LA-UR- 08-7047

Approved for public release; distribution is unlimited.

Title:	Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion.
Author(s):	Diego Martinez, Jean Challacombe, Ingo Morgenstern, David Hibbett, Monika Schmoll, Christian P. Kubicek, Patricia Ferreira, Francisco J. Ruiz-Duenas, Angel T. Martinez, Phil Kersten, Kenneth E. Hammelf Amber Vanden Wymelenberg, Jill Gaskell, Erika Lindquist, Grzegorz Sabat, Sandra Splinter BonDurant, Luis F. Larrondo, Paulo Canessa, Rafael Vicuna,
Intended for:	Proc. Natl. Acad. Sci. USA.

Los Alamos
NATIONAL LABORATORY
EST 1943

Form 836 (7/06)

1. Authors continued

Last	First	Middle	Z No.	Group or Affiliation
Schmoll	Monika		NA	Universitat Wien
Kubicek	Christian	P	NA	Universitat Wien
Ferreira	Patricia		NA	CIB, CSIC, Madrid
Ruiz-Duenase	Francisco	J	NA	CIB, CSIC, Madrid
Martinez	Angel	T	NA	CIB, CSIC, Madrid
Kersten	Phil		NA	Forest Products Lab
Hammel	Kenneth	E	NA	Forest Products Lab
Vanden Wymelenberg	Amber		NA	U. Wisconsin
Gaskell	Jill		NA	Forest Products Lab
Lindquist	Erika		NA	DOE JGI
Sabati	Grzegorz		NA	U. Wisconsin
BonDurant	Sandra	S	NA	U. Wisconsin
Larrondo	Luis	F	NA	U. Catholica de Chile
Canessa	Paulo		NA	U. Catholica de Chile
Vicuna	Rafael		NA	U. Catholica de Chile
Yadavk	Jagjit		NA	U. Cincinatti
Doddapaneni	Harshavardhan		NA	U. Cincinatti
Subramaniank	Venkataramanan		NA	U. Cincinatti
Pisabarro	Antonio	G	NA	Public U. Navarre
Lavín	José	L	NA	Public U. Navarre
Oguiza	José	A	NA	Public U. Navarre
Master	Emma		NA	U. Toronto
Henrissat	Bernard		NA	CNRS, Marseille
Coutinho	Pedro	M	NA	CNRS, Marseille
Harris	Paul		NA	Novozymes, Inc.,
Magnuson	Jon	K	NA	PNNL
Baker	Scott		NA	PNNL
Bruno	Kenneth		NA	PNNL
Kenealy	William		NA	Mascoma, Inc.
Hoegger	Patrik	J	NA	Georg-August-U.
Kües	Ursula		NA	Georg-August-U.
Ramaiya	Preethi		NA	Novozymes, Inc.
Lucas	Susan		NA	DOE JGI
Salamov	Asaf		NA	DOE JGI
Shapiro	Harris		NA	DOE JGI
Tuh	Hank		NA	DOE JGI
Chee	Christine	L	NA	UNM
Misra	Monica		120897	B-6
Xie	Gary		147760	B-6
Teter	Sarah		NA	Novozymes, Inc.
Yaver	Debbie		NA	Novozymes, Inc.
James	Tim		NA	McMaster U.
Mokrejs	Martin		NA	Charles U.
Pospisek	Martin		NA	Charles U.
Grigoriev	Igor		NA	DOE JGI
Brettin	Thomas		120593	B-6
Rokhsar	Dan		NA	DOE JGI
Berka	Randy		NA	Novozymes
Cullen	Dan		NA	Forest Products Lab

Classification: Biological Sciences, Microbiology

Genome, transcriptome, and secretome analysis of wood decay fungus *Postia placenta* supports unique mechanisms of lignocellulose conversion

Diego Martinez[a,b], Jean Challacombe[a], Ingo Morgenstern[c], David Hibbett[c], Monika Schmoll[d], Christian P. Kubicek[d], Patricia Ferreira[e], Francisco J. Ruiz-Duenas[e], Angel T. Martinez[e], Phil Kersten[f], Kenneth E. Hammel[f], Amber Vanden Wymelenberg[g], Jill Gaskell[f], Erika Lindquist[h], Grzegorz Sabat[i], Sandra Splinter BonDurant[i], Luis F. Larrondo[j], Paulo Canessa[j], Rafael Vicuna[j], Jagjit Yadav[k], Harshavardhan Doddapaneni[k], Venkataramanan Subramanian[k], Antonio G. Pisabarro[l], José L. Lavín[l], José A. Oguiza[l], Emma Master[m], Bernard Henrissat[n], Pedro M. Coutinho[n], Paul Harris[o], Jon Karl Magnuson[p], Scott Baker[p], Kenneth Bruno[p], William Kenealy[q], Patrik J. Hoegger[r], Ursula Kües[r], Preethi Ramaiya[o], Susan Lucas[h], Asaf Salamov[h], Harris Shapiro[h], Hank Tu[h], Christine L. Chee[b], Monica Misra[a], Gary Xie[a], Sarah Teter[o], Debbie Yaver[o], Tim James[s], Martin Mokrejs[t], Martin Pospisek[t], Igor Grigoriev[h], Thomas Brettin[a], Dan Rokhsar[h], Randy Berka[o] and Dan Cullen[f,u]

[a]Los Alamos National Laboratory/Joint Genome Institute, PO Box 1663, Los Alamos, New Mexico 87545; [b]University of New Mexico, Albuquerque, NM 87131; [c]Biology Department, Clark University, Worcester, MA 01610; [d]Research Area Gene Technology and Applied Biochemistry, Institute of Chemical Engineering, Technische Universitat Wien, Getreidemarkt 9/166, A-1060 Vienna, Austria; [e]CIB, CSIC, Ramiro de Maeztu 9, E-28040, Madrid, Spain; [f]Forest Products Laboratory, Madison, WI 53726; [g]Department of Bacteriology, University of Wisconsin, Madison, WI 53706; [h]DOE's Joint Genome Institute, 2800 Mitchell Avenue, Walnut Creek, California 94598; [i]University of Wisconsin Biotechnology Center, Madison, WI 53706; [j]Departamento de Genética Molecular y Microbiología, Facultad de Ciencias Biológicas, Pontifica Universidad Católica de Chile and Millennium Institute for Fundamental and Applied Biology, Santiago, Chile; [k]Department of Environmental Health, University of Cincinnati, Cincinnati, Ohio 45267; [l]Genetics and Microbiology Research Group, Public University of Navarre, 31006 Pamplona, Spain; [m]Chemical Engineering, University of Toronto, Toronto, Ontario, Canada; [n]Architecture et Fonction des Macromolécules Biologiques, UMR 6098, CNRS, Universités d'Aix-Marseille I & II, Case 932, 163 Avenue de Luminy, 13288 Marseille, France; [o]Novozymes Inc, 1445 Drew Avenue, Davis, CA 95618; [p]Pacific Northwest National Laboratory, P.O. Box 999, Richland, WA 99352; [q]Mascoma Inc, Lebanon, NH 03766; [r]Molecular Wood Biotechnology and Technical Mycology, Büsgen-Institute, Georg-August-University Göttingen, Germany; [s]Department of Biology, McMaster University, Hamilton, Ontario, Canada and [t]Department of Genetics and Microbiology, Charles University, Prague, Czech Republic.

Author contributions: D.M., J.C., I.G., E.L., S.L., A.S., H.S., H.T., C.L.C., M.M., G.X., T.B., and D.R. produced genome assembly and automated annotations; D.M., D.H, I.M., C.P.K., E.M., B.H., P.M.C., P.H., S.T., D.Y., and R.B. analyzed carbohydrate active enzymes; D.H., I.M., P.F., F.J.R., A.T.M., P.K., K.E.H., L.L., P.C., R.V., P.H., U.K. and D.C. analyzed oxidoreductases; M.S. and U.K. analyzed signal transduction pathways

and mating type; J.Y., H.D. and VS analyzed cytochrome P450s; A.G.P., J.L.L. and J.A.O. analyzed oxidative phosphorylation; J.K.M, W.K. and D.C. analyzed glyoxylate pathway; L.L., and P.C, analyzed iron homeostasis and stress response; S.B., K.B., P.R., T.J., M.M., and M.P. analyzed other gene sets; A.V.W., J.G., G.S., and D.C. acquired and analyzed mass spectrometry data; D.M., A.V.W., J.G., E.L., S.S.B. and D.C. acquired and analyzed gene expression data; D.M., K.E.H., P.K., R.B. and D.C. integrated analyses and wrote the paper.

[u]To whom correspondence should be addressed. Email: dcullen@wisc.edu, FAX: 608-231-9468.

Manuscript Information: Six pages, three figures

Database deposition: The annotated genome is available on an interactive web portal at http://www.jgi.doe.gov/whiterot. Genome and EST sequence have been deposited and assigned GenBank Accession numbers ABWF00000000 and FL595400-FL633513, respectively.

Brown-rot fungi such as *Postia placenta* are common inhabitants of forest ecosystems and are also largely responsible for the destructive decay of wooden structures. Rapid depolymerization of cellulose is a distinguishing feature of brown-rot, but the biochemical mechanisms and underlying genetics are poorly understood. Systematic examination of the *P. placenta* genome, transcriptome and secretome revealed unique extracellular enzyme systems, including an unusual repertoire of extracellular glycoside hydrolases. Genes encoding exocellobiohydrolases and cellulose-binding domains, typical of cellulolytic microbes, are absent in this efficient cellulose-degrading fungus. When *P. placenta* was grown in medium containing cellulose as sole carbon source, transcripts corresponding to many hemicellulases and to a single putative β-1-4 endoglucanase were expressed at high levels relative to glucose grown cultures. These transcript profiles were confirmed by direct identification of peptides by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Also upregulated during growth on cellulose medium were putative iron reductases, quinone reductase, and structurally divergent oxidases potentially involved in extracellular generation of Fe(II) and H_2O_2. These observations are consistent with a biodegradative role for Fenton chemistry in which Fe(II) and H_2O_2 react to form hydroxyl radicals, highly reactive oxidants capable of depolymerizing cellulose. The *P. placenta* genome resources provide unparalleled opportunities for investigating such unusual mechanisms of cellulose conversion. More broadly, the genome offers insight into the diversification of lignocellulose degrading mechanisms in fungi. Comparisons to the closely related white-rot fungus *Phanerochaete chrysosporium* support an evolutionary shift from white-rot to brown-rot during which the capacity for efficient depolymerization of lignin was lost.

\body Lignocellulose in vascular plant cell walls is one of the largest sinks for fixed global carbon and is increasingly eyed as a potential feedstock in biofuels and new biomaterials portfolios (1). Relatively few organisms can efficiently convert the recalcitrant polymer blend in lignocellulose to monomeric components (2). The principal exceptions are basidiomycetes, which attack wood through two main decay types called white-rot and brown-rot. Wood-decaying basidiomycetes are essential contributors to carbon cycling in forest soils, and brown-rot fungi are additionally important because they are a major cause of failure in wooden structures. White-rot fungi degrade all components of plant cell walls, including cellulose, hemicellulose and lignin. Although they cannot grow on lignin alone, they have the unique ability to degrade a large proportion of it completely to CO_2 and H_2O. This biodegradative strategy exposes the structural polysaccharides of plant cell walls, thus making them susceptible to hydrolysis by cellulases and hemicellulases. Brown-rot fungi employ a different approach; although they modify lignin extensively, the products remain *in situ* as a polymeric residue (3, 4). Given the incomplete ligninolysis that occurs during brown-rot, it remains unclear how these fungi gain access to plant cell wall polysaccharides. However, it seems probable that the two decay types share at least some mechanisms, because molecular phylogeny, morphological considerations, and substrate preference suggest that brown-rot fungi have repeatedly evolved from white-rot fungi (5). Indeed, the two major experimental

organisms for studies of brown-rot, *Postia placenta* and *Gloeophyllum trabeum*, are distantly related species that represent independent origins of brown-rot (5). Any similarities in their decay mechanisms must represent either general mechanisms of wood decay common to white-rot and brown-rot species, or convergently-evolved brown-rot mechanisms. Moreover, *P. placenta* is closely related to the model white-rot fungus, *Phanerochaete chrysosporium*, so comparisons between these species may provide insight into the mechanistic basis of transitions from white-rot to brown-rot.

White-rot fungi produce complex ligninolytic systems that are thought to depend in part on extracellular oxidative enzymes, especially peroxidases, laccases and other oxidases. It remains an open question whether brown-rot fungi possess any of these ligninolytic components. White-rot fungi also secrete complete, synergistically acting cellulase systems that include both endo- and exo-acting enzymes. These exocellobiohydrolases and endoglucanases often share architectures that include separate catalytic and cellulose-binding domains. In contrast, relatively few cellulases have been described in brown-rot fungi (6), and it has been long recognized that rapid depolymerization of cellulose appears to occur before the substrate porosity has increased enough to admit cellulases (7). One possibility is that brown-rot fungi attack cellulose with low molecular weight oxidants that act in conjunction with a limited set of relatively small cellulases.

The hydroxyl free radical, generated via Fenton chemistry ($H_2O_2 + Fe^{2+} + H^+ \rightarrow H_2O + Fe^{3+} + \cdot OH$), has long been implicated as one of the small oxidants that contributes to polysaccharide depolymerization during brown-rot. Current models for hydroxyl radical participation have been reviewed (6), and typically involve generation of this highly reactive oxidant at or near the substrate. Key requirements for Fenton systems include mechanisms for extracellular H_2O_2 generation and for reduction of Fe^{3+} to Fe^{2+}, which might be accomplished by extracellular fungal metabolites such as hydroquinones, or by extracellular enzymes such as cellobiose dehydrogenase.

We report here analyses of the *P. placenta* draft genome together with transcript profiles and mass spectrometric identification of extracellular proteins. Consistent with a unique strategy for cellulose degradation, we observed a dramatic absence of conventional cellulase genes and most class II fungal peroxidases, and a rich diversity of genes potentially supporting generation of extracellular reactive oxygen species.

Results

Carbohydrate active enzymes. Given the well known efficiency with which brown-rot fungi rapidly depolymerize and degrade cellulose, the *P. placenta* genome revealed remarkably few, if any, conventional cellulases. Of 17,173 proteins predicted in the dikaryotic genome, 242 unique genes encode potential carbohydrate-active enzymes ((8); http://www.cazy.org), of which 228 (94%) have at least one potential ortholog (BLASTP bit score $\geq$100) in *P. chrysosporium*. These putative CAZY genes include 144 glycoside hydrolases (GH), 10 carbohydrate esterases (CE), 75 glycosyltransferases (GT), 7 expansin-like proteins (EXPN), and 6 polysaccharide lyases (PL) (complete CAZY list in SI Table 1 within NCBI GEO accession 12540. In distinct contrast to all cellulolytic fungal aerobes, exocellobiohydrolases CBH2 (GH6) and CBH1 (GH7), as well as cellulose-binding endoglucanases are missing in the *P. placenta* genome (Fig.1). Also

absent are family 1 carbohydrate binding modules (CBM1). These highly conserved cellulose-binding domains are fused to functionally diverse CAZYs in a wide range of cellulolytic microbes. Surprisingly then, the repertoire of recognizable cellulolytic enzymes in *P. placenta* appears limited to just two potential endoglucanases (1,4-β-glucanases) and several β-glucosidases. In contrast to cellulolytic saprophytes (e.g. *Trichoderma reesei*, *Aspergillus spp.* or *Neurospora crassa*) and aggressive plant pathogens (e.g. *Fusarium graminearum* or *Magnaporthe grisea*), the overall number and distribution of GHs in *P. placenta* are similar to those in the ectomycorrhizal symbiont *Laccaria bicolor*, the human pathogen *Cryptococcus neoformans* and the biotrophic plant pathogen *Ustilago maydis* (SI Table 2). Phylogenetic analyses of *P. placenta* and *P. chrysosporium* genomes indicate that the transition from white-rot to brown-rot has been associated with multiple independent reductions including the GH families 6, 7, 10, 11 and 61 (Figs. 1 and 2; SI Table 2) Thus, the transition from white-rot to brown-rot has been associated with multiple independent reductions in the GH families.

Microarrays representing 12,438 unique alleles were used to examine *P. placenta* transcript levels in basal salts medium containing either glucose or wood-derived microcrystalline cellulose as the sole carbon sources. In total, 290 gene models showed >2 fold transcript accumulation, and of these, 235 increased in cellulose medium and 35 increased in glucose medium (NCBI GEO accession 12540 SI Table 3). Transcripts of 99 GH-encoding genes significantly increased ($P<0.01$) in the cellulose medium, and of these, 18 increased >2 fold (Fig. 3). Twenty-one GH transcripts significantly increased in the glucose-containing medium, but none exceeded a two-fold change. In addition, shotgun liquid chromatography coupled tandem mass spectrometry (LC-MS/MS) identified 26 specific CAZYs in the extracellular fluid of *P. placenta* grown in basal salts supplemented with ball-milled aspen wood, microcrystalline cellulose, or cotton (SI Table 4). The CAZY genes expressed in cellulose included laminarases, chitinases, and various hemicellulases (endoxylanases, β-xylosidases, L-α-arabinofuranosidases, endo-β-mannanases and β-mannosidases). It is unclear whether any of these enzymes could directly attack crystalline cellulose.

Extracellular H_2O_2 generation. Gene models potentially supporting Fenton chemistry through the generation of extracellular H_2O_2 include copper radical oxidases and GMC oxidoreductases (SI Table 5). Results summarized here focus on those genes with expression patterns that are consistent with a role in cellulose depolymerization, and detailed information for all genes is available in SI Table 6 within GEO 12540.

On the basis of overall sequence similarity to *P. chrysosporium* glyoxal oxidase (GLOX) and conservation of catalytic residues (9), three *P. placenta* models were identified as copper radical oxidases (CROs). GLOX is one of 7 CROs in *P. chrysosporium* and physiologically coupled to lignin peroxidase (LiP) via H_2O_2 generation. Of particular relevance to potential Fenton systems, CRO genes encoding proteins Ppl56703 and Ppl130305 are upregulated in microcrystalline cellulose, and Ppl56703 peptides were detected in aspen-grown cultures. The *P. chrysosporium cro1* and *glx1* genes are not closely related, and they do not have orthologs in *P. placenta,* which suggests that there have been two independent losses of these CRO lineages in *Postia* (Fig. 2). Ppl56703 is orthologous to the *cro3-4-5* lineage in *P. chrysosporium*, which therefore represents a *Phanerochaete*-specific expansion of the gene family. As in

the case of the GHs, evolution of brown-rot is associated with a reduced diversity of CROs.

Catalytically distinct from CROs, GMC oxidoreductases (InterPro IPR000172) included various alcohol and sugar oxidases. Among the former, *P. placenta* protein model Ppl118723 is similar to *G. trabeum* methanol oxidase (DQ835989) (> 85% amino acid identity over full length). Recent immunolocalization studies strongly implicate the *G. trabeum* alcohol oxidase as a source of H_2O_2 (10) to support Fenton chemistry. Suggesting a similar role in *P. placenta*, microarray analysis revealed high transcript levels and a sharp increase in transcription of the gene encoding Ppl118723 in cellulose-grown cultures relative to non-cellulolytic cultures. Comparatively high transcript levels in cellulose- and glucose-grown cultures were also observed for genes encoding Ppl128830 and Ppl108489, models tentatively identified as glucose-1-oxidases based on conserved key residues (11). Peptides corresponding to these putative *gox* genes were detected in extracellular filtrates (SI Tables 6 and 11 within GEO 12540). Aryl-alcohol oxidase, an extracellular GMC oxidoreductase cooperating with aryl-alcohol dehydrogenases for continuous peroxide supply in some white-rot fungi (11) does not seem to be involved in cellulose attack by *P. placenta* since the corresponding models were not or only slightly upregulated. Another extracellular GMC oxidoreductase, pyranose-2-oxidase, has been implicated in lignocellulose degradation in *P. chrysosporium* (12), but no orthologs were detected in *P. placenta*.

Iron reduction and homeostasis. Protein model Ppl124517 was identified as a putative quinone reductase (QRD). In the brown-rot fungus *G. trabeum*, a QRD may drive extracellular Fenton systems via redox cycling of secreted fungal quinones (6). Transcription of the *P. placenta* QRD gene was significantly upregulated in cellulose medium (GEO 12540 SI Table 3), which is consistent with a role for cellulolytic Fenton chemistry involving quinone redox-cycling. In this connection, upregulation of the genes encoding phenylalanine ammonia lyase (Ppl112824) and a putative quinate transporter (Ppl44553) may also be relevant by virtue of their respective roles in the biosynthesis and transport of essential quinones.

In addition to hydroquinone-based iron reduction systems, low molecular weight glycoproteins (GLPs) that can act as iron reductases have been hypothesized as components of extracellular Fenton systems in *G. trabeum* and *P. chrysosporium* (13). Four *P. placenta* models show significant similarity (>48% amino acid identity) to *P. chrysosporium glp1* and *glp2*, and the gene encoding Ppl128974 is significantly upregulated on microcrystalline cellulose medium (GEO 12540 SI Table 7). Sequence corresponding to another fungal protein implicated in Fe^{3+} reduction, CDH (6), appears to be absent in *P. placenta*.

In addition to its pivotal role in a wide range of cellular processes, iron homeostasis must play a central role in modulating a functioning Fenton system. The *P. placenta* genome features numerous genes potentially involved in iron transport and redox state (GEO 12540 SI Table 7). In addition to 7 ferric reductases, two iron permeases were identified one of which lies immediately downstream from a canonical yeast ferroxidase ortholog (Fet3). Transcripts of these adjacent genes were among the most highly upregulated in cellulose medium (GEO 12540 SI Table 3).

Modification of lignin and other aromatic compounds. Genes encoding the

class II secretory peroxidases lignin peroxidase (LiP), manganese peroxidase (MnP) and versatile peroxidase (VP) were not detected in the *P. placenta* genome (SI Table 5). Peroxidase model Ppl44056 lacks residues involved in Mn^{2+} binding and oxidation of aromatic compounds (14), and superimposition of protein models strongly suggests that Ppl44056 is a low redox potential peroxidase (SI Fig. 1). Consistent with this structural evidence, phylogenetic analyses of class II peroxidase genes from *Postia, Phanerochaete*, and other fungal genomes suggest that Ppl44056 is not closely related to LiP and MnP, but is part of an assemblage of "basal peroxidases" that includes the novel peroxidase (NoP) of *P. chrysosporium,* and peroxidases from *Coprinopsis cinerea* and *L. bicolor* (Fig. 2) (15). The backbone of the class II peroxidase phylogeny is not strongly supported, but analyses of broadly sampled datasets (15), suggest that the LiP and MnP gene lineages of *P. chrysosporium* were independently derived from the basal peroxidases prior to the divergence of *Postia* and *Phanerochaete*. If so, then the absence of LiP and MnP in *P. placenta* may reflect additional instances of gene loss.

Laccases have been suggested to play a role in lignin modification by white-rot fungi, but have not previously been demonstrated in brown-rot fungi. The precise role of these enzymes remains uncertain, but numerous studies have demonstrated laccase-catalyzed oxidation of phenolic and nonphenolic lignin model substrates particularly in the presence of low molecular weight mediators. The results from *P. placenta* belie the usual picture of brown-rot in that models Ppl62097 and Ppl111314 are likely laccases *sensu stricto* (16) (Fig.2). Transcript levels of the genes encoding Ppl89382 and Ppl111314 appear differentially regulated by decreasing slightly (-1.08-fold) and increasing (+2.29-fold), respectively, on cellulose medium relative to glucose medium (GEO 12540 SI Table 7). These enzymes could contribute to hydroxyl radical generation by oxidizing hydroquinones as described (17). Interestingly, laccase genes are absent from the genome of *P. chrysosporium* (18), suggesting that laccase (*sensu stricto*) is not a core component of fungal wood decay mechanisms, and is certainly not essential for white-rot.

Other upregulated genes potentially involved in quinone redox-cycling, and oxidation of lignin derived products include those encoding "polyphenol oxidase" (Ppl114245), i.e. tyrosinase or catechol oxidase related to typical laccases, and various oxidoreductases of uncertain function (Ppl107061, Ppl28683, Ppl34850, Ppl61437, Ppl24981) (SI Table 3).

