Supporting Information

Martinez *et al*. 10.1073/pnas.0809575106

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.

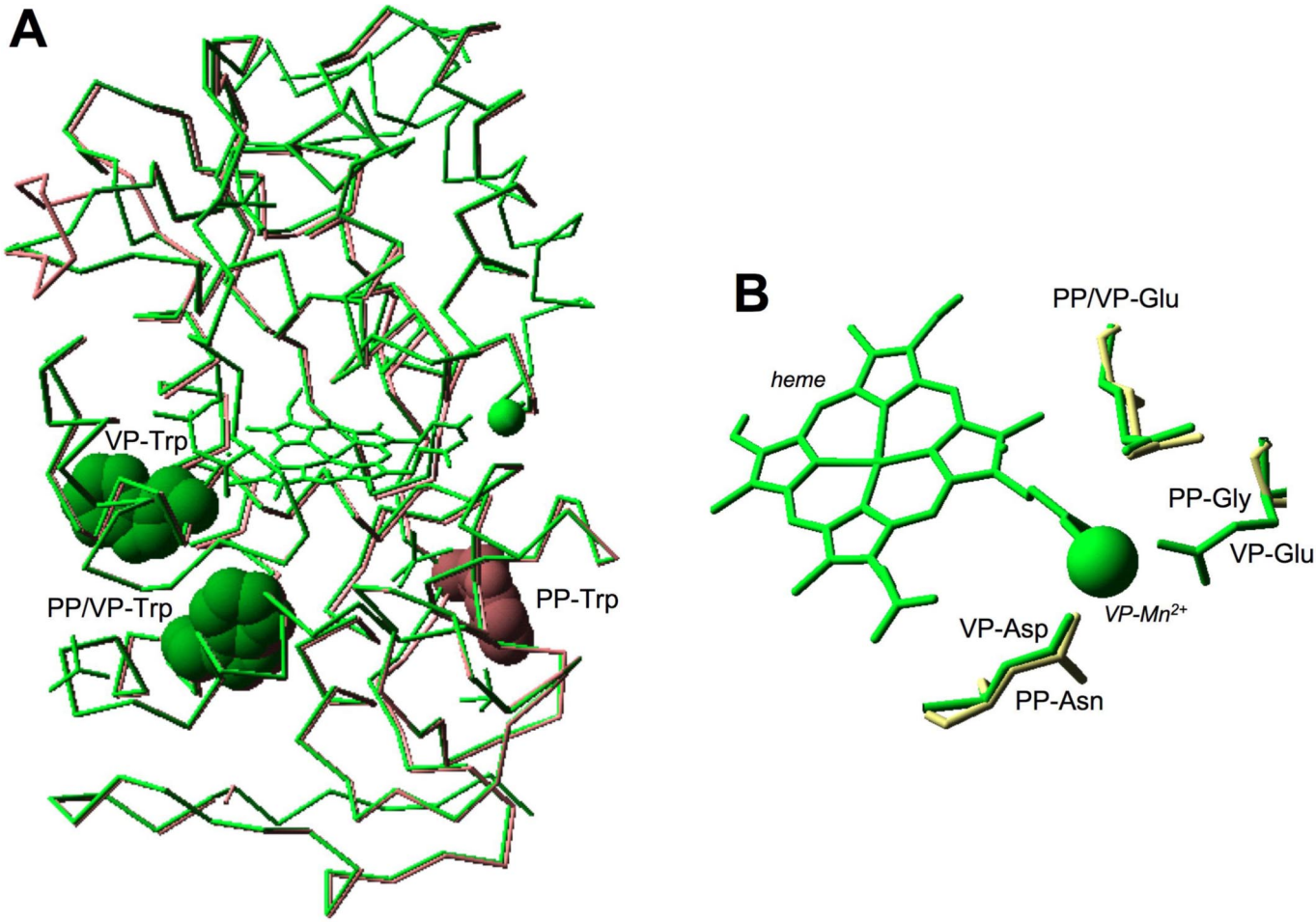

Fig. S1. 3D-model for the hypothetical *Postia placenta* peroxidase Ppl44056 (PP, light brown) obtained by homology modeling, superimposed with the crystal structure of versatile peroxidase (VP) from *Pleurotus eryngii* (green). (*A*) Protein backbones indicating the position of tryptophan residues present in these peroxidases (as van der Waals spheres), one of them being common to both proteins (*Middle*) whereas the two others 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 2 glutamate and 1 aspartate residues (2 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.

