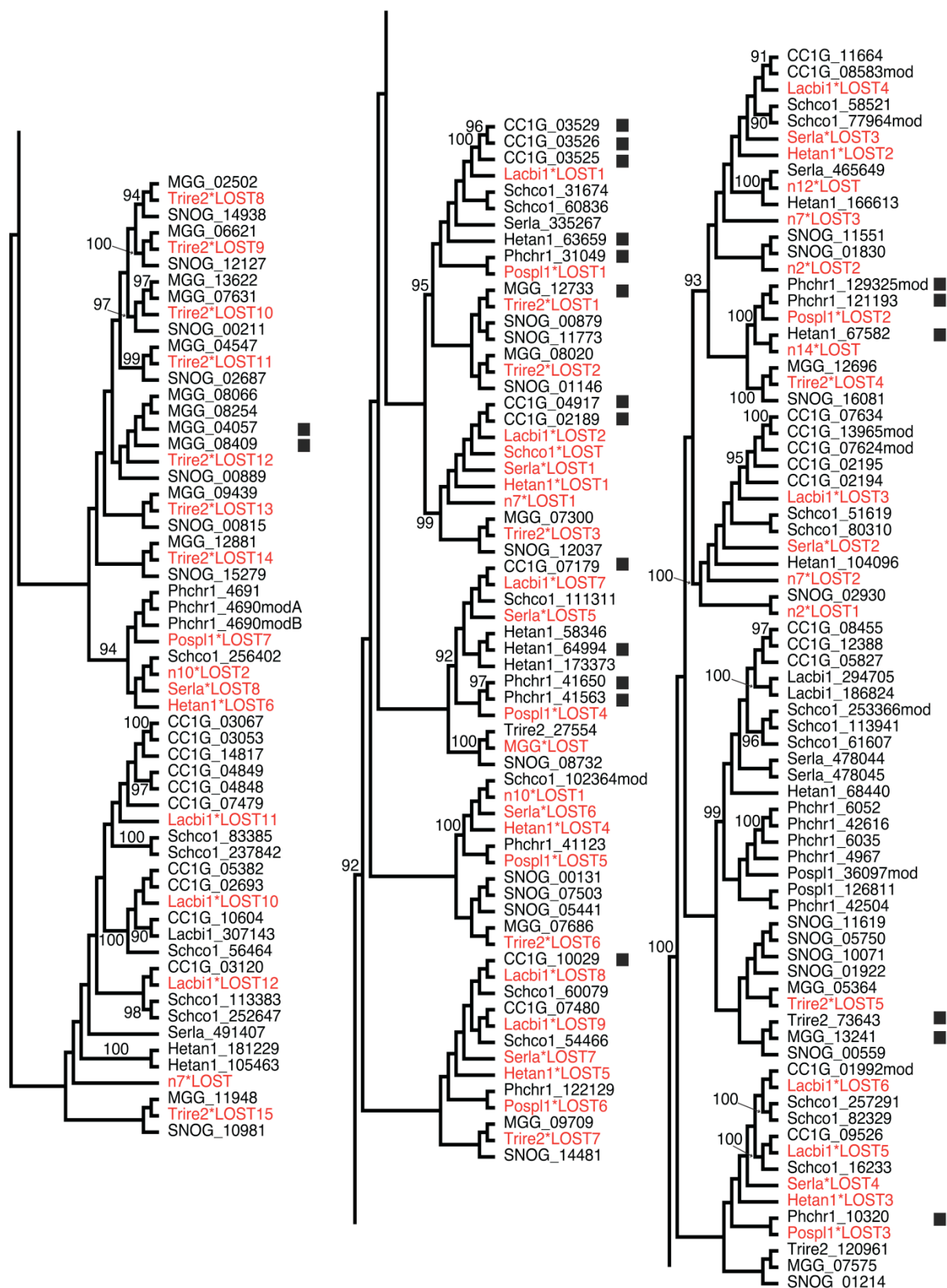
Glycoside hydrolase family 28, EWT 75



Glycoside hydrolase family 28, EWT 90

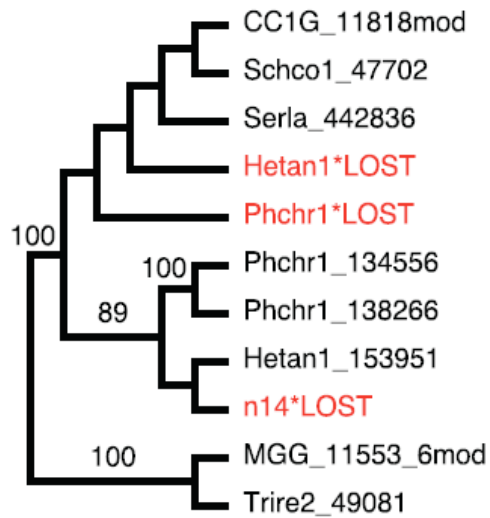




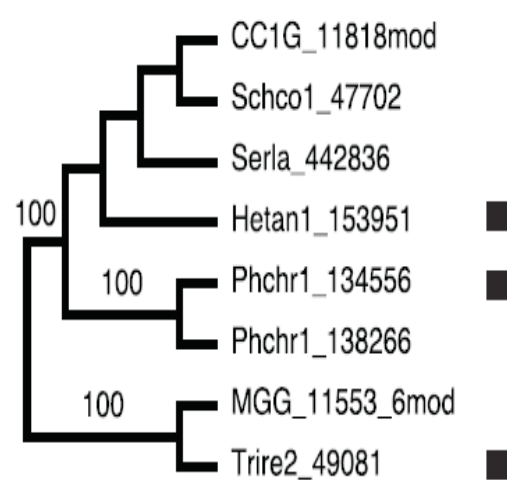


3K.

Glycoside hydrolase family 74, EWT 75

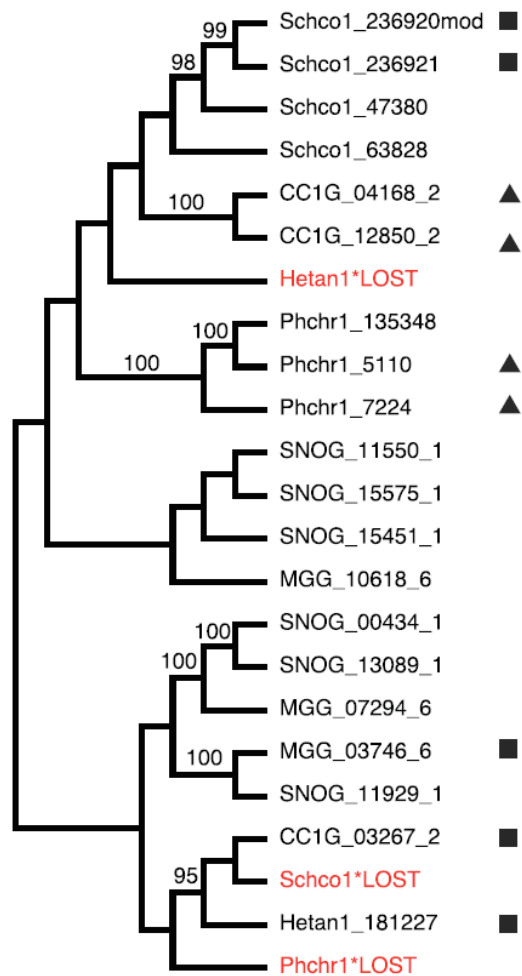


Glycoside hydrolase family 74, EWT 90

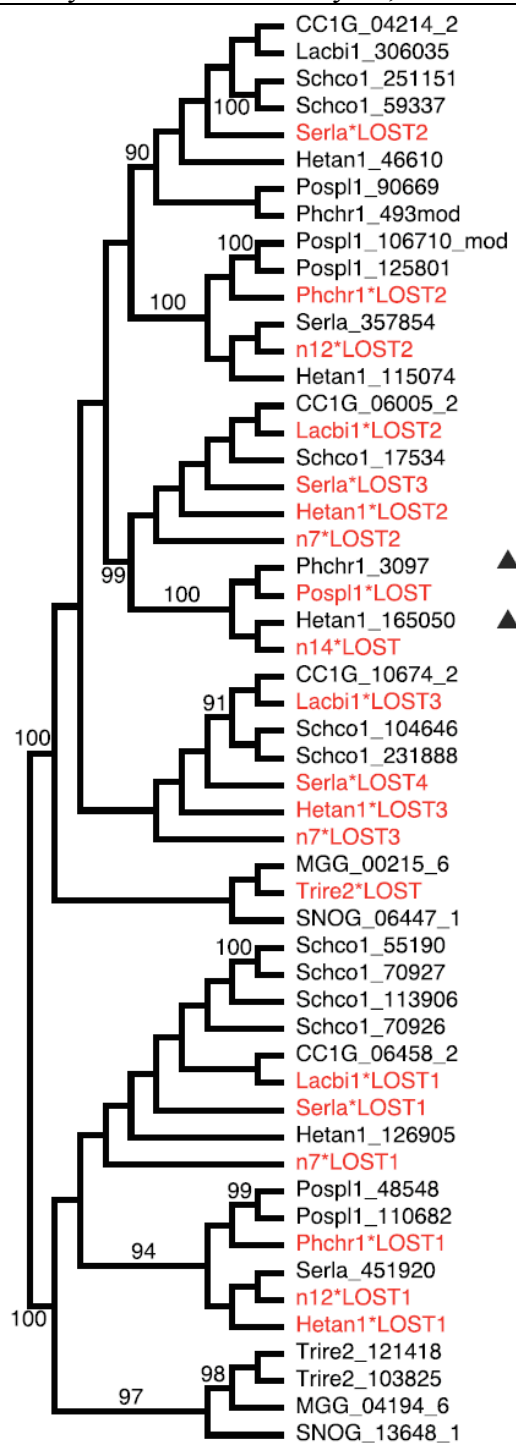


3L.

Carbohydrate esterase family 1, EWT 75 & 90



Carbohydrate esterase family 16, EWT 75



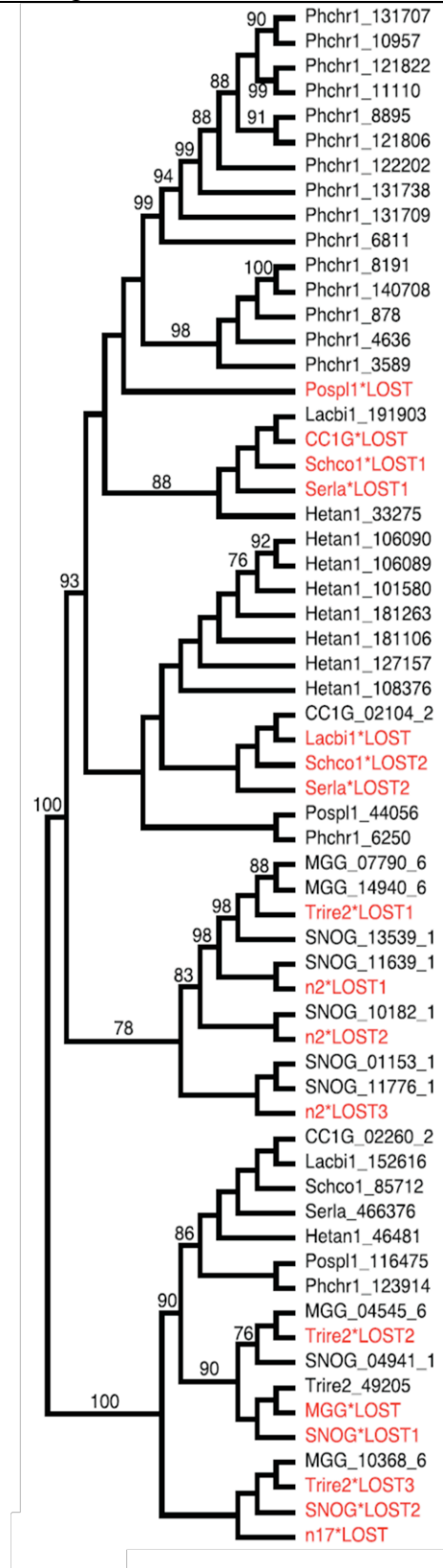
Phylogenetic tree showing relationships between various bacterial strains. The tree is rooted on the left and branches out to the right. Bootstrap values are indicated at the nodes. Strains are listed on the right, with some names in red indicating specific groups of interest.

Strains (from top to bottom):

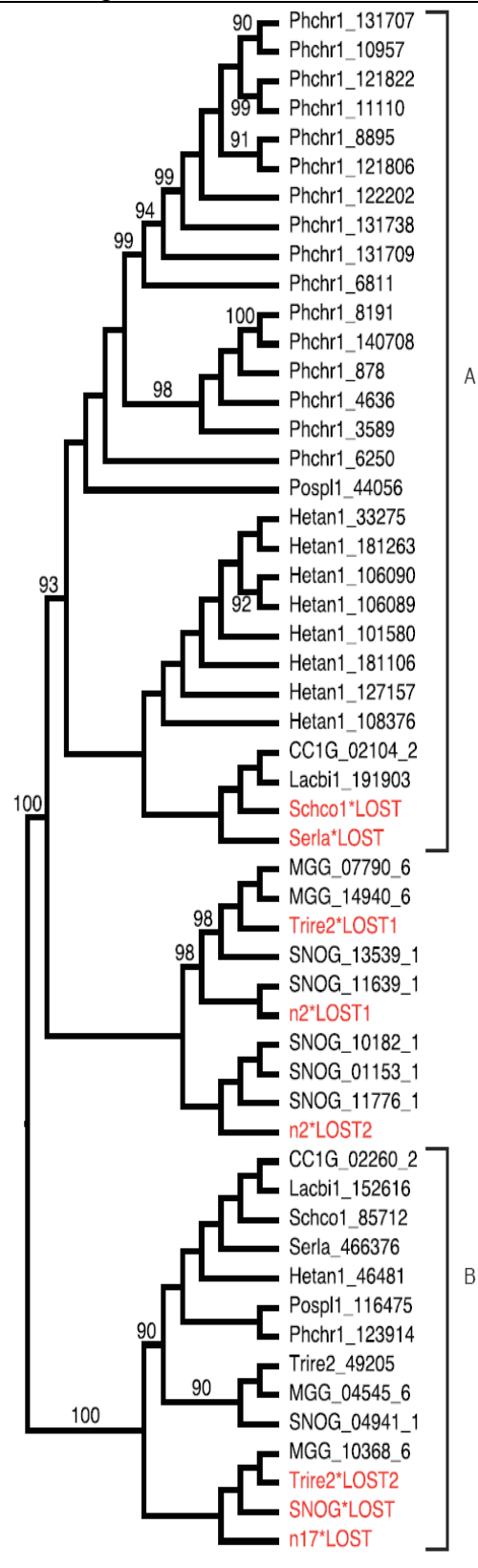
- CC1G_06005_2
- Lacbi1*LOST2
- Schco1_17534
- Serla*LOST3
- Hetan1*LOST3
- n7*LOST3
- Phchr1_3097
- Pospl1*LOST
- Hetan1_165050
- n14*LOST2
- Serla_357854
- n12*LOST1
- Hetan1*LOST2
- Pospl1_106710_mod
- Pospl1_125801
- Phchr1*LOST1
- Hetan1_115074
- n14*LOST1
- n7*LOST2
- MGG_00215_6
- Trire2*LOST
- SNOG_06447_1
- CC1G_04214_2
- Lacbi1_306035
- Schco1_251151
- Schco1_59337
- Serla*LOST1
- Hetan1_46610
- Pospl1_90669
- Phchr1_493mod
- CC1G_10674_2
- Lacbi1*LOST1
- Schco1_104646
- Schco1_231888
- Serla*LOST2
- Hetan1*LOST1
- n7*LOST1
- n4*LOST
- Schco1_55190
- Schco1_70927
- Schco1_113906
- Schco1_70926
- CC1G_06458_2
- Lacbi1*LOST3
- Serla*LOST4
- Hetan1_126905
- n7*LOST4
- Pospl1_48548
- Pospl1_110682
- Phchr1*LOST2
- Serla_451920
- n12*LOST2
- Hetan1*LOST4
- Trire2_121418
- MGG_04194_6
- Trire2_103825
- MGG*LOST
- SNOG_13648_1

3N.