Oxalate metabolism. In addition to pH effects on a wide range of enzymes, extracellular accumulation of oxalate by *P. placenta* may affect ferric iron availability and thereby impact hydroxyl radical formation (19; reviewed in ref. 6). A metabolic shunt between the citric acid and glyoxylate cycles is central to oxalic acid accumulation by the brown-rot fungus *Fomitopsis palustris* (20). Analysis of the *P. placenta* genome demonstrates a functional glyoxylate shunt and substantially extends our understanding of the number, structure, and transcription of key genes (SI Fig.2; SI text; GEO 12540 SI Table 8).

Cytochrome P450 monooxygenases. P450s have various roles in secondary metabolism and thought to be involved in biodegradation of lignin as well as various xenobiotic compounds. The *P. placenta* genome features an impressive set of 236 P450 genes (SI text, GEO 12540 SI Fig. 3), compared to 149 in *P. chrysosporium*, and expansions of

certain families (CYP64, CYP503, CYP5031 and CYP617) were observed. Genes encoding Ppl110015 (CYP53) and Ppl128850 (CYP503) were significantly upregulated in cellulose medium (GEO 12540 SI Table 3). The former is highly conserved in fungi and thought to catalyze benzoate hydroxylation.

Other. The genome was systematically examined for genes involved in oxidative phosphorylation, stress-related genes, signal transduction and regulatory genes, particularly those potentially controlling glycoside hydrolase expression and mating type (complete listings and analysis in SI text, GEO 12540 Figs 5-7).

Discussion

Analysis of the *P. placenta* genome elucidated a repertoire of genes and expression patterns distinct from those of other known cellulose-degrading microbes. The overall number of CAZY-encoding genes in *P. placenta*, 242, is not particularly low, and among these, the number of glycosyl transferases, 75, is fairly typical. However, the genome completely lacks cellulose-binding domains and the number of GHs is relatively low owing in part to the paucity of cellulases. No exocellobiohydrolases and only two potential β-1,4 endoglucanase genes were identified. One putative EG (Ppl115648) is expressed at relatively high levels and may correspond to previously characterized carboxymethylcellulose-active enzymes (21).

Comparisons with genomes of other cellulolytic microbes reveal a strikingly distinct set of glycoside hydrolases in the *P. placenta* genome. Among aerobes, only the cellulolytic gliding bacterium, *Cytophaga hutchinsonii* lacks exocellobiohydrolases and endoglucanases fused to cellulose-binding domains (22). The precise mechanism employed by *C. hutchinsonii* is somewhat mysterious, but it has been suggested that cellulose chains are peeled away from the polymer and transported into the periplasm (23). There, non-processive endoglucanases might readily degrade the cellulose. Such a mechanism seems unlikely in *P. placenta* because all evidence suggests that cellulose depolymerization by brown-rot fungi occurs at a distance from the advancing hyphae. In contrast, *C. hutchinsonii* is in direct contact with cellulose. Also unlike *C. hutchinsonii*, the *P. placenta* genome does not feature any of the GH9 cellulases often associated with processive endoglucanase activity.

Possibly, the CBM-less β-1-4-endoglucanase Ppl115648, which is clearly expressed in cellulose-containing media (Fig. 3), may possess processive activity that enables it to liberate the cellobiose that β-glucosidases then hydrolyze to assimilable glucose. Indeed, the accumulation of putative β-glucosidase transcripts and the corresponding proteins that we observed are consistent with the availability of cellobiose in our cultures. Precedents for crystalline cellulose hydrolysis by β-1,4-endoglucanases within GH family 5 have been reported (24, 25), but it seems unlikely that the Ppl115648 endoglucanase alone can account for the efficient cellulose depolymerization by *P. placenta*. Other GHs and/or hypothetical proteins, perhaps some of those expressed in microcrystalline cellulose cultures (Fig. 3; GEO 12540 SI Table 1), may be necessary for the complete breakdown of cellulose.

Many investigations of white-rot and brown-rot mechanisms have implicated the participation of low molecular weight oxidants, particularly Fenton-generated hydroxyl radicals. As recently reviewed (6), three somewhat overlapping mechanisms of oxidative degradation have been advanced. One view emphasizes the importance of CDH. In the case of *P. placenta*, CDH is absent. Another view invokes the role of low molecular weight glycopeptides that catalyze extracellular iron reduction. Initially identified in *P. chrysosporium* (13), potential orthologs of these glycopeptide-encoding genes were identified in *P. placenta*, and in one case, increased transcript levels were observed in cellulose medium. Accordingly, a role for these glycoproteins in a *P. placenta* Fenton system is possible. The third mechanism involves extracellular quinone redox cycling (26). Evidence supporting this system includes cellulose induction of genes encoding QRD, quinate transporter, phenylalanine ammonia lyase and laccase. However, the importance of hydroquinone-driven Fenton chemistry in *P. placenta* remains unclear because this fungus secretes high levels of oxalate (27), and Fe^{3+}-oxalate chelates are poorly reducible by hydroquinones (28).

The elevated expression in cellulose medium of Fet3 and Ftr1, components of the high affinity iron uptake system, may be at least partially explained by such chelates. While cellulose itself may sequester Fe^{3+} (29), the generation of Fe^{3+}–oxalate and potentially other redox active iron-chelates might also contribute to lower the effective concentration of bioavailable iron that is accessible to the organism. Thus, cellulolytic conditions might turn on the high-affinity iron uptake system to ensure proper levels of intracellular iron.

Also compatible with Fenton mechanisms is the observed cellulose-induced expression of structurally divergent oxidases (e.g. copper radical oxidases, glucose-1-oxidases and methanol oxidases) and putative iron reductases. Given the significant number of secreted hemicellulases observed, wood decay by *P. placenta* likely involves attack by oxidative and hydrolytic mechanisms. Elevated hemicellulase expression may reflect increased substrate exposure and availability, relative to cellulose and lignin, especially early in the decay process. Products of the hydrolytic attack could in turn serve as candidate substrates for copper radical oxidases and GMC oxidoreductases, thereby generating extracellular H_2O_2. Similarly, methanol released via demethoxylation of lignin (3, 4) may play an important role in H_2O_2 generation as a substrate for methanol oxidase. Such a role is consistent with our observed expression patterns and with previous investigations with *G. trabeum* (10). Of course hydroxyl radical may also play an important role early in decay, and it has been demonstrated to preferentially attack hemicellulose in wood (30). Interestingly, •OH attack on cellulose oxidizes chain ends (31) and the depolymerized material becomes less amenable to cellulase action (30), providing a plausible explanation for the lack of exocellobiohydrolase genes in this fungus.

Comparison of the *P. placenta* and *P. chrysosporium* genomes indicates that the derivation of brown-rot is characterized largely by the contraction or loss of multiple gene families that are thought to be important in typical white-rot, such as cellulases, LiPs, MnPs, CROs, CDH, and pyranose-2-oxidase. This general pattern of simplification is consistent with the view (32) that brown-rot fungi, having evolved novel mechanisms for initiating cellulose depolymerization, have cast off much of the energetically costly lignocellulose-degrading apparatus that is retained in white-rot fungi, such as *P.*

chrysosporium.

Materials and Methods

Genome sequencing, assembly and annotation. A pure whole genome shotgun approach was used to sequence *P. placenta* strain MAD-698-R (USDA, Forest Mycology Center, Madison, WI). The 7.2X coverage assembly was produced from sequenced paired reads using JAZZ assembler. Using an array of gene predictors in the JGI annotation pipeline, a total of 17,173 gene models were predicted and annotated for this dikaryotic fungus. Predicted genes, supporting evidence, annotations, and analyses are available through interactive visualization and analysis tools from the JGI genome portal (http://genome.jgi-psf.org/Pospl1/Pospl1.home.html). Detail regarding the assembly, repetitive elements, ESTs and annotation, are provided separately (SI text).

Mass spectrometry. Soluble extracellular protein was concentrated from shake cultures containing ball-milled aspen, microcrystalline cellulose (Avicel) or de-waxed cotton as previously reported (33). Sample preparation and LC-MS/MS analysis were performed as described (www.biotech.wisc.edu/ServicesResearch/MassSpec/ingel.htm.). Peptides were identified using a Mascot search engine (Matrix Science, London, UK) against protein sequences of 17,173 predicted gene models described above. Complete listings of CAZYs and oxidative enzymes, including peptide sequences and scores, are provided in SI text and in NCBI's GEO under series accession GSE12540 SI Table 11).

Expression microarrays. Roche NimbleGen (Madison, WI) arrays were designed to assess expression of 12,438 genes during growth on microcrystalline cellulose or on glucose as sole carbon sources. The corresponding set of coding regions was manually annotated to include only the 'best allelic model' among CAZY-encoding genes (GEO GSE12540 SI Table 1). Methods are detailed in SI text, and all data deposited under GEO accession GSE12540.

ACKNOWLEDGEMENTS. This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory under contract No. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, Los Alamos National Laboratory under contract No. DE-AC02-06NA25396, the University of Wisconsin under grant No. DE-FG02-87ER13712, Forest Products Laboratory under USDA, CREES grant No. 2007-35504-18257, University of New Mexico under National Institute of Health grant GM060201, CIB (Madrid) EU-project NMP2-2006-026456, Ministry of Education Czech Republic grant No. LC06066. We thank Sally Ralph (FPL) for preparation of ball-milled aspen.

References

1. DOE US (2006) *Breaking the Biological Barriers to Cellulosic Ethanol: A Joint Research Agenda, DOE/SC-0095*, (DOE).
2. Eriksson K-EL, Blanchette RA, Ander P (1990) *Microbial and enzymatic degradation of wood and wood components* (Springer-Verlag, Berlin).
3. Niemenmaa O, Uusi-Rauva A, Hatakka A (2007) Demethoxylation of [O^{14}CH$_3$]-labelled lignin model compounds by the brown-rot fungi *Gloeophyllum trabeum* and *Poria (Postia) placenta*. *Biodegradation* 19:555-565.
4. Yelle DJ, Ralph J, Lu F, Hammel KE (2008) Evidence for cleavage of lignin by a brown rot basidiomycete. *Environ Microbiol* 10:1844-1849.
5. Hibbett DS & Donoghue MJ (2001) Analysis of character correlations among wood decay mechanisms, mating systems, and substrate ranges in homobasidiomycetes. *Syst Biol* 50(2):215-242.
6. Baldrian P & Valaskova V (2008) Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiol Rev* 32(3):501-521.
7. Cowling EB & Brown W (1969) Structural features of cellulosic materials in relation to enzymatic hydrolysis. *Cellulases and Their Applications*, eds Hajny GJ & Reese ET (American Chemical Society Advances in Chemistry Series 95, Washington, DC), pp 152-187.
8. Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280 (Pt 2):309-316.
9. Whittaker MM, Kersten PJ, Cullen D, Whittaker JW (1999) Identification of catalytic residues in glyoxal oxidase by targeted mutagenesis. *J Biol Chem* 274(51):36226-36232.
10. Daniel G, *et al.* (2007) Characteristics of *Gloeophyllum trabeum* alcohol oxidase, an extracellular source of H_2O_2 in brown rot decay of wood. *Appl Environ Microbiol* 73(19):6241-6253.
11. Varela E, Martinez JM, Martinez AT (2000) Aryl-alcohol oxidase protein sequence: a comparison with glucose oxidase and other FAD oxidoreductases. *Biochim Biophys Acta* 1481(1):202-208.
12. de Koker TH, Mozuch MD, Cullen D, Gaskell J, Kersten PJ (2004) Pyranose 2-oxidase from *Phanerochaete chrysosporium*: isolation from solid substrate, protein purification, and characterization of gene structure and regulation. *Appl Environ Microbiol* 70:5794-5800.
13. Tanaka H, *et al.* (2007) Characterization of a hydroxyl-radical-producing glycoprotein and its presumptive genes from the white-rot basidiomycete *Phanerochaete chrysosporium*. *J Biotechnol* 128(3):500-511.
14. Martinez AT (2002) Molecular biology and structure-function of lignin-degrading heme peroxidases. *Enzyme Microb Technol* 30:425-444.
15. Morgenstern I, Klopman S, Hibbett DS (2008) Molecular evolution and diversity of lignin degrading heme peroxidases in the Agaricomycetes. *J Mol Evol* 66(3):243-257.
16. Hoegger PJ, Kilaru S, James TY, Thacker JR, Kües U (2006) Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. *FEBS J* 273(10):2308-2326.

17. Guillen F, Gomez-Toribio V, Martinez MJ, Martinez AT (2000) Production of hydroxyl radical by the synergistic action of fungal laccase and aryl alcohol oxidase. *Arch Biochem Biophys* 383(1):142-147.
18. Martinez D, *et al.* (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nature Biotechnol* 22:695-700.
19. Varela E & Tien M (2003) Effect of pH and oxalate on hydroquinone-derived hydroxyl radical formation during brown rot wood degradation. *Appl Environ Microbiol* 69(10):6025-6031.
20. Munir E, Yoon JJ, Tokimatsu T, Hattori T, Shimada M (2001) A physiological role for oxalic acid biosynthesis in the wood-rotting basidiomycete *Fomitopsis palustris*. *Proc Natl Acad Sci USA* 98(20):11126-11130.
21. Clausen CA (1995) Dissociation of the multi-enzyme complex of the brown-rot fungus *Postia placenta*. *FEMS Microbiol Lett* 127:73-78.
22. Xie G, *et al.* (2007) Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*. *Appl Environ Microbiol* 73(11):3536-3546.
23. Wilson DB (2008) Three microbial strategies for plant cell wall degradation. *Ann N Y Acad Sci* 1125:289-297.
24. McCarter SL, *et al.* (2002) Exploration of cellulose surface-binding properties of *Acidothermus cellulolyticus* Cel5A by site-specific mutagenesis. *Appl Biochem Biotechnol* 98:273-287.
25. Tsai CF, Qiu X, Liu JH (2003) A comparative analysis of two cDNA clones of the cellulase gene family from anaerobic fungus *Piromyces rhizinflata*. *Anaerobe* 9(3):131-140.
26. Cohen R, Suzuki MR, Hammel KE (2004) Differential stress-induced regulation of two quinone reductases in the brown rot basidiomycete *Gloeophyllum trabeum*. *Appl Environ Microbiol* 70(1):324-331.
27. Kaneko S, Yoshitake K, Itakura S, Tanaka H, Enoki A (2005) Relationship between production of hydroxyl radicals and degradation of wood, crystalline cellulose, and lignin-related compound or accumulation of oxalic acid in cultures of brown-rot fungi. *J Wood Sci* 51:262-269.
28. Jensen KA, Jr., Houtman CJ, Ryan ZC, Hammel KE (2001) Pathways for extracellular Fenton chemistry in the brown rot basidiomycete *Gloeophyllum trabeum*. *Appl Environ Microbiol* 67(6):2705-2711.
29. Xu G & Goodell B (2001) Mechanisms of wood degradation by brown-rot fungi: chelator-mediated cellulose degradation and binding of iron by cellulose. *J Biotechnol* 87(1):43-57.
30. Ratto M, Ritschkoff A, Viikari L (1997) The effect of oxidative pretreatment on cellulose degradation by *Poria placenta* and *Trichoderma reesei*. *Appl Microbiol Biotechnol* 48:53-57.
31. Kirk TK, Ibach R, Mozuch MD, Conner AH, Highley TL (1991) Characteristics of cotton cellulose depolymerized by a brown-rot fungus, by acid, or by chemical oxidants. *Holzforschung* 45:239-244.
32. Worrall JJ, Anagnost SE, Zabel RA (1997) Comparison of wood decay among diverse lignicolous fungi. *Mycologia* 89:199-219.
33. Vanden Wymelenberg A, *et al.* (2006) Computational analysis of the

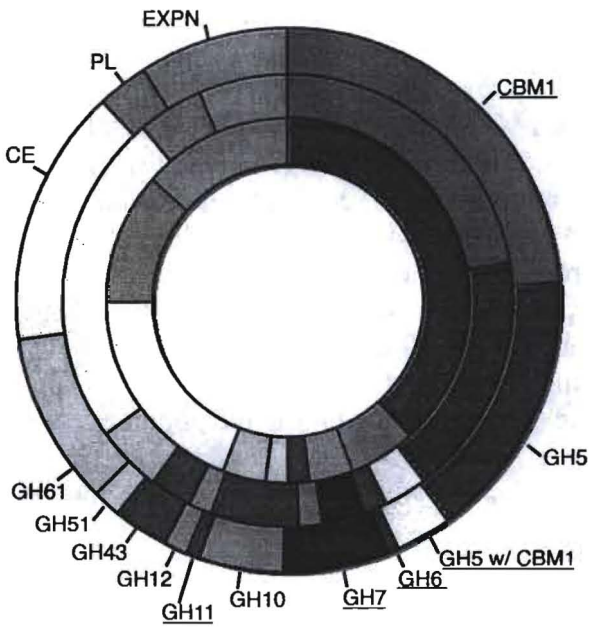
Phanerochaete chrysosporium v2.0 genome database and mass spectrometry identification of peptides in ligninolytic cultures reveals complex mixtures of secreted proteins. *Fungal Genet and Biol* 43:343-356.

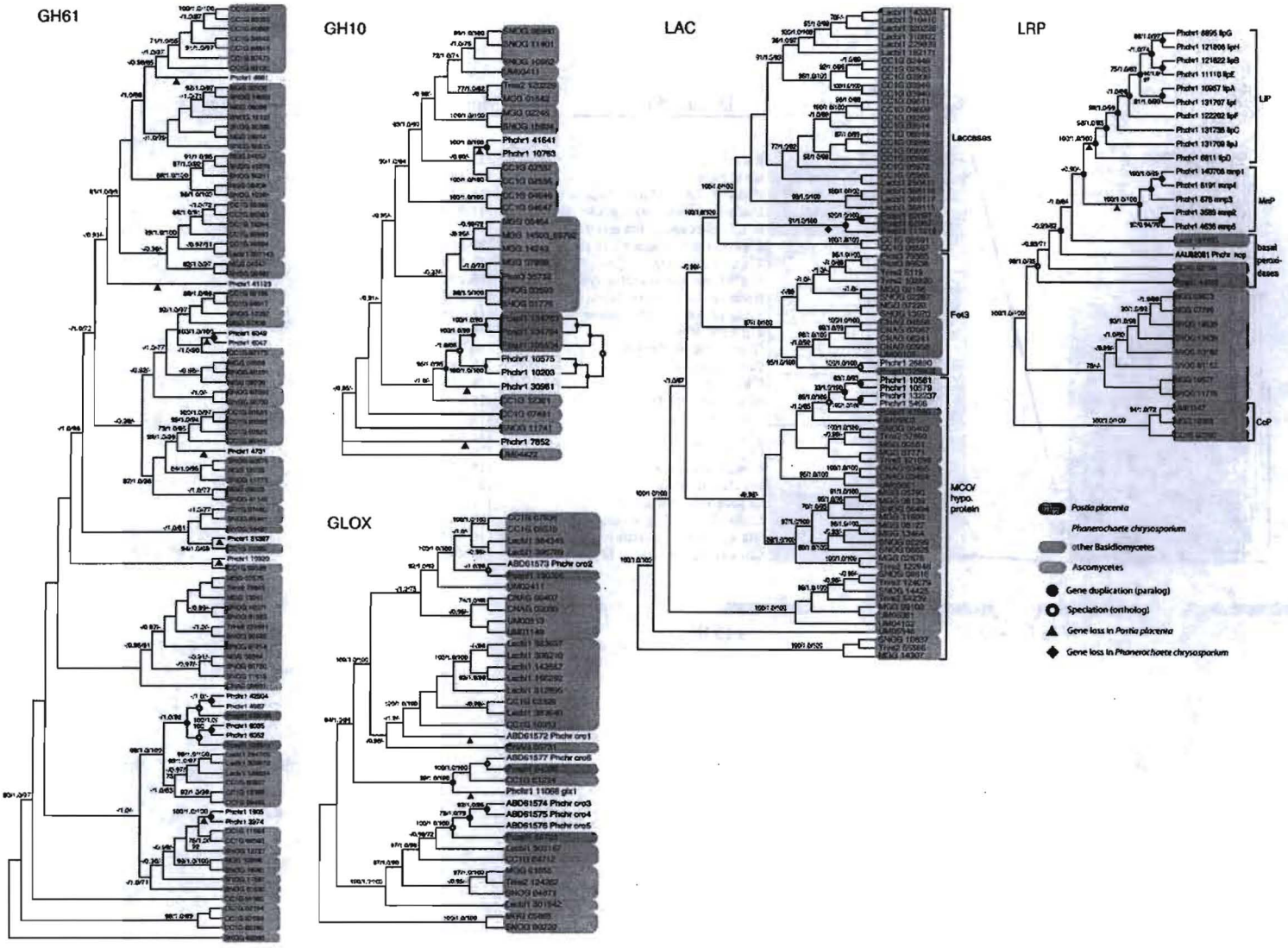
Figure Legends

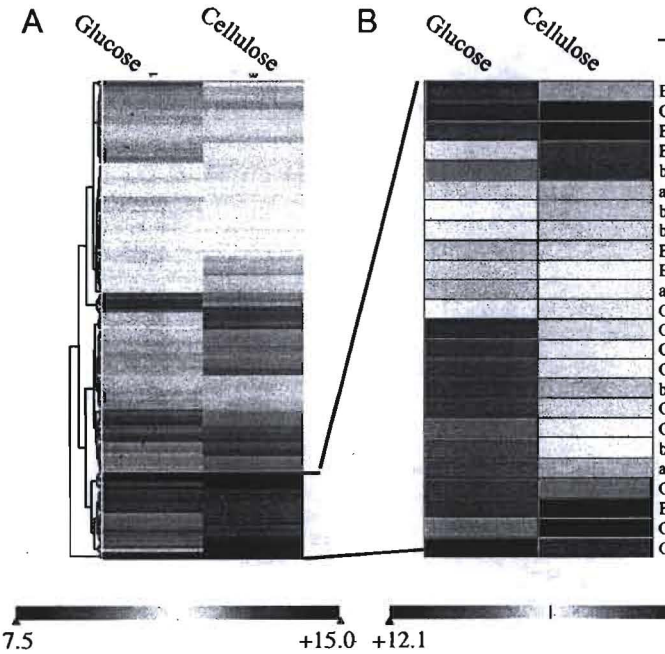
Fig. 1. Distribution of various CAZymes in *P. placenta* (inner ring), *T. reesei* (middle ring), and *P. chrysosporium* (outer ring). Abbreviations: CBM1, family 1 carbohydrate binding modules; GH#, modules within individual glycoside hydrolase families; GH5 (CBM1), glycoside hydrolase family 5 modules associated with family 1 carbohydrate binding modules; GT, glycosyl transferases; CE, carbohydrate esterases; PL, polysaccharide lyases; EXPN, expansin-related proteins. Enzymes not found in *P. placenta* are underlined. Comparisons with additional species are list listed in SI Table 2.

Fig. 2. Phylogenies of glycoside hydrolase (GH 61, GH10), glyoxal oxidase/copper radical oxidase (GLOX), laccase (LAC) and related multicopper oxidase, and low redox peroxidase (LRP) and related class II fungal peroxidases, from complete genomes of *P. placenta* (Pospl1), *P. chrysosporium* (Phchr1), *C. cinerea* (CC1G), *L. bicolor* (Lacbi1), *C. neoformans* (CNAG), *U. maydis* (UM), *M. grisea* (MGG), *Stagonospora nodorum* (SNOG), *T. reesei* (Trire2) and *Pichia stipitis* (Picst3). Datasets were assembled using BLAST (with qUniProtKB query sequences Q5XXE5, O60206, P36218, Q00023, O14405, Q01772 and Q12718), with a cut-off value of E-06. Parsimony (PAUP* 4.0; 10,000 heuristic searches, 1000 bootstrap replicates), maximum likelihood (RAxML; 1000 bootstrap replicates, with models suggested by ProtTest), and Bayesian (MrBayes v3.1.2; two runs of four chains, 10 million generations each, with mixed protein models) support values are indicated in the order MP/PP/ML. Topologies shown are from Bayesian phylogenetic analyses. An alternative topology from parsimony analysis is shown for part of the GH10 phylogeny. Inferred gene losses, duplication events (paralogy), and speciation events (orthology) are indicated within *Postia* and *Phanerochaete* only.