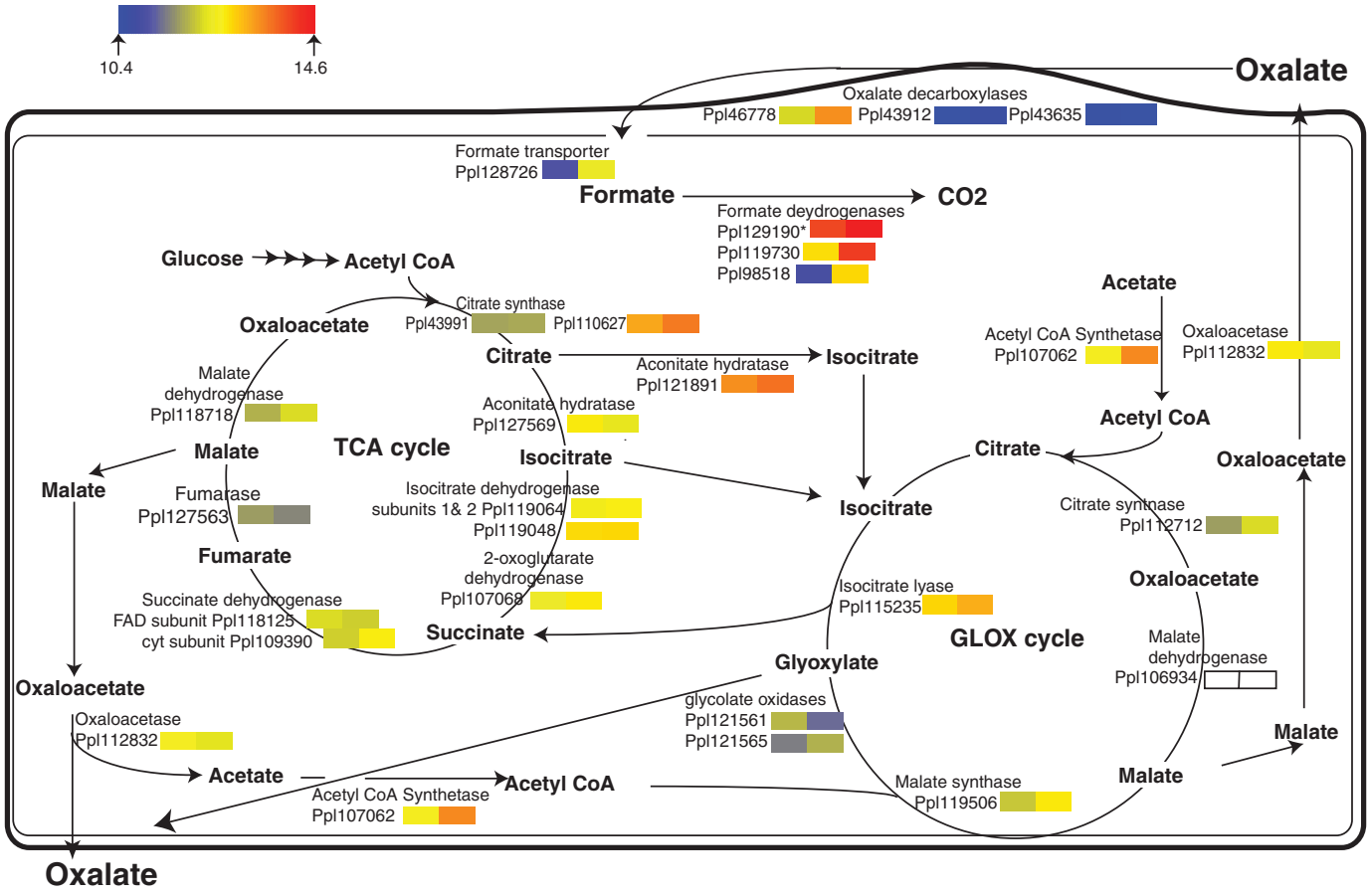

Fig. S2. Putative components of TCA cycle, GLOX cycle, and oxalate metabolism. Colored bars indicate microarray-derived expression results from glucose-grown (*Left*) and cellulose-grown (*Right*) mycelia. Manual inspection of malate dehydrogenase model Ppl106934 revealed several inaccuracies upon which microarray 60mers were based. Reliable expression results are therefore not available. Model details and microarray results are listed under GEO accession no. GSE12540 si_table_8.xls.