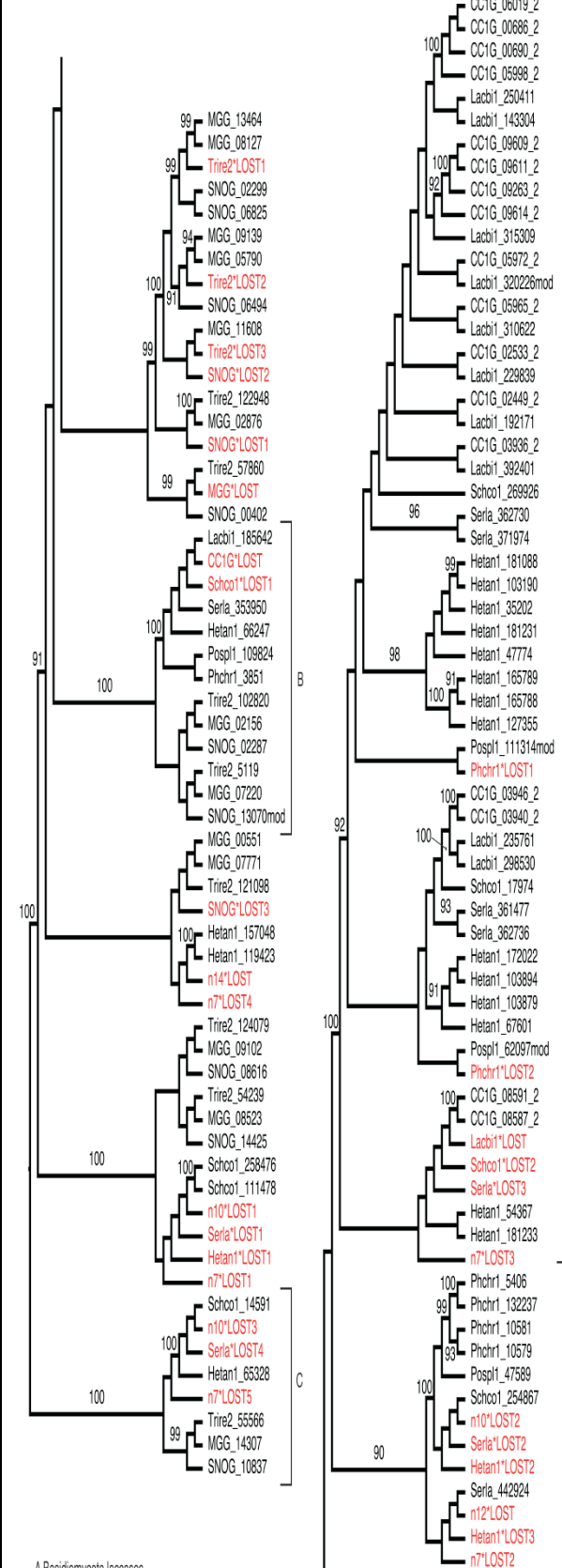
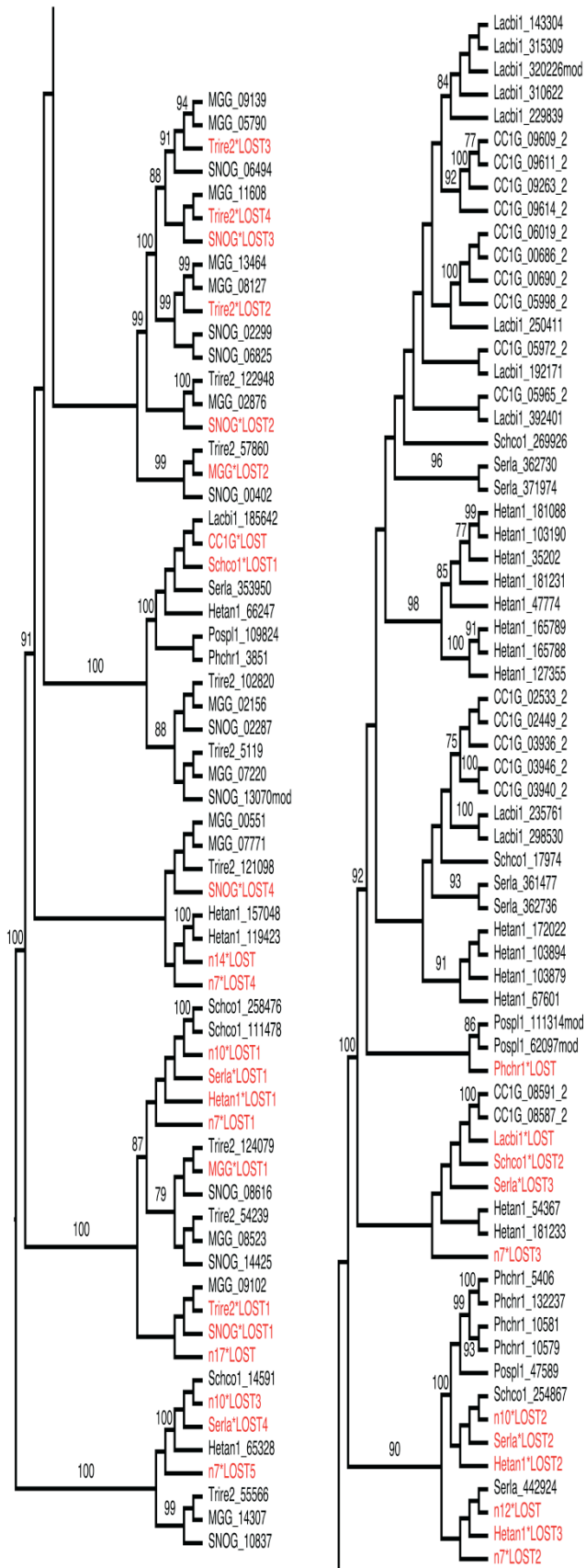
Class II peroxidases, EWT 75



Class II peroxidases, EWT 90



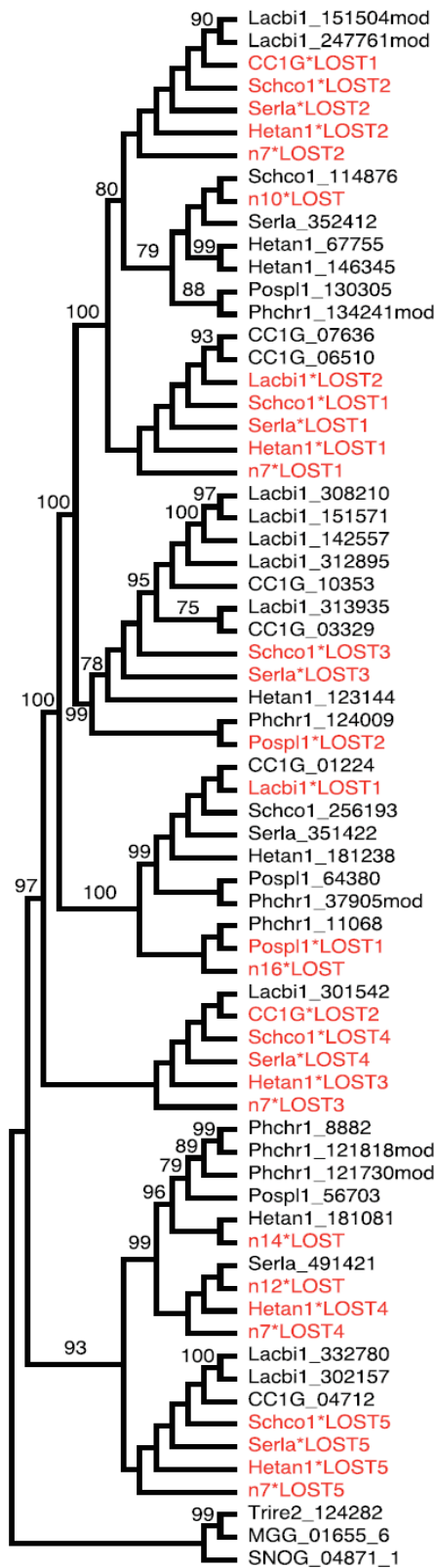
A - Basidiomycete class II peroxidises; B – Cytochrome C peroxidises (not included in duplications-losses counting)



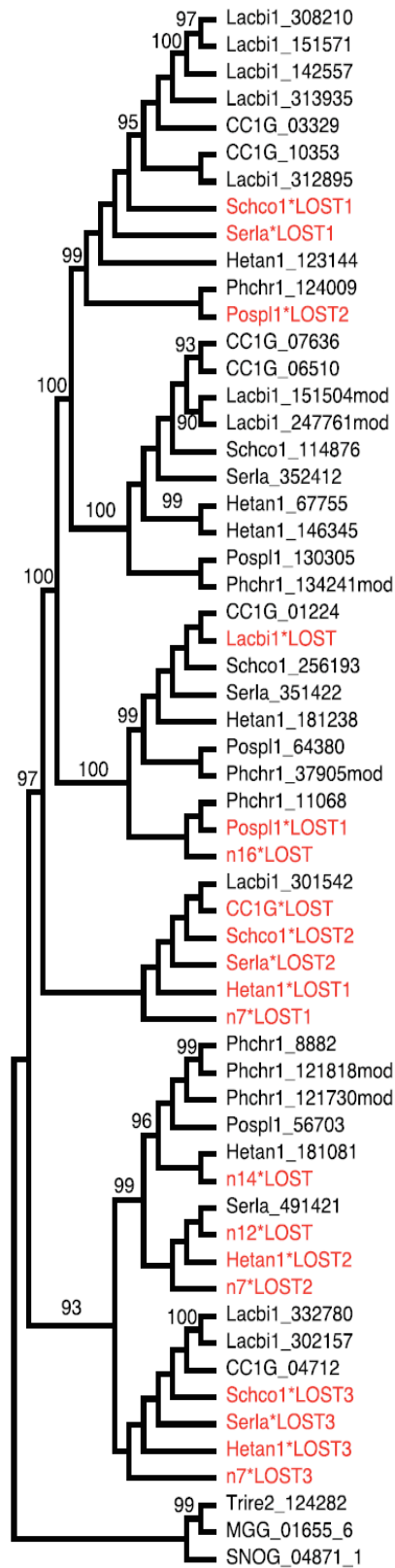
A Basidiomycete laccases
 B Fe(3) ferroxidases
 C L-ascorbate oxidases (not included in the duplications-losses counting)

3P.

Glyoxal oxidases, EWT 75

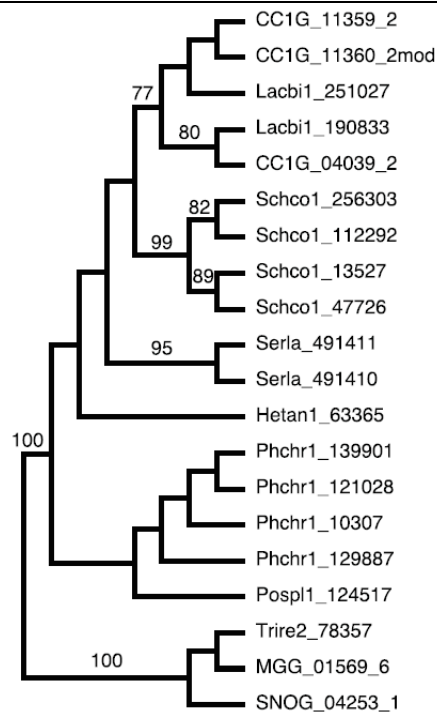


Glyoxal oxidases, EWT 90

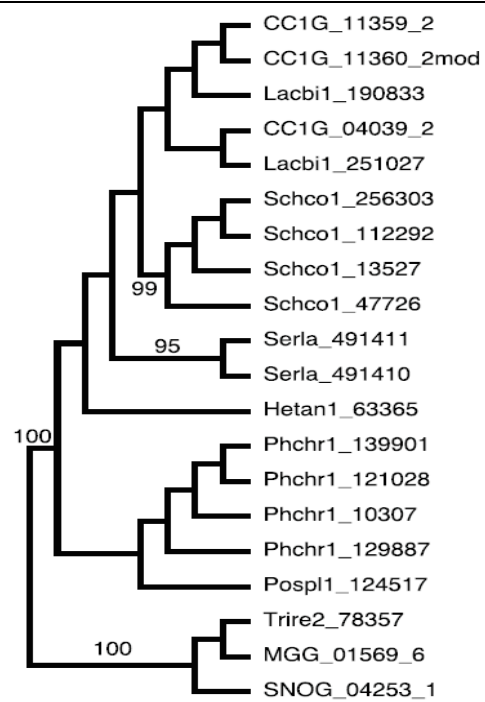


3Q.

