

RESEARCH ARTICLE

The integrative omics of white-rot fungus *Pycnoporus coccineus* reveals co-regulated CAZymes for orchestrated lignocellulose breakdown

Shingo Miyauchi¹, David Navarro¹, Sacha Grisel¹, Didier Chevret², Jean-Guy Berrin¹, Marie-Noelle Rosso^{1*}

1 Aix-Marseille Université, INRA, UMR 1163, Biodiversité et Biotechnologie Fongiques, BBF, Marseille, France, **2** PAPPSO, Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, Jouy-en-Josas, France

* marie-noelle.rosso@inra.fr



OPEN ACCESS

Citation: Miyauchi S, Navarro D, Grisel S, Chevret D, Berrin J-G, Rosso M-N (2017) The integrative omics of white-rot fungus *Pycnoporus coccineus* reveals co-regulated CAZymes for orchestrated lignocellulose breakdown. PLoS ONE 12(4): e0175528. <https://doi.org/10.1371/journal.pone.0175528>

Editor: Daniel Cullen, USDA Forest Service, UNITED STATES

Received: February 15, 2017

Accepted: March 27, 2017

Published: April 10, 2017

Copyright: © 2017 Miyauchi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All RNA-seq files are available from the GEO database (accession number GSE94878).

Funding: The research was funded by the Agence Nationale de la Recherche (ANR-14-CE06-0020). The fund was obtained by MNR. <http://www.agence-nationale-recherche.fr/en/anr-funded-project/> The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Innovative green technologies are of importance for converting plant wastes into renewable sources for materials, chemicals and energy. However, recycling agricultural and forestry wastes is a challenge. A solution may be found in the forest. Saprotrophic white-rot fungi are able to convert dead plants into consumable carbon sources. Specialized fungal enzymes can be utilized for breaking down hard plant biopolymers. Thus, understanding the enzymatic machineries of such fungi gives us hints for the efficient decomposition of plant materials. Using the saprotrophic white-rot fungus *Pycnoporus coccineus* as a fungal model, we examined the dynamics of transcriptomic and secretomic responses to different types of lignocellulosic substrates at two time points. Our integrative omics pipeline (SHIN+GO) enabled us to compress layers of biological information into simple heatmaps, allowing for visual inspection of the data. We identified co-regulated genes with corresponding co-secreted enzymes, and the biological roles were extrapolated with the enriched Carbohydrate-Active Enzyme (CAZymes) and functional annotations. We observed the fungal early responses for the degradation of lignocellulosic substrates including; 1) simultaneous expression of CAZY genes and secretion of the enzymes acting on diverse glycosidic bonds in cellulose, hemicelluloses and their side chains or lignin (i.e. hydrolases, esterases and oxido-reductases); 2) the key role of lytic polysaccharide monoxygenases (LPMO); 3) the early transcriptional regulation of lignin active peroxidases; 4) the induction of detoxification processes dealing with biomass-derived compounds; and 5) the frequent attachments of the carbohydrate binding module 1 (CBM1) to enzymes from the lignocellulose-responsive genes. Our omics combining methods and related biological findings may contribute to the knowledge of fungal systems biology and facilitate the optimization of fungal enzyme cocktails for various industrial applications.

Competing interests: The authors have declared that no competing interests exist.

Introduction

The bioconversion of agricultural, forestry and industrial green wastes into renewable sources of energy with bio-based molecules is highly desirable. Innovations in ecologically-friendly technologies will lead to improvements in green chemistry and biorefinery, and contribute to the circular economy. Lignocellulose, the major component of plant biomass, is highly recalcitrant to enzymatic degradation because the cellulose forms crystalline structures and the lignin sets dense polyaromatic polymers. Great efforts have been made to convert lignocellulosic saccharides into biofuels, biogas, and bio-based chemicals [1,2]. Lignocellulose has enormous potential in the production of phenolics and aromatics-based products [3]. However, the problems are how to break down lignin unamenable to biochemical transformation and to degrade cellulose efficiently.

