RESEARCH

Integrative visual omics of the white-rot fungus *Polyporus brumalis* exposes the biotechnological potential of its oxidative enzymes for delignifying raw plant biomass

Shingo Miyauchi[1,10†], Anaïs Rancon[1†], Elodie Drula[1], Hayat Hage[1], Delphine Chaduli[1,2], Anne Favel[1,2], Sacha Grisel[1], Bernard Henrissat[3,4,5], Isabelle Herpoël-Gimbert[1], Francisco J. Ruiz-Dueñas[6], Didier Chevret[7], Matthieu Hainaut[3,4], Junyan Lin[8], Mei Wang[8], Jasmyn Pangilinan[8], Anna Lipzen[8], Laurence Lesage-Meessen[1,2], David Navarro[1,2], Robert Riley[8], Igor V. Grigoriev[8,9], Simeng Zhou[1,11], Sana Raouche[1] and Marie-Noëlle Rosso[1*]

Abstract

Background: Plant biomass conversion for green chemistry and bio-energy is a current challenge for a modern sustainable bioeconomy. The complex polyaromatic lignin polymers in raw biomass feedstocks (i.e., agriculture and forestry by-products) are major obstacles for biomass conversions. White-rot fungi are wood decayers able to degrade all polymers from lignocellulosic biomass including cellulose, hemicelluloses, and lignin. The white-rot fungus *Polyporus brumalis* efficiently breaks down lignin and is regarded as having a high potential for the initial treatment of plant biomass in its conversion to bio-energy. Here, we describe the extraordinary ability of *P. brumalis* for lignin degradation using its enzymatic arsenal to break down wheat straw, a lignocellulosic substrate that is considered as a biomass feedstock worldwide.

Results: We performed integrative multi-omics analyses by combining data from the fungal genome, transcriptomes, and secretomes. We found that the fungus possessed an unexpectedly large set of genes coding for Class II peroxidases involved in lignin degradation (19 genes) and GMC oxidoreductases/dehydrogenases involved in generating the hydrogen peroxide required for lignin peroxidase activity and promoting redox cycling of the fungal enzymes involved in oxidative cleavage of lignocellulose polymers (36 genes). The examination of interrelated multi-omics patterns revealed that eleven Class II Peroxidases were secreted by the fungus during fermentation and eight of them where tightly co-regulated with redox cycling enzymatic partners.

Conclusion: As a peculiar feature of *P. brumalis*, we observed gene family extension, up-regulation and secretion of an abundant set of versatile peroxidases and manganese peroxidases, compared with other Polyporales species. The orchestrated secretion of an abundant set of these delignifying enzymes and redox cycling enzymatic partners could contribute to the delignification capabilities of the fungus. Our findings highlight the diversity of wood decay mechanisms present in Polyporales and the potentiality of further exploring this taxonomic order for enzymatic functions of biotechnological interest.

Keywords: Solid-state fermentation, Lignin degradation, *Polyporus brumalis*, Plant biomass transformation

*Correspondence: marie-noelle.rosso@inra.fr
†Shingo Miyauchi and Anaïs Rancon contributed equally to this work
[1] Aix Marseille Univ, INRA, UMR 1163, Biodiversité et Biotechnologie Fongiques, BBF, Marseille, France
Full list of author information is available at the end of the article

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 2 of 14

Background

White-rot fungi are a group of wood decay fungi that play a major role in carbon cycling. They are the ultimate degraders of highly recalcitrant raw lignocellulosic substrates in forest ecosystems. From a biotechnological aspect, their natural capacities are particularly suited for plant biomass conversions such as the production of bio-based products from renewable raw materials [1]. White-rot fungi can degrade all plant cell wall polymers through the concerted secretion of complex sets of hydrolytic and oxidative enzymes. These enzymes belong to enzyme families including glycoside hydrolases (GH), carbohydrate esterases (CE), pectate lyases (PL), and auxiliary activity oxidoreductases (AA) as classified in the Carbohydrate Active Enzyme database (CAZy; https://www. CAZy.org; [2]). In particular, the degradation of crystalline cellulose is facilitated by cellobiohydrolases (GH6 and GH7) and lytic polysaccharide monooxygenases (LPMOs; CAZy family AA9), which are often linked to Carbohydrate Binding Modules (CBM1). In addition, genes coding for Class II peroxidases of the peroxidase-catalase superfamily involved in the oxidative breakdown of lignin ([3]; CAZy family AA2) are a hallmark of white-rot fungi. Such enzymes include Lignin Peroxidases (LiP; EC 1.11.1.14), manganese peroxidases (MnP; EC 1.11.1.13) and Versatile Peroxidases (VP; EC 1.11.1.16). Other auxiliary enzymes contribute to lignin breakdown in combination with AA2s, such as copper radical oxidases (CAZy subfamily AA5_1) and glucose–methanol–choline (GMC)-oxidoreductases (AA3). Finally, laccases (AA1_1) and dye-decolorizing peroxidases (DyP) contribute to the further processing of lignin oxidation products [4, 5]. Multi-omics approaches have shown that the transcription of these genes and the secretion of the corresponding enzymes are tightly coordinated during the fungal growth on lignocellulosic substrates [6, 7]. Among white-rot fungi, some species selectively degrade a larger proportion of lignin and hemicelluloses rather than cellulose, thereby keeping C6 saccharides preserved. These species have, therefore, been identified as interesting biological agents for the pretreatment of biomass dedicated to bio-energy [8–10]. A recent screen of 63 white-rot fungal strains for selective delignification of wheat straw under solid-state fermentation (SSF) highlighted *Polyporus brumalis* BRFM 985 as the best performing strain [11]. The cultivation of the fungus for 2 weeks on wheat straw significantly improved the accessibility of cellulase-rich enzymatic cocktails to residual carbohydrates [12].
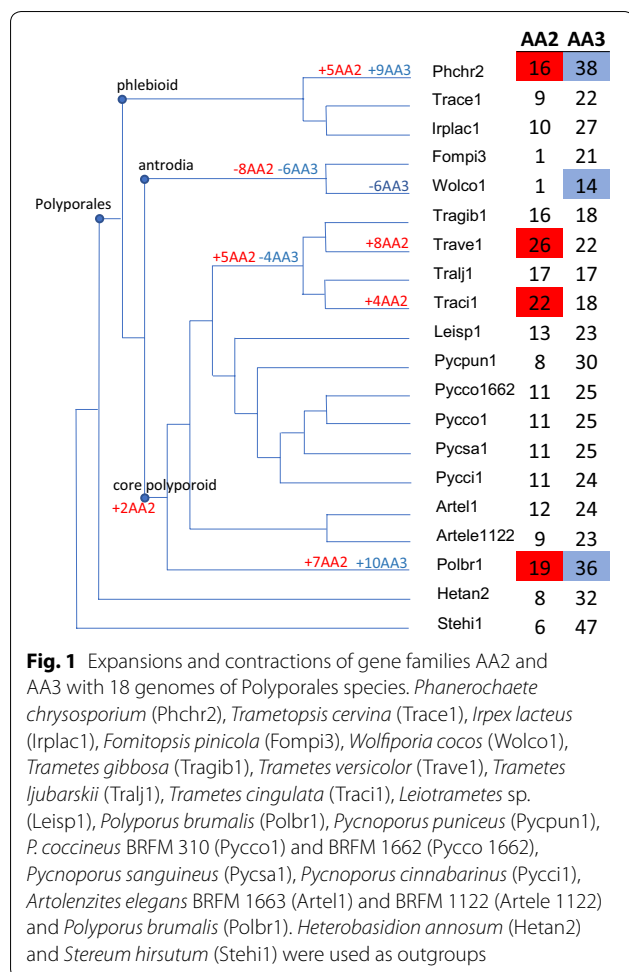
The goal of our study was to understand the lignin-degrading capability of *P. brumalis* using wheat straw in SSF. We conducted an integrative multi-omics analysis by combining data from the genome, transcriptomes, and secretomes. We used the visual multi-omics pipeline SHIN+GO to identify co-regulated genes showing similar transcription patterns throughout the SSF by first integrating time-course transcriptomes with corresponding co-secreted proteins and then converting these data into genome-wide graphical network maps [13]. We found a significant enrichment in genes coding for lignin-active peroxidases in the genome of *P. brumalis*. The enrichment of these genes families enables the rapid deployment of an extraordinary oxidative machinery to drive the efficient decomposition of lignin.

Results

The genome of *P. brumalis* is enriched in genes coding for oxidative enzymes active on lignin

The draft genome sequence (94.9× coverage) of the monokaryotic strain BRFM 1820 is 45.72 Mb large and was assembled into 621 scaffolds and 1040 contigs, with a scaffold L50 of 0.36 Mb. In total, 18,244 protein coding genes were predicted. Expert annotation of the genes coding for enzymes active on lignocellulose revealed typical features of white-rot genomes, such as the presence of GH7 cellobiohydrolases (three genes) and AA9 LPMOs (17 genes) active on crystalline cellulose. The genome of *P. brumalis* also holds a commonly observed suite of enzymes active on hemicelluloses and is rich in enzymes active on acetylated xylooligosaccharides (CE16 acetyl esterases; 14 genes), and on pectin (two GH105 rhamnogalacturonyl hydrolase and 11 GH28 polygalacturonase coding genes). The number of oxidoreductases involved in lignocellulose degradation (AAs, 102 genes) is among the highest compared to sequenced genomes of the taxonomic order Polyporales [14]. We inspected whether CAZy gene family expansions occurred in *P. brumalis* using CAFE, a computational tool that provides statistical analysis of evolutionary changes in gene family size over a phylogenetic tree [15]. The results were consistent with a previous comparison of AA2 gene repertoires in Polyporales genomes, which showed a trend for larger AA2 gene families in the phlebioid clade that contains the deeply studied white-rot fungus *Phanerochaete chrysosporium* (16 AA2 coding genes; Fig. 1, [16]). However, we observed in *P. brumalis* an expansion of the AA2 gene family (19 genes, Additional file 1: Table S1), a feature also found in the core polyporoid clade for *Trametes versicolor* (26 genes) and *Trametes cingulata* (22 genes). AA2s are class II heme-containing peroxidases (PODs), i.e., MnPs, LiPs and VPs able to oxidize lignin [5] and non-ligninolytic Generic Peroxidases (GP: 1.11.1.7). Among them, 16 AA2s from *P. brumalis* were predicted to be secreted and could theoretically participate in the degradation of lignocellulosic substrates in the extracellular space. Remarkably, *P. brumalis* possesses a high number of VPs (nine genes) compared with other Polyporales