Quinone reductase, EWT 75

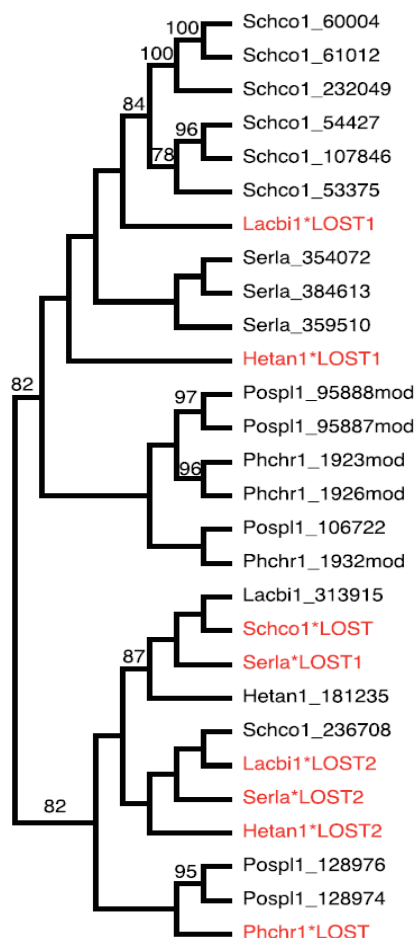


Quinone reductase, EWT 75, EWT 90

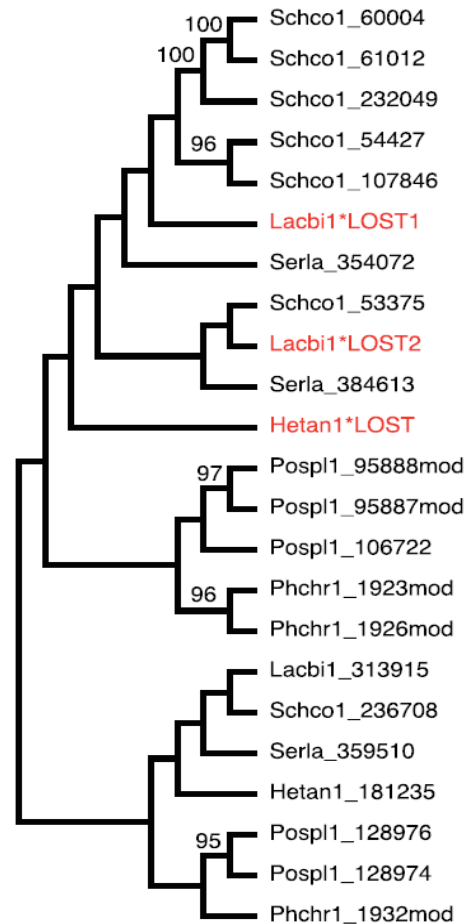


3R.

Iron-binding glycoproteins/reductases,
EWT 75

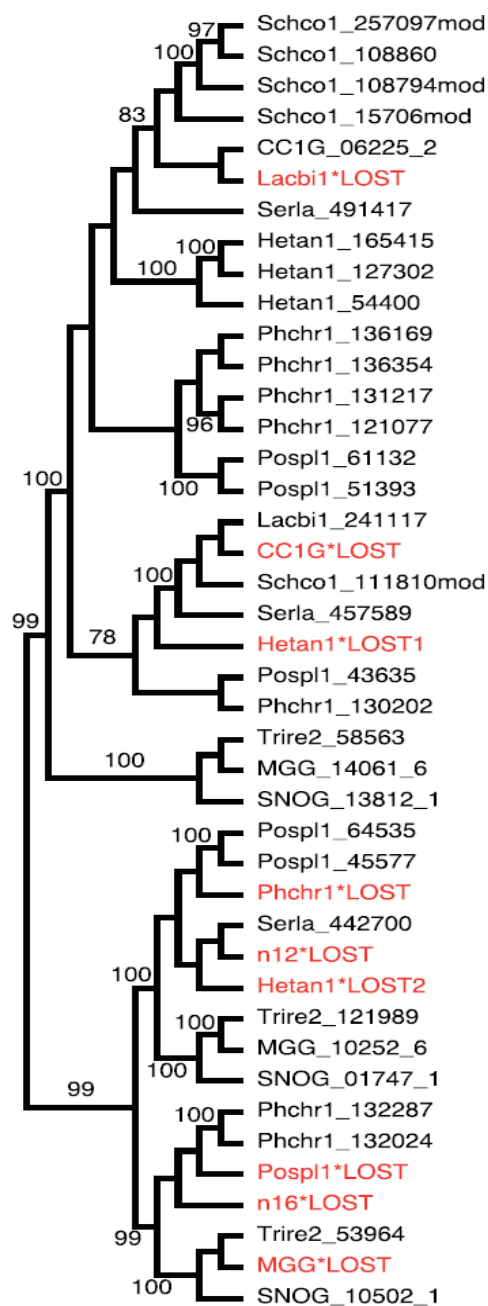


Iron-binding glycoproteins/reductases,
EWT 90

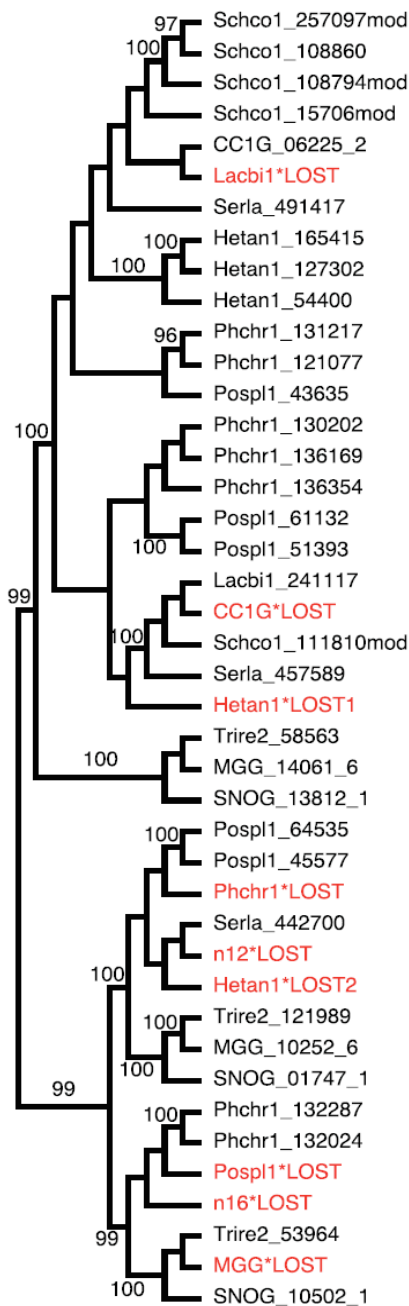


3S.

Oxalate oxidases/decarboxylases, EWT 75



Oxalate oxidases/decarboxylases, EWT 90



3T.
Cellobiose dehydrogenase, EWT 90

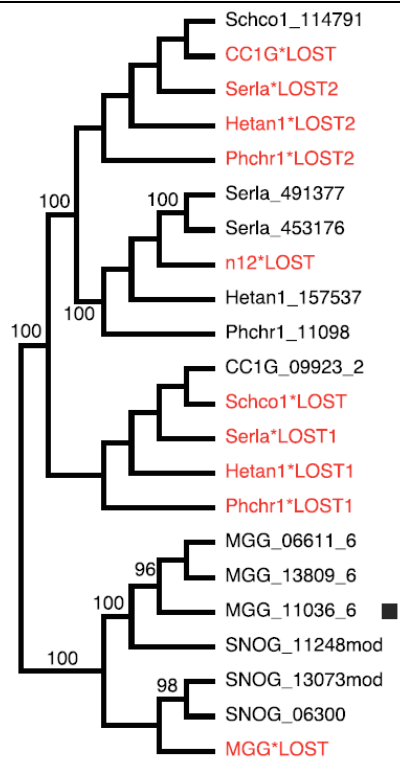


Figure S4. Functional characterisation of *S. lacrymans* transcripts with significant increased regulation (4 fold or greater, ANOVA P<0.01) when grown on wood compared with glucose-based medium (MMN) identified by microarray analysis (n=300 genes). Gene list and heat map of relative expression level is provided for the 30 *S. lacrymans* genes with greatest fold increase in transcript levels on solid wood substrate.

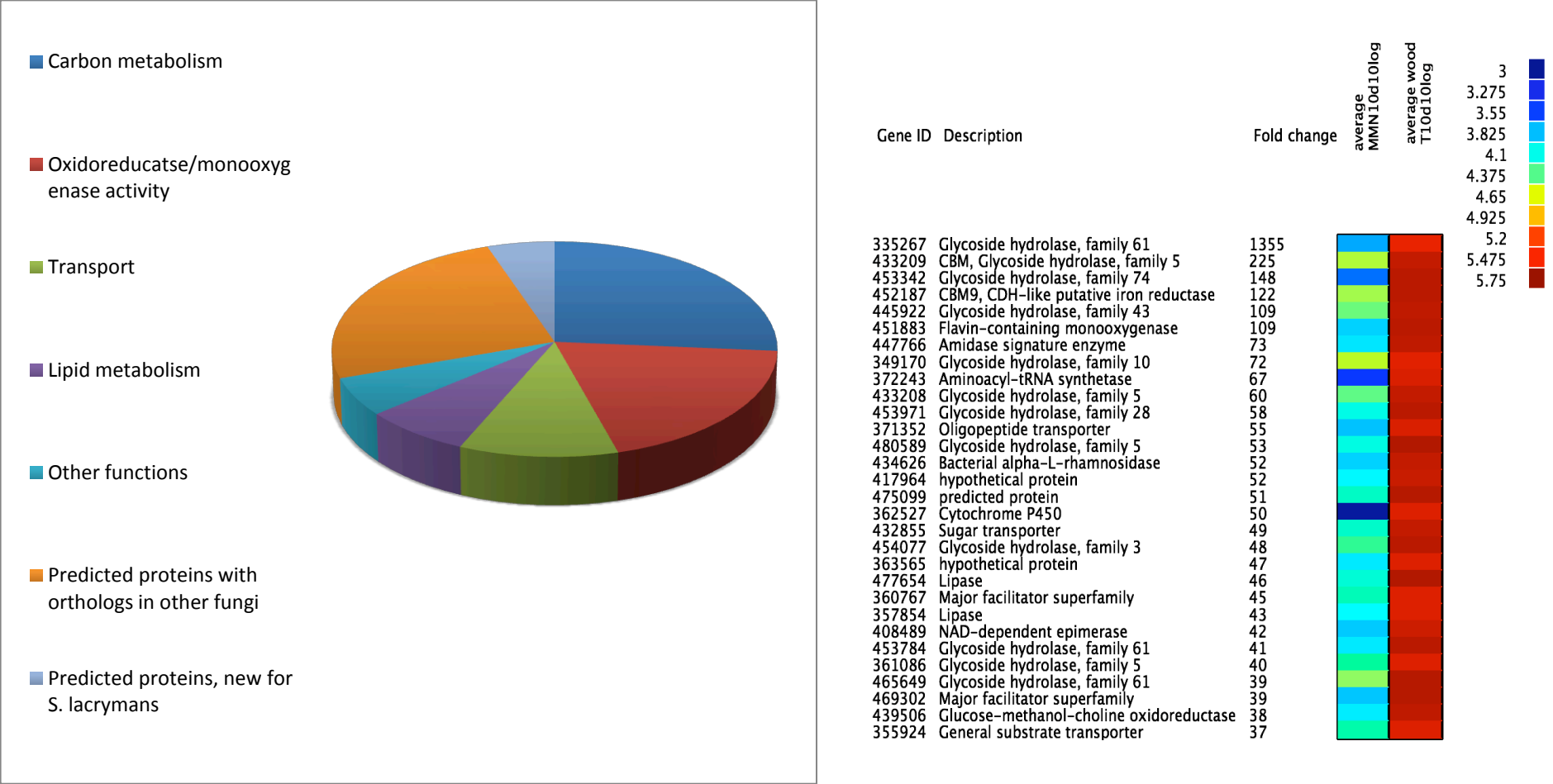


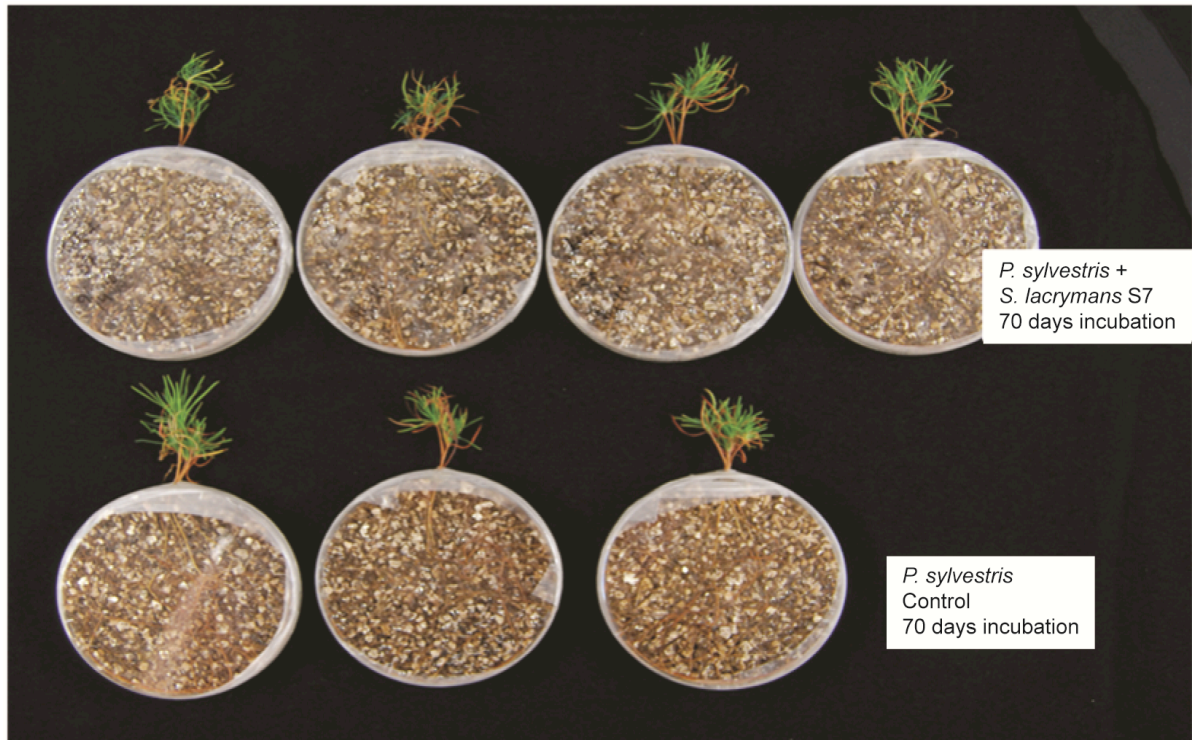
Figure S5. *S. lacrymans* S7.9 protein models with similar iron reductase domain (cd00241), including putatively annotated cellobiose dehydrogenase genes 453175 and 453176.

| Genome Database | Accession Number | Domain Structure |
|---|------------------|----------------------------|
| <i>Serpula lacrymans</i> var. <i>lacrymans</i> S7.9 | 453175 | JUNK SIGN cd00241 LINK CDH |
| <i>Serpula lacrymans</i> var. <i>lacrymans</i> S7.9 | 453176 | SIGN cd00241 LINK CDH |
| <i>Serpula lacrymans</i> var. <i>lacrymans</i> S7.9 | 452187 | SIGN cd00241 LINK CBM |
| <i>Serpula lacrymans</i> var. <i>lacrymans</i> S7.9 | 417465 | SIGN cd00241 |

Iron reductase domain = cd00241; CBM – cellulose binding domain; CDH – cellobiose dehydrogenase oxidoreductase domain; SIGN – signal peptide cleavage motif; LNK – linking sequence without specific function.