Fig. 3. Expression profile of 144 glycoside hydrolase-encoding genes in media containing glucose versus microcrystalline cellulose as sole carbon sources (Part A). In Part B, a cluster of 24 of highly expressed genes is expanded and the color scale re-calibrated to illustrate differences in transcript accumulation. Expression ratios were derived from comparisons of glucose-grown versus cellulose-grown mycelia. Analysis is based on three biological replicates per culture medium. Quantile normalization and robust multi-array averaging (RMA) were applied to the entire dataset. ANOVA showed 120 GH-encoding genes, including all 24 above, were significantly regulated ($P<0.01$). Reciprocals of ratios <1.0 are multiplied by -1. Asterisks indicate proteins identified by LC-MS/MS. Detailed listing of CAZYs with statistical analyses of expression data is presented in SI Table1 under GEO accession GSE12540.


EXPN
PL
CBM1
CE
GH5
GH61
GH51
GH43
GH12
GH11
GH10
GH7
GH6
GH5 w/ CBM1


GH61
GH10
GLOX
LAC
LRP
Laccases
Fet3
MCO/ hypo. protein
LiP
MnP
basal peroxidases
CcP
Postia placenta
Phanerochaete chrysosporium
other Basidiomycetes
Ascomycetes
Gene duplication (paralog)
Speciation (ortholog)
Gene loss in Postia placenta
Gene loss in Phanerochaete chrysosporium



Description	Model	Family	Fold reg.
Endog lucanase	121 191	GH12	-1.51
Chitinase	119 525 *	GH18	+1.35
Endo -b-1,4 -endoglucanase	115 648 *	GH5	+1.43
Endo -b-1,4 -mannosidase	121 831 *	GH5	+1.93
b-1,3 -glucanosyltransfe rase	117 860 *	GH72	+1.37
a-amylase/a-glucan synthase	127 907	GH13	+1.39
b-galactosidase	127 993	GH35	+1.23
b-glucanase/b-glucan synthetase	115 248	GH16	+1.08
Endo -b-1,4 -endoglucanase	117 690	GH5	-1.05
Endo -1,3(4) -b-glucanase	112 923	GH16	+1.13
a-glucosidase	134 924 *	GH31	+1.16
Glycoside hydrolase	116 700	GH16	-1.10
Glucoamylase	117 345 *	GH15	+2.28
Chitinase	126 595 *	GH18	+1.97
Glycoside hydrolase	135 050 *	GH30	+1.89
b-mannosidase	114 395	GH2	+2.27
Glycoside hydrolase	113 893	GH16	+1.59
Glycoside hydrolase	120 595	GH16	+1.48
b-hexosaminidase	112 369	GH20	+1.52
a-galactosidase	134 790	GH2 7	+2.07
Glycoside hydrolase	121 713	GH5	+2.36
Endo -1,3(4) -b-glucanase	112 941 *	GH16	+3.95
Glucan 1,3 -b-glucosidase	116 267 *	GH55	+3.15
Glucan 1,3 -b-glucosidase	105 490	GH55	+5.38

Supporting Information

Figure legends

SI Fig. 1. 3D-model for the hypothetical *Postia placenta* peroxidase Ppl44056 (PP, light brown) obtained by homology modeling, superimposed with the crystal structure of versatile peroxidase (VP) from *Pleurotus eryngii* (green). **A)** Protein backbones indicating the position of tryptophan residues present in these peroxidases (as van der Waals spheres), one of them being common to both proteins (middle) whereas the two other are characteristic of VP (left) and PP (right). The position of heme cofactor in a central pocket is also indicated. **B)** Detail of the propionates side of the heme pocket where the Mn-oxidation site is located in VP formed by two glutamate and one aspartate residues (two of them being substituted by Asn and Gly residues in PP). From **A** it is possible to conclude that *P. placenta* PP lacks the catalytic tryptophan responsible for oxidation of high redox-potential aromatic substrates by *P. eryngii* VP (and *Phanerochaete chrysosporium* lignin peroxidase) (1), located near the heme cofactor at the left side of the figure. From **B** it is possible to conclude that *P. placenta* PP lacks a Mn-oxidation site near the internal propionate of heme as found in *P. eryngii* VP (and *P. chrysosporium* Mn-peroxidase) crystal structure (2). This structural comparison and the corresponding sequence alignments indicate that the *P. placenta* peroxidase is not a ligninolytic peroxidase (as LiP, MnP or VP) but more closely allied to *Coprinopsis cinerea* peroxidase (CIP) and could oxidize low redox-potential dyes and phenols at the edge of the main heme access channel.

SI Fig. 2. Putative components of TCA cycle, GLOX cycle and oxalate metabolism. Colored bars indicate microarray-derived expression results from glucose-grown (left panel) and cellulose-grown (right panel) mycelia. Manual inspection of malate dehydrogenase model Ppl106934 revealed several inaccuracies upon which microarray 60mers were based. Reliable expression results are therefore not available. Model details and microarray results are listed in SI_table_8.xls under NCBI-GEO series accession GSE12540.

SI Fig. 3. Neighbor-Joining tree of the cytochrome P450 contingent (P450ome) of *P. placenta*. The tree shows 254 P450 proteins clustered into 11 fungal clans (indicated by different clan specific color at the branch point). The unrooted phylogenetic tree was constructed using MEGA4 phylogenetic analyses program with 1000 bootstrap replications. Evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset. This large jpeg image is linked as "SI_figure_3.jpg" under NCBI-GEO series accession GSE12540.

SI Fig. 4. Scatter plot showing expression of 12,438 *P. placenta* genes in cultures containing glucose or microcrystalline cellulose as sole carbon source. Data normalized and shown as $\log_2$ scale. Middle solid green line is identity (x=y). Outside green lines delineate 290 genes with >2-fold change in intensity in one data set. Details for these regulated genes in SI Table 3. Dashed purple line is best fit linear regression. Values for

all genes, together with ANOVA results are listed under NCBI-GEO series accession GSE12540.

SI Fig. 5. Analysis of *P. placenta B*-mating locus. Panel A: Schematic representation of the *B*-mating type locus. Clusters and groups within clusters are given according to the phylogenetic proximity of the respective gene models to their orthologs in *C. cinerea*. Closely related pheromone precursors (as revealed by phylogenetic analysis, panel C) are depicted in identical colours and those lacking close neighbours are shown in black. Complete microarray results are listed in NCBI-GEO series accession GSE12540, where Ppl94656 and Ppl51963 are represented by earlier incomplete models Ppl32692 and Ppl32709, respectively. (Oligonucleotide specificity is retained in the improved models.) Panel B: Phylogenetic analysis of G-protein coupled receptors and their orthologs in *C. cinerea*. Sequences used are derived from the *C. cinerea* strain Okayama 7 genome (http://www.broad.mit.edu/annotation/genome/coprinus_cinereus/Home.html). These sequences are: CC1G_02129 = CcSTE3.1, CC1G_02136 = CcSTE3.2a and CcSTE3.2b (previously known as Rcb2 B43, AAQ96345), CC1G_02137 = CcSTE3.3 (previously known as Rcb3 B43; AAQ96346), CC1G_02151 = CcSTE3-2151, CC1G_02153 = CcSTE3-2153 and CC1G_07395 = CcSTE3-7395). Panel C: Phylogenetic analysis of putative a-type peptide pheromone precursors of *P. placenta*. For both trees, sequence alignments were performed using ClustalX (1.81) (3) and the alignment was manually adjusted by the aid of Genedoc. Phylogenetic analysis was performed with MEGA 4.0 (4) using the minimum evolution approach. The reliability of the nodes was estimated by minimum evolution bootstrap percentages obtained after 1000 pseudoreplications.

SI Fig. 6. Phylogenetic analysis of G-protein α subunits. Abbreviation are as follows: *Trichoderma reesei* (*Hypocrea jecorina*) (TR), *Saccharomyces cerevisiae* (SC), *Schizosaccharomyces pombe* (SP), *Neurospora crassa* (NC), *Aspergillus fumigatus* (AF), *Gibberella zeae* (GZ), *Fusarium oxysporum* (FO), *Magnaporthe grisea* (MG), *Cryphonectria parasitica* (CP), *Rattus norvegicus* (RN), *Canis familiaris* (CF), *Drosophila melanogaster* (DM), *Oryza sativa* (OS), *Pisum sativum* (PS), *Glycine max* (GM) and *Ustilago maydis* (UM). The analysis was performed as described with SI Fig. 5.

SI Fig. 7. Schematic representation of putative MAP-kinase pathways in *P. placenta*. Cascades and upstream factors were deduced from Gustin *et al.*, 1998 (5) and similarity of the respective *P. placenta* predicted proteins as suggested by BLASTP-searches restricted to *Saccharomyces cerevisiae*.

SI Text

1. Methods.

Genome sequencing and assembly. A pure whole genome shotgun approach was used to sequence dikaryotic *Postia placenta* strain MAD-698 (USDA, Forest Mycology Center, Madison, WI). Genomic DNA was randomly sheared and size fractionated, and libraries of approximately 3 kb and 8 kb were constructed. The assembly was constructed with the JGI assembler, JAZZ, using paired end sequencing reads at a coverage of 7.23X. After trimming for vector and quality, 574,631 reads assembled into 1243 scaffolds totaling 90.9 Mbp. Completeness of the assembly can be estimated in part by the observation that 98.7% of 38,114 ESTs aligned to the assembly with >50% coverage.

Since *P. placenta* has been shown to be highly polymorphic with a polymorphism rate of 3-4%, this particular assembly uses extra stringent parameters that should only assemble sections of the genome that are more than 99% identical. The entire assembly is available through the JGI portal (http://genome.jgi-psf.org/Pospl1/Pospl1.download.ftp.html) and has been deposited (Accession ABWF00000000). Summarizing assembly features:

Scaffold Total	1243
Scaffold Sequence Total	90.9 MB
Scaffold N50	85
Estimated Depth	7.23 +/- 0.13

EST sequencing and analysis. *P. placenta* MAD-698 total RNA was purified after 5 days growth on three media: rich media (2% w/v malt extract, 2% glucose, 0.5% peptone), basal salts (6) with 0.5% glucose, and basal salts with 0.5% wood-derived microcrystalline cellulose (Avicel). The Qiagen RNeasy Maxi Purification kit was used according to the manufacturer's instructions. cDNA synthesis and cloning was a modified procedure based on the SuperScript plasmid system with Gateway technology for cDNA synthesis and cloning (Invitrogen, Carsbad, CA). 1-2 ug of poly A+ RNA, reverse transcriptase SuperScript II (Invitrogen) and oligo dT-NotI primer (5' -GACTAGTTCTAGATCGCGAGCGGCCGCCCT15VN -3') were used to synthesize first strand cDNA. Second strand synthesis was performed with *E. coli* DNA ligase, polymerase I, and RNaseH followed by end repair using T4 DNA polymerase. The *Sal*I adaptor (5'- TCGACCCACGCGTCCG -3' and 5'- CGGACGCGTGGG -3') was ligated to each cDNA, digested with NotI (New England Biolabs, Ipswich, MA), and subsequently size selected by gel electrophoresis (1.1% agarose). Two size ranges of cDNA from each of the three RNA samples were cut out of the gel, 0.6kb-2kb for the smaller insert size range cDNA libraries and 2kb-10kb for the larger size range cDNA libraries. For the glucose and microcrystalline cellulose samples, the cDNA inserts were directionally ligated into the SalI and NotI digested vector pMCL200-cDNA. Likewise, the rich media

sample was ligated into both pMCL200-cDNA and pCUGI21. The ligations were all transformed into ElectroMAX DH10B cells (Invitrogen).

Library quality was first assessed by randomly selecting 24 clones and PCR amplifying the cDNA inserts with the primers M13-F (5'-GTAAAACGACGGCCAGT--3') and M13-R (5'-AGGAAACAGCTATGACCAT-3') to determine the fraction of clones without inserts. Colonies from each library were plated onto agarose plates (254mm plates from Teknova, Hollister, CA) at a density of approximately 1000 colonies per plate. Plates were grown at 37 C for 18 hours, then individual colonies were picked and each used to inoculate a well containing LB media with appropriate antibiotic in a 384 well plate (Nunc, Rochester, NY). Clones in 384 well plates were grown at 37 C for 18 hours. Contained plasmid DNA for sequencing was produced by rolling circle amplification ((7), Templiphi, GE Healthcare, Piscataway, NJ). Subclone inserts were sequenced from both ends using primers complimentary to the flanking vector sequence (Fw: 5'-AGGAAACAGCTATGACCAT--3', Rv: 5'-GTTTTCCCAGTCACGACGTTGTA-3') using Big Dye terminator chemistry and then run on ABI 3730 instruments (Applied Biosystems, Foster City, CA). A total of 44,520 Expressed Sequence Tags (ESTs) were sequenced from the three RNA samples with an average read length of 710.

The 44,520 ESTs, including 33,792 from rich medium, 1,536 from glucose, and 9,216 from microcrystalline cellulose, were processed through the JGI EST pipeline (ESTs generated in pairs, a 5' and 3' end read from each cDNA clone). The Phred software (8, 9) was used to call the bases and generate quality scores. Vector, linker, adapter, poly-A/T, and other artifact sequences were removed using the Cross_match software, and an internally developed short pattern finder. Low quality regions of the read were identified using internally developed software, masking regions with a combined quality score of less than 15. The longest high quality region of each read was considered the EST. ESTs shorter than 150 bp were removed from the data set. ESTs containing common contaminants such as *E. coli*, common vectors, and sequencing standards were also removed from the data set.

EST clustering was performed *ab-initio*, based on alignments between pairs of trimmed, high quality ESTs yielding 29,763, 7,454, and 897 each from rich media, microcrystalline cellulose, and glucose, respectively (GenBank accessions FL595400 through FL633513). Pair-wise EST alignments were generated using internally developed alignment software. ESTs sharing an alignment of at least 99% identity and 150 bp overlap were assigned to the same cluster. Furthermore, ESTs not sharing alignments were assigned to the same cluster if they were derived from the same cDNA clone. Clusters of ESTs were assembled into consensus sequences, contigs or singlets using Phrap software. A total of 7,251 clusters (including 1,145 singlet EST clusters) were generated and further assembled into 10,716 consensus sequences using Phrap.

Automated Genome Annotation. Following assembly, *P. placenta* genes were predicted and annotated using the JGI annotation pipeline. After masking transposons using RepeatMasker, we used several gene predictors including *ab initio* Fgenesh (10), homology-based, Fgenesh+ (10), and Genewise (11). Fgenesh was trained on a set of available mRNAs, full-length genes derived from EST clusters, and reliable homology-

based gene models. GeneWise models were extended when possible to include start and stop codons. When possible, available ESTs were used to extend the predicted protein coding regions into full-length genes by adding 3' and 5' UTRs. When multiple models were predicted at the same locus, the model with best homology (including coverage in both model and hit sequences) and EST support was selected for the non-redundant set of genes called 'BestModels, v2.0' containing 17,173 gene models. Genes have been annotated and classified according to GO (gene ontology consortium http://www.geneontology.org/), eukaryotic orthologous groups (KOGs (12)), and KEGG metabolic pathways (13).

Annotation summary:

Number gene models	17,173
Models/Mbp	189
Multi-exon genes	94%
Supported by ESTs	29%
KEGG	3,424 (20%)
KOG	9,321 (54%)
GO	8,198 (48%)

Allelic variants. The haplotype assembly presented unique problems in candidate allele pair assignment. An initial manual analysis revealed two potential confounding factors: the first, that there was the possibility that some recent duplication in some gene families may confuse simplistic approaches (e.g., best hit equals allele), the second, a sample of genes are likely not represented as an allele pair (i.e., there is no difference in nucleotide sequence at all between allele pairs). To ensure appropriate stringency, 500 bp of upstream sequence in addition to exon and intron sequence for each gene was aligned using BLASTN. Then mutual best hit pairs that were not on the same scaffold were selected as candidate alleles if they were more than 92% identical (allowing for gaps in the assembly). With this method 4946 allele pairs were detected, resulting in 12,227 total unique genes. To create a final unique gene set, the allele from the largest scaffold was kept, to create the longest continuous regions of haplotype pairs.

Manual curation. Automatically predicted genes and functions were further refined by user community-wide manual curation efforts using web-based tools at www.jgi.doe.gov/postia. The latest version gene set containing manually curated genes is called GeneCatalog.

To detect carbohydrate active genes, each protein model was BLASTed against the library of approximately 100,000 individual CAZY modules using a database size parameter identical to that of the NCBI non-redundant database (14); http://www.cazy.org). All models that gave an e-value better than 0.1 were automatically kept and manually analyzed. Manual analysis involved examination of the alignment of the model with the various members of each family (whether of catalytic or non-catalytic modules), with a search of the conserved signatures/motifs characteristic of each family.

The presence of the catalytic machinery was verified for borderline cases whenever known in the family. The models that displayed the usual features that would lead to their inclusion in the CAZY database were kept for annotation and classified in the appropriate class and family.

Detection of genes potentially involved in peroxide generation, iron homeostasis, lignin modification, and oxalate metabolism involved BLASTP and TBLASTN queries using orthologs from related fungi. Additional genes were detected by their InterPro domains. *P. placenta* genes identified in this manner were used as queries for reiterative BLASTN and TBLASTN searches. To establish allelic relationships, gene sequences plus flanking regions (>400nt) were BLASTed against the genome. Pairs exhibiting >99% identity were mapped as alleles on the browser. Manual inspection of the identity and transcriptional orientation of surrounding genes affirmed allelic pairs (Supplemental Tables).

Repeats. The *P. placenta* genome is particularly rich in repetitive elements, many of which resemble remnants of class I transposable elements (TEs). *P. placenta* specific retroelements (labeled VC_TErepeats#) were discovered by aligning the genome to a manually curated database of known repeats with blastn. Related sequence elements were clustered using MCL (118) . Clustered elements were then aligned with clustalw (119) to produce a consensus sequence that could be later used in repeat masking. Other transposons are available via RepBase (http://www.girinst.org/). All repeats are mapped to the assembly with RepeatMasker (http://www.repeatmasker.org) and appear within a separate track accessible through the JGI Genome Portal at http://jgi.doe.gov/Postia. In aggregate, the simple repeats, known fungal TEs, and 29 extracted retroelements exceed 5% of the genome and all were masked prior to annotation.

Element	# in genome	Total length (bp)
VC_TErepeats1	881	1454578
VC_TErepeats2	669	691844
VC_TErepeats3	292	451090
VC_TErepeats4	461	549715
VC_TErepeats5	160	169497
VC_TErepeats6	277	314361
VC_TErepeats7	119	174741
VC_TErepeats8	127	105777
VC_TErepeats9	88	76440
VC_TErepeats10	25	25093
VC_TErepeats11	74	145248
VC_TErepeats12	36	54510
VC_TErepeats13	114	136888
VC_TErepeats14	33	65118
VC_TErepeats15	19	10335
VC_TErepeats16	32	14762
VC_TErepeats17	67	28030

VC_TErepeats18	37	40330
VC_TErepeats19	23	26218
VC_TErepeats20	56	63654
VC_TErepeats22	35	52828
VC_TErepeats23	40	20172
VC_TErepeats24	54	63869
VC_TErepeats25	15	11531
VC_TErepeats26	13	10010
VC_TErepeats27	21	1376
VC_TErepeats28	17	15166
VC_TErepeats29	1	45
simple repeats	1284	7721
RESTLESS	1	71
MOLLY	1	65
GC_rich	473	17072
CT-rich	175	28284
GA-rich	168	23742
G-rich	110	10372
AT_rich	94	2631
C-rich	93	9808
Helitron-1_AN	12	2644
PCC2	6	2384
T-rich	3	13667
A-rich	3	13341
polypyrimidine	2	70
SKIPPY-I	2	227
MGRL3_I	1	197
FOLYT1	1	44
TPA5_I	1	166
polypurine	1	66
PYRET_I	1	199
5SrRNA_AN	1	71
CRYPT1	1	35

Gypsy and *Copia* retroelements (15-17) are particularly abundant, although the vast majority are inactive remnants and/or present as 'solo' LTRs (long terminal repeats) (18, 19). Typical of repetitive sequences, many of the elements lie at scaffold ends or are excluded from the shotgun assembly. The 1947 nt repeat designated VC-TErepeats 25 is structurally related to *Coprina-Pc1*, a *Penelope*-like retroelement (20). In *C. cinerea* and *P. chrysosporium* these elements are associated with the telomeric repeat, $(TAAACCC)_n$. This telomeric repeat is not observed in *P. placenta*, but (TTAGG)n is associated with the elements and possibly represents a divergent telomeric repeat (21). Analysis of tandem repeats (Tandem Repeat Finder; http://tandem.bu.edu/trf/trf.html) revealed 16 separate locations where the motif's array length exceeded 150 copies, and all arrays were located at scaffold termini or adjacent to gapped regions.

Mass spectrometry. Soluble extracellular protein was concentrated from shake cultures containing ball-milled aspen, wood-derived cellulose (microcrystalline cellulose) or de-waxed cotton as previously reported (22, 23). Following SDS-PAGE fractionation, 'in gel' digestion and mass spectrometric analysis were performed as described (www.biotech.wisc.edu/ServicesResearch/MassSpec/ingel.htm) An in-house licensed Mascot search engine (Matrix Science, London, UK) identified peptides using the dataset of 17,173 gene models described above. To identify potential coding regions not represented in the current gene model set, a second database of all translated ORFs was queried. Complete listing of CAZYs and oxidative enzymes, including peptide sequences and scores are detailed in SI_table_11.xls under NCBI-GEO series accession GSE12540. Throughout, protein similarity scores are based on the Smith-Waterman algorithm (24) using the BLOSUM62 matrix.

Expression microarrays. From a data set of 12,438 unique alleles, each Roche NimbleGen (Madison, WI) array featured 10 unique 60mer probes per gene, all in triplicate. The dataset was manually annotated to include only the 'best allelic model' among CAZY-encoding genes. The listing of all CAZY models (SI_table_1.xls) as well as complete design detail are available as platform GPL7187 within NCBI GEO series GSE12540 (http://www.ncbi.nlm.nih.gov/geo/index.cgi).

Total RNA was purified from 5-day old cultures containing microcrystalline cellulose (Avicel) or glucose as sole carbon source. In short, cultures were harvested by filtering through Miracloth (Calbiochem, EMD Biosciences, Gibbstown, NJ), squeeze dried and snap frozen in liquid nitrogen. Pellets were stored at -80 C until use. Extraction buffer was prepared by combining 10 ml 690 mM para-aminosalicylic acid (sodium salt) (Sigma-Aldrich, St. Louis, MO) with 10 ml 56 mM triisopropylnapthalene sulfonic acid (sodium salt) (Sigma-Aldrich), and placed on ice. To this was added 5 ml 5X RNB (1.0 M Tris, 1.25 M NaCl, 0.25 M EGTA). The pH of the 5X RNB was adjusted to 8.5 with NaOH. The mixture was kept on ice and shaken just before use.

Frozen fungal pellets were ground to a fine powder with liquid nitrogen in an acid washed, pre-chilled mortar and pestle. The ground mycelia were transferred to Falcon 2059 tubes (VWR International, West Chester, PA), and extraction buffer was added to make a thick slurry. The samples were vortexed vigorously and placed on ice until all samples were processed. One half volume TE-saturated phenol (Sigma-Aldrich) and ¼ volume chloroform (Sigma-Aldrich) were added to each sample and vortexed vigorously. Samples were spun at 2940 x *g* in a fixed-angle rotor for 5 minutes. The aqueous layer was removed to a new tube, and phenol:chloroform extractions were repeated until the interface between the aqueous and organic layers was clear. The final aqueous extractions were placed in clean 2059 tubes, to which was added 0.1 volume 3M sodium acetate, pH 5.2, (DEPC-treated) and 2 volumes absolute ethanol. The tubes were shaken vigorously and stored overnight at -20 ^{0}C.

The tubes were spun 1 hour at 2940 x *g*, the supernatants were decanted, and the pellets were resuspended in 4 ml RNase-free H_2O. Total RNA was purified using the RNeasy Maxi kit (Qiagen, Valencia, CA) according to the manufacturer's protocol for RNA cleanup. RNAs were eluted from the RNeasy spin columns using two spins, for a final volume of 2 ml. The eluted RNAs were ethanol precipitated and stored overnight at

-20C. The RNAs were spun 1 hour at 2940 x *g*, washed 1x with 70% ethanol, and resuspended in 50-100ul RNase-free H_2O. Three biological replicates per medium were used (6 separate arrays).