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 3 of 14

Fig. 1 Expansions and contractions of gene families AA2 and AA3 with 18 genomes of Polyporales species. *Phanerochaete chrysosporium* (Phchr2), *Trametopsis cervina* (Trace1), *Irpex lacteus* (Irplac1), *Fomitopsis pinicola* (Fompi3), *Wolfiporia cocos* (Wolco1), *Trametes gibbosa* (Tragib1), *Trametes versicolor* (Trave1), *Trametes ljubarskii* (Tralj1), *Trametes cingulata* (Traci1), *Leiotrametes* sp. (Leisp1), *Polyporus brumalis* (Polbr1), *Pycnoporus puniceus* (Pycpun1), *P. coccineus* BRFM 310 (Pycco1) and BRFM 1662 (Pycco 1662), *Pycnoporus sanguineus* (Pycsa1), *Pycnoporus cinnabarinus* (Pycci1), *Artolenzites elegans* BRFM 1663 (Artel1) and BRFM 1122 (Artele 1122) and *Polyporus brumalis* (Polbr1). *Heterobasidion annosum* (Hetan2) and *Stereum hirsutum* (Stehi1) were used as outgroups

species such as *P. chrysosporium* (which lacks VP genes), *Pycnoporus cinnabarinus* (two genes) and *T. versicolor* (two genes). VPs are characterized by wide substrate specificity due to the presence of three different catalytic sites in their molecular structure (i.e., a main heme access channel, a Mn^{2+} oxidation site, and a catalytic tryptophan exposed to the solvent) [17]. Their catalytic properties enable both the direct oxidation of phenolic and non-phenolic units of lignin, and Mn^{3+}-mediated oxidations. Nine short MnPs complete the repertoire of ligninolytic AA2s. Unlike the long MnPs specific for Mn^{2+} (e.g., typical MnPs from *P. chrysosporium*), the members belonging to the short MnP subfamily can also oxidize phenols in Mn-independent reactions [18].

The AA2 gene family expansion in *P. brumalis* was accompanied by the expansion of the gene family AA3 (36 genes, excluding a unique cellobiose dehydrogenase coding gene that has the modular structure AA8-AA3). AA3 enzymes are flavo-oxidases with a flavin-adenine dinucleotide (FAD)-binding domain. Members of the family AA3 contribute to lignocellulose breakdown, among others through the generation of H_2O_2 that activates peroxidases [19] and the regeneration of AA9 LPMOs active on cellulose via the redox cycling of phenolics and the corresponding quinone mediators [20–22]. An additional role for AA3s could include contributing to the oxidative cleavage of polysaccharides through the generation of H_2O_2, which could subsequently be converted to hydroxyl radical $OH^{\bullet}$ by Fenton reaction as previously described in brown-rot fungi [19]. Twelve of these AA3s were predicted to be secreted and belong to the CAZy subfamily AA3_2 which includes predicted aryl alcohol oxidases (AAO, EC 1.1.3.7), aryl-alcohol quinone oxidoreductases (AAQO), glucose 1-oxidases (GOX, EC 1.1.3.4), glucose dehydrogenases (GDH, EC 1.1.5.9) and pyranose dehydrogenases (PDH; EC 1.1.99.29; [23–25]; Additional file 1: Table S1). There was no gene family expansion observed for AA9s in *P. brumalis*. To conclude, we observed the co-occurrence of gene family expansions for putatively secreted lignin-active peroxidases and auxiliary enzymes contributing to the generation of H_2O_2 and redox cycling of electron donors/acceptors, which could contribute to the oxidative degradation of wheat straw polymers by *P. brumalis*.

Polyporus brumalis selectively degrades lignin during solid state fermentation on wheat straw

The dikaryotic strain *P. brumalis* BRFM 985 was grown under solid-state cultivation with wheat straw as source of nutrient and support for 4, 10 or 15 days. Initial mycelium content in the packed bed bioreactor was of 5.8 mg (dw)/g substrate (dw). *P. brumalis* developed a matrix of hairy mycelia covering the wheat straw at Day 15 (Fig. 2a). The qPCR quantification of the fungal biomass (Fig. 2b) showed significant growth from Day 4 to Day 15 and an important weight loss (about 20%) of the total biomass within the bioreactors (Table 1). This result was consistent with previous work performed by Shi and coworkers [26]. As expected, wheat straw was selectively delignified by *P. brumalis* (Table 1). After 15 days culture, the ratios of lignin loss to cellulose loss (L/C) and lignin loss to holocellulose loss (L/H) reached 2.08 and 2.29, respectively. We investigated whether the changes in biochemical composition of wheat straw during fermentation had an impact on the accessibility of carbohydrates to glycoside hydrolases. This was evaluated using a commercial cocktail of enzymes. As shown in Table 1, the carbohydrate net conversion yields of fermented wheat straw increased over time of culture and reached 35% for cellulose at Day 15. For holocellulose, the carbohydrate net conversion yields were smaller (up to 30% at Day 15), which can be explained by the poor content of the cocktail in enzymes active on hemicelluloses.
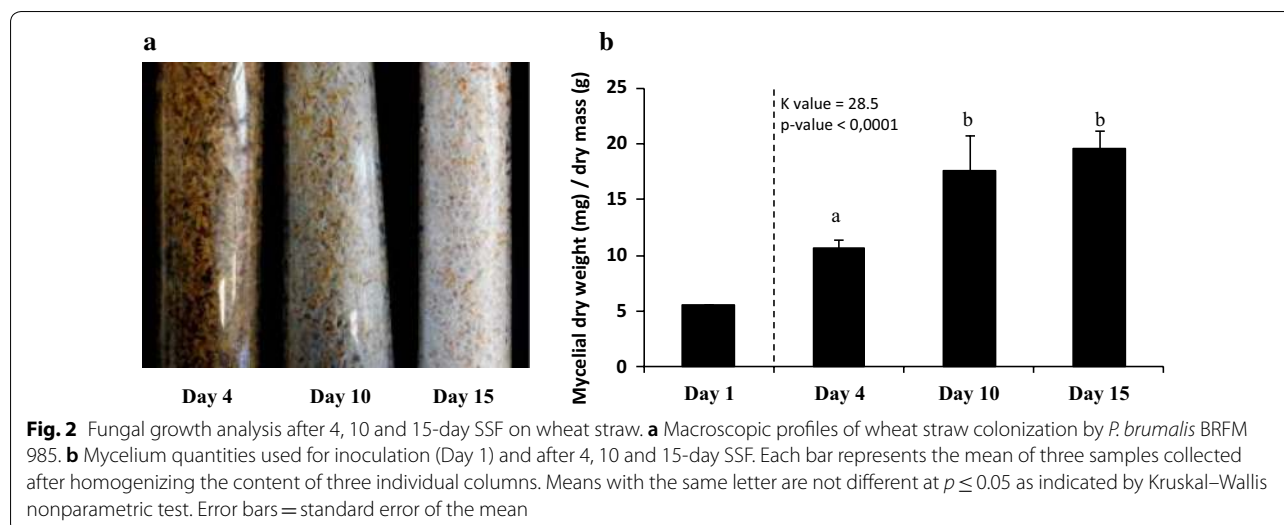
Miyauchi *et al. Biotechnol Biofuels (2018) 11:201*

Page 4 of 14

Fig. 2 Fungal growth analysis after 4, 10 and 15-day SSF on wheat straw. **a** Macroscopic profiles of wheat straw colonization by *P. brumalis* BRFM 985. **b** Mycelium quantities used for inoculation (Day 1) and after 4, 10 and 15-day SSF. Each bar represents the mean of three samples collected after homogenizing the content of three individual columns. Means with the same letter are not different at $p \leq 0.05$ as indicated by Kruskal–Wallis nonparametric test. Error bars = standard error of the mean

Table 1 Weight losses, changes in wheat straw composition, selectivity for lignin degradation and carbohydrate net conversion yields of fermented wheat straw

	Weight loss (%)	Composition loss (%)			Selectivity		Carbohydrate net conversion yield (%)	
		Cellulose	Holocellulose	Lignin	L/C	L/H	Cellulose	Holocellulose
Control	ND	ND	ND	ND	ND	ND	11.1 ± 0.1	10.5 ± 0.1
Day 4	0 ± 1.9	0 ± 1.8	0 ± 2.0	2.4 ± 1.8	ND	ND	19.6 ± 1.0	16.4 ± 0.6
Day 10	10.7 ± 0.1	10.1 ± 0.1	10.3 ± 0.1	18.6 ± 2.2	1.84	1.81	23.6 ± 0.4	20.9 ± 0.3
Day 15	19.6 ± 1.2	17.4 ± 1.2	15.8 ± 1.3	36.2 ± 1.0	2.08	2.29	35.2 ± 1.5	30.0 ± 0.3

ND not determined, *L/C* lignin to cellulose losses ratio, *L/H* lignin to holocellulose losses ratio

Polyporus brumalis activates an extensive arsenal of oxidative enzymes during growth on wheat straw

To identify enzymes responsible for the observed selective degradation of lignin, we investigated the evolution of transcriptomes and secretomes of *P. brumalis* over time by comparing a set of SSF on wheat straw for 4, 10, and 15 days. The robustness of the experimental setup was confirmed by the high consistency of the mean and distributions of the normalized log2 read counts of the transcripts among the biological replicates (Additional file 2: Figure S1). We constructed network maps for the fungal transcriptome and secretome using the visual multi-omics pipeline SHIN + GO, which allowed us to overview the genome-wide transcriptomic and secretomic activities and pinpoint molecular events of interest [7]. We made omics topographies ('Tatami maps') to: (1) visualize nodes containing genes with similar transcription patterns with the corresponding count of secreted proteins; (2) calculate the node-wise mean of the normalized transcript read counts in each condition; (3) identify gene clusters showing high transcription levels and high counts of secreted proteins at specific time points.