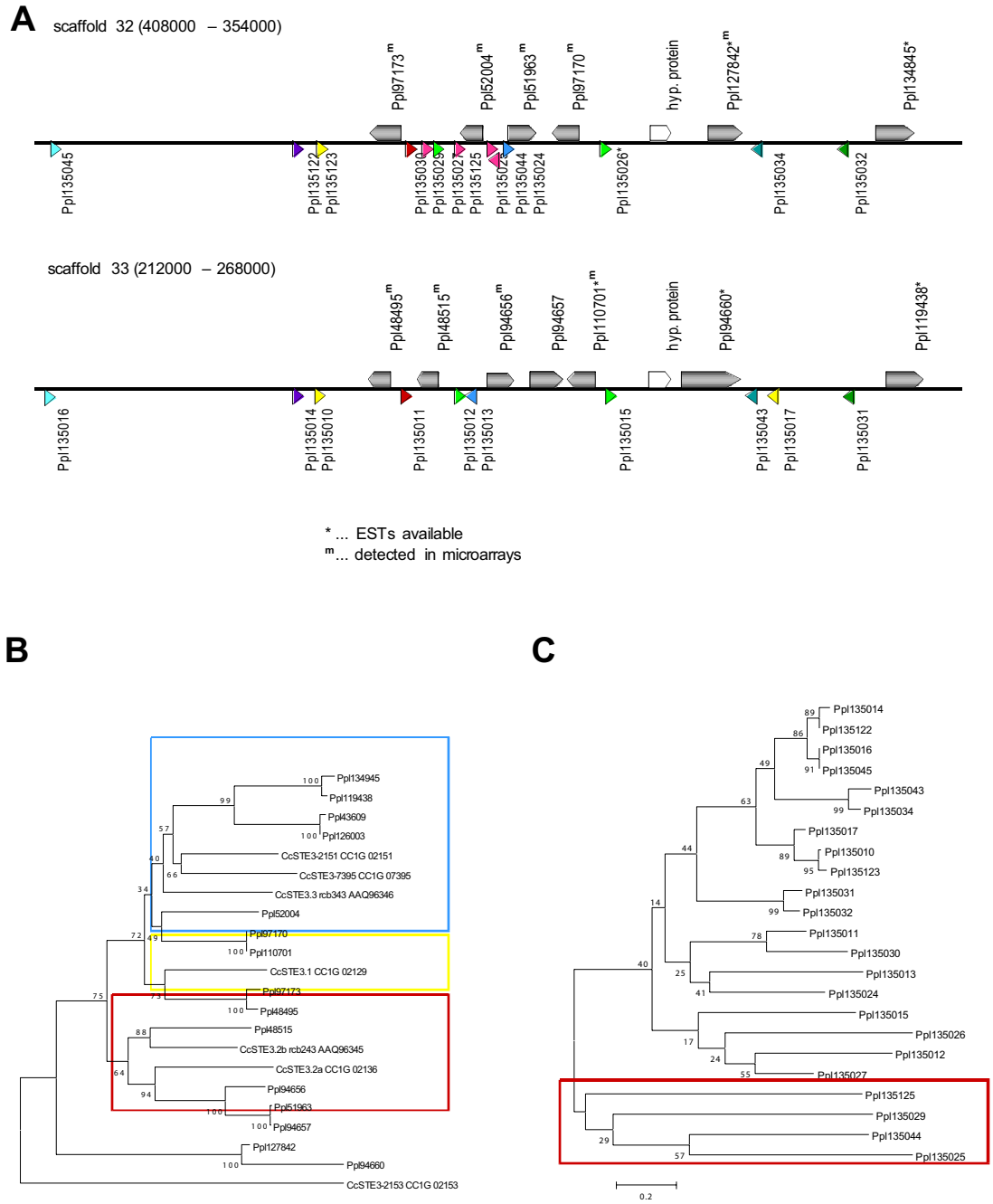

Fig. S3. Analysis of *P. placenta B*-mating locus. *A*: Schematic representation of the *B*-mating type locus. Clusters and groups within clusters are given according to the phylogenetic proximity of the respective gene models to their orthologs in *C. cinerea*. Closely related pheromone precursors (as revealed by phylogenetic analysis, (*C*) are depicted in identical colors and those lacking close neighbors are shown in black. Complete microarray results are listed in GEO accession no. GSE12540, where Ppl94656 and Ppl51963 are represented by earlier incomplete models Ppl32692 and Ppl32709, respectively. (Oligonucleotide specificity is retained in the improved models.) *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). *C*: Phylogenetic analysis of putative a-type peptide pheromone precursors of *P. placenta*. For both trees, sequence alignments were performed by using ClustalX (1.81) (3) and the alignments were manually adjusted by the aid of Genedoc. Phylogenetic analysis was performed with MEGA 4.0 (4) by using the minimum evolution approach. The reliability of the nodes was estimated by minimum evolution bootstrap percentages obtained after 1000 pseudoreplications.