RNA was converted to double-strand cDNA and labeled with the Cy3 fluorophore sample for hybridization to the array by Roche NimbleGen (Iceland). In brief, 10ug of total RNA was incubated with 1X first strand buffer, 10 mM DTT, 0.5mM dNTPs, 100 pM oligo T7 $d(T)_{24}$ primer, and 400U of SuperScript II (Invitrogen) for 60 min at 42C. Second strand cDNA was synthesized by incubation with 1X second strand buffer, 0.2mM dNTPs, 0.07U/ul DNA ligase (Invitrogen), 0.27U/ul DNA polymerase I (Invitrogen), 0.013U/ul RNase H (Invitrogen), at 16C for 2 hours. Immediately following, 10U T4 DNA polymerase (Invitrogen) was added for additional 5 minute incubation at 16C. Double-stranded cDNA was treated with 27ng/ul of RNase A (EpiCenter Technologies) for 10 minutes at 37C. Treated cDNA was purified using an equal volume of phenol:chloroform:isoamyl alcohol (Ambion), ethanol precipitated, washed with 80% ethanol, and resuspended in 20ul water. One ug of each cDNA sample was amplified and labeled with 1 unit per ul of Klenow Fragment (New England BioLabs) and 1 O.D unit of Cy3 fluorophore (TriLink Biotechnologies, Inc.) for 2 hours at 37C. Array hybridization was carried out with 6ug of labeled cDNA suspended in NimbleGen hybridization solution for 17 hours at 42C. Arrays were scanned on the Axon4000B Scanner (Molecular Dynamics) and data was extracted from the scanned image using NimbleScan v2.4. DNASTAR ArrayStar v2.1 (Madison,WI) software was used to quantify and visualize data. A scatterplot of results is shown in SI Fig. 4.. All MIAME compliant (25) microarray expression data has been deposited in NCBI's Gene Expression Omnibus (26) and accessible through GEO Series accession number GSE12540.

2. Extracellular peroxide generation.

We identified a large number of gene models potentially supporting Fenton chemistry through the generation of extracellular H_2O_2 (summarized in SI Table 5) and Fe(II) (see next section). Among the former are members of the GMC oxidoreductase family (InterPro IPR000172) with substantial similarity to known alcohol oxidases (EC 1.1.3.13), aryl-alcohol oxidases (AAO; EC 1.1.3.7) and glucose oxidases (EC 1.1.3.4).

The *P. placenta* protein model Ppl118723 is similar to *G. trabeum* alcohol (methanol) oxidase (DQ835989) and orthologs in *P. chrysosporium* (protein ID 126879) and in *C. cinerea* (CC1G_07964) (> 85% amino acid identity). Unique among known alcohol oxidases, these four sequences contain a conserved 23-24 residue carboxy extension, possibly involved in translocating the proteins to the hyphal surface. As discussed in the main paper, methanol oxidase has been shown as a source of hydrogen peroxide in *G. trabeum* (27). Suggesting a similar role in *P. placenta*, microarray analysis revealed a sharp transcript increase for the gene encoding Ppl118723 in cellulose-grown cultures relative to non-cellulolytic cultures (SI_table_6.xls under GEO accession GSE12540).

Extracellular AAOs together with mycelial aryl alcohol dehydrogenases (AADs) may produce peroxide by redox-cycling of anisyl metabolites in white rot fungi (28, 29).

Overall sequence similarity and conservation of key catalytic residues (30) identified several GMC oxidoreductases as AAO. Amino acid sequences of Ppl55496, Ppl44654 and Ppl54008 are similar to *Pleurotus eryngii* AAO, and less related to the *P. chrysosporium* AAO-type protein. There are also at least 15 putative AADs (EC 1.1.190), four of which are localized within a 33 kb region of scaffold 116. However, arguing against an important role for AAO and AAD, their transcripts are not (or only slightly) increased in cultures containing microcrystalline cellulose relative to glucose-grown cultures (SI_table_6.xls under GEO accession GSE12540).

Eight additional GMC oxidoreductases with similarity to known glucose-1-oxidase genes were identified, and manually inspected alignments showed conservation of key residues (31) in models Ppl44331, Ppl63190, Ppl128558, Ppl128830 and Ppl108489. The latter sequence is most closely related to the glucose-1-oxidase-like protein detected in cellulose-grown *P. chrysosporium* cultures (23). Transcript levels of the genes encoding Ppl128830 and Ppl108489 were relatively high in both glucose- and cellulose-grown cultures and their corresponding peptides were detected in extracellular filtrates (SI_table_6.xls and SI_table_11.xls under GEO accession GSE12540). Pyranose-2-oxidase has been implicated in lignocellulose degradation in *P. chrysosporium* (32), but no orthologs were detected in *P. placenta*.

Genes encoding copper radical oxidases (CROs), catalytically distinct from the FAD-binding GMC oxidoreductases, are present in the *P. placenta* genome. One such CRO enzyme from *P. chrysosporium*, glyoxal oxidase, is physiologically coupled to the peroxidases of *P. chrysosporium* via peroxide generation. Substrates include small carbonyl and α-hydroxycarbonyl compounds (one to three carbons long) that could be derived from oxidative fragmentation of lignocellulose. On the basis of overall sequence similarity to glyoxal oxidase and conservation of five catalytic residues (33, 34), three models were identified as copper radical oxidases. Protein models Ppl64380, Ppl56703, and Ppl130305 have 20-22 amino acid secretion signals, and their deduced amino acid sequences are most closely related to *P. chrysosporium* genes *cro6*, *cro3* and *cro2*, respectively (35). Ppl130305 is also similar to the membrane-bound *U. maydis* protein GLO1, a protein involved in filamentous growth and pathogencity (36). *P. chrysosporium cro2, U. maydis* GLO1, and two hypothetical proteins each from *L. bicolor* and *C. cinerea* grouped together based on phylogenetic analysis (Fig.. 2 in main paper), and all members of this clade feature carboxy terminal transmembrane helices. Like its *P. chrysosporium* orthologs *cro3, cro4* and *cro5*, the N-terminal sequence of Ppl56703 contains direct repeats of a conserved WSC domain (InterPro domain IPR00288). The function of this domain is unknown but may be involved in carbohydrate binding. Of particular relevance to potential Fenton systems, the genes encoding Ppl56703 and Ppl130305 are upregulated in cultures grown on microcrystalline cellulose, and Ppl56703 peptides were detected in aspen-grown cultures along with the abovementioned glucose-1-oxidase-like proteins (SI_table_6.xls under GEO accession GSE12540). As mentioned in the main paper, the *P. chrysosporium cro1* and *glx1* genes are not closely related, and they do not have orthologs in *P. placenta,* which suggests that there have been two independent losses of these CRO lineages in *Postia* (Fig. 2 in main paper). Ppl56703 is orthologous to the *cro3-4-5* lineage in *P. chrysosporium*, which therefore represents a *Phanerochaete*-specific expansion of the gene family. As in the case of the GHs, evolution of brown-rot is associated with a reduced diversity in *P. placenta* CROs

relative to *P. chrysosporium*.

3. Iron reduction and homeostasis.

Several genes potentially involved in the Fe(II) production required for Fenton chemistry were found (SI Table 5 and SI_table_7.xls under GEO accession GSE12540). One of these (present in 2 allelic variants, with protein IDs Ppl124517 and Ppl64069) encodes a putative quinone reductase (QRD; EC 1.6.5.5). In the brown-rot fungus *G. trabeum*, a QRD may drive extracellular Fenton systems via redox cycling of secreted fungal quinones (37). *P. placenta* produces extracellular hydroquinones (38) and its QRD sequence is highly similar to four *P. chrysosporium* genes (39) and to at least two genes from *G. trabeum* (40). Transcription of the *P. placenta* QRD gene was significantly upregulated in cellulose medium (SI_table_3.xls under GEO accession GSE12540), which is consistent with a role for cellulolytic Fenton chemistry that is driven by quinone redox-cycling. In this connection, substantial upregulation of the genes encoding phenylalanine ammonia lyase (Ppl112824) and a putative quinate transporter (Ppl44553) may also be relevant by virtue of their respective roles in the biosynthesis and transport of essential quinones.

A hydroquinone-based iron reduction system may not be the only mechanism driving Fenton chemistry in *P. placenta*. Low molecular weight glycoproteins that can act as iron reductases have been hypothesized as a component of extracellular Fenton systems of *G. trabeum* (41), *P. chrysosporium* (42) and various other lignocellulolytic fungi. Four *P. placenta* models show significant similarity (>48% amino acid identity) to *P. chrysosporium* glycoproteins encoded by *glp1* (AB236889) and *glp2* (AB236890) (42). The *P. placenta* genes are predicted to encode proteins with well-defined secretion signals. The genes corresponding to protein models Ppl128974 and Ppl28976 are tandemly oriented on scaffold 183, and the former is significantly upregulated on microcrystalline cellulose medium (SI_table_7.xls under GEO accession GSE12540). Models Ppl95887 and Ppl95888 lack EST support and the corresponding transcripts show slight decreases in transcript levels under cellulose induction. A gene encoding another fungal protein often implicated in Fe^{3+} reduction, cellobiose dehydrogenase (37, 43), appears to be absent in *P. placenta*.

In addition to its pivotal role in a wide range of cellular processes, iron homeostasis must play a central role in modulating any functioning Fenton system. The *P. placenta* genome features numerous genes potentially involved in iron transport and redox state (SI_table_7.xls under GEO accession GSE12540). Generally, iron homeostasis is generally divided into reductive and non-reductive systems (44-46).

3.1 Reductive pathway

3.1.1. Iron reductases. The manual inspection of the *P. placenta* genome database allows the identification of at least seven putative ferric reductase-encoding genes, and those encoding proteins Ppl98469 and Ppl98470 are tandemly arranged. The gene encoding ferric reductase protein Ppl134648 lies immediately adjacent to a putative copper transporter gene (protein model Ppl134650), an arrangement apparently conserved

in the basidiomycetes *P. chrysosporium, L. bicolor* and *U. maydis*. Such ferric reductases could play a direct role in Fe^{2+} generation for Fenton reactions through the reduction of Fe^{3+} chelates.

Interestingly, clustering can be observed in many of the aforementioned gene models. A clear example lies on Scaffold 52, where 7 out of 8 gene models encoding for putative iron reductases are located. In *U. maydis*, only one iron reductase has been described, while in *S. cerevisiae* one siderophore reductase (FRE 3) and 2 other genes (FRE 4 to 5) probably encode for iron reductases (47, 48). BLASTP searches of the *Phanerochaete* genome revealed at least 11 gene models automatically annotated as iron reductases, all of which show a considerable similarity to the *P. placenta* sequences.

3.1.2. CTR copper transporters. At least one copper transporter gene was identified by sequence similarity to the corresponding yeast gene (49, 50). This putative CTR copper transporter gene lies immediately adjacent to an iron reductase gene on scaffold 171 (SI Table 7). Proteins models for the transporter (Ppl134650) and reductase (Ppl134648) were supported by EST evidence and manually annotated. A second putative CTR copper transporter protein model, Ppl89573, lacks EST support and was not corrected. Typical of CTR-related proteins, Ppl134650 contains 3 transmembrane domains and a GXXXG motif in the third transmembrane domain.

The significance of clustering between the putative iron reductase and CTR copper transporter remains uncertain. Interestingly, the same kind of "metal-gene cluster" is observed in *P. chrysosporium*, where the pairing is observed on two separate scaffolds (numbers 7 and 9). While *Ustilago* also shows a CTR clustered next to a metalloreductase encoding gene, the same was not observed in *C. cinerea.*

3.1.3. Fet3 sequences. Fet3 proteins are multicopper oxidases (MCO) involved in iron homeostasis. Functionally coupled with iron reductases, they oxidize iron (II), which then enters the cell through the Ftr1 iron permease. Model Ppl109824 is a typical ferroxidase exhibiting substantial similarity to known Fet3 proteins and possessing a C-terminal transmembrane domain. The *P. placenta* Fet3 is among the most highly upregulated genes, showing a 5.4-fold increased accumulation in microcrystalline cellulose relative to glucose medium (SI_table_3.xls under GEO accession GSE12540).

3.1.4. Ftr1 sequences. Ftr1 genes encode a membrane-bound iron permease that is critical for iron homeostasis. Transcription and translation occurs in coordination with its functional partner, Fet3. Two protein models, Ppl118556 and Ppl4639, were identified by BLASTP searches with *P. chrysosporium* Ftr1. Both Ftr1-like proteins contain 7 transmembrane domains. The Ftr1 gene encoding protein model Ppl118556 is upregulated in microcrystalline cellulose (3.5-fold). The gene lies downstream of the highly upregulated Fet3 gene (protein model Ppl109824) (SI_table_7.xls under GEO accession GSE12540), an arrangement commonly seen in ascomycete and basidiomycete fet3-ftr1 genes. In contrast, Ftr1 gene (protein model Ppl46394) is slightly downregulated under these same culture conditions.

3.1.5. Other sequences. ATX1 from *S. cerevisiae* is a copper chaperone that plays an important role in the transport of copper to the Golgi apparatus (51). It is a key protein for the acquisition of copper by the Fet3 multicopper oxidase, and thus, it is also implicated

in iron homeostasis. For example, although it plays a role in copper transport, the ATX1 gene is regulated by iron availability. Protein model Ppl134640 is distantly related to *S. cerevisiae* ATX1 (SI_table_7.xls under GEO accession GSE12540). Microarray experiments showed a 1.459-fold increase of the transcripts in microcrystalline cellulose medium relative to glucose medium.

Model Ppl113226 is highly similar (74% amino acid identity over full protein length) to *T. versicolor* ctaA (52), a P-type ATPase involved in copper trafficking. In yeast, the Ccc2 ortholog has been shown to be involved in the iron homeostasis tightly linked to ATX1. CCC2 participates in the transport of copper inside the Golgi apparatus. It is required, as well as ATX1, for copper insertion into Fet3 proteins. CCC2 takes copper from ATX1 (53). Although CCC2 plays an important role in copper transport, the CCC2 gene is also regulated by iron availability (54). We observed a modest accumulation of *P. placenta* transcripts in microcrystalline cellulose (SI_table_7.xls under GEO accession GSE12540).

3.2. Non-reductive pathway

3.2.1. *Sid1*-like genes. The siderophore-iron transporters are major components of non-reductive iron uptake pathways. Sid1 (L-ornithine-N5-monooxygenase) in *U. maydis* encodes for the first committed step (first enzyme) in the hydroxamate siderophore synthesis (55). Orthologs of this gene have been identified in several fungal species, but no gene model was identified in the *Postia* genome. Sid-like non-ribosomal peptide synthetases (NRPS) may include *Postia* protein models Ppl42034, Ppl54642, Ppl95930 and Ppl127746.

3.2.2. Fer-like sequences from *U. maydis*. To date, three fungal peptide synthetases genes involved in siderophore biosynthesis have been functionally characterized: Sid2 from *U. maydis*, SidC from *A. nidulans* and Sib1 from *S. pombe*. In addition, sequences with homology to these siderophore peptide synthetases have been identified in several genomes. These proteins have, in general, lengths close to 4500 aa, and multiple core motifs (composed of e.g. adenylation, peptidyl carrier and condensation domains), which are characteristic for NRPS giving them a modular structure (45).

When Sid2 or Fso1 from *Omphalotus olearius* were BLASTed against the genome, several models showed homology with these sequences. Nevertheless, all encoded putative proteins of about only 1000 aa. For example, protein model Ppl111174 gives the highest E value for Sid2, but it is only 1019 amino acid residues. The model is supported by EST data at the C-terminal region, and NCBI BLASTP revealed that the most closely related proteins were of similar length but designated 'hypothetical' (e.g. EAU85234).

Thorough studies conducted in *U. maydis* have helped to characterize the molecular elements responsible for the non-reductive uptake of iron in fungi. In this organism the *sid1* and *sid2* genes responsible for siderophore synthesis are clustered, while other genes involved in the transport and modification of these molecules are located in a different chromosome but also clustered (47). For example, 8 *fer* (Fe

regulated) genes, potentially involved in the production and transport of ferrichrome A, are clustered. All these genes have been BLASTed against the *Postia* genome and several models were identified. Nevertheless, no clear evidence of clustering of some of these models has been observed. In *O. olearius* the siderophore peptide synthetase is tightly clustered with an acyltransferase and a L-ornithine-N5-monooxygenase (the *sid1*-like gene not identified).

Two siderophores (ferrichrome and ferrichrome A) have been biochemically characterized in *U. maydis*. Ferrichrome-complexed iron is taken up and is believed to serve as an internal reservoir for this metal, while ferrichrome A is thought to present the iron to a membrane-bound iron reductase. Thus, only some of the siderophore-mediated iron uptake is conducted through transporters, while the other one also involves a reduction step (47).

One of the genes present in an "iron-responsive cluster" in *U. maydis* corresponds to *fer8*, which encodes for a conserved protein of unknown function. Interestingly, the *fer8* homolog of *P. placenta* (protein model Ppl55666: 31% identity) is slightly upregulated in microcrystalline cellulose medium (1.26 up). In contrast, the *fer6* homolog (protein model Ppl112613) shows no significant increase in transcript levels.

Thus, although biochemical studies have suggested the presence of siderophores in *P. placenta* (56), no clear homologs of the *U. maydis sid1* or *sid2* (directly involved in siderophore biosynthesis) could be identified in *P. placenta*. While the brown rot *G. trabeum* is one of the few exception in which the isolated siderophores are not hydroxymate-based (57), the nature of the putative iron siderophores has not been yet explored in *P. placenta*. It is plausible then, that phenolates or organic acid-iron complexes could also be biologically relevant as part of the high affinity iron-uptake system in the latter organism.

4. Lignin modification.

Genes encoding the class II secretory peroxidases lignin peroxidase (LiP) and manganese peroxidase (MnP) were not detected in the *P. placenta* genome, although they are typically found in lignin degrading white-rot fungi such as *P. chrysosporium*, which contains ten LiP genes and five MnP genes (SI Table 5). Model Ppl44056 showed over 52% amino acid sequence identity to a *Pleurotus ostreatus* manganese-dependent lignin peroxidase isozyme (MnP3) and to a *Coprinopsis cinerea* peroxidase (CIP; EC 1.11.1.7), but it lacks residues involved in Mn binding and oxidation of aromatic compounds (58), and superimposition of models strongly suggests that Ppl44056 corresponds to a low redox potential peroxidase (SI Fig. 1). Consistent with this structural evidence, phylogenetic analyses of class II peroxidase genes from *Postia, Phanerochaete*, and other fungal genomes suggest that Ppl44056 is not closely related to LiP and MnP, but is part of an assemblage of "basal peroxidases" that includes the "novel peroxidase" of *P. chrysosporium* (59), the *C. cinerea* CIP, and a peroxidase from the mycorrhizal mushroom *L. bicolor* (Fig. 2 in main paper) (60). The backbone of the class II peroxidase phylogeny is not strongly supported, but optimal trees suggest that the LiP and MnP gene lineages of *P. chrysosporium* were independently derived from the basal peroxidases prior to the divergence of *Postia* and *Phanerochaete*. If so, then the absence of LiP and

MnP in *P. placenta* may reflect additional instances of gene loss.

Five *P. placenta* models contained chloroperoxidase (CPO) domains (IPR000028), one of which (Ppl93549) features a putative transmembrane helix and is substantially upregulated in cellulose medium (SI_table_7.xls under GEO accession GSE12540). These sequences were distantly related to the CPO of the ascomycete *Leptoxyphium fumago* (P04963) and more closely related to homologs from the basidiomycetes *P. chrysosporium* and *Agaricus bisporus*. Chlorination of lignin and lignin-related aromatic compounds has been attributed to CPOs from a wide range of fungi (61), and all five predicted *P. placenta* CPOs showed sequence similarity to peptides of *Agrocybe aegerita* haloperoxidase-peroxygenase, an enzyme capable of hydroxylating aromatic substrates (62). Finally, model Ppl116475 was closely related to cytochrome-*c* peroxidase (EC 1.11.1.5) of *S. cerevisiae*, and contains a mitochondrial targeting sequence (TargetP, http://www.cbs.dtu.dk/services/TargetP/).

Laccases have been suggested to play a role in lignin modification by white-rot fungi, but have not previously been demonstrated in brown-rot fungi. The precise role of these enzymes remains uncertain, but numerous studies have demonstrated laccase-catalyzed oxidation of phenolic and nonphenolic lignin model substrates particularly in the presence of low molecular weight mediators. The results from *P. placenta* belie the usual picture of brown-rot in that models Ppl62097 and Ppl111314 are likely laccases *sensu stricto* (63) (Fig. 2 in main paper), showing 62% and 58% identity, respectively, to *T. villosa* laccase 5 (SwissProt 18281739). Transcript levels of the genes encoding Ppl89382 and Ppl111314 appear differentially regulated by decreasing slightly (-1.08-fold) and increasing (+2.29-fold), respectively, on cellulose medium relative to glucose medium (SI_table_7.xls under GEO accession GSE12540). These enzymes could contribute to hydroxyl radical generation by quinone redox-cycling, the semiquinone radicals formed being able to reduce ferric iron and activate oxygen to hydrogen peroxide (64). Interestingly, laccase genes are absent from the genome of *P. chrysosporium* (39, 63), suggesting that laccase is not a core component of fungal wood decay mechanisms, and is certainly not essential for white-rot.

The *P. placenta* genome also contains a single "ferroxidase/laccase" sequence (63), first described in *P. chrysosporium* and, later identified in almost all fungal genomes. Like other members of this group, Ppl47589 contains residues required for ferroxidase activity but lacks the characteristic C-terminal transmembrane domain of Fet3. Based on the significant similarity to *P. chrysosporium* MCO1 (>53% identity) (65), strong ferroxidase activity but weak phenol oxidase activity would be expected for the Ppl47589 protein.

Other upregulated genes potentially involved in quinone redox-cycling (see previous section), and oxidation of lignin-derived products include those encoding a "polyphenol oxidase" (Ppl114245), i.e. tyrosinase (EC 1.14.18.1) or catechol oxidase (EC 1.10.3.1) related to typical laccases, and various oxidoreductases of uncertain function (Ppl107061, Ppl28682, Ppl34850, Ppl61437, Ppl24981) (SI_table_3.xls under GEO accession GSE12540).

5. Regulatory aspects of cellulose degradation by *P. placenta*

Carbon catabolite repression is a central mechanism by which microbes adjust gene expression in relation to the complexity of degradation of a carbon source. In the model ascomycete *Aspergillus nidulans*, this regulatory circuit is known to involve the C2H2-type zinc finger transcriptional repressor CreA; the C-terminal deubiquinating peptidase CreB; the WD40 repeat protein CreC; and the arrestin domains and PY motifs containing protein CreD (66, 67). In addition, the serine/threonine protein kinase SnfA has been shown to be necessary for activation of CreA (68). Putative orthologs of all these genes could be identified in the genome of *P. placenta,* thus suggesting that carbon catabolite repression is also likely to occur in this (and other) basidiomycete(s). In support of this, we detected the consensus element for binding of CreA (5'-SYGGRG-3') in several of the promoters of genes encoding glycosyl hydrolases, although - in contrast to ascomycetous genes – in most cases only as single sites. Putative orthologs include CreA, Ppl101488; CreB, Ppl63810; CreC, Ppl105086; Ppl90866; CreD, Ppl130178 and Snf1, Ppl49893.

Transcriptional regulation of cellulase and hemicellulase formation has been best studied in the ascomycete *Hypocrea jecorina* (=*T. reesei*), where it involves the binding of two transcriptional activators (XYR1, ACE2) and one transcriptional repressor (ACE1) (69). An ortholog of XYR1 (=XlnR) has also been shown to function in *Aspergillus niger*, whereas orthologs for *ace2* are absent from other fungal genera, and *ace1* – while present in other genera – has not yet been studied. A screening for the presence of genes orthologous to *xyr1, ace1* and *ace2* in *P. placenta* failed to detect any candidates, and this absence is also supported by the lack of detection of consensus binding sites for XYR1/ACE2 (5'-GGNTAA-3') in 12 promoters of glycosyl hydrolases which were upregulated more than 2-fold on cellulose (SI_table_3.xls under GEO accession GSE12540). The 12 upregulated GH-encoded genes encode proteins Ppl105490, Ppl112941, Ppl59600, Ppl54405, Ppl105534, Ppl116267, Ppl109743, Ppl129476, Ppl121713, Ppl128099, Ppl135021 and Ppl90501.

A comparison of 1000 bp of the same genes by RSA-tools (http://rsat.ulb.ac.be/rsat) identified the consensus sequence 5'-GCNTNA-3' to be present in all of them with an average of four copies per promoter. We conclude that *P. placenta* has likely developed different transcriptional proteins and DNA-binding motifs for the regulation of its cellulose degrading genes.