To analyze early adaptive responses of the fungus, we compared Tatami maps made from the SSF and liquid conditions. Triplicates of 10-day stationary liquid cultures on malt extract were compared with those from 4-day SSF on wheat straw with an initial 6-day liquid culture on malt extract. The change of the conditions from the liquid culture to SSF on wheat straw triggered rapid shifts in the transcriptome (Fig. 3). A total of 727 genes were differentially highly transcribed after 4-day SSF (Additional file 3: Table S2; Additional file 4: Dataset S2). According to KOG predictions [27], these genes were mainly related to metabolism, cellular processes and signaling, information storage and processing. The importance of intracellular signaling in this adaptive response was highlighted by the induction of 14 genes coding for predicted protein kinases and five for putative transcription factors. In addition, genes involved in detoxification and excretion were upregulated at the early time point including twenty, five, and four genes coding for
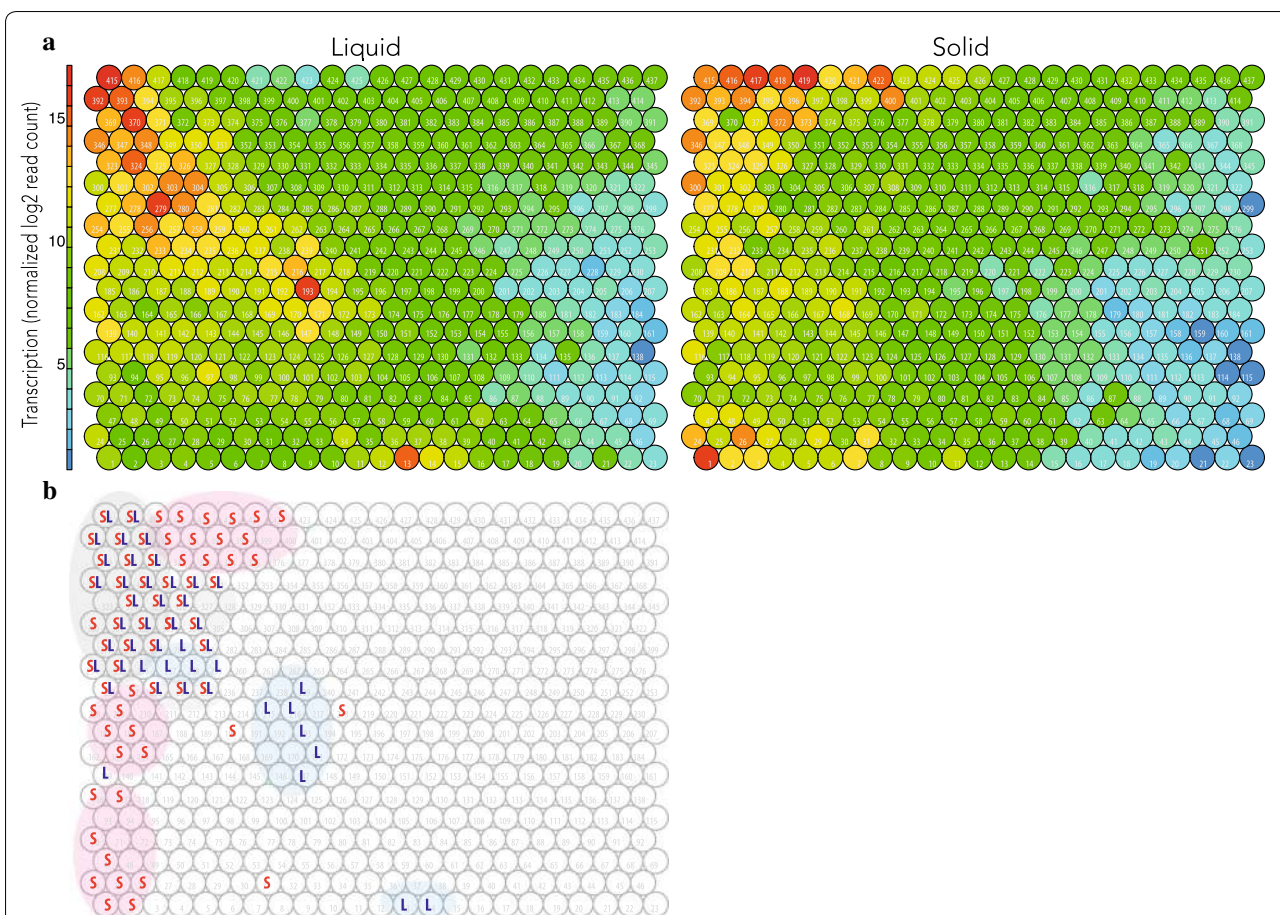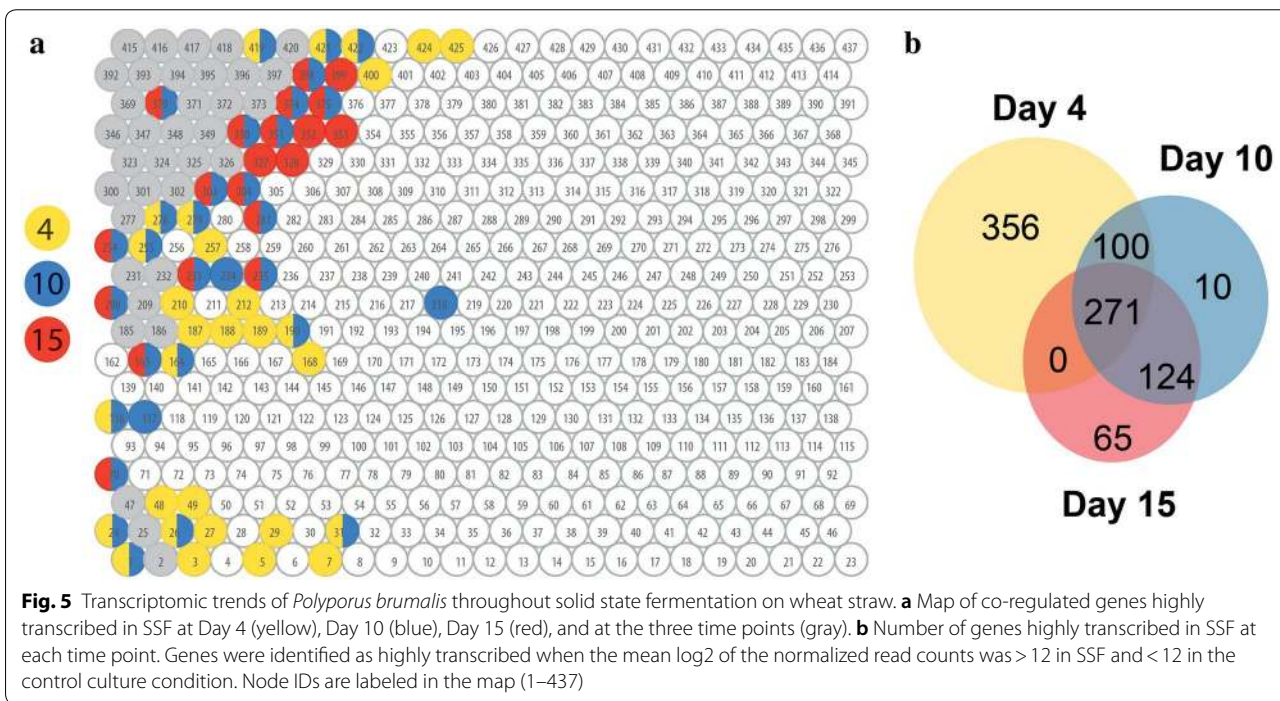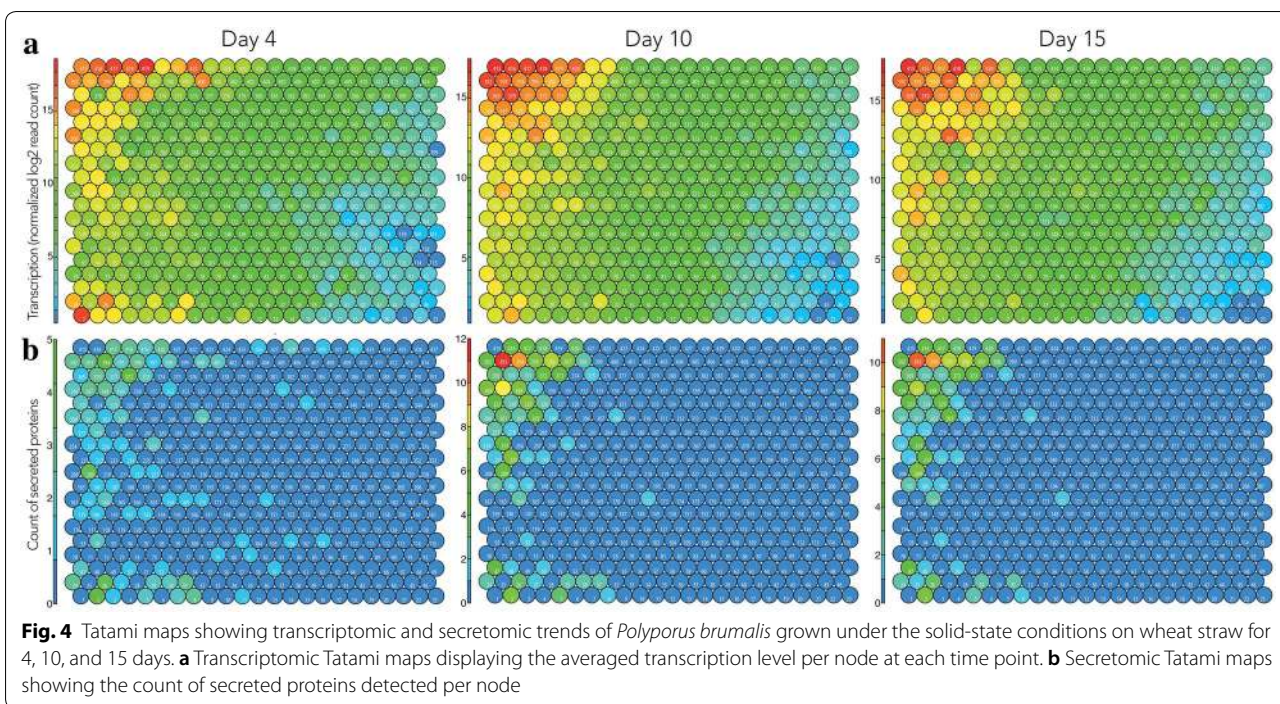
Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 5 of 14

Fig. 3 Tatami maps showing transcriptomic patterns of the averaged biological replicates grown under the solid-state and liquid conditions. Node IDs are labeled in the maps (1–437). **a** Liquid: 10 continuous days of liquid cultivation on malt extract. Solid: 6 days of liquid cultivation on malt extract with 4 days of solid-state cultivation on wheat straw. **b** Nodes with > mean 12 log2 normalized read counts per node in response to the solid-state (S) and liquid (L) conditions

predicted Cytochrome P450, Glutathione S-Transferases (GSTs), and ABC transporters, respectively, suggesting the active detoxification of compounds released from the biomass or newly generated during the wheat straw degradation (Additional file 5: Figure S2). The upregulation of 53 CAZyme coding genes at Day 4 was indicative of the early adaptive response of the fungus to solid-state cultivation and to wheat straw as a carbon source. There were three short MnPs and six VPs identified in the secretome at this early time point. The secretion of delignifying enzymes was accompanied with transcription upregulation and secretion of enzymes active on cellulose, such as Lytic Polysaccharide Monooxygenases (five AA9 LPMOs identified in the secretome), endo- and exo-β-1,4 glucanases (one GH5_5, one GH5_22, one GH6 and three GH7s) and a GH131 β-1,3;1,4-endoglucanase. The coding genes of a CE1 (acetyl-esterase), a GH10 (xylanase), and a GH28 (polygalacturonase) were upregulated and the enzymes were present in the secretome, suggesting the concurrent depolymerization of hemicelluloses and pectin (Additional file 6: Table S3), despite pectin being present in very small quantities in wheat straws [28].