Figure S6. Interaction between *S. lacrymans* and roots of *Picea sylvestris*, A: sterile *P. sylvestris* seedlings grown in vermiculite in the presence of *S. lacrymans* and uninoculated control showing *S. lacrymans* growing on and around the roots, B: Magnification (X24) of *P. sylvestris* short laterally-branched root and *S. lacrymans* mycelium, C: Section (X400 magnification) through *P. sylvestris* root surrounded by *S. lacrymans* cells, while close interaction is observed, hartig net and true mantle usually associated with ectomycorrhizal associations are not observed. Red arrow = fungal tissue/incipient mycorrhiza, green arrow = plant cells

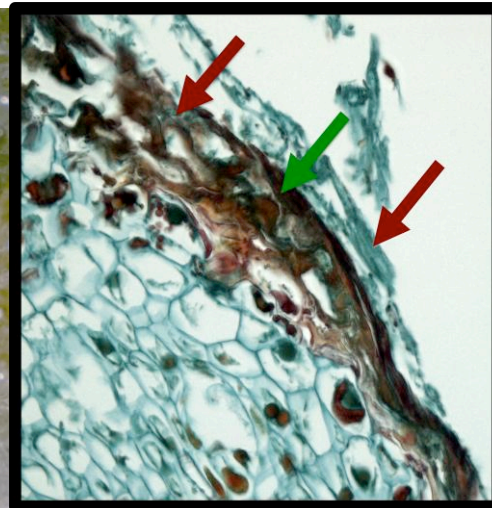
A.



B:



C:



5. Supplementary Tables

Table S1. Genomic libraries included in the *Serpula lacrymans* genome assembly and their respective assembled sequence coverage levels in the final release.

| Library Type | Average Insert Size | Number of Reads | Assembled Sequence Coverage (X) |
|----------------|---------------------|-----------------|---------------------------------|
| S7.9: | | | |
| Sanger, 3kb | 2,552 | 234,528 | 3.59 |
| Sanger, 8kb | 6,426 | 291,744 | 3.98 |
| Sanger, fosmid | 39,854 | 29,952 | 0.44 |
| Total | | 556,224 | 8.01 |
| S7.3: | | | |
| 454, std | N/A | 2,489,569 | 24.06 |
| Total | | 2,489,569 | 24.06 |

Table S2. Summary statistics of the output of the *S. lacrymans* whole genome shotgun assembly S7.9 before screening and removal of organelles and contaminating scaffolds. The table shows total contigs and total assembled basepairs for each set of scaffolds greater than the given size.

| Size | Number | Contigs | Scaffold Size | Basepairs | % Non-gap Basepairs |
|-----------|--------|---------|---------------|------------|---------------------|
| 5,000,000 | 1 | 51 | 5,733,305 | 5,701,910 | 99.45% |
| 2,500,000 | 8 | 245 | 27,090,049 | 26,889,349 | 99.26% |
| 1,000,000 | 15 | 350 | 38,261,814 | 37,960,259 | 99.21% |
| 500,000 | 20 | 388 | 42,140,175 | 41,797,878 | 99.19% |
| 250,000 | 20 | 388 | 42,140,175 | 41,797,878 | 99.19% |
| 100,000 | 22 | 399 | 42,458,091 | 42,101,004 | 99.16% |
| 50,000 | 25 | 405 | 42,692,124 | 42,298,518 | 99.08% |
| 25,000 | 25 | 405 | 42,692,124 | 42,298,518 | 99.08% |
| 10,000 | 34 | 418 | 42,825,041 | 42,429,394 | 99.08% |
| 5,000 | 43 | 432 | 42,886,961 | 42,488,184 | 99.07% |
| 2,500 | 67 | 468 | 42,968,800 | 42,566,449 | 99.06% |
| 1,000 | 68 | 469 | 42,970,512 | 42,568,161 | 99.06% |
| 0 | 68 | 469 | 42,970,512 | 42,568,161 | 99.06% |

Table S3. *S. lacrymans* final assembly statistics.

| | S7.9 | S7.3 |
|-------------------------------------|--------------------|---------------------|
| Assembly method | Arachne (Sanger) | Newbler (454/Roche) |
| Main genome scaffold total | 46 | 2133 |
| Main genome contig total | 434 | 3303 |
| Main genome scaffold sequence total | 42.8 Mb | 47.0 Mb |
| Main genome contig sequence total | 42.4 Mb (0.9% gap) | 41.2 Mb (12.4% gap) |
| Main genome scaffold N/L50 | 6/2.9 MB | 7/2.7 Mb |
| Main genome contig N/L50 | 62/228.0 KB | 143/86.6 kb |
| Number of scaffolds > 50 KB | 24 | 49 |
| % main genome in scaffolds > 50 KB | 99.60% | 95.10% |

Table S4. Predicted gene models and supporting lines of evidence

| | S7.9 | S7.3 |
|---|--------|--------|
| # gene models | 12917 | 14495 |
| % complete (with start and stop codons) | 87% | 82% |
| % genes with homology support | 69% | 68% |
| % genes with Pfam domains | 42% | 40% |
| % with 100% EST support | 33.00% | 26.00% |
| % with > 50% EST support | 74.00% | 70.00% |

Table S5. Characteristics of predicted gene models.

| | S7.9 | S7.3 |
|-------------------------|----------------|----------------|
| Avg. gene length, bp | 1600 | 1501 |
| Avg. protein length, aa | 339 | 322 |
| Avg. exon frequency | 5.6 exons/gene | 5.3 exons/gene |
| Avg. exon length, bp | 222 | 226 |
| Avg. intron length, bp | 77 | 75 |

Table S6. Functional annotation of proteins.

| | S7.9 | S7.3 |
|---------------------------------|------------|------------|
| Proteins assigned to a KOG | 5475 (42%) | 5908 (41%) |
| KOG categories genome-wide | 3011 | 3112 |
| Proteins assigned a GO term | 5148 (40%) | 5423 (37%) |
| GO terms genome-wide | 2114 | 2154 |
| Proteins assigned an EC number | 1843 (14%) | 2027 (14%) |
| EC numbers genome-wide | 619 | 647 |
| Proteins assigned a Pfam domain | 5421 (42%) | 5781 (40%) |
| Pfam domains genome wide | 2152 | 2250 |

Table S7. Comparison of putative CAZy enzymes from genome sequenced fungi with differing nutritional modes

| Mode & Species | GH | GH3 | GH5 | GH6 | GH7 | GH10 | GH11 | GH12 | GH28 | GH43 | GH51 | GH61 | GH74 | GT | PL | CE | CE1 | CE16 | CBM | CBM1 | EXP |
|---|-----|-----|--------|-----|-----|------|------|------|------|------|------|------|------|-----|----|----|-----|------|-----|------|-----|
| Brown rot | | | | | | | | | | | | | | | | | | | | | |
| <i>S. lacrymans</i> S7.9 (Boletales) | 154 | 10 | 20 (3) | 1 | 0 | 1 | 0 | 1 | 7 | 2 | 1 | 5 | 1 | 61 | 5 | 12 | 1 | 4 | 23 | 8 | 8 |
| <i>P. placenta</i> * (Polyporales) | 248 | 9 | 36 (0) | 0 | 0 | 4 | 0 | 4 | 11 | 1 | 3 | 4 | 0 | 102 | 8 | 25 | 0 | 9 | 28 | 0 | 19 |
| White rot | | | | | | | | | | | | | | | | | | | | | |
| <i>Ph. chrysosporium</i> (Polyporales) | 182 | 11 | 19 (4) | 1 | 9 | 6 | 1 | 2 | 4 | 4 | 2 | 15 | 4 | 66 | 4 | 17 | 5 | 1 | 48 | 31 | 11 |
| <i>Sc. commune</i> (Agaricales) | 236 | 12 | 16 (0) | 1 | 2 | 5 | 1 | 1 | 3 | 19 | 2 | 22 | 1 | 75 | 16 | 30 | 9 | 4 | 30 | 5 | 19 |
| Leaf litter/ Coprophilous | | | | | | | | | | | | | | | | | | | | | |
| <i>C. cinerea</i> (Agaricales) | 211 | 7 | 26 (2) | 5 | 7 | 5 | 6 | 1 | 3 | 4 | 1 | 33 | 1 | 71 | 13 | 51 | 5 | 2 | 89 | 46 | 15 |
| Ectomycorrhizal | | | | | | | | | | | | | | | | | | | | | |
| <i>Laccaria bicolor</i> (Agaricales) | 163 | 2 | 21 (1) | 0 | 0 | 0 | 0 | 3 | 6 | 0 | 0 | 8 | 0 | 88 | 7 | 17 | 1 | 1 | 24 | 1 | 12 |
| Plant pathogen | | | | | | | | | | | | | | | | | | | | | |
| <i>U. maydis</i> (Ustilaginales) | 94 | 3 | 12 (0) | 0 | 0 | 2 | 1 | 0 | 1 | 3 | 2 | 0 | 0 | 58 | 1 | 13 | 1 | 0 | 7 | 0 | 10 |
| Ascomycete Saprotrophs | | | | | | | | | | | | | | | | | | | | | |
| <i>T. reesei</i> (Hypocreales) | 192 | 13 | 8 (2) | 1 | 2 | 1 | 4 | 2 | 4 | 2 | 0 | 3 | 1 | 92 | 6 | 17 | 3 | 2 | 46 | 15 | 4 |
| <i>A. nidulans</i> (Eurotiales) | 264 | 21 | 16 (1) | 2 | 3 | 3 | 2 | 1 | 10 | 18 | 3 | 9 | 2 | 97 | 22 | 31 | 4 | 3 | 47 | 7 | 3 |
| Ascomycete Plant pathogen | | | | | | | | | | | | | | | | | | | | | |
| <i>M. grisea</i> (Magnaporthales) | 268 | 20 | 13 (1) | 3 | 5 | 7 | 5 | 3 | 3 | 20 | 3 | 23 | 1 | 105 | 5 | 53 | 10 | 1 | 86 | 22 | 4 |
| <i>S. nodorum</i> (Pleosporales) | 284 | 16 | 18 (0) | 4 | 5 | 7 | 7 | 4 | 4 | 15 | 2 | 30 | 0 | 95 | 10 | 53 | 11 | 2 | 75 | 13 | 4 |

*A full list of gene identification numbers is provided in additional file 1, additional table 1

†*P. placenta* figures are for the sequenced dikaryon; Figures in parentheses glycoside hydrolase family 5 with putative carbohydrate-binding module 1 (cellulose). GH, glycoside hydrolases; GT, glycosyl transferases; PL, polysaccharide lyases; CE, carbohydrate esterases; CMB, carbohydrate-binding module; CBM1, cellulose-binding module; EXP, plant expansin-like.

Table S8. Comparison of putative oxidoreductase enzymes from genome sequenced fungi with differing nutritional modes*

| Mode & Species | POX | MCO | CDH | AAOX | GlyOX | PryOX | GluOX | QR | AOX | OXO | IGP |
|---|-----|-----|-----|------|-------|-------|-------|----|-----|-----|-----|
| Brown rot | | | | | | | | | | | |
| <i>S. lacrymans</i> S7.9 (Boletales) | 0 | 6 | 2 | 6 | 3 | 0 | 0 | 2 | 1 | 3 | 3 |
| <i>P. placenta</i> † (Polyporales) | ? | 8 | 0 | 8 | 5 | 0 | 0 | 2 | 2 | 10 | 8 |
| White rot | | | | | | | | | | | |
| <i>Ph. chrysosporium</i> (Polyporales) | 16 | 5 | 1 | 3 | 7 | 1 | 1 | 4 | 1 | 7 | 3 |
| <i>Sc. commune</i> (Agaricales) | 0 | 5 | 1 | 1 | 2 | 1 | 4 | 4 | 1 | 5 | 7 |
| Leaf litter/ Coprophilous | | | | | | | | | | | |
| <i>C. cinerea</i> (Agaricales) | 1 | 17 | 1 | 14 | 6 | 0 | 1 | 3 | 2 | 1 | 0 |
| Ectomycorrhizal | | | | | | | | | | | |
| <i>Laccaria bicolor</i> (Agaricales) | 1 | 11 | 0 | 4 | 10 | 0 | 2 | 2 | 2 | 1 | 1 |
| Plant pathogen | | | | | | | | | | | |
| <i>U. maydis</i> (Ustilaginales) | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 |
| Saprotrophic Ascomycete | | | | | | | | | | | |
| <i>T. reesii</i> (Hypocreales) | 0 | 7 | 0 | 1 | 1 | 0 | 1 | 1 | 2 | 3 | 0 |
| <i>A. nidulans</i> (Eurotiales) | 0 | 11 | 1 | 5 | 0 | 1 | 1 | 0 | 2 | 2 | 0 |
| Ascomycete Plant pathogen | | | | | | | | | | | |
| <i>M. grisea</i> (Magnaporthales) | 2 | 12 | 3 | 1 | 1 | 0 | 2 | 1 | 2 | 2 | 0 |
| <i>S. nodorum</i> (Pleosporales) | 5 | 8 | 3 | 2 | 1 | 0 | 2 | 1 | 3 | 3 | 0 |

* A full list of gene identification numbers is provided in additional file 1, additional table 2.