6. Mating system.

Sexual development in basidiomycetes is governed by two types of homeodomain transcription factors with distinct DNA-binding domains (HD1 and HD2) and by a pheromone and pheromone receptor system (70). In tetrapolar species, these functions are encoded in two unlinked mating-type loci (mostly called *A* and *B*), whereas in bipolar species these two classes of genes are either closely linked in one inheritable mating-type locus (e.g., (71)) or only one of the two loci has retained mating type function (e.g., (72)). *P. placenta* is reported to have a bipolar mating system (73). Searches with mating type proteins from other Agaricomycete species such as *C. cinerea* or *L. bicolor* revealed a candidate *B* mating type locus with pheromone and pheromone receptor genes (see section on G-protein coupled receptors below) but no *A* mating type locus with genes for

homeodomain transcription factors. In other species, the *A* mating type locus has been shown to be flanked by a *mip* gene for a mitochondrial intermediate peptidase and by a gene *β-fg* for a conserved fungal protein of unknown function (74-76). Allelic copies of a *mip* gene and a *β-fg* gene were found in the *P. placenta* genome at a distance from each other of about 7 kb on scaffold 8 and about 20 kb on scaffold 166, respectively. Sequence analysis of the enlarged region in scaffold 166 suggests that the larger size of the locus is caused by an insertion of a retrotransposon of a type occurring in multiple copies in the *P. placenta* genome. Insertion of a transposon has not been reported previously in any *A* mating type locus. Indeed, the *A* mating type locus and the surrounding regions in *L. bicolor* and *C. cinerea* appear to be somehow protected against such intrusion (76).

Ab initio models in the regions in between the *mip* and the *β-fg* alleles of *P. placenta* did not resemble *HD1* or *HD2* genes from other fungi. Because of low sequence conservation between genes of species and between alleles of the same species, in *L. bicolor* and *C. cinerea* various *ab initio* models of *A* mating type genes were incorrectly annotated or overlooked (76). Therefore, we tested the sequences in between the *mip* and *β-fg* genes for the presence of homeodomain transcription factor genes by translating the DNA into protein sequences and manually searching for introns using positional information from genes of other species.

Using this approach, one *HD2* gene was found in both allelic loci (scaffolds 8 and 166) of *P. placenta* but the absence of overlapping EST sequences leaves the 3′ ends of the genes uncertain. The N-termini of the protein models Ppl135105 (652 aa) from scaffold 166 and Ppl135106 (508 aa) from scaffold 8 have over a length of about 300 - 400 aa about 40% similarity to the N-terminal regions of HD2 proteins from *C. cinerea*, *Coprinellus disseminatus* and *Pleurotus djamor*. This level of similarity is common between *A* mating type proteins of different species and also between allelic proteins of the same species (77). Regions of similarity cover the HD2 homeodomain and the N-terminal domain determining allele specificity of mating type proteins (78).

Evidence for sequences related to known *HD1* genes of other basidiomycetes was also found in both putative *A* mating-type loci in between the *HD2* genes and the *β-fg* genes in the divergently transcribed arrangement typical of *HD1-HD2* gene pairs in the basidiomycetes (79). We were however unable to deduce from the available sequences of the two scaffolds complete protein models containing a HD1 homeodomain in their N-terminal half as found in HD1 transcription factors of other species. Nevertheless, translated sequences from the 5′end of the proposed genes align perfectly over the length with up to 55 % similarity to the N-terminal domains of HD1 proteins of *L. bicolor* and *C. cinerea*, the region responsible for protein specificity (78). The translated 3′ sequences also align over large regions with 40 – 48 % similarity with the C-terminal sequences of the *L. bicolor* and *C. cinerea* proteins. These C-terminal regions have been shown to function in *C. cinerea* as transactivation domains (78). Currently, we do not know whether the genes encoding HD1 homeodomains in *P. placenta* are degenerate or functional. In the future, sequences of the putative *A* mating-type locus will have to be verified and mating type function tested. Such experimental verification was reported for the bipolar species *Coprinellus disseminatus* by progeny analysis and by functional analysis of subcloned genes in heterologous transformable hosts such as *C. cinerea* (72).

Finally, one interesting observation regarding the putative *A* mating type locus of *P. placenta* concerns the relative arrangement of mating type genes in comparison to *mip*

and *β-fg*. In all published *A* mating type loci of the Agaricales, the *mip* gene situates tail-to-tail next to an *HD1* gene (74, 76). As in *P. chrysosporium* (unpublished observation), the *P. placenta* the *HD2* gene is positioned tail-to-tail with the *mip* gene (ignoring in case of scaffold 166 the inserted retrotransposon). Possibly, there has been an inversion of the locus since the divergence of the Agaricales and Polyporales.

With regard to the *B* mating type locus in tetrapolar species (76, 80), there is one locus within the haploid genome of *P. placenta* that contains several STE3-like pheromone receptor genes and mating type pheromone genes similar to the a-type pheromone precursors of *S. cerevisiae* (SI Fig. 5).

The genome includes linked genes for STE3-type G-protein coupled pheromone receptors (GPCRs) on allelic scaffolds 32 (6 genes) and 33 (7 genes) (SI Fig. 5, panel A). These likely present a locus of *B* mating type function in mating and/or sexual reproduction (compare the analysis of the *B* mating-type-like genes in the bipolar species *C. disseminatus*; (72)). The clustering of STE3 GPCRs has also been observed in *C. cinerea* (81), *Laccaria bicolor* (76), and *Phanerochaete chrysosporium* (39). Although this pheromone receptor gene cluster is likely to function in mating, its presence in bipolar species, such as *P. chrysosporium* and *P. placenta*, suggests that the genes may not be part of the mating type locus. Two additional STE3-like GPCRs, which most likely represent allelic variants, have been detected on scaffold 39 (protein models Ppl43609 and Ppl126003). Comparison with GPCR representatives from the three paralogous *B* mating type gene groups in the *B* mating type locus of *C. cinerea* suggests that the GPCRs in *P. placenta* are ancestral duplications of proteins with demonstrated function in mating (SI Fig. 5, panel B).

B mating type function in *C. cinerea* is governed by three paralogous clusters of one gene for a pheromone receptor and various genes for pheromones (80). Phylogeny of all STE3-type GPCRs of *P. placenta* along with orthologs in *C. cinerea* (SI Fig. 5, panel B) revealed that the STE3-type GPCRs of *P. placenta* predominantly belong to the three independent subfamilies of STE3-like receptors known to be involved in mating in *C. cinerea* (80). The allelic pair corresponding to protein models Ppl127842 and Ppl94660 could not be grouped into the abovementioned subfamilies. Therefore, the products of these genes could represent either new subtypes of receptors for mating type determination or receptors not involved in mating. Evidence for GPCR genes of non-mating type function are available in other Agaricomycetes (72, 74, 76). Strikingly, the pheromone receptor genes on scaffold 39 have no pheromone genes in close vicinity (3 kb up- and downstream). EST mapping (www.jgi.doe.gov/postia) and microarray analysis (GEO accession GSE12540 and SI Fig. 6) demonstrated transcription of all these GPCR genes, except the gene encoding Ppl94657. (Microarray design did not include the Ppl94657 as a target.)

Phylogenetic analysis of the predicted lipopeptide pheromone precursors of *P. placenta* (SI Fig. 5, panel C) reveals the corresponding allelic variants for some of these pheromone precursors and suggests possible duplications within the same locus (SI Fig. 5). In all cases except the gene encoding Ppl94657, putative pheromone precursor genes are located in close proximity to the pheromone receptors. Besides pheromone genes for products showing characteristic features of lipopeptide pheromone precursors of other Agaricomycetes (similarity of the outermost N-terminal precursor sequence, protease processing site, CaaX-motif), we also detected genes for small CaaX-domain proteins

within this locus which are more distantly related (protein models Ppl135125, Ppl135029, Ppl135044 and Ppl135025).

7. Signal transduction.

Based on *in silico* analysis of the genome, *P. placenta* appears to sense its environment in a manner profoundly different from what is known from other fungi (82). The complexity of G-protein coupled receptors is relatively low. However, the characteristics of the members of the two component signal transduction systems suggest that this signaling mechanism might at least partially have compensated for the lack of G-protein complexity. Also, the increased number of members of the MAPkinase cascades suggests a higher flexibility within these signaling cascades relative to other fungi.

7.1. G-protein coupled receptors

The most conspicuous finding is that one type of G-protein coupled receptor (GPCR) has predominantly been identified in the genome of *P. placenta*; specifically, the STE3-type pheromone receptors located within the *B*-locus (scaffolds 32 and 33) and additionally on scaffold 39 (see above). Otherwise, only one further candidate with distant similarity to mPR-like GPCRs was detected, protein model Ppl127104 (allelic variant =Ppl130083).

7.2. Heterotrimeric G-proteins

Interestingly, the *P. placenta* system of heterotrimeric G-proteins seems typical among fungi (83), with three G-α subunits (Ppl122782, Ppl5128 and Ppl113593), one G-β subunit (Ppl117250) and one G-γ subunit (Ppl59950).

In addition to the three G-α subunits, which share considerable homologies to other fungal sequences, 12 additional models with the characteristic G-α protein subunits were detected. All showed highest similarity to the *U. maydis* G-protein α subunit Gpa4 (84). The function(s) of this G-α subunit are unclear.

Phylogenetic analysis of the putative G-α subunits of *P. placenta* confirmed the prediction of the three G-α subunits with homologs in other fungi. However, the additional putative G-α subunits clustered outside of this group and most of them only with low bootstrap support even to the *U. maydis* G-α protein (SI Fig. 6).

As with *U. maydis*, all these proteins show alterations in both the GTP binding domains as well as the GTPase domains. In one case, Ppl118240, the respective alterations include exchanges of glycine for alanine (both neutral, uncharged amino acids) within the GTP binding domain and glutamine for serine (both uncharged, polar residues) within the GTPase domain. It therefore seems likely that this protein functions as G-α subunit.

Also detected were five WD40 repeat proteins with considerable similarity to the *Podospora anserina* vegetative incompatibility protein HET-E2C (Ppl57682, Ppl57670, Ppl128228, Ppl54617 and Ppl25239). Intriguingly, expansions of the number of these

proteins might have taken place, since loci with duplicated genes have been found on at least three scaffolds. Considering that the abovementioned group of pheromone receptors has also likely expanded, processes related to mating may be of particular importance to *P. placenta*. Further analysis along these lines will require careful manual annotation and cDNA sequencing to improve model confidence.

7.3. MAPKinases

The MAPKinase machinery of *Postia* seems comparable to other fungi (5, 85), although there might be more than three pathways. However, these pathways do not correspond to the five pathways detected for *S. cerevisiae*. An STE7 homolog (MAPKK, pheromone response MAPK pathway) was not detected (SI Fig. 7).

Three FUS3-homologs, predicted to have a function in pheromone response (86), were detected; an observation again in accordance with an increased importance of this process in *P. placenta*. Also, the other MAP-kinase pathways are likely to be extended, thereby enhancing the flexibility of this signaling mechanism. Moreover, the two component phosphorelay pathway likely to act upstream of the osmosensing HOG1-pathway comprises 4 potential group VI hybrid histidine kinases, two histidine phosphotransferases and the response regulator homolog Ppl134813. Thus, the mechanisms acting upstream of the MAP-kinases also reflect the increased complexity of this branch of the signal transduction cascade.

7.4. Two component phosphorelay systems

Signaling via two component phosphorelay systems in response to an environmental signal is initiated by ATP-dependent autophosphorylation of the histidine kinase (HK) at a conserved histidine residue. This phosphate is transferred to a conserved aspartic acid within a response regulator (RR) domain, which ultimately causes a change in transcription of the respective target gene or regulation of a mitogen-activated protein kinase pathway (87). Two different types of two component signaling have been described: with the simple histidine kinases the sensor histidine kinase and the regulator receiver are separate proteins, whereas hybrid histidine kinases contain both HK and RR domains on the same protein and generally require additional rounds of phosphorelay through a histidine phosphotransferase (HPt) and another RR protein. To date, only hybrid histidine kinases have been identified in fungi, and in general are predominant in eukaryotes, while simple HKs are predominant in prokaryotes.

The *P. placenta* genome contains 6 predicted histidine kinases, all of them hybrid histidine kinases, 3 predicted response regulator receiver proteins and a histidine phosphotransferase. Summarizing on the basis of the classification system of Catlett *et al.* (88):

Type	ID	Group	Transmembrane

			domains
Histidine kinase	Ppl101025	I	5
	Ppl89890	III	none
	Ppl106926	VI	5
	Ppl126760	VI	5
	Ppl51297	VIII	none
	Ppl128265	X	none
Response regulator receiver	Ppl107196	related to Rim15p	
	Ppl127182	related to Skn7p	
	Ppl134813	related to Ssk1p	
Histidine phosphotransferase	Ppl45959		

The histidine kinases reveal several interesting features. The study of Catlett *et al.*, (88) described a core set of histidine kinases found in every euascomycete analyzed (groups III, V, VI, VIII, IX and X). The *P. placenta* genome has only four groups (III, VI, VIII, and X). Additionally, one member of group I, the function of which has not yet been determined, has been detected. Interestingly, 3 of the 6 histidine kinases of *P. placenta* can be predicted to have a function in osmosensing (members of groups III and VI). In group VI, two putative members have been detected, which comprise partial MYHT-domains (IPR005330) as well as 5 transmembrane domains. These MYHT-domains are thought to function as sensor domains in bacterial signaling proteins. They consist of six predicted transmembrane domains. The MYHT domain has been found in several phylogenetically distinct bacteria, either as separate, single domain or in combination with other domains, similar to the signaling histidine kinase domains for the two proteins described here. Thus, this appears be the first report of MYHT domains in fungi. Moreover, this finding would provide an explanation for the lack of GPCRs other than pheromone receptors. Since both group VI histidine kinases can be predicted to have transmembrane domains, these proteins could fulfill the task of sensing extracellular stimuli instead of GPCRs. This group of histidine kinases is also likely to target the HOG1-MAP kinase cascade, which is more complex and interesting than expected. While the function of the remaining group X histidine kinase is not known, the group VIII histidine kinase is predicted to represent the phytochrome of *P. placenta* and thus indicates that this fungus is able to perceive and react to red light. With regard to the response regulator receiver (RR) proteins, *P. placenta* contains homologs to all three expected RR proteins (Ssk1p, Skn7 and Rim15p). Finally, our initial analyses identified a putative histidine phosphotransferase gene (protein model Ppl45959 and allelic variant Ppl55453). A closely related model (Ppl124576), lacking a clear allele, was located as sole gene model on scaffold 2623. Although multiple histidine phosphotransferases are known in *E. coli* and *A. thaliana* (88), fungi typically contain a single gene. If both genes are confirmed as functional, this may indicate a more sophisticated relay network in *P. placenta*. Firm conclusions regarding the function of these putative signalling proteins would require additional study.

7.5. PAS domain proteins/Light response

In order to react to the harmful effects of light, *P. placenta* apparently relies on a single DNA photolyase (Ppl119113). It does not contain cryptochromes. A potential homolog to *A. nidulans* protein veA (89), involved in red light sensing and numerous other processes, has been identified (Ppl104508), but the similarity between these two proteins is rather weak.

In addition to the putative the phytochrome listed above (Ppl51297), screening the genome for potential photoreceptors yielded a protein with predicted PAS-domain, model Ppl105827. This sequence is distantly related to the *Phycomyces blakesleeanus* photoreceptor White Collar A. Other than Ppl105827, no protein likely to function as photoreceptor was detected. No close homologs of the *Neurospora crassa* PAS-domain proteins White Collar 2 (WC-2) or VIVID were detected by BLASTP. However, TBLASTN queries revealed potential translations distantly related to *L. bicolor* white collar photoreceptor like protein (EDR14883) and *N. crassa* WC-2 (P78714) on scaffold 83 (between coordinates 259975-261292) and in the corresponding allelic region on scaffold 281 (between coordinates 49826-51146). Model Ppl113058 features a PAS domain as well as a zinc finger domain, but since it did not show similarities to characterized proteins, a function could not be assigned.

7.6. Calcium signaling

One calcium calmodulin protein (CAM1) and three closely related sequences (Ppl112896, Ppl129253 and Ppl63783) were detected. Moreover, *P. placenta* has one calcium calmodulin dependent protein phosphatase regulatory subunit (CAR1, calcineurin-family) and at least 6 calcium calmodulin dependent phosphatase catalytic subunits of the calcineurin family.

7.7. cAMP metabolism

cAMP metabolism did not reveal striking peculiarities, except that there was only a weak homolog for adenylyl cyclase (Ppl19012). Two hypothetical adenylyl cyclase associated proteins (Ppl127413 and Ppl95708), one cAMP-dependent protein kinase regulatory subunit (Ppl123625) and three cAMP-dependent protein kinase catalytic subunits were identified.

7.8. Other groups

A β-arrestin domain protein (Ppl130178), may function as a regulator of G-protein coupled receptors. Phosducin-like proteins were not detected.

8. Oxidative phosphorylation (OXPHOS) nuclear genes in *P. placenta*

The OXPHOS system, consisting of the mitochondrial respiratory pathway and ATP

synthase, is the main mitochondrial function and one of the most ancient and conserved functions of all aerobic organisms (90). The OXPHOS system is unique because it contains large enzyme complexes formed by a combination of subunits encoded by both the mitochondrial and nuclear genomes (91, 92). The OXPHOS system is organized in five multipolypeptide complexes embedded in the lipid bilayer of the inner mitochondrial membrane: NADH dehydrogenase (Complex I), succinate dehydrogenase (Complex II), cytochrome *bc1* (Complex III), cytochrome *c* oxidase (Complex IV), and ATP synthase (Complex V). Mitochondrial genomes encode only a limited number of the essential components of Complexes I, III, IV and V (93, 94). Fungal mitochondria frequently have additional OXPHOS components such as an alternative oxidase (AOX) and/or alternative NAD(P)H dehydrogenases (95).

The complete set of nuclear genes coding for mitochondrial proteins that participate in OXPHOS (Complexes I-V, alternative oxidases and alternative NAD(P)H dehydrogenases) have been identified in *P. placenta.* The approach used to identify OXPHOS nuclear genes in *P. placenta* was as described (96). Briefly, detection of orthologs was determined by BLASTP based on the reciprocal best hits of the *P. placenta* nuclear genome against the *S. cerevisiae, N. crassa, C. neoformans* and *Yarrowia lipolytica* OXPHOS proteins used as query. When an OXPHOS ortholog was lacking in *P. placenta,* the nuclear genome sequence was investigated using TBLASTN.

P. placenta contains at least 59 nuclear genes coding for OXPHOS components. In addition to the 7 central subunits making up the eukaryotic core Complex I encoded in the nucleus (NDUFS1, NDUFS2, NDUFS3, NDUFS7, NDUFS8, NDUFV1 and NDUFV2), this fungus has 21 accessory subunits that are nuclear-encoded (NDUFS4, NDUFS6, NDUFA1/MWFE, NUDFA2, NDUFA4, NDUFA5, NDUFA6, NDUFA8, NDUFA9, NDUFA11, NDUFA12/DAP13, NDUFA13/GRIM19, NDUFAB1/ACP, NDUFB7, NDUFB8, NDUFB9, NDUFB11, NUXM, NUWM, NUZM and NI9M). NDUFA4 appears to be a Complex I accessory subunit specific to basidiomycetes and zygomycetes (96). The accessory subunits NURM, NUVM and 10.4 are fungus-specific subunits of Complex I (97) but they are absent from all the available nuclear genomes of Basidiomycetes (*U. maydis, L. bicolour, P. chrysosporium, Cryptococcus neoformans* and *P. placenta*). The accessory subunit NUWM is only present in the basidiomycete *P. placenta* and in three Saccharomycotina species (*P. stipitis, Debaryomyces hansenii* and *Y. lipolytica*).

Nuclear genes coding for cytochrome *c* (CYTC) and the four subunits of Complex II (SDH1-SDH4) are present in *P. placenta.* The core proteins QCR1 and QCR2, the cytochrome c_1 (CYT1), the Ryeske iron-sulfur protein (RIP1) and the additional subunits QCR6-QCR9 of Complex III are also nuclearly encoded in *P. placenta.* Orthologs of 6 nuclearly encoded subunits of Complex IV (COX4, COX5A, COX5B, COX6A, COX6B and COX9) are also present in *P. placenta.* However, *P. placenta* is the unique basidiomycete that lacks a nuclear gene coding for the QCR10 additional subunit of Complex III.

Complex V separates into F_0 and F_1 subcomplexes (98). Four essential subunits of each subcomplex (ATP4, ATP5, ATP7 and ATP17 of subcomplex F_0; and ATP1, ATP2, ATP3 and ATP16 of subcomplex F_1) are nuclearly encoded in *P. placenta.* However, the nuclear genes coding for the essential subunits ATP14 (subcomplex F_0) are absent from all the basidiomycetes. *C. neoformans* and *P. placenta* are the unique basidiomycetes that

lack a nuclear gene coding for the essential subunit ATP15 (subcomplex F_1). In addition, two additional subunits of Complex V (ATP18 and ATP20) are nuclearly encoded in *P. placenta*, and the additional subunits ATP19, STF1 and STF2 are absent from all the basidiomycetes.

In addition, the nuclear genome of *P. placenta* contains an alternative AOX and an alternative NAD(P)H dehydrogenase. It is noteworthy that there is only one NAD(P)H dehydrogenase, whereas in other basidiomycetes there are more (two in *C. cinerea* and three in the others).

Summarizing this analysis, the complete set of OXPHOS nuclear genes have been identified in *P. placenta*. This fungus contains at least 59 OXPHOS components encoded in the nucleus. Comparison of *P. placenta* with other basidiomycetes reveals a high conservation of the OXPHOS system in this fungal phylum. In particular, specific subunits of OXPHOS complexes are absent in all the nuclear genomes of basidiomycetes: the NURM, NUVM and 10.4 accessory subunits of Complex I, the essential subunits ATP14 and the additional subunits ATP19, STF1 and STF2 of Complex V.

9. Oxalate metabolism.

In addition to pH effects on a wide range of enzymes, extracellular accumulation of oxalate by *P. placenta* may affect ferric iron availability and thereby impact hydroxyl radical formation (99) reviewed in (37, 100). A metabolic shunt between the citric acid and glyoxylate cycles is central to oxalic acid accumulation by the brown-rot fungus *Fomitopsis palustris* (101). Key enzyme components are compartmentalized in mitochondria, peroxisomes and the cytosol, and the *F. palustris* gene encoding peroxisomal isocitrate lyase was recently characterized (102). Our analysis of the *P. placenta* genome demonstrates a functional glyoxylate shunt and substantially extends our understanding of the number, structure and transcription of key genes (SI_table_8.xls under GEO accession GSE12540 and SI Fig.2). Based on overall sequence similarity to known genes and on the presence of a C-terminal tripeptide (~SKL) targeting motif, we identified sequences encoding putative peroxisomal enzymes directly involved in the GLOX cycle (malate dehydrogenase Ppl106934; citrate synthase Ppl112712; malate synthase, Ppl119506). All these protein models have EST support, as did a putative glycolate oxidase (Ppl121561) which may provide an additional route to oxalate (103). The other route, experimentally supported in *F. palustris*, involves cytosolic oxaloacetase (Ppl112832), to which over 127 filtered EST reads were mapped. High OXA activity would generate acetate, and consistent with these observations, a highly transcribed cytosolic acetyl-CoA synthetase (Ppl107062) was found. The highly expressed glycolate oxidase and oxaloacetase genes lie approximately 1.3kb apart, and this arrangement appears conserved in *P. chrysosporium*. Separate genes with putative mitochondrial targeting sequences include those encoding malate dehydrogenase (Ppl118718), citrate synthases (Ppl43991, Ppl110627), fumarase (Ppl127563), isocitrate dehydrogenases (Ppl119064, Ppl119048) and 2-oxoglutarate dehydrogenase (Ppl107068). A mitochondrial-targeted aconitate dehydratase with minimal EST support was also observed, as was a putative cytosolic form, Ppl121891, represented by 43 EST reads. Mitochondrial and cytosolic forms of aconitase have been characterized in *S. cerevisiae*,

where they are thought to participate in the TCA and glyoxylate shunt, respectively. Interestingly, the yeast aconitase isozymes are encoded by a single gene, *aco1*, followed by uneven cellular distribution (104).