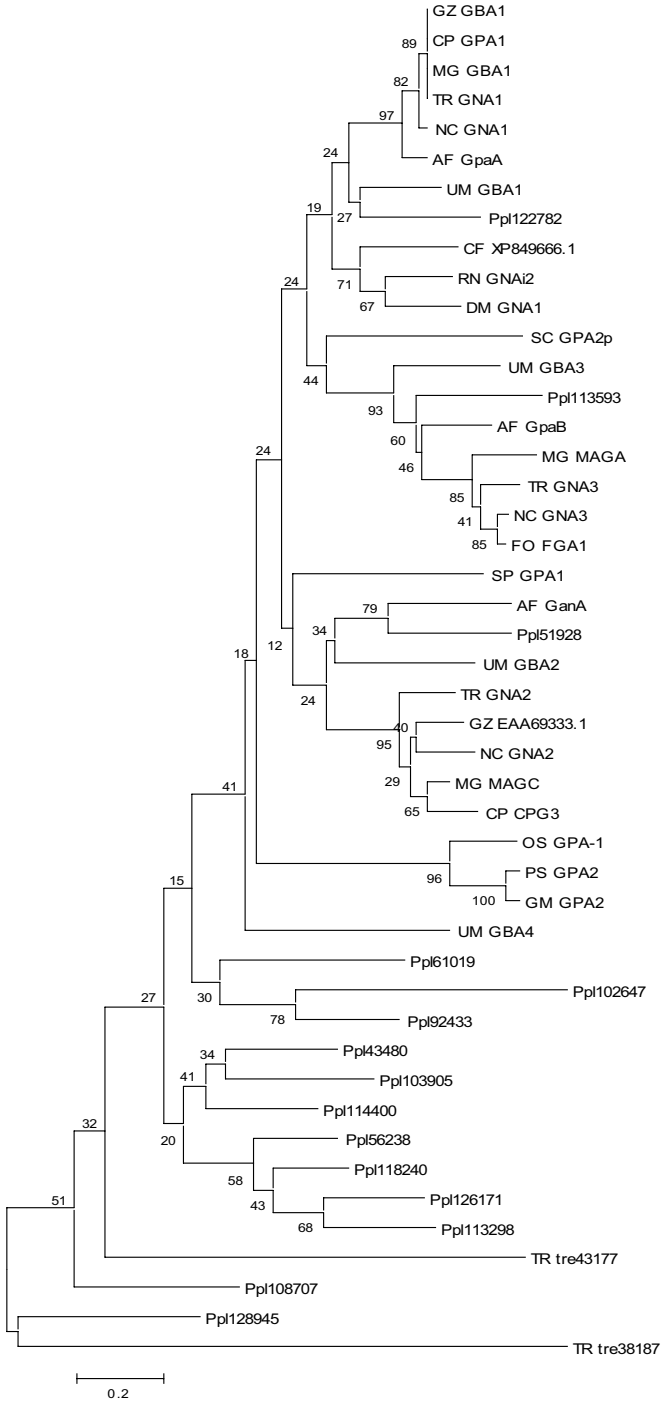

Fig. S4. Phylogenetic analysis of G protein α subunits. The analysis was performed as described with Fig. S3 *Trichoderma reesei (Hypocrea jecoring)* (TR), *Saccharomyces cerevisiae* (SC), *Schizosaccharomyces pombe* (SP), *Neurospora crassa* (NC), *Aspergillus fumigatus* (AF), *Gibberella zeae* (GZ), *Fusarium oxysporum* (FO), *Magnaportne 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).