† *P. placenta* figures are for the sequenced dikaryon. POX, class II peroxidases; MCO, multicopper oxidases; CDH, cellobiose dehydrogenases; AAOX, aryl alcohol oxidases; GlyOX, glyoxal oxidases; PyrOX, pyranose oxidases; GluOX, glucose oxidases; QR Quinone reductases; AOX, alcohol oxidases; OXO, oxalate oxidases /decarboxylases; IGP, iron-binding glycoproteins.

? *Postia placenta* candidate class II peroxidase candidate was identified, but appears to lack the enzymatic machinery for oxidation; database description: “Structural characteristics suggest a low redox-potential peroxidase (EC 1.11.1.7) lacking a Mn(II)-oxidation site (as in MnP and VP) and an exposed tryptophan responsible for high redox-potential substrate oxidation (as in LiP and VP)”

*, †, ‡, §, ||, ¶.

Table S9. Gene copy distribution for the gene families and organisms used in reconciliation analysis*. Red = Gene families absent in genome, blue – highest number of gene copies per family for Basidiomycete genomes only. GH – glycoside hydrolase, CE – carbohydrate esterase, † - GH family 5 endoglucanases (EC3.2.1.4) & mannanases (EC3.2.1.78) only, ‡ - CE1 excluding S-formylglutathione hydrolases, § - CE16 cinnamoyl esterases only, no lysophospholipases.

| Gene Family | <i>Coprinopsis cinerea</i> | <i>Laccaria bicolor</i> | <i>Schizophyllum commune</i> | <i>Serpula lacrymans</i> | <i>Heterobasidion irregulare</i> | <i>Postia placenta</i> | <i>Phanerochaete chrysosporium</i> | <i>Trichoderma reesei</i> | <i>Magnaporthe grisea</i> | <i>Stagonospora nodorum</i> |
|----------------------------------|----------------------------|-------------------------|------------------------------|--------------------------|----------------------------------|------------------------|------------------------------------|---------------------------|---------------------------|-----------------------------|
| CAZs | | | | | | | | | | |
| GH Family 3 | 7 | 2 | 12 | 10 | 12 | 5 | 9 | 13 | 18 | 16 |
| GH Family 5† | 6 | 3 | 3 | 8 | 6 | 5 | 5 | 3 | 4 | 5 |
| GH Family 6 | 5 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 3 | 4 |
| GH Family 7 | 4 | 0 | 2 | 0 | 1 | 0 | 7 | 2 | 5 | 5 |
| GH Family 10 | 5 | 0 | 5 | 1 | 2 | 2 | 6 | 1 | 6 | 7 |
| GH Family 11 | 6 | 0 | 1 | 0 | 0 | 0 | 1 | 2 | 5 | 7 |
| GH Family 12 | 1 | 2 | 1 | 1 | 4 | 2 | 2 | 1 | 3 | 4 |
| GH Family 28 | 3 | 3 | 3 | 7 | 8 | 7 | 4 | 3 | 2 | 4 |
| GH Family 61 | 30 | 3 | 22 | 5 | 10 | 2 | 16 | 3 | 21 | 27 |
| GH Family 74 | 1 | 0 | 1 | 1 | 1 | 0 | 2 | 1 | 1 | 0 |
| CE Family 1‡ | 3 | 0 | 9 | 0 | 1 | 0 | 3 | 0 | 3 | 6 |
| CE Family 16§ | 4 | 1 | 4 | 2 | 4 | 5 | 2 | 2 | 2 | 2 |
| Oxidoreductases | | | | | | | | | | |
| Class II peroxidases | 1 | 1 | 0 | 0 | 8 | ? | 16 | 0 | 2 | 5 |
| Multicopper oxidases | 17 | 11 | 5 | 6 | 17 | 4 | 5 | 7 | 12 | 8 |
| Cellobiose dehydrogenases | 1 | 0 | 1 | 2 | 1 | 0 | 1 | 0 | 3 | 3 |
| Glyoxal oxidases | 6 | 10 | 2 | 3 | 5 | 3 | 7 | 1 | 1 | 1 |
| Quinone reductases | 3 | 2 | 4 | 2 | 1 | 1 | 4 | 1 | 1 | 1 |
| Oxalate oxidases/ decarboxylases | 1 | 1 | 5 | 3 | 3 | 5 | 7 | 3 | 2 | 3 |
| Iron binding glycopeptides | 0 | 1 | 7 | 3 | 1 | 5 | 3 | 0 | 0 | 0 |
| CAZY/oxidoreductases /total | 75/29/104 | 14/26/40 | 64/24/88 | 36/19/55 | 50/36/86 | 28/19/47 | 58/43/101 | 32/12/44 | 73/21/94 | 87/21/108 |

* A full list of gene identification numbers is provided in additional file 1, additional table 3.

? *Postia placenta* candidate class II peroxidase candidate was identified, but appears to lack the enzymatic machinery for oxidation; database description: “Structural characteristics suggest a low redox-potential peroxidase (EC 1.11.1.7) lacking a Mn(II)-oxidation site (as in MnP and VP) and an exposed tryptophan responsible for high redox-potential substrate oxidation (as in LiP and VP)”.

Table S10. Number of of orthologs (% amino acid identity) between *Serpula* strains and with other Agaricomycetes

| | S7.9 | S7.3 |
|-------------------------------------|--------------|--------------|
| <i>Serpula lacrymans</i> (non-self) | 10378(98.5%) | 10378(98.5%) |
| <i>Agaricus bisporus</i> | 5354(58%) | 5616(58%) |
| <i>Coprinopsis cinerea</i> | 5585(57%) | 5859(57%) |
| <i>Laccaria bicolor</i> | 5578(60%) | 5840(60%) |
| <i>Phanerochaete chrysosporium</i> | 5138(60%) | 5341(60%) |
| <i>Postia placenta</i> | 5097(61%) | 5330(61%) |

Table S11. Posterior probability distribution for divergence times (in millions of years before present) for major lineages in the Agaricomycetes and the *Serpula-Austropaxillus* split using relaxed molecular clock analyses with normal distribution.

| Node | tMRCA | Posterior means | 95% HPD |
|------|---------------------------------------|-----------------|---------------|
| 1 | Ingroup (-Fomitiporia) | 180.7 | 219.4 – 140.6 |
| 2 | Russulales | 114.4 | 178.1 – 66.9 |
| 3 | Polyporales | 98.4 | 150.9 – 57.6 |
| 4 | Agaricomycetidae | 166.1 | 189.8 – 126.5 |
| 5 | core Agaricomycetidae (-Jaapia) | 143.9 | 174.2 – 119.3 |
| 6 | Agaricales | 124.0 | 155.2 – 109.4 |
| 7 | Boletales (inc. <i>S. lacrymans</i>) | 113.4 | 140.5 – 87.3 |
| 8 | Amylocorticiales | 93.7 | 127.9 – 43.4 |
| 9 | Atheliales | 77.7 | 118.5 – 33.2 |
| 10 | Marasmioid clade | 92.0 | 93.9 – 90.0 |
| 11 | Suillineae | 52.0 | 54.0 – 50.0 |
| 12 | <i>Serpula-Austropaxillus</i> | 34.9 | 53.1 – 15.0 |

Table S12. *S. lacrymans* transcriptomic comparison from cultures grown on either glucose or wood wafers. 50 genes with the greatest differential expression when grown on wood compared with glucose cultures are shown.











































| Sequence ID | Average glucose 10 days | Average wood 10 days | Ratio wood 10d/glucose 10D | Signal peptide | Interpro definition |
|-------------|-------------------------|----------------------|----------------------------|----------------|--|
| 335267 | 4 | 5935 | 1355,044 | * | Glycoside hydrolase, family 61 |
| 433209 | 150 | 33665 | 224,987 | * | Cellulose-binding region, fungal ; Glycoside hydrolase, family 5 ; Cellulose-binding region, fungal ; Cellulose-binding region, fungal ; Cellulose-binding region, fungal ; Glycoside hydrolase, family 5 ; Glycoside hydrolase, catalytic core ; Cellulose-binding region, fungal |
| 453342 | 29 | 4307 | 148,172 | * | Glycoside hydrolase, family GH74 with BNR repeat |
| 452187 | 260 | 31801 | 122,372 | * | Cellulose-binding region, fungal ; Cellulose-binding region, fungal ; Cellulose-binding region, fungal ; Cellulose-binding region, fungal ; Carbohydrate-binding family 9/cellobiose dehydrogenase, cytochrome ; Cellulose-binding region, fungal |
| 445922 | 234 | 25504 | 109,013 | * | Glycoside hydrolase, family GH43 ; β -galactosidase |
| 451883 | 72 | 7797 | 108,642 | | Flavin-containing monooxygenase FMO ; FAD-dependent pyridine nucleotide-disulphide oxidoreductase |
| 447766 | 121 | 8898 | 73,354 | | Amidase signature enzyme ; Amidase signature enzyme |
| 349170 | 493 | 35476 | 71,977 | * | Glycoside hydrolase, family 10 ; Glycoside hydrolase, family 10 ; Glycoside hydrolase, family 10 ; Glycoside hydrolase, family 10 ; Glycoside hydrolase, catalytic core |
| 372243 | 52 | 3510 | 67,380 | | Aminoacyl-tRNA synthetase, class I, conserved site ; Glycoside hydrolase, catalytic core |
| 433208 | 409 | 24635 | 60,182 | | Glycoside hydrolase, family 5 ; Glycoside hydrolase, family 5 ; Glycoside hydrolase, catalytic core |
| 453971 | 228 | 13133 | 57,688 | * | Glycoside hydrolase, family 28 ; Glycoside hydrolase, family 28 ; Pectin lyase fold/virulence factor |
| 371352 | 127 | 6979 | 55,084 | | Oligopeptide transporter OPT superfamily ; Tetrapeptide transporter, OPT1/isp4 ; Oligopeptide transporter OPT superfamily |

| | | | | | |
|--------|-----|-------|--------|---|---|
| 480589 | 250 | 13356 | 53,330 | * | Glycoside hydrolase, family 5 ; Serine/threonine protein kinase, active site ; DNA topoisomerase, type IIA, subunit B or N-terminal ; Glycoside hydrolase, catalytic core |
| 434626 | 150 | 7823 | 52,128 | * | Bacterial alpha-L-rhamnosidase ; Six-hairpin glycosidase-like |
| 417964 | 199 | 10315 | 51,742 | * | |
| 475099 | 316 | 16233 | 51,434 | | |
| 362527 | 22 | 1088 | 49,779 | | Cytochrome P450 ; Cytochrome P450, E-class, group I ; Cytochrome P450 ; Cytochrome P450 ; Cytochrome P450 |
| 432855 | 325 | 15793 | 48,544 | * | Sugar transporter ; General substrate transporter ; Sugar transporter ; Major facilitator superfamily ; Sugar transporter, conserved site ; MFS general substrate transporter |
| 454077 | 464 | 22294 | 48,035 | | Glycoside hydrolase, family 3, N-terminal ; Glycoside hydrolase, family 3, C-terminal ; Glycoside hydrolase, catalytic core ; Glycoside hydrolase, family 3, C-terminal |
| 363565 | 194 | 9197 | 47,348 | * | |
| 477654 | 321 | 14645 | 45,655 | * | Lipase, GDSL ; Esterase, SGNH hydrolase-type |
| 360767 | 399 | 18001 | 45,166 | | Major facilitator superfamily MFS-1 ; Major facilitator superfamily ; MFS general substrate transporter |
| 357854 | 253 | 10882 | 43,047 | * | Lipase, GDSL Esterase, SGNH hydrolase-type |
| 408489 | 173 | 7248 | 41,901 | | NAD-dependent epimerase/dehydratase ; NAD(P)-binding |
| 453784 | 218 | 9017 | 41,425 | * | Glycoside hydrolase, family 61 |
| 361086 | 541 | 21583 | 39,897 | * | Glycoside hydrolase, family 5 ; Glycoside hydrolase, catalytic core |
| 465649 | 742 | 28838 | 38,854 | * | Glycoside hydrolase, family 61 |
| 469302 | 186 | 7172 | 38,560 | | Major facilitator superfamily MFS-1 ; Major facilitator superfamily ; MFS general substrate transporter |
| 439506 | 249 | 9491 | 38,100 | | Glucose-methanol-choline oxidoreductase, N-terminal ; Glucose-methanol-choline oxidoreductase, C-terminal ; Glucose-methanol-choline oxidoreductase ; Glucose-methanol-choline oxidoreductase, N-terminal ; Glucose-methanol-choline oxidoreductase, N-terminal |
| 355924 | 530 | 19410 | 36,588 | | General substrate transporter ; Sugar transporter ; Major facilitator |