However, the suggestion that oxalic acid production is due to low levels of 2-oxoglutarate dehydrogenase and isocitrate dehydrogenases in brown-rot fungi (101) is not directly supported by the presence and EST patterns of the corresponding *P. placenta* genes. However, the compartmentalization of the components of the glyoxylic acid shunt and the extent to which TCA cycle components can be transported out of the mitochondria could diminish the overall activity of the TCA cycle and result in the excretion of oxalic acid.

Subsequent metabolism of extracellular oxalate remains unclear. Models Ppl46778, Ppl43912 and Ppl43635 showed high similarity (>51% amino acid identity) to known oxalate oxidases or decarboxylases of the bicupin family. The one glutamate and three histidines that coordinate the manganese ion are conserved in the cupin domains of all three translated genes (105). Oxalate oxidase (EC 1.2.3.4) uses oxygen to convert oxalate to two carbon dioxide molecules and hydrogen peroxide, whereas oxalate decarboxylase (EC 4.1.1.2) cleaves oxalate into formate and carbon dioxide. Based on amino acid sequence alone it is difficult to distinguish the oxalate oxidases from the oxalate decarboxylases. However, the similarity of an important pentapeptide S(E/D)DST following the second cupin motif to known oxalate decarboxylases suggests that the three genes in *P. placenta* all encode oxalate decarboxylases (106). Previous reports suggest these putative oxalate decarboxylases are associated with the mycelium (107). SignalP predicts with high confidence an N-terminal secretion signal of 20-24 amino acids for all of the gene products. Transcripts of the gene encoding Ppl46778 substantially increased in cellulose media relative to glucose (+2.41-fold). Suggesting a physiological connection, genes involved in formate metabolism (putative formate transporter Ppl128726; formate dehydrogenases Ppl129190, Ppl119730, Ppl98518) were also upregulated.

10. Cytochrome P450 monooxygenases.

Cytochrome P450 monooxygenases represent a superfamily of heme-thiolate proteins distributed across phyla. In fungi, P450 proteins act as terminal oxidases of the monooxygenase system and play an important role in the metabolism of a wide variety of endogenous compounds. In filamentous fungi, P450s are part of the secondary metabolism pathways and are involved in the synthesis of biologically active compounds such as aflatoxins and gibberellins. P450s in saprophytic fungi such as wood-rotting basidiomycetes *P. chrysosporium* and *P. placenta*, are believed to be involved in the biodegradation process of lignin, one of the most recalcitrant polymers found in nature, and in the degradation of anthropogenic and xenobiotic compounds such as polycyclic aromatic hydrocarbons (PAHs) and alkanes, among others.

Available genome sequence information on four basidiomycetous fungi has revealed 149 P450 genes in *P. chrysosporium* (39, 108), 101 in *C. cinerea*, 30 in *Puccinia graminis*, and 3 (to date) in *U. maydis*. Here, we report the P450 genes and their alleles in the fifth basidiomycete fungus, *P. placenta* (brown-rot) and discuss the phylogenomic

aspects of its P450ome.

10.1. Methodology for P450 analysis

A total of 326 DNA and protein sequences were downloaded from the JGI website (http://jgi.doe.gov/Postia) that were initially predicted as P450s. A custom Perl script was written that identified the two conserved domains of cytochrome P450 proteins, namely ERR triad motif (EXXR) and the oxygen-binding motif (CXG) to reconfirm their prediction. There were 274 P450s that had these signature domains. A multiple alignment file was generated for the entire set of the 326 P450 protein sequences using the CLUSTALX program with the program default options. Next, a sequence identity matrix file was created from this alignment file using the sequence editor program BioEdit. Sequences were grouped based on a set cut-off percent identity of their aa sequences for family classification. This kind of family grouping was further refined by generating phylogenetic trees using the MEGA 4 software (109). For this, Neighbor-Joining method of clustering was used with 1000 bootstrap replications (110). Sites with alignment gaps and missing data were not included in the analysis. The above two motifs were in the expected relative positions to each other in only 254 P450s, which were finally used as an input for the phylogenetic tree generating program. For the fused P450 proteins (fused P450 monooxygenase and P450 reductase polypeptides) only the P450 portion of the protein sequence was used. Similarly, family/clan names were assigned to the *Postia* P450ome based on their protein homology to the P450s of the white rot fungus, *P. chrysosporium*.

The P450 proteins were grouped based on their aa similarity using three cut-off percentages, 70%, 80% and 90% identity scores. Next, using the mRNA-to-genomic DNA alignment program Spidey, the intron-exon junction and the number of exons were compared for the grouped P450s. In addition, the 100 bp 3' UTR sequences (downstream of the predicted P450 CDS) for each of these P450s were extracted from the whole genome sequence and were aligned to compare their percent similarity. Highly similar pair of P450s based on the above three comparison criteria were defined as allelic pairs only when they had the same structural organization (number and length of exons and introns), belonged to different scaffolds on the genome, and had their UTR regions matching 90-100%.

10.2. Gene numbers.

The 326 gene sequences that were initially identified as "P450-like" based on BLAST homology searches were further analyzed for confirmation and P450 classification based on (i) domain and phylogeny analysis, and (ii) amino acid homology criteria. Based on the domain analysis, only 274 genes were found to have the two conserved functional domains ERR triad motif (EXXR) and the oxygen binding motif (CXG) that are typical of a P450 protein. Of these 274 P450s, only 254 genes showed optimal alignment based on the relative position of the above two functional domains and thus were used for generation of the phylogenetic tree (SI_figure_3.jpg under GEO accession GSE12540)

for family/clan level classification as discussed below. Subsequently, based on the conventional amino acid homology criterion for family classification (genes showing >40% homology belong to the same P450 family), an additional 14 P450 genes including 5 from the 274 list and 9 from the other genes in the original 326 P450-like sequences could also be sorted into P450 families/clans, raising the number of classifiable genes to a total of 268 (254 + 14 = 268). The remaining 15 of the 274 P450s (274-254-5 = 15) could not be sorted out for classification based on the above criteria. In essence, our overall analysis based on a combination of the above criteria confirmed 283 genes (268 + 15) as P450s out of the 326 P450-like gene sequences originally detected in the assembled genome based on BLAST homology search. Of the abovementioned P450 genes, 47 genes have identifiable alleles.

	Allele 1		Allele 2	
	Protein ID	Gene model	Protein ID	Gene model
1	105546	fgenesh3_pg.138__3	91027	fgenesh3_pg.8__29
2	89741	fgenesh3_pm.9__21	87960	fgenesh3_pm.129__4
3	121666	estExt_Genewise1.C_530047	124807	estExt_Genewise1.C_450096
4	52903	e_gw1.65.31.1	112332	estExt_Genewise1Plus.C_940018
5	52821	e_gw1.65.24.1	110751	estExt_Genewise1Plus.C_1630051
6	52780	e_gw1.65.79.1	119475	estExt_Genewise1.C_1630018
7	112203	estExt_Genewise1Plus.C_150097	60038	e_gw1.70.15.1
8	57109	e_gw1.9.31.1	62768	e_gw1.75.15.1
9	44435	e_gw1.112.9.1	62479	e_gw1.95.36.1
10	105719	fgenesh3_pg.10__27	92340	fgenesh3_pg.80__39
11	45371	e_gw1.80.17.1	105726	fgenesh3_pg.10__34
12	48082	e_gw1.193.16.1	56013	e_gw1.2.64.1
13	44960	e_gw1.48.75.1	44724	e_gw1.26.81.1
14	48638	e_gw1.163.64.1	52830	e_gw1.65.118.1
15	89499	fgenesh3_pm.55__7	98989	fgenesh3_pg.245__12
16	20385	gw1.96.11.1	51235	e_gw1.5.44.1
17	126035	estExt_fgenesh3_pg.C_10061	109545	estExt_Genewise1Plus.C_600011
18	123864	estExt_Genewise1.C_3520008	46863	e_gw1.18.87.1
19	100048	fgenesh3_pg.36__76	110469	estExt_Genewise1Plus.C_430068
20	125663	estExt_Genewise1.C_240104	126564	estExt_fgenesh3_pg.C_900031
21	87307	fgenesh3_kg.144__2__3669380:1	22365	gw1.208.7.1
22	97267	fgenesh3_pg.15__57	102772	fgenesh3_pg.171__16
23	97921	fgenesh3_pg.160__15	127097	estExt_fgenesh3_pg.C_3200003
24	126573	estExt_fgenesh3_pg.C_900043	130416	estExt_fgenesh3_pg.C_240078
25	93392	fgenesh3_pg.41__44	60471	e_gw1.4.65.1|Pospl1
26	97505	fgenesh3_pg.94__12	97728	fgenesh3_pg.65__49
27	115931	estExt_Genewise1Plus.C_570065	97538	fgenesh3_pg.94__45
28	89096	fgenesh3_pm.65__14	21309	gw1.107.11.1
29	43853	e_gw1.1.157.1|Pospl1	54877	gw1.162.29.1
30	108714	estExt_Genewise1Plus.C_10138	49051	e_gw1.126.34.1
31	44433	e_gw1.112.8.1	105138	fgenesh3_pg.95__46
32	51753	e_gw1.117.34.1	63327	e_gw1.173.20.1
33	49676	e_gw1.133.4.1	113167	estExt_Genewise1Plus.C_340126
34	51017	e_gw1.154.10.1	35865	gw1.1716.1.1
35	52482	e_gw1.65.83.1	48612	e_gw1.163.17.1
36	48648	e_gw1.163.51.1	52844	e_gw1.65.75.1
37	48656	e_gw1.163.47.1	52906	e_gw1.65.23.1
38	127929	estExt_fgenesh3_pg.C_3090005	114275	estExt_Genewise1Plus.C_1770038
39	56875	e_gw1.27.40.1	94692	fgenesh3_pg.33__90
40	31744	gw1.34.105.1	95454	fgenesh3_pg.133__7

41	55867	e_gw1.36.51.1	48020	e_gw1.43.59.1
42	60489	e_gw1.4.36.1	46717	e_gw1.41.17.1
43	116084	estExt_Genewise1Plus.C_450099	112932	estExt_Genewise1Plus.C_530041
44	95422	fgenesh3_pg.478__1	99629	fgenesh3_pg.142__25
45	89660	fgenesh3_pm.21__12	55398	e_gw1.142.6.1
46	51586	e_gw1.46.29.1	47044	e_gw1.157.11.1
47	93354	fgenesh3_pg.41__6	103594	fgenesh3_pg.4__206

10.3. Gene characteristics.

The average P450 coding sequence size was 1.0 – 1.8 kb; however, for certain gene models such as fgenesh3_pg.281__11, estExt_fgenesh3_pg.C_950043 and estExt_fgenesh3_pg.C_500027, the apparent coding sequence spanned as long as 4-5 kb, and needs further analysis to detect any fused sequences. The number of introns varied widely, ranging from as low as no intron (gw1.18.227.1) or a single intron (e_gw1.168.41.1, fgenesh3_pg.43__49, gw1.36.69.1) to as high as 31 introns (estExt_fgenesh3_pg.C_500027). Most ranged from 7 to 13 introns per gene.

10.4. Homology grouping and phylogenetic analysis (tree construction SI Fig. 3).

We performed an overall sequence similarity-based initial grouping of the 326 *P. placenta* P450 genes with the *P. chrysosporium* P450ome from a previous study (111). This led to the identification of a strong grouping for the core fungal P450 clans CYP51, CYP52, CYP61, and CYP62. This pattern for Clan- and Family- based classification became apparent when the phylogenetic tree was analyzed. In *Postia*, 11 fungal Clans (CYP51, CYP52, CYP53, CYP54, CYP61, CYP62, CYP64, CYP67, CYP503, CYP505, CYP547) and 29 Families were identified. CYP 64 was the largest Clan as in the *P. chrysosporium* genome; a new family CYP530, with multiple members that show high homology to the members in Ascomycetous fungi (*Aspergillus spp.*, *Magnaporthe grisea* and *Fusarium graminearum*), was identifiable as a part of this clan.

P450 Clan family	Member families in each clan (arbitrarily numbered)	Number genes in each family
CYP51	I	2
CYP52	I	8
CYP53	I	1
	II	3
	III	9
	IV	4
CYP54	I	21
	II	2
CYP61	I	2
CYP62	I	2
CYP64	I	39
	II	1

	III	1
	IV	3
	V	3
	VI	17
	VII	14
	VIII	24
	IX	6
	X	22
CYP67	I	5
CYP505	I	3
CYP534	I	5
	II	3
CYP547	I	11
	II	14
	III	4
	IV	36
Unresolved	I	3

10.5. Brown-rot vs White-rot.

Overall, the *P. placenta* genome revealed a higher number of P450 genes (283 genes including 47 alleles) than *P. chrysosporium* (149 genes). Similar to *P. chrysosporium*, clustering of the P450 genes on the *Postia* genome was observed, with a similar distribution pattern. The number of P450s in fungal families such as CYP51, CYP52, CYP61 and CYP62, was the same as *P. chrysosporium*. However, families such as CYP505 (Foxy, fused proteins), CYP58 (CYP53 clan) and CYP67 show reductions compared to *P. chrysosporium*. On the other hand, families such as CYP64, CYP503, and CYP5031 and CYP617 (clan 547) showed expansion when compared to *P. chrysosporium*.

The genes encoding P450 proteins Ppl110015 and Ppl128850 were upregulated 2.49-fold and 2.23-fold, respectively, in cellulose grown cultures relative to glucose grown cells (SI_table_3.xls under GEO accession GSE12540). Ppl110015 shares >66% amino acid identity with CYP53s (benzoate-para-hydroxylase) that are known to be upregulated by benzoate in *P. chrysosporium* and *Aspergillus*. Ppl128850 has limited sequence similarity to *P. chrysosporium* P450s, but is more closely related to members of the CYP503 clan. This clan contains P450 genes associated with secondary metabolism in fungi, including white rot fungus gene CYP512A1 in *T. versicolor* (111). Upregulation of these genes in cellulose and their homology with the P450 genes in the lignin-degrading white rot fungi suggest potential role(s) in lignocellulose conversions.

11. Stress related genes

11.1. Superoxide dismutases.

Superoxide dismutase (SOD) is thought to be an important component of the oxidative stress response system of *P. chrysosporium* (112, 113). Three putative SOD-encoding genes were detected in *P. placenta*. Ppl111797 belongs to the Cu/Zn SOD group (InterPro 001424) and its full-length amino acid sequence is 68% identical to *P. chrysosporium* protein model number 128732. Transcripts accumulate 1.404-fold in microcrystalline cellulose medium relative to glucose medium (GEO accession GSE12540). Ppl120789 and Ppl130297 are classified as Mn-containing SODs (InterPro 001189) and show substantial sequence similarity to predicted proteins from *Taiwanofungus camphorates* (70% identity to AAQ16628) and from *P. chrysosporium* (87% identity to AAK82369), respectively. N-terminus sequence analysis (www.cbs.dtu.dk/services/TargetP/) strongly predicts mitochondrial targeting of Ppl130297, and transcript levels increase 1.541 fold in cellulose medium. No targeting is predicted for Ppl120789, and transcript levels decrease 1.314-fold in cellulose (under GEO accession GSE12540).

11.2. Catalases.

Also associated with oxidative stress (114), four putative catalase-encoding genes were identified in *P. placenta*. On the basis of sequence similarity, these correspond to 4 genes previously characterized in *C. neoformans* and classified as spore specific (CAT1 and CAT3), peroxisomal (CAT2) and cytosolic (CAT4). The 4 *Postia* catalases are: Ppl116667 (CAT1/CAT3-like), Ppl99098 (CAT1/CAT3-like), Ppl112311 (CAT2-like) and Ppl123169 (CAT4-like). None were substantially upregulated in microcrystalline cellulose medium.

11.3. NADPH Oxidases.

The NOX family of NADPH oxidases function in electron transport and the generation of reactive oxygen species (ROS). First characterized in mammalian organisms, NOX enzymes are widely distributed. While they are not found in prokaryotes and most unicellular eukaryotes, they are present in fungi, plants and animals. It has been proposed that these enzymes are a signature of multicellular organisms, having a crucial role in developmental processes. Thus, though they are present in many ascomycetes they are absent in *Candida* or *S. cerevisiae* yeast (115, 116). They have generated increasing attention as a clear and abundant source of reactive oxygen species, and appear to be developmentally regulated in *P. anserina* (117). Two models, both with substantial EST support, were identified in *P. placenta*; Ppl109097 and Ppl110233. In microcrystalline cellulose medium, transcripts increased 1.446-fold and 2.496-fold, respectively. The presence of two NOX genes has also been observed in *P. anserina* and *N. crassa*.

References

1. Perez-Boada M, *et al.* (2005) Versatile peroxidase oxidation of high redox potential aromatic compounds: site-directed mutagenesis, spectroscopic and crystallographic investigation of three long-range electron transfer pathways. *J Mol Biol* 354(2):385-402.
2. Ruiz-Duenas FJ, *et al.* (2007) Manganese oxidation site in *Pleurotus eryngii* versatile peroxidase: a site-directed mutagenesis, kinetic, and crystallographic study. *Biochemistry* 46(1):66-77.
3. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, & Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25(24):4876-4882.
4. Kumar S, Tamura K, Jakobsen IB, & Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17(12):1244-1245.
5. Gustin MC, Albertyn J, Alexander M, & Davenport K (1998) MAP kinase pathways in the yeast *Saccharomyces cerevisiae. Microbiol Mol Biol Rev* 62(4):1264-1300.
6. Highley TL (1973) Influence of carbon source on cellulase activity of white rot and brown rot fungi. *Wood Fiber* 5:50-58.
7. Detter JC, *et al.* (2002) Isothermal strand-displacement amplification applications for high-throughput genomics. *Genomics* 80(6):691-698.
8. Ewing B & Green P (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res* 8(3):186-194.
9. Ewing B, Hillier L, Wendl M, & Green P (1998) Base-calling of automated sequencer traces using Phred.I. Accuracy assessment. *Genome Res* 8:175-185.
10. Salamov AA & Solovyev VV (2000) *Ab initio* gene finding in *Drosophila* genomic DNA. *Genome Res* 10(4):516-522.
11. Birney E & Durbin R (2000) Using GeneWise in the *Drosophila* annotation experiment. *Genome Res* 10(4):547-548.
12. Koonin EV, *et al.* (2004) A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol* 5(2):R7.
13. Kanehisa M, Goto S, Kawashima S, Okuno Y, & Hattori M (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res* 32(Database issue):D277-280.
14. Henrissat B (1991) A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J* 280 (Pt 2):309-316.
15. Daboussi MJ & Capy P (2003) Transposable elements in filamentous fungi. *Annu Rev Microbiol* 57:275-299.
16. Kempken F & Kück U (1998) Transposons in filamentous fungi-facts and perspectives. *BioEssays* 20:652-659.
17. Wostemeyer J & Kreibich A (2002) Repetitive DNA elements in fungi (Mycota): impact on genomic architecture and evolution. *Curr Genet* 41(4):189-198.
18. Goodwin TJ & Poulter RT (2000) Multiple LTR-retrotransposon families in the asexual yeast *Candida albicans*. *Genome Res* 10(2):174-191.
19. Kim JM, Vanguri S, Boeke JD, Gabriel A, & Voytas DF (1998) Transposable elements and genome organization: a comprehensive survey of retrotransposons

revealed by the complete *Saccharomyces cerevisiae* genome sequence. *Genome Res* 8(5):464-478.
20. Gladyshev EA & Arkhipova IR (2007) Telomere-associated endonuclease-deficient Penelope-like retroelements in diverse eukaryotes. *Proc Natl Acad Sci U S A* 104(22):9352-9357.
21. Robertson HM & Gordon KH (2006) Canonical TTAGG-repeat telomeres and telomerase in the honey bee, *Apis mellifera*. *Genome Res* 16(11):1345-1351.
22. Vanden Wymelenberg A, *et al.* (2006) Computational analysis of the *Phanerochaete chrysosporium* v2.0 genome database and mass spectrometry identification of peptides in ligninolytic cultures reveals complex mixtures of secreted proteins. *Fungal Genetics and Biology* 43:343-356.
23. Vanden Wymelenberg A, *et al.* (2005) The *Phanerochaete chrysosporium* secretome: database predictions and initial mass spectrometry peptide identifications in cellulose-grown medium. *J Biotechnol* 118(1):17-34.
24. Smith TF & Waterman MS (1981) Identification of common molecular subsequences. *J Mol Biol* 147(1):195-197.
25. Brazma A, *et al.* (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 29(4):365-371.
26. Edgar R, Domrachev M, & Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 30(1):207-210.
27. Daniel G, *et al.* (2007) Characteristics of *Gloeophyllum trabeum* alcohol oxidase, an extracellular source of H_2O_2 in brown rot decay of wood. *Appl Environ Microbiol* 73(19):6241-6253.
28. Gutierrez A, Caramelo L, Prieto A, Martinez MJ, & Martinez AT (1994) Anisaldehyde production and aryl-alcohol oxidase and dehydrogenase activities in ligninolytic fungi of the genus *Pleurotus*. *Appl Environ Microbiol* 60(6):1783-1788.
29. Guillen F & Evans CS (1994) Anisaldehyde and veratraldehyde acting as redox cycling agents for H_2O_2 production by *Pleurotus eryngii*. *Appl Environ Microbiol* 60(8):2811-2817.
30. Ferreira P, Ruiz-Duenas FJ, Martinez MJ, van Berkel WJ, & Martinez AT (2006) Site-directed mutagenesis of selected residues at the active site of aryl-alcohol oxidase, an H_2O_2-producing ligninolytic enzyme. *FEBS J* 273(21):4878-4888.
31. Varela E, Martinez JM, & Martinez AT (2000) Aryl-alcohol oxidase protein sequence: a comparison with glucose oxidase and other FAD oxidoreductases. *Biochim Biophys Acta* 1481(1):202-208.
32. de Koker TH, Mozuch MD, Cullen D, Gaskell J, & Kersten PJ (2004) Pyranose 2-oxidase from *Phanerochaete chrysosporium*: isolation from solid substrate, protein purification, and characterization of gene structure and regulation. *Appl. Environ Microbiol* 70:5794-5800.
33. Kersten P & Cullen D (1993) Cloning and characterization of a cDNA encoding glyoxal oxidase, a peroxide-producing enzyme from the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. *Proc Nat Acad Sci USA* 90:7411-7413.
34. Whittaker MM, Kersten PJ, Cullen D, & Whittaker JW (1999) Identification of

catalytic residues in glyoxal oxidase by targeted mutagenesis. *J Biol Chem* 274(51):36226-36232.