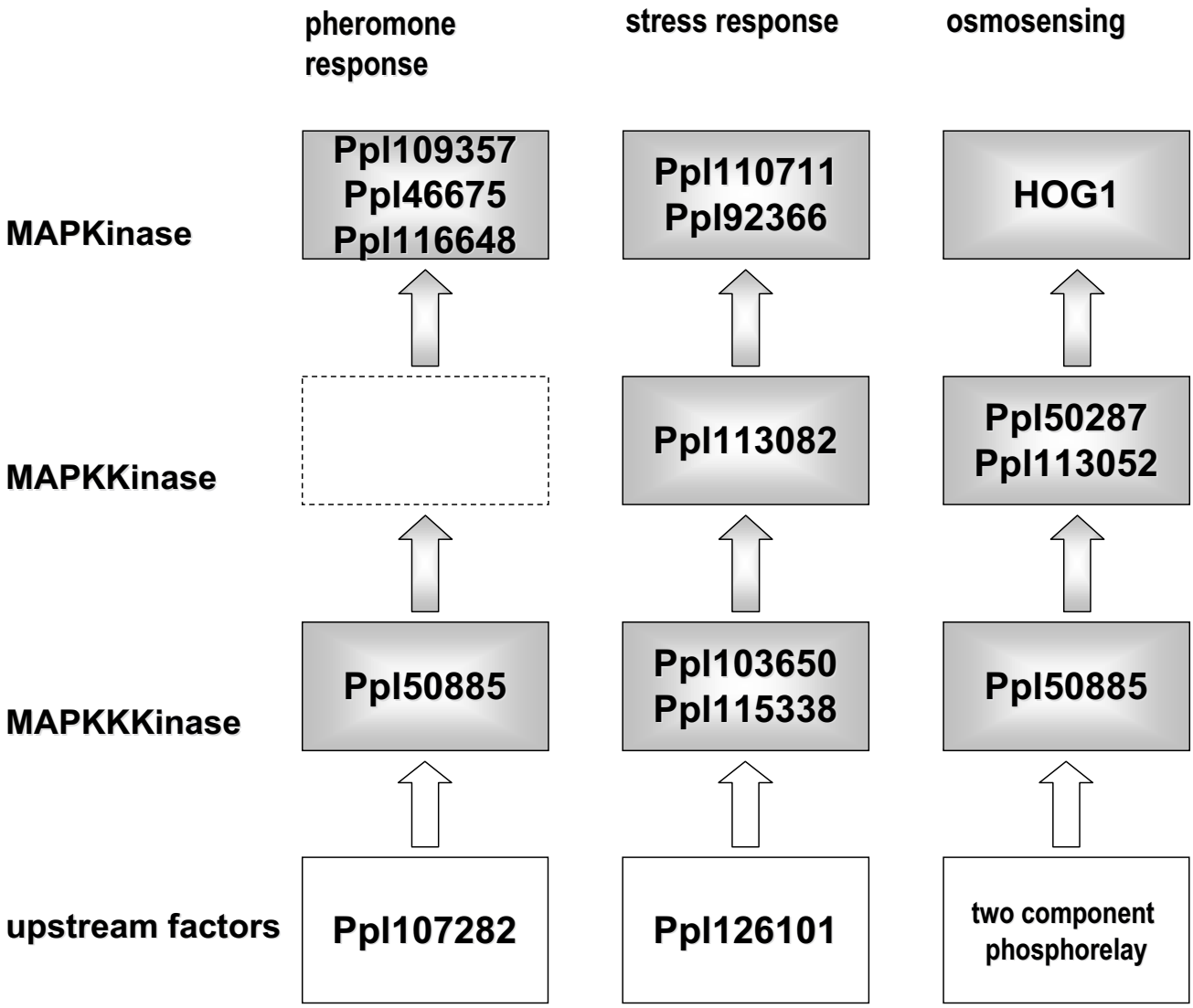

Fig. S5. Schematic representation of putative MAP-kinase pathways in *P. placenta*. Cascades and upstream factors were deduced from Gustin *et al*. (5) and similarity of the respective *P. placenta* predicted proteins as suggested by BLASTP-searches restricted to *Saccharomyces cerevisiae*.