| | | | | | |
|--------|------|-------|--------|---|---|
| | | | | | superfamily ; MFS general substrate transporter |
| 471097 | 80 | 2910 | 36,452 | | |
| 440027 | 192 | 6884 | 35,800 | * | Aromatic-ring hydroxylase ; Monooxygenase, FAD-binding |
| 433131 | 366 | 12605 | 34,395 | * | Carboxylesterase, type B ; Carboxylesterase, type B |
| 458151 | 863 | 29183 | 33,796 | * | Chitin-binding, domain 3 |
| 442919 | 833 | 27457 | 32,945 | * | Carboxylesterase, type B ; Carboxylesterase, type B |
| 478367 | 148 | 4738 | 32,100 | * | Glycoside hydrolase, family 12 ; Glycoside hydrolase, family 12 ; Concanavalin A-like lectin/glucanase |
| 371376 | 177 | 5554 | 31,458 | * | Glycoside hydrolase, family 28 ; Parallel beta-helix repeat ; Glycoside hydrolase, family 28 ; Pectin lyase fold/virulence factor |
| 434546 | 829 | 25656 | 30,931 | * | Glycoside hydrolase, family 3, N-terminal ; Glycoside hydrolase, family 3, N-terminal ; Glycoside hydrolase, family 3, C-terminal ; Glycoside hydrolase, catalytic core ; Glycoside hydrolase, family 3, C-terminal |
| 439032 | 204 | 5949 | 29,134 | * | Pectinesterase, catalytic ; Pectin lyase fold/virulence factor |
| 451920 | 425 | 12053 | 28,332 | * | Lipase, GDSL ; Esterase, SGNH hydrolase-type |
| 469301 | 258 | 7265 | 28,193 | | MFS general substrate transporter |
| 355683 | 783 | 21495 | 27,442 | * | Cellulose-binding region, fungal ; Glycoside hydrolase, family 5 ; Cellulose-binding region, fungal ; Cellulose-binding region, fungal ; Cellulose-binding region, fungal ; Cellulose-binding region, fungal ; Glycoside hydrolase, family 5 ; Glycoside hydrolase, catalytic core ; Cellulose-binding region, fungal |
| 408717 | 754 | 20156 | 26,742 | * | Carboxylesterase, type B ; Carboxylesterase, type B ; Carboxylesterase, type B |
| 362272 | 507 | 13172 | 25,956 | | Glycoside hydrolase, family 5 ; Glycoside hydrolase, catalytic core |
| 391341 | 174 | 4291 | 24,628 | | |
| 450352 | 1232 | 30308 | 24,597 | | |
| 443184 | 232 | 5572 | 24,024 | | Cytochrome P450, E-class, group I ; Cytochrome P450 ; Cytochrome P450 ; Cytochrome P450 |
| 445039 | 882 | 21161 | 23,998 | | Glycoside hydrolase, family 1 ; Glycoside hydrolase, family 1 ; Glycoside hydrolase, family 1 ; Glycoside hydrolase, catalytic core |

| | | | | | |
|--------|-----|-------|--------|---|--|
| 441797 | 454 | 10901 | 23,989 | * | Histidine acid phosphatase ; Histidine acid phosphatase, eukaryotic ; Histidine acid phosphatase ; Histidine acid phosphatase |
| 453756 | 212 | 5089 | 23,987 | | |

Table S13. Extracellular proteins from *S. lacrymans* (dikaryon) grown in solid state wood culture. Protein Nr refers to protein number in annotated genomes of *S. lacrymans* monokayrons S7.9 and S7.3; A – all models, F – filtered models

| Protein # | ProteinName | Protein Nr | Protein in | | | Signal peptide | NCBIInr | Pfam | Pfam domain | CDD | CDD domain |
|---------------------|---|--------------|---|---|---|----------------|--------------|-------------------------------|---|-------------------------------|---|
| | | | with Ca | without Ca | both | | | | | | |
| Glycosyl hydrolases | | | | | | | | | | | |
| 1 | beta-glucosidase, GH1 | 79-A-433206 |  |  |  | Q | Q25BW4 | PF00232 | Glycosyl hydrolase family 1 | cl01046 | Glycosyl hydrolase family 1 |
| 2 | beta-mannosidase GH2, (beta-galactosidase/beta-glucuronidase) | 732-F-107546 |  |  |  | S | Q2UN00 | PF00703 | Glyco_hydro_2 | COG3250 | Beta-galactosidase/beta-glucuronidase |
| 3 | beta-glucosidase, GH3 | 732-F-103008 |  |  |  | S | Q4WGT3 | PF00933 PF01915 | Glycosyl hydrolase family 3 N terminal domain Glycosyl hydrolase family 3 C terminal domain | cl03393 cl07971 | Glyco_hydro_3_C super family Glycosyl hydrolase family 3 N terminal domain |
| 4 | beta-glucosidase, GH3 | 732-F-117918 | |  | | Q | A1CA51 | PF00933 PF01915 PF07691 | Glycosyl hydrolase family 3 N terminal domain Glycosyl hydrolase family 3 C terminal domain PA14 domain | cl03393 cl07971 | Glyco_hydro_3_C super family Glycosyl hydrolase family 3 N terminal domain |
| 5 | exo-1,4-beta-xylosidase, GH3 | 732-A-78770 |  |  |  | S | B8NYD8 | PF00933 PF01915 | Glycosyl hydrolase family 3 N terminal domain Glycosyl hydrolase family 3 C terminal domain | cl03393 cl07971 | Glyco_hydro_3_C super family Glycosyl hydrolase family 3 N terminal domain |
| 6 | beta-glucosidase, GH3 | 79-A-406110 |  |  |  | S | Q4WGT3 | PF00933 PF01915 | Glycosyl hydrolase family 3 N terminal domain Glycosyl hydrolase family 3 C terminal domain | cl03393 cl07971 | Glyco_hydro_3_C super family Glycosyl hydrolase family 3 N terminal domain |
| 7 | endo-1,4-beta-glucanase, (cellulase) GH5 | 732-A-106926 |  |  |  | Q | XP_001838381 | PF00150 | Cellulase (glycosyl hydrolase family 5) | cl12144 | Cellulase super family |
| 8 | endo-1,4-beta-xylanase, (xylanase) GH10 | 79-F-349170 |  |  |  | S | O60206 | PF00331 | Glycosyl hydrolase family 10 | cl01495 | Glyco_hydro_10 super family |
| 9 | alpha-glucosidase, GH31 | 79-F-447930 |  |  |  | S | Q9P999 | PF01055 | Glycosyl hydrolases family 31 | cl11402 PRK10658 | GH31 super family alpha-xylosidase YicI |
| 10 | alpha-glucosidase, (maltase) GH31 | 79-A-411290 |  | | | S | P29064 | PF01055 | Glycosyl hydrolases family 31 | cd06602 COG1501 | GH31_MGAM_SI_GAA Alpha-glucosidases, GH family 31 |
| 11 | beta-galactosidase, (lactase) GH35 | 79-F-453704 |  |  |  | Q | P29853 | PF01301 PF10435 | Glycosyl hydrolases family 35 Beta-galactosidase. domain 2 | cl03154 cl11086 | Glyco_hydro_42 Beta-galactosidase, domain 2 |
| 12 | beta-galactosidase, (lactase) GH35 | 732-F-167797 |  |  |  | S | Q700S9 | PF01301 PF10435 | Glycosyl hydrolases family 35 Beta-galactosidase. domain 2 | cl03154 cl11086 COG1874 | Glyco_hydro_42 Beta-galactosidase, domain 2 LacA |
| 13 | alpha-1,2-mannosidase, GH47 | 732-F-102103 |  |  |  | S | Q12563 | PF01532 | Glycosyl hydrolase family 47 | cl08327 | Glyco_hydro_47 super family |
| 14 | alpha-N-arabinofuranosidase, GH51 | 79-F-412847 | |  | | S | Q8NK90 | PF06964 | Alpha-L-arabinofuranosidase C-terminus | cl01412 | Alpha-L-AF_C super family |
| 15 | alpha-1,2-mannosidase, GH92 | 732-F-158402 |  | | | Q | EFI98309 | PF07971 | Glycosyl hydrolase family 92 | cl14007 | aman2_put super family |
| 16 | beta-glucuronidase, GH79 | 79-F-468146 |  |  |  | S | B0DA80 | no Pfam | no Pfam | no CDD | no CDD |
| Oxidoreductases | | | | | | | | | | | |
| 17 | polyphenol oxidase (tyrosinase) | 732-F-71478 |  |  |  | Q | Q00024 | PF00264 | Common central domain of tyrosinase | cl02830 | Tyrosinase super family |
| 18 | multicopper oxidase (laccase) | 79-F-362730 | |  | | S | Q12719 | PF00394 PF07731 PF07732 | Multicopper oxidase Multicopper oxidase Multicopper oxidase | cl11412 TIGR03388 | Cu-oxidase super family ascorbase |
| 19 | MOX/AOX, alcohol oxidase | 79-F-471949 | |  | | Q | A8DPS4 | PF00732 PF05199 | GMC oxidoreductase GMC oxidoreductase | cl02950 cl08434 COG2303 | GMC_oxred_N super family GMC_oxred_C super family BetA |
| 20 | secreted FAD-oxidoreductase | 79-F-459686 |  | | | S | EFI92548 | PF01565 PF08031 | FAD binding domain Berberine and berberine like | cl10516 cl10516 | FAD_binding_4 super family FAD_binding_4 super family |






































| Protein # | ProteinName | Protein Nr | Protein in | | | Signal peptide | NCBIInr | Pfam | Pfam domain | CDD | CDD domain |
|---------------------------------|---|--------------|---|---|---|----------------|--------------|--------------------|---|--|--|
| | | | with Ca | without Ca | both | | | | | | |
| Peptidases | | | | | | | | | | | |
| 21 | metallopeptidase (M20/M25/M40) | 732-F-191527 |  |  |  | Q | Q9P6I2 | PF01546 PF07687 | Peptidase family M20/M25/M40 Peptidase dimerisation domain | cl09126 COG0624 | M20_dimer super family ArgE |
| 22 | aspartyl protease | 732-F-159520 |  |  |  | S | P17576 | PF00026 | Eukaryotic aspartyl protease | cd05471 | pepsin_like |
| 23 | aspartyl protease | 732-F-82569 |  | | | S | P17576 | PF00026 | Eukaryotic aspartyl protease | cd05471 | pepsin_like |
| 24 | dipeptidyl-peptidase | 79-F-470352 | |  | | S | P0C959 | PF00326 | Prolyl oligopeptidase family | cl12031 COG1506 | Esterase_lipase super family DAP2 - Dipeptidyl aminopeptidases/ acylaminoacyl-peptidases |
| 25 | tripeptidyl-peptidase | 79-F-446956 |  |  |  | S | Q70J59 | PF00082 PF09286 | Subtilase family Pro-kumamolisin. activation domain | cd04056 cl07889 | Peptidases_S53 Pro-kuma_activ super family |
| 26 | serine peptidase (S28) | 79-F-354447 | |  | | Q | B0CZX3 | PF05577 | Serine carboxypeptidase S28 | cl12031 | Esterase_lipase super family |
| Lipases/Esterases/Phosphatases | | | | | | | | | | | |
| 27 | carboxylesterase | 732-F-78161 |  |  |  | S | XP_002470650 | PF00135 | Carboxylesterase | cl12031 COG2272 | Esterase_lipase super family Carboxylesterase type B |
| 28 | neutral ceramidase | 79-F-462762 |  | | | S | B0DCM2 | PF04734 | Neutral/alkaline non-lysosomal ceramidase | cl04712 | Ceramidase_alk super family |
| 29 | carboxylesterase | 79-A-433735 |  |  |  | Q | XP_002470650 | PF00135 | Carboxylesterase | cl02423 cl12031 COG2272 | LRR_RI super family Esterase_lipase super family Carboxylesterase type B |
| 30 | histidine acid phosphatase (3-phytase A) | 79-F-441797 |  |  |  | S | Q9C1T1 | PF00328 | Histidine acid phosphatase | cd07061 | HP_HAP_like |
| 31 | carboxylesterase | 732-F-116990 |  |  |  | S | P20261 | PF00135 | Carboxylesterase | cl12031 COG2272 | Esterase_lipase super family Carboxylesterase type B |
| 32 | carboxylesterase | 79-F-464062 |  | | | S | P32949 | PF00135 | Carboxylesterase | cl12031 COG2272 | Esterase_lipase super family Carboxylesterase type B |
| 33 | carboxylesterase | 79-F-415739 |  |  |  | S | Q5XG92 | PF00135 | Carboxylesterase | | |
| Other proteins/Unknown proteins | | | | | | | | | | | |
| 34 | unknown protein | 79-F-363565 |  | | | S | XP_001834028 | no Pfam | no Pfam | no CDD | no CDD |
| 35 | amidase | 732-F-159561 |  | | | S | Q9URY4 | PF01425 PF01494 | Amidase FAD binding domain | cl09931 cl11426 PRK07538 PRK08137 | NADB_Rossmann super family GGCT_like super family PRK07538 amidase, Provisional |
| 36 | similar to putative secreted protein from <i>L. bicolor</i> | 79-F-462086 |  | | | S | XP_001876668 | PF03227 | Gamma interferon inducible lysosomal thiol reductase (GILT) | cl03944 | GILT super family |
| 37 | putative agmatinase | 79-F-473708 |  | | | S | A2R3Y9 | PF00491 | Arginase family | cl00306 | Arginase super family |
| 38 | similar to phosphatidylglycerol/ phosphatidylinositol transfer protein | 79-F-467241 |  | | | S | Q5KIR9 | PF02221 | ML domain | cd00917 | PG-PI_TP |
| 39 | PNGase A | 79-F-443896 |  |  |  | S | O43119 | PF12222 | Peptide N-acetyl-beta-D-glucosaminyl asparaginase amidase A | cl13632 | PNGaseA super family |

Table S14. Comparison of proteins identified as expressed on solid state wood culture in the proteomic study with transcript expression in the transcriptomic study.