The time-course dynamics of the transcriptome and secretome during SSF were further investigated at Days 10 and 15. We made transcriptome-based Tatami maps at each time point showing the averaged transcription of the replicates (Fig. 4a) and the corresponding secretome-based Tatami maps displaying the count of secreted protein IDs detected per node (Fig. 4b). We observed a few groups of genes (5% of the analyzed protein coding genes) highly transcribed at Day 4 and later on downregulated, and global consistency between transcriptome maps at Days 10 and 15, reflecting that regulations related to the early adaptive response to cultivation under SSF were followed by a global stabilization of the transcription regulations for most of the nodes (Fig. 5). We also observed an apparent time lag in the intracellular processes between transcription and protein secretion.

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 6 of 14

Fig. 4 Tatami maps showing transcriptomic and secretomic trends of *Polyporus brumalis* grown under the solid-state conditions on wheat straw for 4, 10, and 15 days. **a** Transcriptomic Tatami maps displaying the averaged transcription level per node at each time point. **b** Secretomic Tatami maps showing the count of secreted proteins detected per node

Fig. 5 Transcriptomic trends of *Polyporus brumalis* throughout solid state fermentation on wheat straw. **a** Map of co-regulated genes highly transcribed in SSF at Day 4 (yellow), Day 10 (blue), Day 15 (red), and at the three time points (gray). **b** Number of genes highly transcribed in SSF at each time point. Genes were identified as highly transcribed when the mean log2 of the normalized read counts was > 12 in SSF and < 12 in the control culture condition. Node IDs are labeled in the map (1–437)

The correlation of the transcriptomes and secretomes showed that the transcriptomes at Day 4 were more positively correlated with the secretomes at Day 10 and 15. The same trend was seen for Day 10 transcriptome with Day 15 secretome (Table 2).

We determined a core set of 271 highly transcribed genes throughout SSF based on the criteria that the mean normalized read counts per node are more than 12 log2 in SSF and less than 12 log2 in the control condition (Additional file 7: Table S4). Of these, 14 genes encoded

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 7 of 14

Table 2 Correlations between transcriptome and secretome at three time points (*p* value < 0.001)

	Day 4 transcriptome	Day 10 transcriptome	Day 15 transcriptome
Day 4 secretome	0.53	0.48	0.44
Day 10 secretome	*0.64*	0.66	0.62
Day 15 secretome	*0.64*	*0.67*	0.64

Italics: Higher correlation coefficients observed between Day 4 vs Day 10/15 and Day 10 vs Day 15

predicted permeases and transporters of the major facilitator superfamily, four encoded GSTs and 11 encoded P450s. These genes were located in neighboring nodes in three areas of the Tatami map, indicating that subsets of GSTs, P450s, transporters and permeases could have concomitant physiological functions during fungal growth on wheat straw. Six P450 families were upregulated throughout the fermentation: CYP53, CYP512, CYP5144, CYP5136, CYP5137 and CYP 5150. The core set of highly transcribed genes also coded for enzymes active on cellulose (e.g., one GH7 cellobiohydrolase, one GH5_22 endo-β-1,4-glucanase, two AA9 LPMOs), hemicelluloses (e.g., one GH30 and two GH10 endo-1,4-β-xylanases), and seven AA2 peroxidases (three short MnPs and four VPs) that were all detected in the secretomes (Additional file 7: Table S4). We observed the expected

trend that the number of total secreted CAZymes increased from Day 4 (47 secreted CAZymes consisting of 24 GHs, five CEs, 18 AAs) to Day 15 (76 secreted CAZymes consisting of 34 GHs, 12 CEs, two PLs, 26 AAs). We found four nodes of highly transcribed genes during SSF that shared similar transcription profiles and contained a high number of Auxiliary Activity enzymes detected in the secretomes (nodes 417, 418, 419 and 420). These included seven AA2 peroxidases (two short MnPs and five VPs), two AA5_1 copper radical oxidases (Prot-IDs #713930 and #1557562 sharing 71 and 74% identity with the glyoxal oxidase Q01772.1 from *P. chrysosporium*, respectively), and one AA3_3 predicted alcohol oxidase (ProtID #1364967, 88% identity with the characterized alcohol oxidase CDG66232 from *P. chrysosporium* [29]; Table 3). The neighboring node 416 contained an AA3_2 GMC oxidoreductase (ProtID #1422328). The tight co-regulation of these genes over time suggested that these specific AA2, AA3_2, AA3_3 and AA5_1 isoforms participated cooperatively to the oxidative attack of the substrate. In-depth in situ co-localization analyses would be necessary to confirm this hypothesis. Four additional AA2s, two AA3_2s, and one AA5_1 were secreted during SSF (Table 3). The regulated AA3_2 genes (Prot IDs #1399288 and #1343556) were predicted to code for aryl alcohol oxidases (AAOs) as deduced from a phylogenetic analysis which included biochemically characterized AA3_2 (Additional file 8: Figure S3). These AAOs and

Table 3 Genes coding for auxiliary activity enzymes highly transcribed and secreted during SSF on wheat straw

Prot ID	Node ID	Detected in the secretome	Log2 normalized read counts				Predicted function
			Control	Day 4	Day 10	Day 15	
1399288	209	D10, D15	11.43	*13.63*	*13.92*	*13.34*	AA3_2 GMC oxidoreductase
855582	267	D10, D15	6.47	9.20	8.51	8.46	AA2 MnP-short
1353775	331	D15	6.91	9.39	11.21	8.98	AA5_1 copper radical oxidase
1343556	349	D4, D10, D15	*12.06*	*12.84*	*13.27*	*13.28*	AA3_2 GMC oxidoreductase
1486819	372	D4, D10, D15	8.52	*16.89*	*14.66*	*12.89*	AA2 MnP-short
1359988	400	D4, D10, D15	6.73	*17.36*	10.29	7.41	AA2 VP
1422328	416	D10, D15	*15.27*	*15.94*	*16.52*	*14.83*	AA3_2 GMC oxidoreductase
1364967	417	D10, D15	10.76	*17.12*	*16.14*	*15.71*	AA3_3 GMC oxidoreductase
1487275	417	D4, D10, D15	9.84	*18.95*	*17.82*	*15.66*	AA2 VP
713930	418	D10, D15	8.38	*17.61*	*18.31*	*17.41*	AA5_1 copper radical oxidase
897918	418	D4, D10, D15	8.06	*15.10*	*16.76*	*16.49*	AA2 VP
1412926	418	D4, D10, D15	7.42	*14.95*	*17.04*	*16.86*	AA2 VP
1557562	418	D4, D10, D15	6.06	*16.87*	*17.30*	*16.83*	AA5_1 copper radical oxidase
918032	419	D4	7.31	*17.99*	*14.98*	10.42	AA2 VP
1360396	419	D4, D10, D15	6.40	*18.95*	*14.78*	*12.07*	AA2 VP
1185543	420	D4, D10, D15	7.00	*12.95*	*15.69*	*16.08*	AA2 MnP-short
1347226	420	D4, D10, D15	6.63	*12.72*	*15.16*	*16.27*	AA2 MnP-short
1487292	422	D10, D15	5.49	*15.94*	*12.94*	8.21	AA2 MnP-short

Control/Day 4/Day 10/Day 15: average log2 read count of combined biological replicates per condition. Log2 normalized read counts > 12, considered here as high transcription levels, are indicated with italics

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 8 of 14

AA5_1 copper radical oxidase isoforms could be involved in generating H_2O_2 required for AA2s peroxidase activity during growth on wheat straw.

Discussion

Polyporus brumalis is a white-rot fungus found on dead wood of leaf trees in nature and is able to grow on pine wood in laboratory conditions [30]. Recently, the strain BRFM 985 has attracted attention due to its ability to delignify raw wheat straw with moderate consumption of the polysaccharides, which makes this fungus particularly suitable for plant biomass pretreatment in bioprocesses aimed at the valorization of plant-derived saccharides [31]. We examined the genome of *P. brumalis*, explored the time-course series of transcriptomes and secretomes of the fungus grown on wheat straw, and identified the enzymes responsible for the efficient lignin degradation. Besides the set of cellulolytic enzymes commonly found in white-rot fungi, we observed in the genome of *P. brumalis* the expansion of: (1) the Class II peroxidase gene family (AA2, 19 gene copies) including short MnPs and VPs able to oxidize all units of lignin directly or in a Mn^{2+}-mediated reaction; and (2) GMC oxidoreductases/dehydrogenases (AA3, 36 genes excluding the cellobiose dehydrogenase coding gene) assisting lignocellulose breakdown by generating H_2O_2 and by recycling the electron donors and acceptors required for oxidative attack of the polymers. This ligninolytic machinery was completed by an important arsenal of genes coding for laccases (nine gene copies) and AA6 1,4-benzoquinone reductases (three gene copies), which contribute, respectively, to the fractionation of lignin and the generation of extracellular hydroxyl radical [5]. The delignification ability of the fungus was demonstrated by the massive transcription and the secretion of such enzymes during the growth on wheat straw. We identified 11 AA2s (five MnPs and six VPs) that were highly expressed and secreted during the fermentation and seven AA2s (three MnPs and four VPs) that were co-regulated at a consistently high transcription level from Day 4 to Day 15. To the best of our knowledge, such an orchestrated enzymatic delignification system is unprecedented. Four AA2s (one MnP and three LiPs) were detected in the secretome of *Phanerochaete chrysosporium* during SSF on artichoke stalks [32], up to five MnPs were detected in the secretome of *Ceriporiopsis subvermispora* during growth on aspen [33], and six AA2s (three MnPs and three LiPs) were detected in the secretome of *Phlebia radiata* during growth on pine wood [6]. Remarkably, all these fungal species belong to the phlebioid clade of Polyporales which is known to have many lignin-degrading enzymes [34]. We showed that *P. brumalis*, in belonging to the core polyporoids within Polyporales, challenges the view that the phlebioids uniquely contain a lignolytic enzymatic arsenal in their genomes.