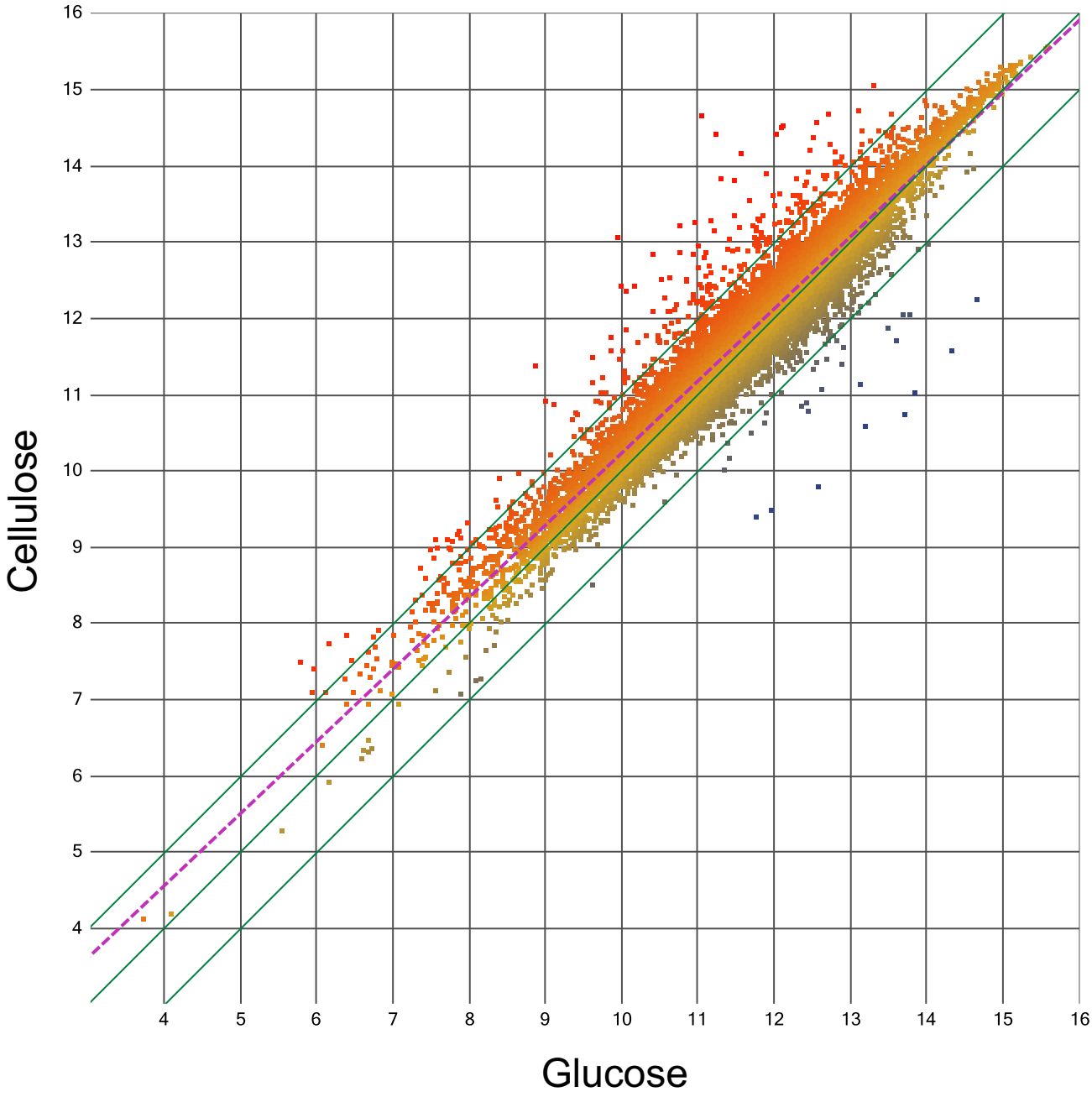

Fig. S6. Scatter plot showing expression of 12,438 *P. placenta* genes in cultures containing glucose or microcrystalline cellulose as sole carbon source. Data normalized and shown as $\log_2$ scale. Middle solid green line is identity (x = y). Outside green lines delineate 290 genes with >2-fold change in intensity in one dataset. Details for these regulated genes appear in si_table_3.xls under GEO accession no. GSE12540. Dashed purple line is best fit linear regression. Values for all genes, together with ANOVA results are listed under GEO accession no. GSE12540.

Table S1. Distribution of putative CAZyme modules in sequenced fungal genomes

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

As defined by Henrissat and coworkers (http://afmb.cnrs.mrs.fr/CAZY/index.html). Family 1 carbohydrate binding modules confer cellulose binding. 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.

Table S2. Extracellular CAZY peptides sequenced by LC-MS/MS

Alleles*	Description	ESTs† Total/Avicel	Microarray‡ *P* value	Microarray‡ Fold change	LC-MS/MS peptides§ total/unique/score 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	0.000182	1.14	4/3/237	–	–
97540/115929	GH37 trehalase	7/2	0.0086	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	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

*Protein model identification number. Underlined model number selected as microarray target. Alternative models shown parenthetically.

†Total number of ESTs derived from all media/number of ESTs from medium with microcrystalline cellulose (Avicel) as sole carbon source.

‡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.

§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.

Table S3. Summary of oxidoreductases potentially involved in lignocellulose degradation by *P. placenta* (Ppl) and *P. chrysosporium* (Pch)

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)

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 ($P < 0.01$; >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).

Other Supporting Information Files

SI Appendix