| Proteomic Protein # | S7.9 Protein number | Protein Name | # peptides identified | Transcript fold change | Relative transcript level | Ranking based on expression† |
|---------------------|--------------------------------------|-------------------------|-----------------------|----------------------------|--------------------------------|---|
| 1 | 433206 | Glycoside Hydrolase 1 | 13 | No Probe | | |
| 2 | 361341 | Glycoside Hydrolase 2 | 9 | 6.9 | 19573 | 225 (1.9) |
| 3 | 434546 354470 359881 452594 | Glycoside Hydrolase 3 | 13 | 30.9 8.3 0.79 1.3 | 25656 19873 3964 3684 | 59 (0.5) 212 (1.8) 3832 (28.7) 4058 (34.4) |
| 4 | 362919 451868 446869 | Glycoside Hydrolase 3 | 13 | 12.9 6.8 2.0 | 15220 26195 2851 | 505 (4.2) 53 (0.45) 4846 (41.1) |
| 5 | 454077 | Glycoside Hydrolase 3 | 7 | 48.0 | 22294 | 116 (0.98) |
| 6 | 406110 | Glycoside Hydrolase 3 | 13 | No Probe | | |
| 7 | 465929 362272 | Glycoside Hydrolase 5 | 7 | 8.4 26.0 | 23026 13172 | 102 (0.9) 689 (5.8) |
| 8 | 349170 | Glycoside Hydrolase 10 | 3 | 72.0 | 35476 | 8 (0.07) |
| 9 | 447930 | Glycoside Hydrolase 31 | 10 | 22.6 | 15957 | 438 (3.7) |
| 10 | 411290 | Glycoside Hydrolase 31 | 4 | No Probe | | |
| 11 | 453704 | Glycoside Hydrolase 35 | 7 | 2.6 | 9763 | 1190 (10.1) |
| 12 | 415055 | Glycoside Hydrolase 35 | 7 | 9.0 | 12872 | 725 (6.1) |
| 13 | 445448 | Glycoside Hydrolase 47 | 5 | 1.6 | 25242 | 68 (0.6) |
| 14 | 412847 | Glycoside Hydrolase 51 | 2 | 8.8 | 7527 | 1820 (15.4) |
| 15 | 413209 | Glycoside Hydrolase 92 | 6 | 2.0 | 11074 | 953 (8.1) |
| 16 | 468146 | Glycoside Hydrolase 79 | 7 | 8.1 | 22796 | 106 (0.9) |
| 17 | 435855 462697 | Polyphenol Oxidase | 15 | 0.98 0.92 | 20491 1042 | 180 (1.5) 7340 (62.2) |
| 18 | 362730 | Multicopper Oxidase | 4 | 1.5 | 8202 | 1596 (13.5) |
| 19 | 471949 | MOX/AOX Alcohol Oxidase | 4 | 7.6 | 26661 | 50 (0.4) |
| 20 | 459686 | FAD-Oxidoreductase | 3 | No Probe | | |
| 21 | 463883 | Metallopeptidase | 6 | 1.3 | 15994 | 433 (3.7) |

| | | | | | | |
|----|------------------|-------------------------------|---|--------------|--------------|-----------------------------|
| 22 | 470882 * | Aspartyl Peptidase | 3 | 0.95 | 13689 | 636 (5.4) |
| 23 | 470882 * | Aspartyl Peptidase | 3 | NA | | |
| 24 | 470352 | Dipeptidyl Peptidase | 2 | 1.2 | 13460 | 660 (5.6) |
| 25 | 446956 | Tripeptidyl Peptidase | 3 | 6.6 | 15243 | 502 (4.3) |
| 26 | 354447 | Serine Peptidase | 5 | 0.43 | 4831 | 3161 (26.8) |
| 27 | 443197 | Carboxylesterase | 9 | 15.9 | 17016 | 359 (3.0) |
| 28 | 462762 | Neutral Ceramidase | 2 | 1.1 | 6932 | 2059 (17.4) |
| 29 | 433735 | Carboxylesterase | 6 | No Probe | | |
| 30 | 441797 | Histidine Acid Phosphatase | 4 | 24.0 | 27457 | 41 (0.35) |
| 31 | 442919 | Carboxylesterase | 6 | 33.0 | 10901 | 982 (8.3) |
| 32 | 464062 | Carboxylesterase | 5 | 7.5 | 37155 | 5 (0.04) |
| 33 | 415739 | Carboxylesterase | 3 | 16.0 | 20823 | 161 (1.4) |
| 34 | 363565 | Unknown | 5 | 47.4 | 9197 | 1319 (11.2) |
| 35 | 447766 447765 | Amidase | 6 | 73.4 14.1 | 8898 5367 | 14.02 (11.9) 2862 (24.3) |
| 36 | 462086 | Thiol Reductase | 5 | 0.70 | 4436 | 3456 (29.3) |
| 37 | 473708 | Agmatinase | 3 | 2.8 | 6909 | 2123 (18.0) |
| 38 | 467241 | Phosphatidyl Transfer Protein | 4 | 0.86 | 15310 | 494 (4.2) |
| 39 | 443896 | PNGase A | 3 | 9.1 | 16970 | 362 (3.1) |

* no homolog identified in S7.9, e-value similarity to S7.9 470882 for proteins 21 and 22 were 1×10^{-103} and 1×10^{-99} respectively. †Ranking of gene based on relative transcript levels in 10 day wood cultures, based on 11804 probes. Figure in brackets indicate the position as a percentage, i.e. 5.0 = transcript is in the top 5% highly expressed genes. Proteins 10, 20, 23, 28, 32, 34, 35, 36, 37 and 38 were associated with the presence of Calcium. Only proteins 37 and 38 were shown to be differentially regulated in the transcriptomic study.

Table S15. Comparison of secondary metabolic genes in sequenced basidiomycetes

| Species | Polyketide synthases | Terpene cyclases | Prenyl transferases |
|--------------------------|----------------------|------------------|---------------------|
| <i>S. lacrymans</i> | 5 | 10 | 2 |
| <i>C. cinerea</i> | 2 | 3 | 1 |
| <i>L. bicolor</i> | 4 | 3 | 2 |
| <i>Sc. commune</i> | 1 | 2 | 1 |
| <i>Ph. chrysosporium</i> | 1 | 3 | 0 |
| <i>P. placenta</i> | 4 | 5 | 1 |
| <i>C. neoformans</i> | 1 | 1 | 0 |
| <i>U. maydis</i> | 1 | 1 | 0 |

References and Notes

1. F. Martin *et al.*, Sequencing the fungal tree of life. *New Phytol.* **190**, 818 (2011). [doi:10.1111/j.1469-8137.2011.03688.x](https://doi.org/10.1111/j.1469-8137.2011.03688.x) [Medline](#)
2. F. Martin, in *Biology of the Fungal Cell*, R. J. Howard, N. A. R. Gow, Eds. (Springer Berlin Heidelberg, 2007), vol. 8, pp. 291-308.
3. R. L. Gilbertson, Wood rotting fungi of North America. *Mycologia* **72**, 1 (1980). [doi:10.2307/3759417](https://doi.org/10.2307/3759417)
4. D. S. Hibbett, M. J. Donoghue, Analysis of character correlations among wood decay mechanisms, mating systems, and substrate ranges in homobasidiomycetes. *Syst. Biol.* **50**, 215 (2001). [doi:10.1080/10635150151125879](https://doi.org/10.1080/10635150151125879) [Medline](#)
5. K.-E. Eriksson, R. A. Blanchette, P. Ander, *Microbial and enzymatic degradation of wood and wood components*. (Springer-Verlag, Berlin; New York, 1990).
6. B. D. Lindahl *et al.*, Spatial separation of litter decomposition and mycorrhizal nitrogen uptake in a boreal forest. *New Phytol.* **173**, 611 (2007). [doi:10.1111/j.1469-8137.2006.01936.x](https://doi.org/10.1111/j.1469-8137.2006.01936.x) [Medline](#)
7. R. R. Northup, Z. Yu, R. A. Dahlgren, K. A. Vogt, Polyphenol control of nitrogen release from pine litter. *Nature* **377**, 227 (1995). [doi:10.1038/377227a0](https://doi.org/10.1038/377227a0)
8. B. Goodell *et al.*, Low molecular weight chelators and phenolic compounds isolated from wood decay fungi and their role in the fungal biodegradation of wood. *J. Biotechnol.* **53**, 133 (1997). [doi:10.1016/S0168-1656\(97\)01681-7](https://doi.org/10.1016/S0168-1656(97)01681-7)
9. T. Shimokawa, M. Nakamura, N. Hayashi, M. Ishihara, Production of 2,5-dimethoxyhydroquinone by the brown-rot fungus *Serpula lacrymans* to drive extracellular Fenton reaction. *Holzforschung* **58**, 305 (2005). [doi:10.1515/HF.2004.047](https://doi.org/10.1515/HF.2004.047)
10. V. Arantes, A. M. Milagres, T. R. Filley, B. Goodell, Lignocellulosic polysaccharides and lignin degradation by wood decay fungi: the relevance of nonenzymatic Fenton-based reactions. *J. Ind. Microbiol. Biotechnol.* **38**, 541 (2011). [doi:10.1007/s10295-010-0798-2](https://doi.org/10.1007/s10295-010-0798-2) [Medline](#)
11. S. F. Curling, C. A. Clausen, J. E. Winandy, Experimental method to quantify progressive stages of decay of wood by basidiomycete fungi. *Int. Biodeterior. Biodegradation* **49**, 13 (2002). [doi:10.1016/S0964-8305\(01\)00101-9](https://doi.org/10.1016/S0964-8305(01)00101-9)
12. M. Binder, D. S. Hibbett, Molecular systematics and biological diversification of Boletales. *Mycologia* **98**, 971 (2006). [doi:10.3852/mycologia.98.6.971](https://doi.org/10.3852/mycologia.98.6.971) [Medline](#)
13. D. Martinez *et al.*, 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 (2009). [doi:10.1073/pnas.0809575106](https://doi.org/10.1073/pnas.0809575106) [Medline](#)
14. F. Martin *et al.*, Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* **464**, 1033 (2010). [doi:10.1038/nature08867](https://doi.org/10.1038/nature08867) [Medline](#)
15. F. Martin *et al.*, The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* **452**, 88 (2008). [doi:10.1038/nature06556](https://doi.org/10.1038/nature06556) [Medline](#)

16. D. Martinez *et al.*, Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nat. Biotechnol.* **22**, 695 (2004). [doi:10.1038/nbt967](https://doi.org/10.1038/nbt967) [Medline](#)
17. M. Binder, K. H. Larsson, P. B. Matheny, D. S. Hibbett, Amylocorticiales ord. nov. and Jaapiales ord. nov.: early diverging clades of agaricomycetidae dominated by corticioid forms. *Mycologia* **102**, 865 (2010). [doi:10.3852/09-288](https://doi.org/10.3852/09-288) [Medline](#)
18. A. J. Eckert, B. D. Hall, Phylogeny, historical biogeography, and patterns of diversification for Pinus (Pinaceae): phylogenetic tests of fossil-based hypotheses. *Mol. Phylogenet. Evol.* **40**, 166 (2006). [doi:10.1016/j.ympev.2006.03.009](https://doi.org/10.1016/j.ympev.2006.03.009) [Medline](#)
19. H. Kauserud *et al.*, Asian origin and rapid global spread of the destructive dry rot fungus *Serpula lacrymans*. *Mol. Ecol.* **16**, 3350 (2007). [doi:10.1111/j.1365-294X.2007.03387.x](https://doi.org/10.1111/j.1365-294X.2007.03387.x) [Medline](#)
20. O. Schmidt, Molecular methods for the characterization and identification of the dry rot fungus *Serpula lacrymans*. *Holzforschung* **54**, 221 (2000). [doi:10.1515/HF.2000.038](https://doi.org/10.1515/HF.2000.038)
21. Materials and methods are available as supporting material on *Science Online*.
22. B. L. Cantarel *et al.*, The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* **37**, (Database issue), D233 (2009). [doi:10.1093/nar/gkn663](https://doi.org/10.1093/nar/gkn663) [Medline](#)
23. G. Vaaje-Kolstad *et al.*, An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science* **330**, 219 (2010). [doi:10.1126/science.1192231](https://doi.org/10.1126/science.1192231) [Medline](#)
24. P. Schneider, S. Bouhired, D. Hoffmeister, Characterization of the atromentin biosynthesis genes and enzymes in the homobasidiomycete *Tapinella panuoides*. *Fungal Genet. Biol.* **45**, 1487 (2008). [doi:10.1016/j.fgb.2008.08.009](https://doi.org/10.1016/j.fgb.2008.08.009) [Medline](#)
25. M. Tlalka, M. Fricker, S. Watkinson, Imaging of long-distance α -aminoisobutyric acid translocation dynamics during resource capture by *Serpula lacrymans*. *Appl. Environ. Microbiol.* **74**, 2700 (2008). [doi:10.1128/AEM.02765-07](https://doi.org/10.1128/AEM.02765-07) [Medline](#)
26. A. Vanden Wymelenberg *et al.*, Comparative transcriptome and secretome analysis of wood decay fungi *Postia placenta* and *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* **76**, 3599 (2010). [doi:10.1128/AEM.00058-10](https://doi.org/10.1128/AEM.00058-10) [Medline](#)
27. T. Nilsson, J. Ginns, Cellulolytic activity and the taxonomic position of selected brown-rot fungi. *Mycologia* **71**, 170 (1979). [doi:10.2307/3759230](https://doi.org/10.2307/3759230)
28. B. O. Lindahl, A. F. S. Taylor, R. D. Finlay, Defining nutritional constraints on carbon cycling in boreal forests - Towards a less 'phytcentric' perspective. *Plant Soil* **242**, 123 (2002). [doi:10.1023/A:1019650226585](https://doi.org/10.1023/A:1019650226585)
29. R. Vasiliauskas, A. Menkis, R. D. Finlay, J. Stenlid, Wood-decay fungi in fine living roots of conifer seedlings. *New Phytol.* **174**, 441 (2007). [doi:10.1111/j.1469-8137.2007.02014.x](https://doi.org/10.1111/j.1469-8137.2007.02014.x) [Medline](#)
30. J. Sambrook, D. W. Russell, 3rd Edn (Cold Spring Harbor Laboratory Press, NY, 2001).