The time-course analysis of transcriptional regulations from Day 4 to Day 15 showed the rapid adaptation of the fungus to the substrate. At Day 4, a complete set of genes coding for enzymes active on cellulose, hemicelluloses, pectin, and lignin was strongly up-regulated and the corresponding enzymes were detected in the secretome. These regulations could be partly related to differences in growth condition between static liquid cultures and SSF, as previously observed in *Pleurotus ostreatus* [35, 36]. In most cases, however, the transcript levels were high until Day 15. There was an exception that genes coding for GH5_5 (beta-1,4-endoglucanases), GH7 (cellobiohydrolases), and AA9 (LPMOs) showed decreasing transcript levels from Day 4 to Day 15. As a result, the set of highly transcribed genes coding for cellulose-active enzymes at Day 15 was reduced to one GH5_22 (ProtID # 1412845), one GH7 (#1347545) and three LPMOs (#1339229, 1403153, 1452362). Therefore, we propose that the observed selective delignification could be a result of orchestrated fungal mechanics composed of: (1) the expression and secretion of numerous lignin-active peroxidases with associated oxidoreductases; and (2) the down-regulation of cellulose-degrading enzymes between Day 4 and Day 15.

Degradation of lignin produces by-products such as polyaromatic and phenolic compounds that must be metabolized or detoxified by the fungus. We found that the secretion of lignin-active enzymes was associated with strong up-regulation of genes involved in signal transduction (kinases) or detoxification (Cytochrome P450s, GSTs and ABC transporters), which could play a role in the fungal adaptation to lignin breakdown. Six P450 families were up-regulated throughout fermentation: CYP53, CYP512, CYP5144, CYP5136, CYP5137 and CYP5150. Strikingly, these six gene families are expanded in the genomes of wood decay basidiomycetes [37–39]. CYP512, CYP5136, and CYP5150 families can oxidize a wide range of hydrocarbons, including plant chemicals and key metabolic intermediates in fungi such as steroids, suggesting that they play a role both in oxidation of xenobiotic compounds and fungal primary and secondary metabolism [38]. CYP53 family members oxidize benzoate and benzoate derivatives [40]. The conservation of CYP53 in the genome of fungal plant pathogens suggested that CYP53 could play a key role during colonization of plant tissues. Jawallapersand et al. [41] suggested that CYP53 could have a role in the generation of the secondary metabolite veratryl alcohol, a cofactor for lignin-active peroxidases. Altogether, these findings suggested that CYP53

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 9 of 14

could contribute to the degradation of lignocellulosic polymers and/or to the detoxification of anti-fungal compounds released during plant material degradation.

Conclusion

In this study, we used self-organizing maps to overlay genome annotations with transcriptomics and secretomics data obtained in the time course of *P. brumalis* fermentation on wheat straw. The clustering of co-regulated and co-secreted enzymes showed that the delignification capabilities of the fungus was associated with the orchestrated secretion of an abundant set of delignifying enzymes including several VPs and short MnPs. The secretion by the fungus of a rich set of these enzymes, together with the mobilization of elaborated intracellular detoxification pathways could be the essential determinants for the ability of the fungus to efficiently degrade lignin.

Methods

Fungal strain and growth conditions

Polyporus brumalis BRFM 985 was obtained from the International Center of Microbial Resources (CIRM; https://www6.inra.fr/cirm). The fungus was maintained on 2% (w/v) malt extract, 2% (w/v) agar at 4 °C. Inoculum preparation and SSF procedures were performed as described by Zhou et al. [12]. Briefly, the fungal inoculum was prepared by grinding mats (Ultra-Turrax, 60 s, 9500 rpm) obtained from 7-day-old Roux flasks (malt extract 2%, w/v). Solid-state fermentations were performed in packed bed bioreactors (glass columns) on 20 g dw (dry weight) sterilized wheat straw (≈ 4 mm, Haussmann soft wheat, VIVESCIA, Reims, France) impregnated with 6 mg dw mycelium/g dw substrate, and 1 mL metal solution ($CuSO_4$, $FeSO_4$, and $MnSO_4$ 0.9 μmol each/g dw substrate). The glass-column reactors were incubated at 28 °C for 4, 10 or 15 days. Air stream was filtered, wetted and flow rate set to 0.5/v.v/min. Control wheat straw was incubated under the same conditions without inoculum. Assays were performed in triplicate.

Wheat straw biochemical characterization

Fermented and control wheat straw from three replicates were pooled and homogenized. Dry weights were measured from 1-g (wet weight) aliquots after drying overnight at 105 °C. Five grams (dw) aliquots were washed with deionized water (5% w dw/v) for 24 h at 4 °C, with shaking at 300 rpm, and filtered through GF/F filters (Whatman). The recovered solid extracts were freeze dried and stored at 4 °C for biochemical composition and enzymatic hydrolysis analyzes. The recovered water extracts were stored at −20 °C for secretome analyzes.

Biochemical composition (klason lignin, cellulose and hemicelluloses) of the fermented wheat straw was determined in duplicate according to the NERL method [42]. Carbohydrate accessibility to enzymatic degradation was evaluated by measuring carbohydrate net conversion yields as described by Zhou et al. [12]. Briefly, control or fermented wheat straw was subjected to mild alkali treatment, followed by a 72 h enzymatic hydrolysis with 12 FPU/g substrate (dw) of cellulases from *Trichoderma reesei* (GC220, Genencor Danisco) and 60 U/g substrate (dw) of β-glucosidase from *Aspergillus niger* (Novozyme SP188, Sigma). The released glucose and reducing sugars were quantified, respectively, using the Glucose RTU kit (Biomérieux) and the dinitrosalicylic acid method [43].

Fungal biomass quantification

Fungal biomass was quantified using a PCR-based method as described by Zhou et al. [44]. Briefly, genomic DNA was extracted from fermented wheat straw samples using the NucleoSpin PlantII kit (Macherey–Nagel, France) and quantified by qPCR amplification of a 150-bp fragment in the 5.8S conserved sequence. Standard curves were established using known quantities of genomic DNA from *P. brumalis* BRFM 985. The PCR cycle was as follows: 30 s at 95 °C, and then 5 s at 95 °C, 5 s at 60 °C for 39 cycles, followed by a melt curve step (65–95 °C, with 0.5 °C steps). The presence of a single amplicon was checked on the melting curve. All reactions were performed in triplicate, and all qPCR runs included a negative control without template. Quantification cycles (Cq) were determined using the regression mode of the Bio-Rad CFX Manager™ software (v 3.0).

Genome sequencing

To facilitate genome sequencing and assembly, the monokaryotic strain BRFM 1820 was obtained by de-dikaryotization [45] from the dikaryotic strain BRFM 985. The method for protoplast isolation was adapted from Alves et al. [46]. All steps were performed at 30 °C. Fifteen 7-day-old cultures (9 cm in diameter) of BRFM 985 grown on 2% (w/v) Malt extract, 2% (w/v) agar were recovered and homogenized for 20 s with a Waring blender in 100 mL of YM medium (10 g/L glucose, 5 g/L bactopeptone, 3 g/L yeast extract and 3 g/L malt broth extract). The homogenate was grown for 24 h with agitation at 130 rpm. The culture was again homogenized 20 s, diluted twice in YM, and grown for an additional 24 h with agitation. After washing with 0.5 M $MgSO_4 \cdot 7H_2O$, the mycelium was harvested by centrifugation, weighted and resuspended in lysis buffer (0.5 M $MgSO_4$, $7H_2O$, 0.03 M maleic acid pH 5.8 with 15 mg Caylase C4 per mL and per 0.25 g wet weight mycelium). The mycelium was incubated with gentle shaking (80 rpm) for 3 h and

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 10 of 14

filtered through Miracloth. Protoplasts were collected by centrifugation at low speed and gently resuspended in 1 M sorbitol (10^7/mL). Dilutions (10^4/mL) were incubated overnight without shaking in regeneration medium (0.5 M $MgSO_4,7H_2O$, 20 g/L glucose, 0.46 g/L KH_2PO_4, 1 g/L K_2HPO_4, 2 g/L bactopeptone and 2 g/L yeast extract) and spread on MA2 medium. After culturing for 3–7 days, slow growing colonies obtained from the regenerated protoplasts were examined by microscopy to select the ones without clamp-connection. Fragments of mycelium were stained with DAPI to verify they were monokaryotic. From the 20 protoplast-derived monokaryons isolated, the monokaryotic line BRFM 1820, phenotypically similar to its dikaryotic parent, was selected for genome sequencing.

For genome, two (270 bp fragment and long 4 kb mate-pair) libraries were sequenced. For 270 bp fragments, 100 ng of DNA was sheared to 300 bp using the Covaris LE220 and size selected using SPRI beads (Beckman Coulter). For the 4 kb library, 5 μg of DNA was sheared using the Covaris g-TUBE™ (Covaris) and gel size selected for 4 kb. The sheared DNA was treated with end repair and ligated with biotinylated adapters containing loxP. The adapter ligated DNA fragments were circularized via recombination by a Cre excision reaction (NEB) and randomly sheared using the Covaris LE220 (Covaris). Sheared DNA fragments were processed for ligation to Illumina compatible adapters (IDT, Inc) using the KAPA-Illumina library creation kit (KAPA biosystems). For transcriptome, stranded cDNA libraries were generated using the Illumina Truseq Stranded RNA LT kit. Sequencing was performed on the Illumina HiSeq 2500 sequencer using HiSeq TruSeq SBS sequencing kits, v4, following a 2×150 bp (or 2×100 bp for 4 kb library) indexed run recipe. Each fastq file was QC filtered for artifact/process contamination. Genomic reads were assembled with AllPathsLG version R49403 [47]. RNA-Seq reads were assembled using Rnnotator v. 3.3.2 [48].