35. Vanden Wymelenberg A, *et al.* (2006) Structure, organization, and transcriptional regulation of a family of copper radical oxidase genes in the lignin-degrading basidiomycete *Phanerochaete chrysosporium. Appl Environ Microbiol* 72:4871-4877.
36. Leuthner B, *et al.* (2005) A H_2O_2-producing glyoxal oxidase is required for filamentous growth and pathogenicity in *Ustilago maydis. Mol Genet Genomics* 272:639-650.
37. Baldrian P & Valaskova V (2008) Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiol Rev* 32(3):501-521.
38. Cohen R, Jensen KA, Houtman CJ, & Hammel KE (2002) Significant levels of extracellular reactive oxygen species produced by brown rot basidiomycetes on cellulose. *FEBS Lett* 531(3):483-488.
39. Martinez D, *et al.* (2004) Genome sequence of the lignocellulose degrading fungus *Phanerochaete chrysosporium* strain RP78. *Nature Biotechnol* 22:695-700.
40. Cohen R, Suzuki MR, & Hammel KE (2004) Differential stress-induced regulation of two quinone reductases in the brown rot basidiomycete *Gloeophyllum trabeum. Appl Environ Microbiol* 70(1):324-331.
41. Enoki A, Hirano T, & Tanaka H (1992) Extracellular substance from the brown-rot basidiomycete *Gloeophyllum trabeum* that produces and reduces hydrogen peroxide. *Mater Org* 27:247-261.
42. Tanaka H, *et al.* (2007) Characterization of a hydroxyl-radical-producing glycoprotein and its presumptive genes from the white-rot basidiomycete *Phanerochaete chrysosporium. J Biotechnol* 128(3):500-511.
43. Henriksson G, Johansson G, & Pettersson G (2000) A critical review of cellobiose dehydrogenases. *J Biotechnol* 78(2):93-113.
44. Haas H (2003) Molecular genetics of fungal siderophore biosynthesis and uptake: the role of siderophores in iron uptake and storage. *Appl Microbiol Biotechnol* 62(4):316-330.
45. Johnson L (2008) Iron and siderophores in fungal-host interactions. *Mycol Res* 112(Pt 2):170-183.
46. Kosman DJ (2003) Molecular mechanisms of iron uptake in fungi. *Mol Microbiol* 47(5):1185-1197.
47. Eichhorn H, *et al.* (2006) A ferroxidation/permeation iron uptake system is required for virulence in *Ustilago maydis. Plant Cell* 18(11):3332-3345.
48. Yun CW, Bauler M, Moore RE, Klebba PE, & Philpott CC (2001) The role of the FRE family of plasma membrane reductases in the uptake of siderophore-iron in *Saccharomyces cerevisiae. J Biol Chem* 276(13):10218-10223.
49. Dumay QC, Debut AJ, Mansour NM, & Saier MH, Jr. (2006) The copper transporter (Ctr) family of Cu+ uptake systems. *J Mol Microbiol Biotechnol* 11(1-2):10-19.
50. Laliberte J & Labbe S (2008) The molecular bases for copper uptake and distribution: lessons from yeast. *Med Sci (Paris)* 24(3):277-283.
51. Field LS, Luk E, & Culotta VC (2002) Copper chaperones: personal escorts for

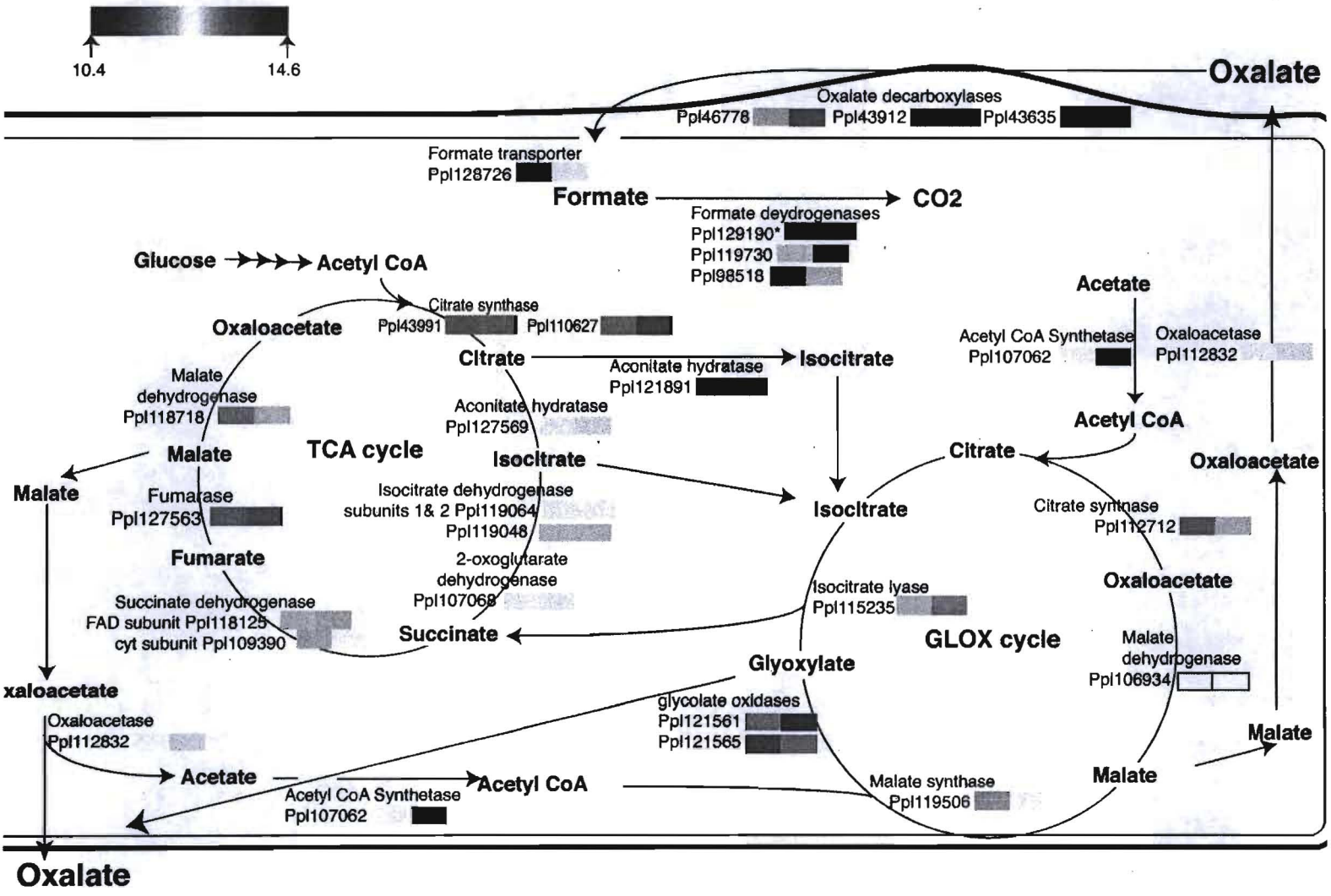
metal ions. *J Bioenerg Biomembr* 34(5):373-379.
52. Uldschmid A, Dombi R, & Marbach K (2003) Identification and functional expression of ctaA, a P-type ATPase gene involved in copper trafficking in *Trametes versicolor*. *Microbiology* 149(Pt 8):2039-2048.
53. Banci L, Bertini I, Chasapis CT, Rosato A, & Tenori L (2007) Interaction of the two soluble metal-binding domains of yeast Ccc2 with copper(I)-Atx1. *Biochem Biophys Res Commun* 364(3):645-649.
54. Lin SJ, Pufahl RA, Dancis A, O'Halloran TV, & Culotta VC (1997) A role for the *Saccharomyces cerevisiae* ATX1 gene in copper trafficking and iron transport. *J Biol Chem* 272(14):9215-9220.
55. Mei B, Budde AD, & Leong SA (1993) *sid1*, a gene initiating siderophore biosynthesis in *Ustilago maydis*: molecular characterization, regulation by iron, and role in phytopathogenicity. *Proc Natl Acad Sci U S A* 90(3):903-907.
56. Fekete FA, Chandhoke V, & Jellison J (1989) Iron-binding compounds produced by wood-decaying basidiomycetes. *Appl Environ Microbiol* 55(10):2720-2722.
57. Goodell B, *et al.* (1997) Low molecular weight chelators and phenolic compounds isolated from wood decay fungi and their role in the fungal biodegradation of wood. *J Biotechnol* 53:133-162.
58. Martinez AT (2002) Molecular biology and structure-function of lignin-degrading heme peroxidases. *Enzyme Microb Technol* 30:425-444.
59. Larrondo L, Gonzalez A, Perez Acle T, Cullen D, & Vicuna R (2005) The *nop* gene from *Phanerochaete chrysosporium* encodes a peroxidase with novel structural features. *Biophys Chem* 116(2):167-173.
60. Morgenstern I, Klopman S, & Hibbett DS (2008) Molecular evolution and diversity of lignin degrading heme peroxidases in the Agaricomycetes. *J Mol Evol* 66(3):243-257.
61. Ortiz-Bermudez P, Hirth KC, Srebotnik E, & Hammel KE (2007) Chlorination of lignin by ubiquitous fungi has a likely role in global organochlorine production. *Proc Natl Acad Sci U S A* 104(10):3895-3900.
62. Anh DH, *et al.* (2007) The coprophilous mushroom *Coprinus radians* secretes a haloperoxidase that catalyzes aromatic peroxygenation. *Appl Environ Microbiol* 73(17):5477-5485.
63. Hoegger PJ, Kilaru S, James TY, Thacker JR, & Kües U (2006) Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. *FEBS J* 273(10):2308-2326.
64. Guillen F, Gomez-Toribio V, Martinez MJ, & Martinez AT (2000) Production of hydroxyl radical by the synergistic action of fungal laccase and aryl alcohol oxidase. *Arch Biochem Biophys* 383(1):142-147.
65. Larrondo L, Salas L, Melo F, Vicuna R, & Cullen D (2003) A novel extracellular multicopper oxidase from *Phanerochaete chrysosporium* with ferroxidase activity. *Appl Environ Microbiol* 69:6257-6263.
66. Boase NA & Kelly JM (2004) A role for creD, a carbon catabolite repression gene from *Aspergillus nidulans*, in ubiquitination. *Mol Microbiol* 53(3):929-940.
67. Roy P, Lockington RA, & Kelly JM (2008) CreA-mediated repression in *Aspergillus nidulans* does not require transcriptional auto-regulation, regulated intracellular localisation or degradation of CreA. *Fungal Genet Biol* 45(5):657-

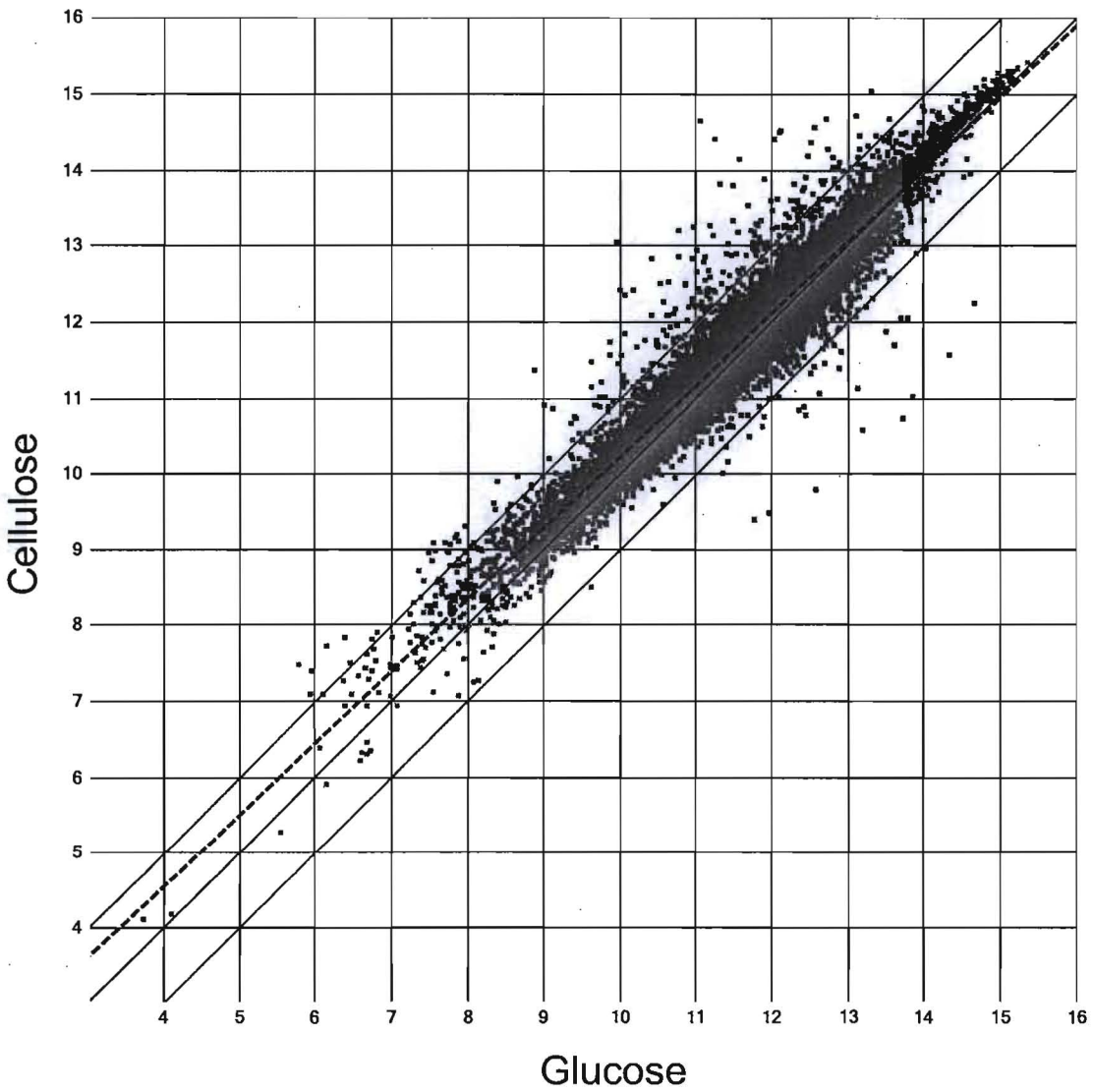
670.
68. Vautard-Mey G & Fevre M (2000) Mutation of a putative AMPK phosphorylation site abolishes the repressor activity but not the nuclear targeting of the fungal glucose regulator CRE1. *Curr Genet* 37(5):328-332.
69. Stricker AR, Mach RL, & de Graaff LH (2008) Regulation of transcription of cellulases- and hemicellulases-encoding genes in *Aspergillus niger* and *Hypocrea jecorina (Trichoderma reesei). Appl Microbiol Biotechnol* 78(2):211-220.
70. Casselton LA & Kües U (2007) The origin of multiple mating types in the model mushrooms *Coprinopsis cinerea* and *Schizophyllum commune. Sex in Fungi*, eds Heitman J, Kronstad JW, Taylor JW, & Casselton LA (ASM Press, Washington), pp 283-300.
71. Fraser JA & Heitman J (2004) Evolution of fungal sex chromosomes. *Mol Microbiol* 51(2):299-306.
72. James TY, Srivilai P, Kües U, & Vilgalys R (2006) Evolution of the bipolar mating system of the mushroom *Coprinellus disseminatus* from its tetrapolar ancestors involves loss of mating-type-specific pheromone receptor function. *Genetics* 172(3):1877-1891.
73. Nobles MK (1943) A contribution toward a clarification of the *Trametes serialis* complex. *Can J Res* 21:211-234.
74. James TY, Kües U, Rehner SA, & Vilgalys R (2004) Evolution of the gene encoding mitochondrial intermediate peptidase and its cosegregation with the *A* mating-type locus of mushroom fungi. *Fungal Genet Biol* 41(3):381-390.
75. Kües U, James TY, Vilgalys R, & Challen MP (2001) The chromosomal region containing *pab-1, mip*, and the *A* mating type locus of the secondarily homothallic homobasidiomycete *Coprinus bilanatus*. *Curr Genet* 39(1):16-24.
76. Niculita-Hirzel H, *et al.* (2008) Gene organization of the mating type regions in the ectomycorrhizal fungus *Laccaria bicolor* reveals distinct evolution between the two mating type loci. *New Phytol.*
77. Hiscock SJ & Kües U (1999) Cellular and molecular mechanisms of sexual incompatibility in plants and fungi. *Int Rev Cytol* 193:165-295.
78. Banham AH, *et al.* (1995) An N-Terminal dimerization domain permits homeodomain proteins to choose compatible partners and initiate sexual development in the mushroom *Coprinus cinereus*. *Plant Cell* 7(6):773-783.
79. Kües U & Casselton LA (1993) The origins of multiple mating types in mushrooms. *J Cell Sci* 104:227-230.
80. Riquelme M, Challen MP, Casselton LA, & Brown AJ (2005) The origin of multiple *B* mating specificities in *Coprinus cinereus*. *Genetics* 170(3):1105-1119.
81. Halsall JR, Milner MJ, & Casselton LA (2000) Three subfamilies of pheromone and receptor genes generate multiple *B* mating specificities in the mushroom *Coprinus cinereus*. *Genetics* 154(3):1115-1123.
82. Lafon A, Han KH, Seo JA, Yu JH, & d'Enfert C (2006) G-protein and cAMP-mediated signaling in aspergilli: a genomic perspective. *Fungal Genet Biol* 43(7):490-502.
83. Li L, Wright SJ, Krystofova S, Park G, & Borkovich KA (2007) Heterotrimeric G protein signaling in filamentous fungi. *Annu Rev Microbiol* 61:423-452.
84. Regenfelder E, *et al.* (1997) G proteins in *Ustilago maydis*: transmission of

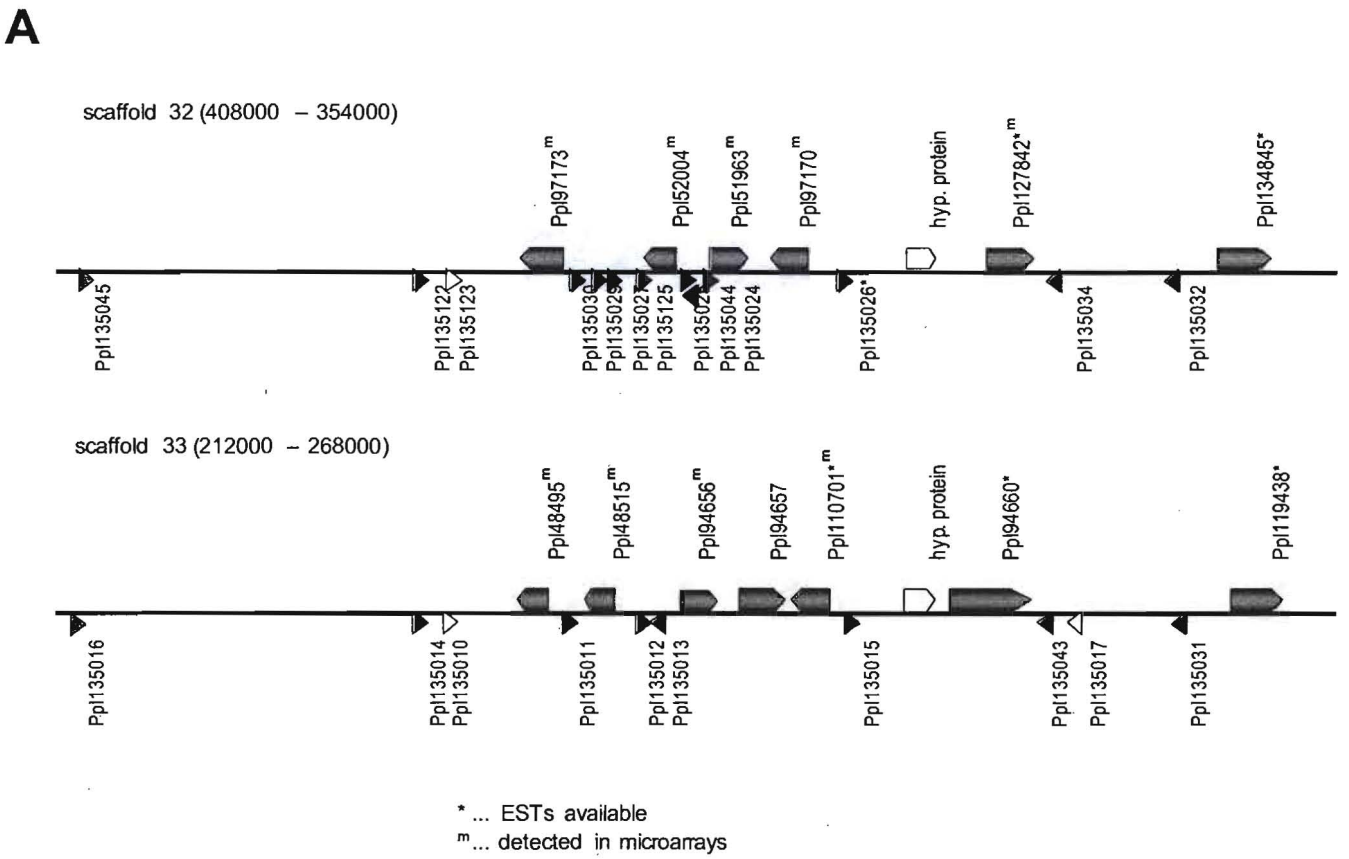
multiple signals? *Embo J* 16(8):1934-1942.
85. Borkovich KA, *et al.* (2004) Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. *Microbiol Mol Biol Rev* 68(1):1-108.
86. Bardwell L (2004) A walk-through of the yeast mating pheromone response pathway. *Peptides* 25(9):1465-1476.
87. Chang C, Clark K, Wang X, & Stewart R (1998) 'Two-component' ethylene signaling in *Arabidopsis*. *Symp Soc Exp Biol* 51:59-64.
88. Catlett NL, Yoder OC, & Turgeon BG (2003) Whole-genome analysis of two-component signal transduction genes in fungal pathogens. *Eukaryot Cell* 2(6):1151-1161.
89. Calvo AM (2008) The VeA regulatory system and its role in morphological and chemical development in fungi. *Fungal Genet Biol* 45(7):1053-1061.
90. Saraste M (1999) Oxidative phosphorylation at the fin de siecle. *Science* 283(5407):1488-1493.
91. D'Elia D, *et al.* (2006) The MitoDrome database annotates and compares the OXPHOS nuclear genes of *Drosophila melanogaster*, *Drosophila pseudoobscura* and *Anopheles gambiae*. *Mitochondrion* 6(5):252-257.
92. Garesse R & Vallejo CG (2001) Animal mitochondrial biogenesis and function: a regulatory cross-talk between two genomes. *Gene* 263(1-2):1-16.
93. Burger G, Lang BF, Reith M, & Gray MW (1996) Genes encoding the same three subunits of respiratory complex II are present in the mitochondrial DNA of two phylogenetically distant eukaryotes. *Proc Natl Acad Sci U S A* 93(6):2328-2332.
94. Gray MW, Burger G, & Lang BF (2001) The origin and early evolution of mitochondria. *Genome Biol* 2(6):REVIEWS1018.
95. Joseph-Horne T, Hollomon DW, & Wood PM (2001) Fungal respiration: a fusion of standard and alternative components. *Biochim Biophys Acta* 1504(2-3):179-195.
96. Lavin JL, Oguiza JA, Ramirez L, & Pisabarro A (2008) Comparative genomics of oxidative phosphorylation systems in fungi. *Fungal Genet Biol* 45:1248-1256.
97. Abdrakhmanova A (2005) Accessory subunits of complex I from *Yarrowia lipolytica*. PhD (Johann Wolfgang Goethe-Universität, Frankfurt).
98. Velours J & Arselin G (2000) The *Saccharomyces cerevisiae* ATP synthase. *J Bioenerg Biomembr* 32(4):383-390.
99. Varela E & Tien M (2003) Effect of pH and oxalate on hydroquinone-derived hydroxyl radical formation during brown rot wood degradation. *Appl Environ Microbiol* 69(10):6025-6031.
100. Goodell B (2003) Brown rot fungal degradation of wood: our evolving view. *Wood deterioration and preservation*, eds Goodell B, Nicholas D, & Schultz T (American Chemical Society, Washington, DC), pp 97-118.
101. Munir E, Yoon JJ, Tokimatsu T, Hattori T, & Shimada M (2001) A physiological role for oxalic acid biosynthesis in the wood-rotting basidiomycete *Fomitopsis palustris*. *Proc Natl Acad Sci USA* 98(20):11126-11130.
102. Sakai S, *et al.* (2006) Subcellular localization of glyoxylate cycle key enzymes involved in oxalate biosynthesis of wood-destroying basidiomycete *Fomitopsis palustris* grown on glucose. *Microbiology* 152(Pt 6):1857-1866.