31. D. B. Jaffe *et al.*, Whole-genome sequence assembly for mammalian genomes: Arachne 2. *Genome Res.* **13**, 91 (2003). [doi:10.1101/gr.828403](https://doi.org/10.1101/gr.828403) [Medline](#)
32. S. Kurtz *et al.*, Versatile and open software for comparing large genomes. *Genome Biol.* **5**, R12 (2004). [doi:10.1186/gb-2004-5-2-r12](https://doi.org/10.1186/gb-2004-5-2-r12) [Medline](#)
33. X. Huang, A. Madan, CAP3: A DNA sequence assembly program. *Genome Res.* **9**, 868 (1999). [doi:10.1101/gr.9.9.868](https://doi.org/10.1101/gr.9.9.868) [Medline](#)
34. A. F. A. Smit, R. Hubley, P. Green, *RepeatMasker Open-3.0*. 1996-2010.
35. J. Jurka *et al.*, *Genome Res.* **110**, 462 (2005).
36. T. M. Lowe, S. R. Eddy, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955 (1997). [doi:10.1093/nar/25.5.955](https://doi.org/10.1093/nar/25.5.955) [Medline](#)
37. A. A. Salamov, V. V. Solovyev, Ab initio gene finding in Drosophila genomic DNA. *Genome Res.* **10**, 516 (2000). [doi:10.1101/gr.10.4.516](https://doi.org/10.1101/gr.10.4.516) [Medline](#)
38. K. Isono, J. D. McIninch, M. Borodovsky, Characteristic features of the nucleotide sequences of yeast mitochondrial ribosomal protein genes as analyzed by computer program GeneMark. *DNA Res.* **1**, 263 (1994). [doi:10.1093/dnares/1.6.263](https://doi.org/10.1093/dnares/1.6.263) [Medline](#)
39. E. Birne, R. Durbin, *Genome Res.* **10**, 547 (2000). [Medline](#) [doi:10.1101/gr.10.4.547](https://doi.org/10.1101/gr.10.4.547)
40. W. J. Kent, BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656 (2002). [Medline](#)
41. H. Nielsen, J. Engelbrecht, S. Brunak, G. von Heijne, Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**, 1 (1997). [doi:10.1093/protein/10.1.1](https://doi.org/10.1093/protein/10.1.1) [Medline](#)
42. K. Melén, A. Krogh, G. von Heijne, Reliability measures for membrane protein topology prediction algorithms. *J. Mol. Biol.* **327**, 735 (2003). [doi:10.1016/S0022-2836\(03\)00182-7](https://doi.org/10.1016/S0022-2836(03)00182-7) [Medline](#)
43. E. M. Zdobnov, R. Apweiler, InterProScan—an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* **17**, 847 (2001). [doi:10.1093/bioinformatics/17.9.847](https://doi.org/10.1093/bioinformatics/17.9.847) [Medline](#)
44. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403 (1990). [Medline](#)
45. M. Kanehisa *et al.*, KEGG for linking genomes to life and the environment. *Nucleic Acids Res.* **36**, (Database issue), D480 (2008). [doi:10.1093/nar/gkm882](https://doi.org/10.1093/nar/gkm882) [Medline](#)
46. E. V. Koonin *et al.*, A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol.* **5**, R7 (2004). [doi:10.1186/gb-2004-5-2-r7](https://doi.org/10.1186/gb-2004-5-2-r7) [Medline](#)
47. A. J. Enright, S. Van Dongen, C. A. Ouzounis, An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* **30**, 1575 (2002). [doi:10.1093/nar/30.7.1575](https://doi.org/10.1093/nar/30.7.1575) [Medline](#)

48. T. De Bie, N. Cristianini, J. P. Demuth, M. W. Hahn, CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**, 1269 (2006). [doi:10.1093/bioinformatics/btl097](https://doi.org/10.1093/bioinformatics/btl097) [Medline](#)
49. D. R. Maddison, W. P. Maddison, MacClade 4 Sunderland, Massachusetts: Sinauer Associates (2005).
50. A. J. Drummond, A. Rambaut, BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007). [doi:10.1186/1471-2148-7-214](https://doi.org/10.1186/1471-2148-7-214) [Medline](#)
51. M. Binder *et al.*, The phylogenetic distribution of resupinate forms across the major clades of mushroom-forming fungi (Homobasidiomycetes). *Syst. Biodivers.* **3**, 113 (2005). [doi:10.1017/S1477200005001623](https://doi.org/10.1017/S1477200005001623)
52. P. B. Matheny *et al.*, Major clades of Agaricales: a multilocus phylogenetic overview. *Mycologia* **98**, 982 (2006). [doi:10.3852/mycologia.98.6.982](https://doi.org/10.3852/mycologia.98.6.982) [Medline](#)
53. D. S. Hibbett, D. Grimaldi, M. J. Donoghue, Fossil mushrooms from Miocene and Cretaceous ambers and the evolution of Homobasidiomycetes. *Am. J. Bot.* **84**, 981 (1997). [doi:10.2307/2446289](https://doi.org/10.2307/2446289) [Medline](#)
54. B. A. LePage, R. S. Currah, R. A. Stockey, G. W. Rothwell, Fossil Ectomycorrhizae from the Middle Eocene. *Am. J. Bot.* **84**, 470 (1997). [doi:10.2307/2446014](https://doi.org/10.2307/2446014)
55. A. Rambaut, A. J. Drummond, Tracer v1.4 <http://beast.bio.ed.ac.uk/Tracer> (2007).
56. K. Katoh, K. Kuma, H. Toh, T. Miyata, MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511 (2005). [doi:10.1093/nar/gki198](https://doi.org/10.1093/nar/gki198) [Medline](#)
57. F. Abascal, R. Zardoya, D. Posada, ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**, 2104 (2005). [doi:10.1093/bioinformatics/bti263](https://doi.org/10.1093/bioinformatics/bti263) [Medline](#)
58. A. Stamatakis, P. Hoover, J. Rougemont, A Rapid Bootstrap Algorithm for the RAxML Web Servers. *Syst. Biol.* **75**, 758 (2008). [doi:10.1080/10635150802429642](https://doi.org/10.1080/10635150802429642)
59. T. Y. James *et al.*, Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443**, 818 (2006). [doi:10.1038/nature05110](https://doi.org/10.1038/nature05110) [Medline](#)
60. D. Durand, B. V. Halldórsson, B. Vernot, A hybrid micro-macroevolutionary approach to gene tree reconstruction. *J. Comput. Biol.* **13**, 320 (2006). [doi:10.1089/cmb.2006.13.320](https://doi.org/10.1089/cmb.2006.13.320) [Medline](#)
61. B. Vernot, M. Stolzer, A. Goldman, D. Durand, Reconciliation with non-binary species trees. *J. Comput. Biol.* **15**, 981 (2008). [doi:10.1089/cmb.2008.0092](https://doi.org/10.1089/cmb.2008.0092) [Medline](#)
62. Y. Benjamini, Y. Hochberg, *J. R. Statist. Soc. B* **57**, 289 (1995).
63. A. Bensadoun, D. Weinstein, Assay of proteins in the presence of interfering materials. *Anal. Biochem.* **70**, 241 (1976). [doi:10.1016/S0003-2697\(76\)80064-4](https://doi.org/10.1016/S0003-2697(76)80064-4) [Medline](#)
64. H. Y. Yeang, F. Yusof, L. Abdullah, Protein purification for the Lowry assay: acid precipitation of proteins in the presence of sodium dodecyl sulfate and other biological detergents. *Anal. Biochem.* **265**, 381 (1998). [doi:10.1006/abio.1998.2893](https://doi.org/10.1006/abio.1998.2893) [Medline](#)

65. D. Fagner, M. Zomorodi, U. Kües, A. Majcherczyk, Optimized protocol for the 2-DE of extracellular proteins from higher basidiomycetes inhabiting lignocellulose. *Electrophoresis* **30**, 2431 (2009). [doi:10.1002/elps.200800770](https://doi.org/10.1002/elps.200800770) [Medline](#)
66. B. J. Cargile, J. L. Bundy, T. W. Freeman, J. L. Stephenson, Jr., Gel based isoelectric focusing of peptides and the utility of isoelectric point in protein identification. *J. Proteome Res.* **3**, 112 (2004). [doi:10.1021/pr0340431](https://doi.org/10.1021/pr0340431) [Medline](#)
67. J. Krijgsveld, S. Gauci, W. Dormeyer, A. J. Heck, In-gel isoelectric focusing of peptides as a tool for improved protein identification. *J. Proteome Res.* **5**, 1721 (2006). [doi:10.1021/pr0601180](https://doi.org/10.1021/pr0601180) [Medline](#)
68. J. Havlis, H. Thomas, M. Sebela, A. Shevchenko, Fast-response proteomics by accelerated in-gel digestion of proteins. *Anal. Chem.* **75**, 1300 (2003). [doi:10.1021/ac026136s](https://doi.org/10.1021/ac026136s) [Medline](#)
69. J. Rappsilber, Y. Ishihama, M. Mann, Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* **75**, 663 (2003). [doi:10.1021/ac026117i](https://doi.org/10.1021/ac026117i) [Medline](#)
70. C. L. Chepanoske, B. E. Richardson, M. von Rechenberg, J. M. Peltier, Average peptide score: a useful parameter for identification of proteins derived from database searches of liquid chromatography/tandem mass spectrometry data. *Rapid Commun. Mass Spectrom.* **19**, 9 (2005). [doi:10.1002/rcm.1741](https://doi.org/10.1002/rcm.1741) [Medline](#)
71. I. Shadforth, T. Dunkley, K. Lilley, D. Crowther, C. Bessant, Confident protein identification using the average peptide score method coupled with search-specific, ab initio thresholds. *Rapid Commun. Mass Spectrom.* **19**, 3363 (2005). [doi:10.1002/rcm.2203](https://doi.org/10.1002/rcm.2203) [Medline](#)
72. M. Gill, W. Steglich, *Prog. Chem. Org. Nat. Prod.* **51**, 1 (1987).
73. L. L. Stookey, Ferrozine---a new spectrophotometric reagent for iron. *Anal. Chem.* **42**, 779 (1979). [doi:10.1021/ac60289a016](https://doi.org/10.1021/ac60289a016)
74. D. W. Malloch, K. A. Pirozynski, P. H. Raven, Ecological and evolutionary significance of mycorrhizal symbioses in vascular plants (A Review). *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2113 (1980). [doi:10.1073/pnas.77.4.2113](https://doi.org/10.1073/pnas.77.4.2113) [Medline](#)
75. J. W. G. Cairney, Evolution of mycorrhiza systems. *Naturwissenschaften* **87**, 467 (2000). [doi:10.1007/s001140050762](https://doi.org/10.1007/s001140050762) [Medline](#)
76. X.-Q. Wang, D. C. Tank, T. Sang, Phylogeny and divergence times in Pinaceae: evidence from three genomes. *Mol. Biol. Evol.* **17**, 773 (2000). [Medline](#)
77. D. S. Hibbett, P. B. Matheny, The relative ages of ectomycorrhizal mushrooms and their plant hosts estimated using Bayesian relaxed molecular clock analyses. *BMC Biol.* **7**, 13 (2009). [doi:10.1186/1741-7007-7-13](https://doi.org/10.1186/1741-7007-7-13) [Medline](#)
78. A. M. Bresinsky, M. Jarosch, M. Fischer, I. Schönberger, B. Wittmann-Bresinsky, Phylogenetic Relationships within Paxillus s. l. (Basidiomycetes, Boletales): Separation of a Southern Hemisphere Genus. *Plant Biol.* **1**, 327 (1999). [doi:10.1111/j.1438-8677.1999.tb00260.x](https://doi.org/10.1111/j.1438-8677.1999.tb00260.x)

79. F. Florindo, A. K. Cooper, P. E. O'Brien, Introduction to 'Antarctic Cenozoic palaeoenvironments: geologic record and models'. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **198**, 1 (2003). [doi:10.1016/S0031-0182\(03\)00405-X](https://doi.org/10.1016/S0031-0182(03)00405-X)
80. L. A. Lawyer, L. M. Gahagan, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **198**, 11 (2003). [doi:10.1016/S0031-0182\(03\)00392-4](https://doi.org/10.1016/S0031-0182(03)00392-4)
81. L. G. Cook, M. D. Crisp, Not so ancient: the extant crown group of Nothofagus represents a post-Gondwanan radiation. *Proc. Biol. Sci.* **272**, 2535 (2005). [doi:10.1098/rspb.2005.3219](https://doi.org/10.1098/rspb.2005.3219) [Medline](#)
82. R. A. Ohm *et al.*, Genome sequence of the model mushroom *Schizophyllum commune*. *Nat. Biotechnol.* **28**, 957 (2010). [doi:10.1038/nbt.1643](https://doi.org/10.1038/nbt.1643) [Medline](#)
83. D. R. Schmidhalter, G. Canevascini, Purification and characterization of two exo-cellobiohydrolases from the brown-rot fungus *Coniophora puteana* (Schum ex Fr) Karst. *Arch. Biochem. Biophys.* **300**, 551 (1993). [doi:10.1006/abbi.1993.1076](https://doi.org/10.1006/abbi.1993.1076) [Medline](#)
84. T. Kajisa, K. Igarashi, M. Samejima, The genes encoding glycoside hydrolase family 6 and 7 cellulases from the brown-rot fungus *Coniophora puteana*. *J. Wood Sci.* **55**, 376 (2009). [doi:10.1007/s10086-009-1042-4](https://doi.org/10.1007/s10086-009-1042-4)
85. S. Nagendran, H. E. Hallen-Adams, J. M. Paper, N. Aslam, J. D. Walton, Reduced genomic potential for secreted plant cell-wall-degrading enzymes in the ectomycorrhizal fungus *Amanita bisporigera*, based on the secretome of *Trichoderma reesei*. *Fungal Genet. Biol.* **46**, 427 (2009). [doi:10.1016/j.fgb.2009.02.001](https://doi.org/10.1016/j.fgb.2009.02.001) [Medline](#)
86. A. Bresinsky, *Z. Naturforsch.* **28c**, 627 (1973).
87. P. Aqueveque, T. Anke, O. Sterner, *Z. Naturforsch.* **57**, 257 (2002).
88. I. W. Farrell, J. W. Keeping, M. G. Pellatt, V. Thaller, Natural acetylenes. Part XLI. Polyacetylenes from fungal fruiting bodies. *J. Chem. Soc. Perkin Trans.* **22**, 2642 (1973). [doi:10.1039/p19730002642](https://doi.org/10.1039/p19730002642)
89. D. Hoffmeister, N. P. Keller, Natural products of filamentous fungi: enzymes, genes, and their regulation. *Nat. Prod. Rep.* **24**, 393 (2007). [doi:10.1039/b603084j](https://doi.org/10.1039/b603084j) [Medline](#)

Acknowledgements: Drs J. Schilling, U Minnesota and Dr D. Barbara, U Warwick critically reviewed the manuscript, Tony Marks designed graphics, B Wackler and M. Zomorodi gave technical assistance. Assembly and annotations of *S.lacrymans* genomes are available at <http://www.jgi.doe.gov/Serpula> and DDBJ/EMBL/GenBank, accessions AECQB000000000 and AEq. C000000000. Complete microarray expression dataset is available at the Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>) accession GSE27839. The work conducted by the U.S. Department of Energy Joint Genome Institute and supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Further financial support is acknowledged as supporting material on *Science* online.