Gene functional annotation

The genome was annotated using the Joint Genome Institute (JGI) Annotation Pipeline and made publicly available via JGI fungal genome portal MycoCosm [49]. Proteins predicted to be secreted by SignalP 4.1. (threshold 0.34; [50]), TargetP 1.1 [51] or Phobius 1.01 [52] were analyzed with pscan to withdraw proteins targeted at the endoplasmic reticulum. The retrieved proteins with no transmembrane domain according to TMHMM Server 2.0. or a single transmembrane domain corresponding to the signal peptide were identified as predicted secreted. For CAZyme gene expert annotation, all putative proteins were compared to the entries in the CAZy database [2] using BLASTP. The proteins with *E* values

smaller than 0.1 were further screened by a combination of BLAST searches against individual protein modules belonging to the AA, GH, GT (Glycosyl Transferases), PL, CE and CBM classes (https://www.CAZy.org/). HMMer 3 [53] was used to query against a collection of custom-made hidden Markov model (HMM) profiles constructed for each CAZy family. All identified proteins were manually curated. Within families, subfamilies were manually defined according to their homology relationships between members of the focal family. Class II ligninolytic peroxidases (AA2s) were annotated as LiP, MnP or VP on the basis of the presence or absence of specific amino acid residues at the substrate oxidation sites [17] after homology modeling using crystal structures of related peroxidases as templates and programs implemented by the automated protein homology modeling server "SWISS-MODEL" [54].

Proteins classified as KOG "Cytochrome P450" were subject to BLASTP comparisons against the Fungal CYP Database (http://p450.riceblast.snu.ac.kr, last accessed June 10, 2018) [55]. These predicted CYPs were assigned to the corresponding family types based on their highest sequence similarity (at least 40%) against all fungal CYPs as followed by the International P450 Nomenclature Committee.

Gene family expansions/contractions

CAZyme gene family expansions or contractions were analyzed using the CAFE Software v3.1 (Computational Analysis of gene Family Evolution; [15]) and the numbers of genes coding for CAZymes from 18 Polyporales genomes. The selected public genomes were those from *Pycnoporus cinnabarinus* [56], *Phanerochaete chrysosporium* [57], *Fomitopsis pinicola*, *Wolfiporia cocos* and *Trametes versicolor* [58]. Genomes from *Artolenzites elegans* BRFM 1663 and BRFM 1122, *Irpex lacteus* CCBAS Fr. 238 617/93, *Leiotrametes* sp. BRFM 1775, *Pycnoporus coccineus* BRFM 310 and BRFM 1662, *P. puniceus* BRFM 1868, *P. sanguineus* BRFM 1264, *Trametopsis cervina* BRFM 1824, *Trametes cingulata* BRFM 1805, *T. gibbosa* BRFM 1770 and *T. ljubarskii* BRFM 1659 were newly sequenced at the JGI. The Russulales *Stereum hirsutum* [58] and *Heterobasidion annosum* [59] were used as outgroups. The protein sequences deduced from these genomes were downloaded from the Mycocosm portal (https://genome.jgi.doe.gov/programs/fungi/index.jsf), with authorization of the principal investigators when not yet published. Groups of orthologs were obtained from these genomes using orthoMCL version 2.0.9 [60] with default parameters and using NCBI BLAST version 2.4.0 + [61] in conjunction with MCL version 14-137 (http://micans.org/mcl/) under default settings. A set of 20 uni-copy genes was selected among the 25 best

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 11 of 14

performing genes for Polyporales phylogenetics ([62]; Additional file 9: Dataset S1). The concatenated 20-gene dataset in each genome had a mean number of 28,996,9 amino acids, the shortest being for Pycci1 (28144 aa) and the longest for Stehi1 (30313 aa). Alignments were done using Mafft v7.271 [63] and filtered with GBLOCKS version 0.91b [64]. The final tree was generated with RaxML version 8.2.4 [65] and PROTGAMMAWAG as substitution model.

Identification of the fungal proteins secreted during fermentation

The water extracts collected after 4, 10 and 15 days of culture were filtered using 0.22-μm pore-size polyethersulfone membranes (Vivaspin, Sartorius), diafiltered with 50 mM sodium acetate (pH 5.0) and concentrated using a Vivaspin polyethersulfone membrane with a 10-kDa cutoff (Sartorius). LC–MS/MS analysis of the secretomes was performed as described by Navarro et al. [66]. Briefly, 10 μg of proteins was in-gel digested using a standard trypsinolysis protocol. For protein identification, online analysis of peptides was performed with a Q-exactive mass spectrometer (Thermo Fisher Scientific), using a nanoelectrospray ion source. MS/MS data were queried against the catalog of predicted proteins from the *P. brumalis* genome, and an in-house contaminant database, using the X!Tandem software (X!Tandem Cyclone, Jouy en Josas, France). Peptides that matched with an E value < 0.05 were parsed with the X!Tandem pipeline software. Proteins identified with at least two unique peptides and a log (E value) < -2.6 were validated.

Transcriptome and secretome network maps

Total RNA was extracted from 100 mg tissue as described in Couturier et al. [67]. RNA quantity and quality were determined using the Experion RNA Std-Sens kit (QIAGEN). RNASeq was done using 150 bp long paired end reads obtained on Illumina HiSeq-2500 (Beckman Coulter Genomics). The quality of raw fastq reads was checked with FastQC (Simon Andrews, Babraham Bioinformatics, 2011; https://www.bioinformatics.babraham. ac.uk/projects/) and reads were cleaned with Sickle Tool (Joshi and Fass, 2011; https://github.com/najoshi/sickle) with the following criteria: quality threshold of 20; length threshold of 20. RNA reads were aligned to the genome of *P. brumalis* BRFM 1820 using TopHat2 with only unique mapping allowed. Read counts were determined by HTSeq, normalized using ddsNorm from the DESeq 2 Bioconductor package [68] and log2 transformed.

Genome-wide profiles with integrated transcriptome and secretome data were constructed using the SHIN part of the pipeline SHIN + GO [7]. A single self-organizing map (SOM) was generated with the normalized log2 read count of the genes from all biological replicates. The map units used were 437. The number of iterations used was 43,700 (437×100) times to minimize the Euclidian distances between the nodes for the optimal convergence. Two sets of omics topographies (Tatami maps) were generated based on the master SOM. Transcriptomic maps showing node-wise mean transcription were created with the averaged read count of the genes clustered in each node. Secretomic maps were made with the count of detected proteins. For each protein ID, the presence of the protein in the secretome was counted as one. The total count of secreted proteins was calculated node-wise according to the trained SOM. A summary map indicating condition-specific highly transcribed genes was also made. Genes in the nodes with a mean read count > 12 log2 were considered as highly transcribed in each condition at each time point. Such genes constituted approximately above 75th percentile of the entire gene population.

Spearman's rank correlations of fungal transcriptome and secretome were estimated at individual time points using the node-wise mean transcription and the node-wise count of secreted proteins identified. Correlations were calculated for: (1) the entire genes; and (2) the subset of genes coding for predicted secreted proteins.

Additional files

Additional file 1: Table S1. List of predicted Auxiliary Activity enzymes from CAZy families AA2 and AA3 encoded in the genome of *Polyporus brumalis* BRFM 1820. Expert annotations for AA3 sub-families, versatile peroxidases (VP), manganese peroxidases (MnP) and generic peroxidases (GP) are indicated. The Cellobiose Dehydrogenase with the modular structure AA8-AA3_1, ProtID #1364243, is not indicated.

Additional file 2: Figure S1. Distributions of the normalized log2 transformed read count of all biological replicates. Liq: Liquid cultivation on malt extract for 10 days. Day4: Solid-state cultivation on wheat straw for 4 days after liquid cultivation on malt extract for 6 days. Day10: Solid-state cultivation on wheat straw for 10 days after liquid cultivation on malt extract for 6 days. Day15: Solid-state cultivation on wheat straw for 15 days after liquid cultivation on malt extract for 6 days. A: Box plot showing the mean and the range of the read count of the genes. B: Density plot showing distributions of log2 read count of the genes.

Additional file 3: Table S2. The number of highly transcribed genes in the selected nodes (> mean 12 log2 normalized read count per node) in response to the solid and liquid conditions. Total: The total number of genes including the unique and shared genes. Unique: specifically highly transcribed in each condition. Shared: Highly transcribed in both conditions. Node IDs containing such genes are provided (Additional file 7: Table S4).

Additional file 4: Dataset S2. Mean transcript read counts and numbers of proteins detected in the secretomes for each node of the Tatami map.

Additional file 5: Figure S2. KOG classification of the genes differentially highly transcribed after 4-day growth in SSF on wheat straw as compared to control liquid cultures. The identified genes were clustered into nodes of co-regulated genes with mean Log2 of the normalized read counts > 12 in SSF and < 12 in control liquid cultures. The number of up-regulated

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 12 of 14

genes in each KOG class is indicated for the KOG groups Cellular Processes and Signaling, Information storage and Processing and Metabolism.

Additional file 6: Table S3. CAZyme coding genes present in nodes with mean log2 read counts < 12 in control liquid cultures and mean log2 read counts > 12 at Day 4 in SSF.

Additional file 7: Table S4. List of the core set of genes differentially highly transcribed at Day 4, 10 and 15 of SSF on wheat straw.

Additional file 8: Figure S3. Phylogeny of the AA3 predicted proteins from Polyporus brumalis and biochemically characterized fungal AA3s. Sequence descriptions include accession numbers from JGI Mycocosm, NBCI or PDB. The tree was constructed with a strategy similar to Sützl et al. (Sützl et al. 2018 [22]). Sequences were aligned using M-coffee (Wallace et al. [69]) with default settings. Phylogeny was inferred using PhyML (Guindon et al. [70]) and WAG amino acid substitution model (Whelan and Goldman [71]). Branch support was calculated by 500 bootstrap repetitions. The tree was visualized in iTOL (Letunic and Bork [72]).