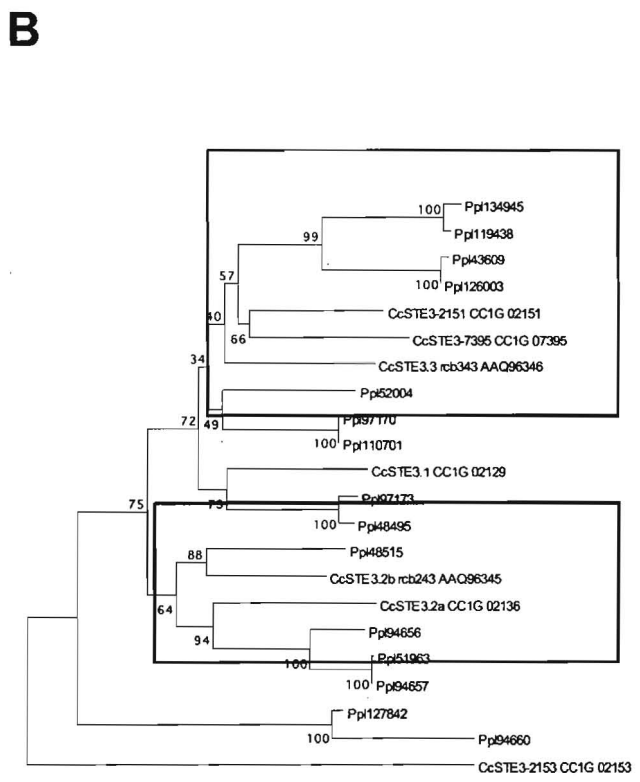
103. Akamatsu Y & Shimada M (1994) Partial purification and characterization of glyoxylate oxidase from the brown-rot basidiomycete *Tyromyces palustris*. *Phytochemistry* 37:649-653.
104. Regev-Rudzki N, Karniely S, Ben-Haim NN, & Pines O (2005) Yeast aconitase in two locations and two metabolic pathways: seeing small amounts is believing. *Mol Biol Cell* 16(9):4163-4171.
105. Escutia MR, *et al.* (2005) Cloning and sequencing of two *Ceriporiopsis subvermispora* bicupin oxalate oxidase allelic isoforms: implications for the reaction specificity of oxalate oxidases and decarboxylases. *Appl Environ Microbiol* 71(7):3608-3616.
106. Burrell MR, *et al.* (2007) Oxalate decarboxylase and oxalate oxidase activities can be interchanged with a specificity switch of up to 282,000 by mutating an active site lid. *Biochemistry* 46(43):12327-12336.
107. Micales J (1997) Localization and induction of oxalate decarboxylase in the brown-rot decay fungus *Postia placenta*. *Int Biodeter Biodegrad* 39:125-132.
108. Yadav JS, Doddapaneni H, & Subramanian V (2006) P450ome of the white rot fungus *Phanerochaete chrysosporium*: structure, evolution and regulation of expression of genomic P450 clusters. *Biochem Soc Trans* 34(Pt 6):1165-1169.
109. Tamura K, Dudley J, Nei M, & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24(8):1596-1599.
110. Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4(4):406-425.
111. Doddapaneni H, Chakraborty R, & Yadav JS (2005) Genome-wide structural and evolutionary analysis of the P450 monooxygenase genes (P450ome) in the white rot fungus *Phanerochaete chrysosporium*: Evidence for gene duplications and extensive gene clustering. *BMC Genomics* 6(1):92.
112. Matityahu A, Hadar Y, Dosoretz CG, & Belinky PA (2008) Gene silencing by RNA interference in the white-rot fungus *Phanerochaete chrysosporium*. *Appl Environ Microbiol* 74:5359-5365.
113. Belinky PA, Flikshtein N, Lechenko S, Gepstein S, & Dosoretz CG (2003) Reactive oxygen species and induction of lignin peroxidase in *Phanerochaete chrysosporium*. *Appl Environ Microbiol* 69(11):6500-6506.
114. Giles SS, *et al.* (2006) The *Cryptococcus neoformans* catalase gene family and its role in antioxidant defense. *Eukaryot Cell* 5(9):1447-1459.
115. Kawahara T, Quinn MT, & Lambeth JD (2007) Molecular evolution of the reactive oxygen-generating NADPH oxidase (Nox/Duox) family of enzymes. *BMC Evol Biol* 7:109.
116. Takemoto D, Tanaka A, & Scott B (2007) NADPH oxidases in fungi: diverse roles of reactive oxygen species in fungal cellular differentiation. *Fungal Genet Biol* 44(11):1065-1076.
117. Malagnac F, Lalucque H, Lepere G, & Silar P (2004) Two NADPH oxidase isoforms are required for sexual reproduction and ascospore germination in the filamentous fungus *Podospora anserina*. *Fungal Genet Biol* 41(11):982-997.
118. Enright AJ, Van Dongen S, & Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30(7):1575-1584.

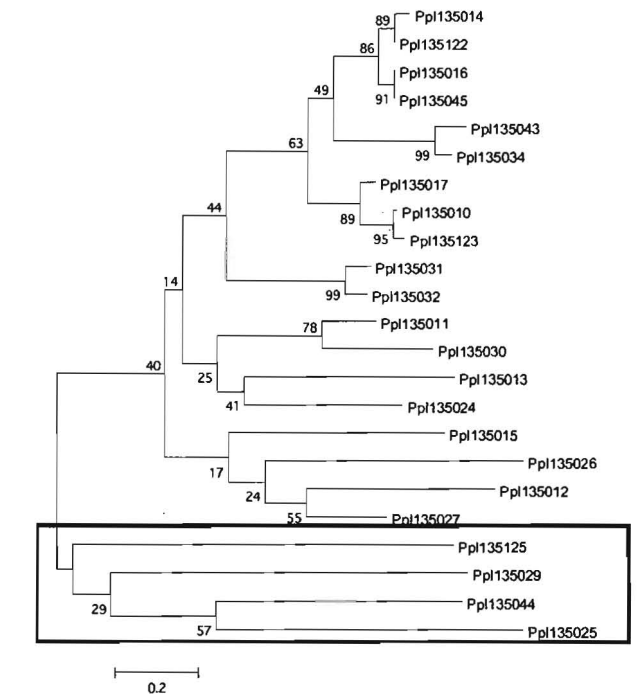
119. Thompson JD, Higgins DG, & Gibson TJ (1994) ClustalW improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:2552-2556.

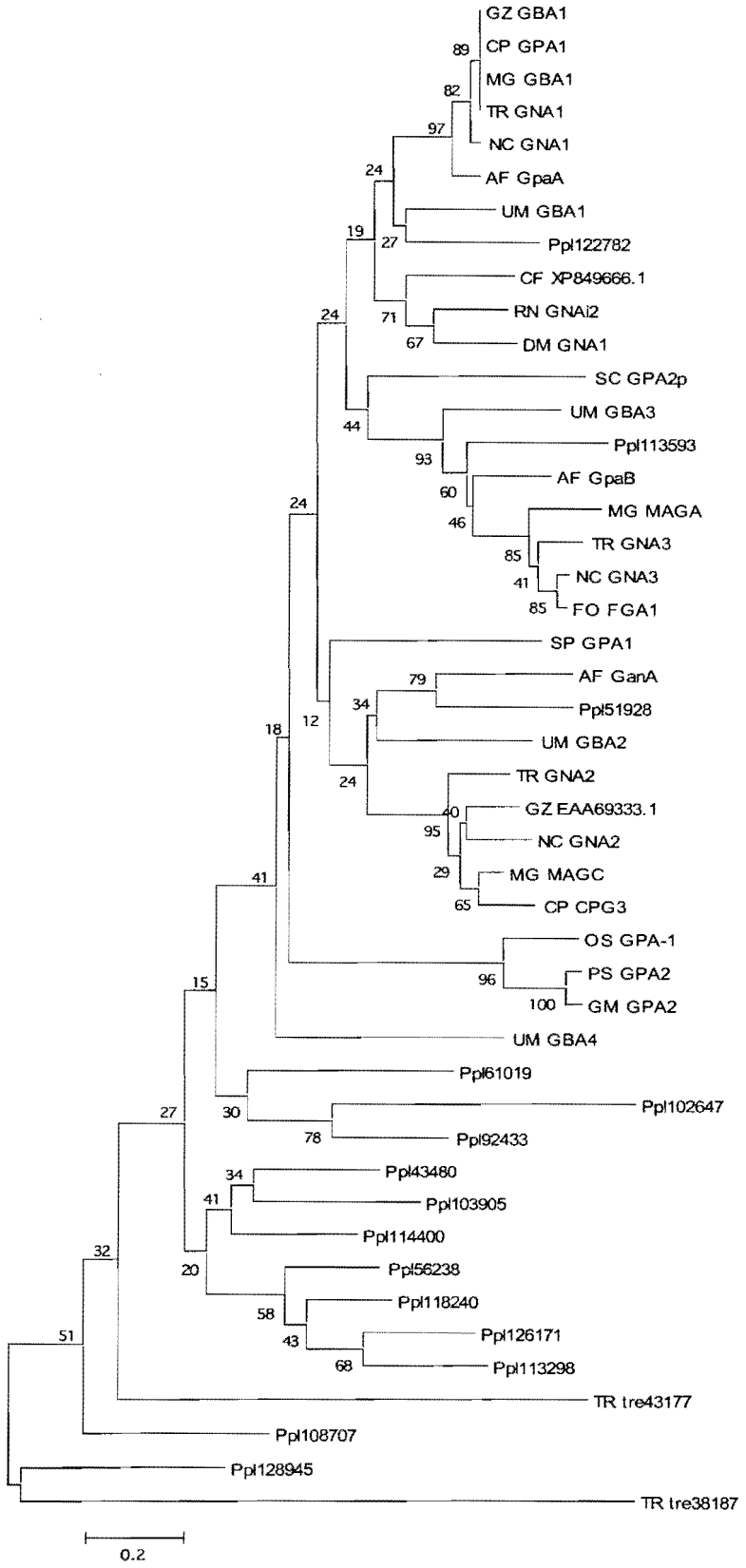

10.4
14.6
Oxalate
Oxalate decarboxylases
Ppl46778
Ppl43912
Ppl43635
Formate transporter
Ppl128726
Formate
CO2
Formate deydrogenases
Ppl129190*
Ppl119730
Ppl98518
Glucose
Acetyl CoA
Acetate
Oxaloacetate
Citrate synthase
Ppl43991
Ppl110627
Acetyl CoA Synthetase
Ppl107062
Oxaloacetase
Ppl112832
Citrate
Isocitrate
Aconitate hydratase
Ppl121891
Malate dehydrogenase
Ppl118718
Aconitate hydratase
Ppl127569
Acetyl CoA
Malate
TCA cycle
Isocitrate
Citrate
Oxaloacetate
Malate
Fumarase
Ppl127563
Isocitrate dehydrogenase
subunits 1& 2 Ppl119064
Ppl119048
Isocitrate
Citrate synthase
Ppl112712
Fumarate
2-oxoglutarate
dehydrogenase
Ppl107068
Oxaloacetate
Succinate dehydrogenase
FAD subunit Ppl118125
cyt subunit Ppl109390
Isocitrate lyase
Ppl115235
Succinate
GLOX cycle
Malate
dehydrogenase
Ppl106934
xaloacetate
Glyoxylate
glycolate oxidases
Ppl121561
Ppl121565
Oxaloacetase
Ppl112832
Malate
Acetate
Acetyl CoA
Malate synthase
Ppl119506
Malate
Acetyl CoA Synthetase
Ppl107062
Oxalate

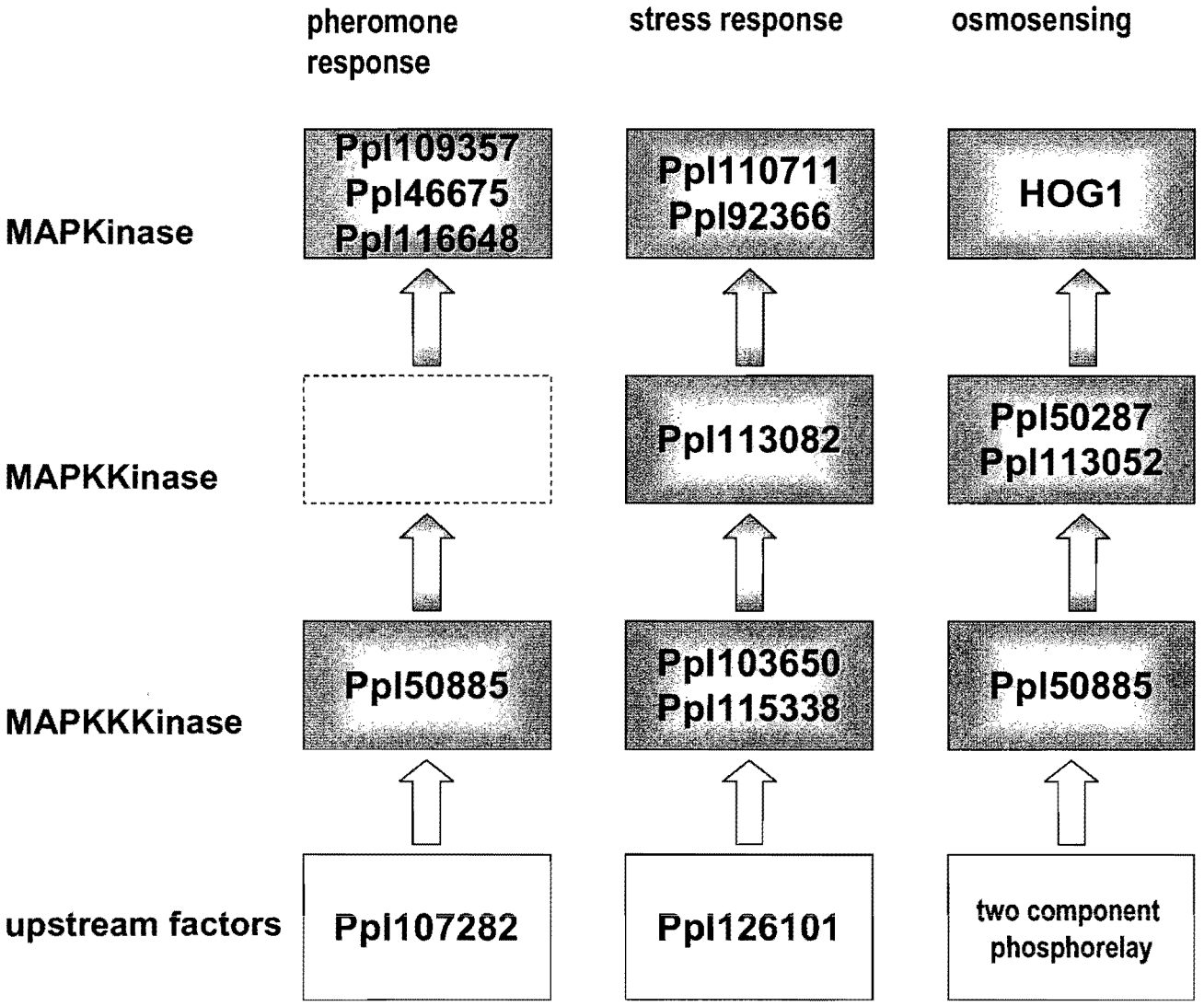

Cellulose
Glucose
16
15
14
13
12
11
10
9
8
7
6
5
4
4
5
6
7
8
9
10
11
12
13
14
15
16


A
scaffold 32 (408000 – 354000)
Ppl97173m
Ppl52004m
Ppl51963m
Ppl97170m
hyp. protein
Ppl127842*m
Ppl134845*
Ppl135045
Ppl135122
Ppl135123
Ppl135026*
Ppl135034
Ppl135032
scaffold 33 (212000 – 268000)
Ppl48495m
Ppl48515m
Ppl94656m
Ppl94657
Ppl110701*m
hyp. protein
Ppl94660*
Ppl119438*
Ppl135016
Ppl135014
Ppl135010
Ppl135011
Ppl135012
Ppl135013
Ppl135015
Ppl135043
Ppl135017
Ppl135031
* ... ESTs available
m ... detected in microarrays


B
Ppl134945
Ppl119438
Ppl43609
Ppl126003
CcSTE3-2151 CC1G 02151
CcSTE3-7395 CC1G 07395
CcSTE3.3 rcb343 AAQ96346
Ppl52004
Ppl97170
Ppl110701
CcSTE3.1 CC1G 02129
Ppl97173
Ppl48495
Ppl48515
CcSTE3.2b rcb243 AAQ96345
CcSTE3.2a CC1G 02136
Ppl94656
Ppl51963
Ppl94657
Ppl127842
Ppl94660
CcSTE3-2153 CC1G 02153


C
Ppl135014
Ppl135122
Ppl135016
Ppl135045
Ppl135043
Ppl135034
Ppl135017
Ppl135010
Ppl135123
Ppl135031
Ppl135032
Ppl135011
Ppl135030
Ppl135013
Ppl135024
Ppl135015
Ppl135026
Ppl135012
Ppl135027
Ppl135125
Ppl135029
Ppl135044
Ppl135025
0.2


GZ GBA1
CP GPA1
MG GBA1
TR GNA1
NC GNA1
AF GpaA
UM GBA1
Ppl122782
CF XP849666.1
RN GNAi2
DM GNA1
SC GPA2p
UM GBA3
Ppl113593
AF GpaB
MG MAGA
TR GNA3
NC GNA3
FO FGA1
SP GPA1
AF GanA
Ppl51928
UM GBA2
TR GNA2
GZ EAA69333.1
NC GNA2
MG MAGC
CP CPG3
OS GPA-1
PS GPA2
GM GPA2
UM GBA4
Ppl61019
Ppl102647
Ppl92433
Ppl43480
Ppl103905
Ppl114400
Ppl56238
Ppl118240
Ppl126171
Ppl113298
TR tre43177
Ppl108707
Ppl128945
TR tre38187
0.2


pheromone response
stress response
osmosensing
MAPKinase
MAPKKinase
MAPKKKinase
upstream factors
Ppl109357
Ppl46675
Ppl116648
Ppl110711
Ppl92366
HOG1
Ppl113082
Ppl50287
Ppl113052
Ppl50885
Ppl103650
Ppl115338
Ppl50885
Ppl107282
Ppl126101
two component phosphorelay

SI Table 2. Distribution of putative CAZyme modules in sequenced fungal genomes[1]

Species	GH	GH5	GH5 (CBM1)	GH6	GH7	GH10	GH11	GH12	GH43	GH51	GH61	GT	CBM	CBM1	CE	PL	EXPN
M. oryzae	231	13	1	3	6	5	5	3	19	3	17	92	60	21	47	1	1
G. zeae	243	15	1	1	2	5	3	2	16	2	13	102	61	12	42	20	5
T. reesei	192	11	2	1	2	1	4	1	2	0	3	87	35	14	15	3	4
C. neoformans	75	12	0	0	0	0	0	0	0	1	1	68	10	0	9	3	1
U. maydis	98	12	0	0	0	2	1	0	4	2	0	64	9	0	19	1	8
L. bicolor	163	22	1	0	0	0	0	3	0	0	8	88	26	1	19	7	12
P. chrysosporium	180	20	4	1	8	6	1	2	4	2	13	68	45	30	19	4	11
C. cinereus	211	27	1	5	7	5	6	1	4	1	33	72	89	46	51	13	14
P. placenta	144	20	0	0	0	3	0	2	1	1	2	75	6	0	10	6	7

[1]As defined by Henrissat and coworkers (http://afmb.cnrs.mrs.fr/CAZY/index.html). Abbreviations: GH, total number glycoside hydrolase modules; GH#, modules within individual glycoside hydrolase families; GH5 (CBM1), glycoside hydrolase family 5 modules associated with family 1 carbohydrate binding module; GT, glycosyl transferase modules; CBM, carbohydrate binding modules; CBM1, family 1 carbohydrate binding modules; CE, carbohydrate esterases; PL, polysaccharide lyases; EXPN, expansins. Family 1 carbohydrate binding modules confer cellulose binding.

SI Table 4. Extracellular CAZY peptides sequenced by LC-MS/MS

Alleles[1]	Description	ESTs[2]	Microarray[3]		LC-MS/MS Peptides[4] Total/unique/score		
		Total/Avicel	*P* value	Fold change	Aspen	Avicel	Cotton
88470/90501	GH not yet assigned to family	0	2.24E-7	2.24	1/1/46	-	-
105534/none	GH10 xylanase	9/0	1.93E-11	3.69	5/4/299	-	-
113670/134787 (90657)	GH10 xylanase	2/2	1.84E-6	1.45	2/2/90	-	-
113112/117345	GH15 glucoamylase	2/1	9.25E-10	2.28	6/4/390	-	-
112941/61809	GH16 endo-1,3-β-glucanase	5/4	1.98E-12	3.95	-	2/1/181	2//1/141
124498/126595	GH18 chitinase	16/8	1.50E-8	1.97	1/1/79	-	-
119525/120960	GH18 chitinase	38/12	2.45E-6	1.35	1/1/44	-	-
56576/57564	GH2 β -mannosidase	0	0.0093	1.074	1/1/49	-	-
130398/134907 (134894)	GH20 β -hexosaminidase	2/0	0.0063	-1.20	4/2/227	-	-
128150/98662	GH27 α -galactosidase	2/2	0.00023	1.27	3/2/198*	-	-
107557/none	GH3 β -glucosidase	3/3	0.00014	1.15	12/5/803	5/3/385	2/1/144
127469/51213	GH3 β -xylosidase	6/2	8.75E-6	1.18	7/5/382		
135050/45962 (122151)	GH30	12/3	1.46E-9	2.30	1/1/46	-	-
60599/93878 (134924)	GH31 α -glucosidase (maltase)	0	0.00003	1.16	4/3/241	-	-
127993/128101	GH35 β -galactosidase	12/2	0.00007	1.23	5/4/264	3/2/155*	1/1/48
61292/none	GH37 trehalase	1/0	.000182	1.14	4/3/237	-	-
97540/115929	GH37 trehalase	7/2	.0086	1.06	-	2/2/95	1/1/49
115593/134925	GH47 α -mannosidase	1/1	2.94E-7	-1.89	3/3/149	-	-
121831/134772 (57321)	GH5 endo- β -1,4-mannosidase	6/6	4.13E-9	1.93	7/2/551	-	-
115648/108962	GH5 endo- β -1,4-endoglucanase	6/4	7.80E-7	1.43	4/3/232	-	-

127046/100251	GH51 (α - arabinofuranosidase)	8/0	.0082	1.10	3/1/239*	2/1/134	5/1/491*
119394/105490	GH55 1,3-glucanase	4/4	2.14E-12	5.38	18/6/1233	4/2/393*	6/3/487
108648/116267	GH55 1,3-glucanase	16/14	2.07E-11	3.15	14/5/1103	2/2/203*	13/3/1053
117860/118950	GH72 β-1,3-glucanosyltransferase	50/17	2.00E-6	1.37	1/1/64		
126692/111332	GH79 endo- β - glucuronidase	3/3	2.06E-7	1.86	7/4/461	-	-
112047/116992	GH92 α -1,2-mannosidase	4/4	0.0052	-1.15	3/2/198	-	2/1/148

[1]Protein model identification number. Underlined model number selected as microarray target. Alternative models shown parenthetically. [2]Total number of ESTs derived from all media/number of ESTs from medium with microcrystalline cellulose (Avicel) as sole carbon source. [3]Expression ratios derived from comparisons of glucose-grown versus cellulose-grown mycelia. Analysis of variance P values based on 3 full biological replicates per culture medium. Quantile normalization and robust multi-array averaging (RMA) applied to entire dataset. Reciprocals of ratios <1.0 are multiplied by -1. [4]Total soluble extracellular protein fractionated and analyzed by one-dimensional SDS-PAGE and LC-MS/MS. Total number of peptides/number of unique peptides/Mascot score. Asterisks indicate identification of all allele-specific peptide sequences, all of which are listed with scoring in SI_table _11.xls under GEO accession 12540.

SI Table 5. Summary of oxidoreductases potentially involved in lignocellulose degradation by *P.placenta* (Ppl) and *P.chrysosporium* (Pch)[1]

Putative function	EC class	Ppl	Pch	Gene; (Ref.)
Peroxide generation				
Methanol oxidase	1.1.3.13	1*	1	*mox1*; (25)
Aryl alcohol oxidase	1.1.3.7	3	3	*aox*; (28)
Glucose oxidase	1.1.3.14	≥5^{+}	≥4^{+}	*gox*; (29)
Pyranose-2-oxidase	1.1.3.10	0	1	*pox1*; (30)
Copper radical oxidase	-	3*$^{+}$	5^{+}	*cro*; (33)
Iron reduction and homeostasis				
Quinone reductase	1.6.5.5	1*	4	*qrd*; (37)
Glycoprotein iron reductase	-	4*	2	*glp*; (40)
Cellobiose dehydrogenase	1.1.99.18	0	1^{+}	*cdh1*; (37)
Iron ferroxidase	1.16.3.1	1*	1	*fet1*; (48)
Ligin modification				
Lignin peroxidase	1.11.1.13	0	10	*lip*; (37)
Manganese peroxidase	1.11.1.16	0	5	*mnp*; (37)
Low redox peroxidase	1.11.1.7	1	1	*lrp1*; (44)
Chloroperoxidase	1.11.1.10	5*	3	*cpo*; (43)
Laccase	1.10.3.2	3*	0	*lac*; (48)

[1]Detailed information on all genes in SI_table_6.xls and SI_table_7.xls under GEO accession 12540.

*Gene, or member(s) of the gene family, with significant (P<.01; >2-fold) transcript accumulation in cultures of *P. placenta* containing microcrystalline cellulose as sole carbon source relative to glucose-grown cultures. (Comparable microarray data not available for *P. chrysosporium.*)

$^{+}$Peptides identified in culture supernatants by LC-MS/MS (SI_table_11.xls under GEO accession 12540).