Additional file 9: Dataset S1. Uni-copy genes selected for the construction of the phylogenetic species tree including 16 Polyporales species; *Artolenzites elegans, Fomitopsis pinicola* (Floudas et al. [58]), *Irpex lacteus, Leiotrametes* sp., *Phanerochaete chrysosporium* (Ohm et al. [57]), *Polyporus brumalis, Pycnoporus cinnabarinus* (Levasseur et al. [56]), *P. coccineus, P. puniceus, P. sanguineus, Trametopsis cervina, Trametes cingulata, T. gibbosa, T. ljubarskyi, T. versicolor* (Floudas et al. [58]), *Wolfiporia cocos* (Floudas et al. [58]) and two Russulales species: *Heterobasidion annosum* (Olson et al. [59]), *Stereum hirsutum* (Floudas et al.[58]). For each gene, the Mycocosm genome name, the putative yeast ortholog, and the protID in Mycocosm (https://genome.jgi.doe.gov/programs/fungi/index.jsf) are provided.

Abbreviations
AA: auxiliary activity; AAO: aryl alcohol oxidase; CAZyme: carbohydrate-active enzyme; CBM: carbohydrate-binding module; CE: carbohydrate esterase; CIRM: International Center of Microbial Resources; DAPI: 4′,6-diamidino-2-phenylindole; DyP: dye-decolorizing peroxidase; FAD: flavin-adenine dinucleotide; FPU: filter paper cellulase unit; GH: glycoside hydrolase; GMC: glucose methanol choline; GOX: glucose 1-oxidase; GP: generic peroxidase; GST: glutathione S-transferase; GT: glycosyltransferase; JGI: Joint Genome Institute; LiP: lignin peroxidase; LPMO: lytic polysaccharide monooxygenase; MA: malt agar; MCL: Markov cluster; MnP: manganese peroxidase; MS: mass spectrometry; NCBI: National Center for Biotechnology Information; NREL: National Renewable Energy Laboratory; PCR: polymerase chain reaction; PL: polysaccharide lyase; POD: class II heme-containing peroxidase; SOM: self organizing map; SSF: solid-state fermentation; VP: versatile peroxidase; YM: yeast mold.

Authors' contributions
AR, SG and SZ carried out fungal cultures and prepared biological material for transcriptomics and secretomics analyses; AR and SM analyzed the transcriptomics and secretomics data; AF, DC, DN, LLM prepared the material for genome sequencing; ED and RR performed the CAFÉ analysis; DC performed the proteomic analysis; RR, carried out functional annotations; BH and MH performed the expert annotation of CAZymes; FJRD performed the expert annotation of class II peroxidases; HH performed the phylogenetic analysis of AA3s; AL, JP, ML and MW performed genome and transcriptome sequencing and assembly; IVG supervised genome sequencing, assembly and annotation; MNR, IHG and SR conceived the study, participated in its design and coordination; AR, SM, SR and MNR wrote the manuscript. All authors read and approved the final manuscript.

Author details
[1] Aix Marseille Univ, INRA, UMR 1163, Biodiversité et Biotechnologie Fongiques, BBF, Marseille, France. [2] CIRM-CF, UMR1163, INRA, Aix-Marseille Univ, Marseille, France. [3] UMR 7257, CNRS, Aix-Marseille Univ, Marseille, France. [4] INRA, USC 1408, AFMB, Marseille, France. [5] Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia. [6] Centro de Investigaciones Biológicas, CSIC, Madrid, Spain. [7] INRA, UMR1319, Micalis, Plateforme d'Analyse Protéomique de Paris Sud-Ouest, Jouy-en-Josas, France. [8] US Department of Energy Joint Genome Institute, Walnut Creek, CA, USA. [9] Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, CA, USA. [10] Present Address: Laboratoire d'Excellence ARBRE, UMR 1136, INRA-Université de Lorraine 'Interactions Arbres/Microorganismes', Champenoux, France. [11] Present Address: Institut des Sciences Moléculaires de Marseille, UMR 7313, CNRS, Aix-Marseille Université, Marseille, France.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Wan C, Li Y. Fungal pretreatment of lignocellulosic biomass. Biotechnol Adv. 2012;30:1447–57.
2. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014;42:D490–5.
3. Zámocký M, Hofbauer S, Schaffner I, Gasselhuber B, Nicolussi A, Soudi M, Pirker KF, Furtmüller PG, Obinger C. Independent evolution of four heme peroxidase superfamilies. Arch Biochem Biophys. 2015;574:108–19.
4. Riley R, Salamov A, Brown D, Nagy L, Floudas D, Held B, Levasseur A, Lombard V, Morin E, Otillar R, Lindquist E, Sun H, LaButti K, Schmutz J, Jabbour D, Luo H, Baker S, Pisabarro A, Walton J, Blanchette R, Henrissat B, Martin F, Cullen D, Hibbett D, Grigoriev I. Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. Proc Natl Acad Sci USA. 2014;111:9923–8.
5. Martínez AT, Speranza M, Ruiz-Dueñas FJ, Ferreira P, Camarero S, Guillén F, Martínez MJ, Gutiérrez A, del Río JC. Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. Int Microbiol. 2005;8:195–204.
6. Kuuskeri J, Häkkinen M, Laine P, Smolander O-P, Tamene F, Miettinen S, Nousiainen P, Kemell M, Auvinen P, Lundell T. Time-scale dynamics of proteome and transcriptome of the white-rot fungus *Phlebia radiata*: growth on spruce wood and decay effect on lignocellulose. Biotechnol Biofuels. 2016;9:192.

Miyauchi *et al. Biotechnol Biofuels* (2018) 11:201

Page 13 of 14

7. Miyauchi S, Navarro D, Grisel S, Chevret D, Berrin J-G, Rosso M-N. The integrative omics of white-rot fungus *Pycnoporus coccineus* reveals co-regulated CAZymes for orchestrated lignocellulose breakdown. PLoS ONE. 2017;12:e0175528.

8. Liu J, Wang ML, Tonnis B, Habteselassie M, Liao X, Huang Q. Fungal pretreatment of switchgrass for improved saccharification and simultaneous enzyme production. Bioresour Technol. 2013;135:39–45.

9. Vasco-Correa J, Li Y. Solid-state anaerobic digestion of fungal pretreated *Miscanthus sinensis* harvested in two different seasons. Bioresour Technol. 2015;185:211–7.

10. Zhao L, Cao G-L, Wang A-J, Ren H-Y, Dong D, Liu Z-N, Guan X-Y, Xu C-J, Ren N-Q. Fungal pretreatment of cornstalk with *Phanerochaete chrysosporium* for enhancing enzymatic saccharification and hydrogen production. Bioresour Technol. 2012;114:365–9.

11. Zhou S, Raouche S, Grisel S, Navarro D, Sigoillot J-C, Herpoël-Gimbert I. Solid-state fermentation in multi-well plates to assess pretreatment efficiency of rot fungi on lignocellulose biomass. Microb Biotechnol. 2015;8:940–9.

12. Zhou S, Herpoël-Gimbert I, Grisel S, Sigoillot JC, Sergent M, Raouche S. Biological wheat straw valorization: multicriteria optimization of *Polyporus brumalis* pretreatment in packed bed bioreactor. Microbiologyopen. 2017;7(1):e00530. https://doi.org/10.1002/mbo3.530.

13. Miyauchi S, Navarro D, Grigoriev IV, Lipzen A, Riley R, Chevret D, Grisel S, Berrin J-G, Henrissat B, Rosso M-N. Visual comparative omics of fungi for plant biomass deconstruction. Front Microbiol. 2016;7:1335.

14. Hori C, Gaskell J, Igarashi K, Samejima M, Hibbett D, Henrissat B, Cullen D. Genomewide analysis of polysaccharides degrading enzymes in 11 white- and brown-rot Polyporales provides insight into mechanisms of wood decay. Mycologia. 2013;105:1412–27.

15. Han MV, Thomas GWC, Lugo-Martinez J, Hahn MW. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. Mol Biol Evol. 2013;30:1987–97.

16. Ruiz-Dueñas FJ, Lundell T, Floudas D, Nagy LG, Barrasa JM, Hibbett DS, Martínez AT. Lignin-degrading peroxidases in Polyporales: an evolutionary survey based on 10 sequenced genomes. Mycologia. 2013;105:1428–44.

17. Ruiz-Dueñas FJ, Morales M, García E, Miki Y, Martínez MJ, Martínez AT. Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot. 2009;60:441–52.

18. Fernández-Fueyo E, Acebes S, Ruiz-Dueñas FJ, Martínez MJ, Romero A, Medrano FJ, Guallar V, Martínez AT. Structural implications of the C-terminal tail in the catalytic and stability properties of manganese peroxidases from ligninolytic fungi. Acta Crystallogr D. 2014;70:3253–65.

19. Ferreira P, Carro J, Serrano A, Martínez AT. A survey of genes encoding H2O2-producing GMC oxidoreductases in 10 Polyporales genomes. Mycologia. 2015;107:1105–19.

20. Kracher D, Scheiblbrandner S, Felice AKG, Breslmayr E, Preims M, Ludwicka K, Haltrich D, Eijsink VGH, Ludwig R. Extracellular electron transfer systems fuel cellulose oxidative degradation. Science. 2016;352:1098–101.

21. Garajova S, Mathieu Y, Beccia MR, Bennati-Granier C, Biaso F, Fanuel M, Ropartz D, Guigliarelli B, Record E, Rogniaux H, Henrissat B, Berrin J-G. Single-domain flavoenzymes trigger Lytic Polysaccharide Monooxygenases for oxidative degradation of cellulose. Sci Rep. 2016;6:28276.

22. Sützl L, Laurent CVFP, Abrera AT, Schütz G, Ludwig R, Haltrich D. Multiplicity of enzymatic functions in the CAZy AA3 family. Appl Microbiol Biotechnol. 2018;102:2477–92.

23. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol Biofuels. 2013;6:41.

24. Piumi F, Levasseur A, Navarro D, Zhou S, Mathieu Y, Ropartz D, Ludwig R, Faulds CB, Record E. A novel glucose dehydrogenase from the white-rot fungus *Pycnoporus cinnabarinus*: production in *Aspergillus niger* and physicochemical characterization of the recombinant enzyme. Appl Microbiol Biotechnol. 2014;98:10105–18.

25. Mathieu Y, Piumi F, Valli R, Aramburu JC, Ferreira P, Faulds CB, Record E. Activities of secreted aryl alcohol quinone oxidoreductases from *Pycnoporus cinnabarinus* provide insights into fungal degradation of plant biomass. Appl Environ Microbiol. 2016;82:2411–23.

26. Shi J, Sharma-Shivappa RR, Chinn M, Howell N. Effect of microbial pretreatment on enzymatic hydrolysis and fermentation of cotton stalks for ethanol production. Biomass Bioenergy. 2009;33:88–96.

27. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DW, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. The COG database: an updated version includes eukaryotes. BMC Bioinf. 2003;4:41.

28. Collins SRA, Wellner N, Martinez Bordonado I, Harper AL, Miller CN, Bancroft I, Waldron KW. Variation in the chemical composition of wheat straw: the role of tissue ratio and composition. Biotechnol Biofuels. 2014;7:121.

29. Linke D, Lehnert N, Nimtz M, Berger RG. An alcohol oxidase of *Phanerochaete chrysosporium* with a distinct glycerol oxidase activity. Enzyme Microb Technol. 2014;61–62:7–12.

30. Lee JW, Gwak KS, Park JY, Park MJ, Choi DH, Kwon M, Choi IG. Biological pretreatment of softwood *Pinus densiflora* by three white rot fungi. J Microbiol. 2007;45:485–91.

31. Zhou S, Raouche S, Grisel S, Sigoillot JC, Gimbert I. Efficient biomass pretreatment using the white-rot fungus *Polyporus brumalis*. Fungal Genom Biol. 2017;7:1–6.

32. Zhu N, Liu J, Yang J, Lin Y, Yang Y, Ji L, Li M, Yuan H. Comparative analysis of the secretomes of *Schizophyllum commune* and other wood-decay basidiomycetes during solid-state fermentation reveals its unique lignocellulose-degrading enzyme system. Biotechnol Biofuels. 2016;9:42.

33. Hori C, Gaskell J, Igarashi K, Kersten P, Mozuch M, Samejima M, Cullen D. Temporal alterations in the secretome of the selective ligninolytic Fungus *Ceriporiopsis subvermispora* during growth on aspen wood reveal this organism's strategy for degrading lignocellulose. Appl Environ Microbiol. 2014;80:2062–70.

34. Justo A, Miettinen O, Floudas D, Ortiz-Santana B, Sjökvist E, Lindner D, Nakasone K, Niemelä T, Larsson K-H, Ryvarden L, Hibbett DS. A revised family-level classification of the Polyporales (*Basidiomycota*). Fungal Biol. 2017;121:798–824.

35. Castanera R, Pérez G, Omarini A, Alfaro M, Pisabarro AG, Faraco V, Amore A, Ramírez L. Transcriptional and enzymatic profiling of *Pleurotus ostreatus* laccase genes in submerged and solid-state fermentation cultures. Appl Environ Microbiol. 2012;78:4037–45.

36. Alfaro M, Castanera R, Lavín JL, Grigoriev IV, Oguiza JA, Ramírez L, Pisabarro AG. Comparative and transcriptional analysis of the predicted secretome in the lignocellulose degrading basidiomycete fungus *Pleurotus ostreatus*. Environ Microbiol. 2016;18:4710–26.

37. Fernandez-Fueyo E, Ruiz-Dueñas FJ, Ferreira P, Floudas D, Hibbett DS, Canessa P, Larrondo LF, James TY, Seelenfreund D, Lobos S, Polanco R, Tello M, Honda Y, Watanabe T, Watanabe T, Ryu JS, Kubicek CP, Schmoll M, Gaskell J, Hammel KE, St. John FJ, Vanden Wymelenberg A, Sabat G, Splinter BonDurant S, Syed K, Yadav JS, Doddapaneni H, Subramanian V, Lavín JL, Oguiza JA, et al. Comparative genomics of *Ceriporiopsis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. Proc Natl Acad Sci USA. 2012;109:5458–63.

38. Syed K, Shale K, Pagadala NS, Tuszynski J. Systematic identification and evolutionary analysis of catalytically versatile cytochrome P450 monooxygenase families enriched in model basidiomycete fungi. PLoS ONE. 2014;9(1):e86683.

39. Qhanya L, Matowane G, Chen W, Sun Y, Letsimo E, Parvez M, Yu J, Mashele S, Syed K. Genome-wide annotation and comparative analysis of Cytochrome P450 Monooxygenases in basidiomycete biotrophic plant pathogens. PLoS ONE. 2015;10(11):e0142100.

40. Crešnar B, Petrič S. Cytochrome P450 enzymes in the fungal kingdom. Biochim Biophys Acta. 2011;1814:29–35.

41. Jawallapersand P, Mashele SS, Kovačič L, Stojan J, Komel R, Pakala SB, Kraševec N, Syed K. Cytochrome P450 monooxygenase CYP53 family in fungi: comparative structural and evolutionary analysis and its role as a common alternative anti-fungal drug target. PLoS ONE. 2014;9:e107209.

42. Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker D. Determination of structural carbohydrates and lignin in biomass: Laboratory analytical procedures, NREL. vol, 1617; 2008. https://permanent.access.gpo.gov/lps94089/42618.pdf. Issued 25 Apr 2008.

43. Miller GL. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem. 1959;31:426–8.

44. Zhou S, Grisel S, Herpoël-Gimbert I, Rosso M-N. A PCR-based method to quantify fungal growth during pretreatment of lignocellulosic biomass. J Microbiol Methods. 2015;115:67–70.

45. Miyazaki K, Maeda H, Sunagawa M, Tamai Y, Shiraishi S. Screening of heterozygous DNA markers in shiitake (*Lentinula edodes*) using

Miyauchi *et al. Biotechnol Biofuels (2018) 11:201*

Page 14 of 14

de-dikaryotization via preparation of protoplasts and isolation of four meiotic monokaryons from one basidium. J Wood Sci. 2000;46:395–400.

46. Alves AMCR, Record E, Lomascolo A, Scholtmeijer K, Asther M, Wessels JGH, Wösten HAB. Highly efficient production of laccase by the Basidiomycete *Pycnoporus cinnabarinus*. Appl Environ Microbiol. 2004;70:6379–84.

47. Gnerre S, MacCallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci USA. 2011;108:1513–8.

48. Martin J, Bruno VM, Fang Z, Meng X, Blow M, Zhang T, Sherlock G, Snyder M, Wang Z. Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. BMC Genomics. 2010;11:663.

49. Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I. MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res. 2014;42:D699–704.

50. Nielsen H. Predicting Secretory Proteins with SignalP. Methods Mol Biol. 2017;1611:59–73.

51. Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol Mol. 2000;300:1005–16.

52. Käll L, Krogh A, Sonnhammer ELL. A combined transmembrane topology and signal peptide prediction method. J Mol Biol Mol. 2004;338:1027–36.

53. Finn RD, Clements J, Eddy SR. HMMER Web Server: interactive sequence similarity searching. Nucleic Acids Res. 2011;39:W29–37.

54. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. Nat Protoc. 2008;4:1.

55. Moktali V, Park J, Fedorova-Abrams ND, Park B, Choi J, Lee YH, Kang S. Systematic and searchable classification of cytochrome P450 proteins encoded by fungal and oomycete genomes. BMC Genomics. 2012;13:525.

56. Levasseur A, Lomascolo A, Chabrol O, Ruiz-Dueñas FJ, Boukhris-Uzan E, Piumi F, Kues U, Ram A, Murat C, Haon M, Benoit I, Arfi Y, Chevret D, Drula E, Kwon M, Gouret P, Lesage-Meessen L, Lombard V, Mariette J, Noirot C, Park J, Patyshakuliyeva A, Sigoillot J, Wiebenga A, Wosten H, Martin F, Coutinho P, de Vries R, Martinez A, Klopp C, Pontarotti P, Henrissat B, Record E. The genome of the white-rot fungus *Pycnoporus cinnabarinus*: a basidiomycete model with a versatile arsenal for lignocellulosic biomass breakdown. BMC Genomics. 2014;15:486.

57. Ohm RA, Riley R, Salamov A, Min B, Choi I-G, Grigoriev IV. Genomics of wood-degrading fungi. Fungal Genet Biol. 2014;72:82–90.

58. Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS, Aerts A, Benoit I, Boyd A, Carlson A, Copeland A, Coutinho PM, de Vries RP, Ferreira P, Findley K, Foster B, Gaskell J, Glotzer D, Górecki P, Heitman J, Hesse C, Hori C, Igarashi K, Jurgens JA, Kallen N, Kersten P, et al. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science. 2012;336:1715–9.

59. Olson Å, Aerts A, Asiegbu F, Belbahri L, Bouzid O, Broberg A, Canbäck B, Coutinho PM, Cullen D, Dalman K, Deflorio G, van Diepen LTA, Dunand C, Duplessis S, Durling M, Gonthier P, Grimwood J, Fossdal CG, Hansson D, Henrissat B, Hietala A, Himmelstrand K, Hoffmeister D, Högberg N, James TY, Karlsson M, Kohler A, Kües U, Lee Y-H, Lin Y-C, et al. Insight into trade-off between wood decay and parasitism from the genome of a fungal forest pathogen. New Phytol. 2012;194:1001–13.

60. Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 2003;13:2178–89.

61. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25:3389–402.

62. Binder M, Justo A, Riley R, Salamov A, Lopez-Giraldez F, Sjökvist E, Copeland A, Foster B, Sun H, Larsson E, Larsson K-H, Townsend J, Grigoriev IV, Hibbett DS. Phylogenetic and phylogenomic overview of the Polyporales. Mycologia. 2013;105:1350–73.

63. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30:3059–66.

64. Talavera G, Castresana J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007;56:564–77.

65. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3.

66. Navarro D, Rosso M-N, Haon M, Olivé C, Bonnin E, Lesage-Meessen L, Chevret D, Coutinho PM, Henrissat B, Berrin J-G. Fast solubilization of recalcitrant cellulosic biomass by the basidiomycete fungus *Laetisaria arvalis* involves successive secretion of oxidative and hydrolytic enzymes. Biotechnol Biofuels. 2014;7:143.

67. Couturier M, Navarro D, Chevret D, Henrissat B, Piumi F, Ruiz-Dueñas FJ, Martinez AT, Grigoriev IV, Riley R, Lipzen A, Berrin J-G, Master ER, Rosso M-N. Enhanced degradation of softwood versus hardwood by the white-rot fungus *Pycnoporus coccineus*. Biotechnol Biofuels. 2015;8:216.

68. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

69. Wallace IM, O'Sullivan O, Higgins DG, Notredame C. M-Coffee: combining multiple sequence alignment methods with T-Coffee. Nucleic Acids Res. 2006;34(6):1692–9.

70. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 2010;59(3):307–21.

71. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol. 2001;18(5):691–9.

72. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016;44(W1):W242–5.