A solution to such technical challenges may be found in the nature. Some forest-dwelling mushrooms are evolved to break down plant materials and release (poly)saccharides and (poly)aromatics from lignocellulose. Saprotrophic white-rot fungi are highly specialized to decompose dead plants and convert them into consumable carbon sources [4]. Phylogenetic reconstructions showed that white-rot fungi evolutionarily gained specific groups of enzymes to break down recalcitrant lignocelluloses [5,6]. The gene repertoires of saprotrophic white-rot fungi are enriched in enzymes for oxidative degradation of lignocellulose, called auxiliary activity enzymes (AA) classified in the CAZy database [7]. There are different types of Carbohydrate Active Enzymes (CAZymes) in the CAZy database including glycoside hydrolases (GH), carbohydrate esterases (CE), polysaccharide lyases (PL), glycosyltransferases (GT), and carbohydrate binding modules (CBM) [8]. There are some known synergistic activities among the different families of CAZymes. For example, synergies between endoglucanases, cellobiohydrolases and beta-glucosidases result in an efficient cellulolytic enzyme system, which collective activity is higher than the sum of the activities of the individual enzymes [9]. Other synergies exist between oxido-reductases for the oxidative cleavage of lignocellulose. GMC oxidoreductases including cellobiose dehydrogenases (CDH; CAZy families AA3_1 and AA8) and AA3-2 donate electrons to lytic polysaccharide monooxygenases (LPMO, family AA9) for the oxidative cleavage of cellulose polymers [10–12]. Copper radical oxidases (CAZy family AA5) give hydrogen peroxide to manganese peroxidases (MnP; family AA2) for the oxidative cleavage of lignin [13]. There may be as-yet-to-be-known synergies between the fungal enzymes. Thus, it is necessary to explore transcriptional and proteomic regulations of CAZymes.

Our general understanding of the regulation of fungal enzymes in lignocellulose deconstruction is limited. On one hand, ascomycetes fungi have been studied mainly for the regulation of genes coding for cellulases and hemicellulases [14]. On the other hand, studies of wood-decay basidiomycetes fungi have been focused on the role of oxidative enzymes and their interplay during lignocellulose breakdown [15–17]. Meanwhile, white-rot fungi from the genus *Pycnoporus* are known to degrade cellulose, hemicelluloses and lignin, and the versatility of the genus for biotechnological applications is recognized [18,19]. The *Pycnoporus* strain *P. coccineus* CIRM-BRFM310 was previously studied regarding the roles of cellulases, hemicellulases, esterases and lignin active peroxidases during cultivation on pine and aspen [16]. However, the enzymatic mechanisms for the alleviation of plant biomass recalcitrance by this fungus remained to be elucidated.

Fungal responses to the source of carbon are dynamic. Fungi utilize specific groups of genes to execute lignocellulose breakdown, metabolization, and detoxification of molecules during the degradation process. Such orchestrated molecular machineries are required to convert plant biomass into a carbon source. Time-course investigations of fungal gene co-regulations

and corresponding protein co-secretions induced by plant biomass are well suited to examine the systematic fungal depolymerization of lignocellulose. However, genome-wide transcriptomic and secretomic activities are complex. Capturing just a single time point of fungal transcriptomic activity involves over ten thousands genes showing various transcription levels. The number of observations increases exponentially when we add the number of biological replicates, different growth conditions, and time points. The addition of secretomic information gives an extra layer of complexity.

To extract biologically meaningful patterns from such high-dimensional omics data, we have developed the multi-omics profiling platform, Self-organizing map Harboring Informative Nodes with Gene Ontology (SHIN+GO). Genome-wide omics models constructed with the platform are designed to pinpoint biological activities of interest that would otherwise be buried in the high-dimensional data. A prototype of the platform was previously developed as Applied Biomass Conversion Design for Efficient Fungal Green Technology (ABCDEFGT), which had a limited capacity of processing large-scale omics data [20]. As one of the key components of this platform, Self-organizing map (SOM) is an algorithm constructing a neural network with given input data in an unsupervised manner [21]. SOM reduces the number of features in high-dimensional data by grouping similar items and forming clusters. It has a unique property of making two-dimensional maps suitable for large-scale data visualization. The method has been used to generate neural networks of genome-wide genes and identify condition-specific responses in transcriptomes of newly sequenced three fungal species with limited gene annotations [20].

In this study, we used the improved platform (SHIN+GO) to generate dynamic genome-wide integrative omics models with two time points. We profiled the early transcriptomic and secretomic responses of *P. coccineus* CIRM-BRFM310 at day 3 and 7. The fungal strain was grown with three types of recalcitrant lignocellulosic substrates (i.e. cereal straw, softwood, and hardwood) to capture an overview of the fungal adaptive responses to plant biomasses with different compositions of cellulose, hemicelluloses, and lignin. The selected plant substrates are considered as sustainable resources for green technologies [1,22]. We built integrative omics models in an unsupervised machine learning manner and used them as a guide to explore biologically interesting parts of the large-scale omics data (i.e. omics hotspots). We interpreted the findings in terms of fungal transcriptomic, secretomic, and metabolomic responses to the lignocellulosic substrates.

Results and discussion

Fungal omics models for the lignocellulosic substrates

We investigated co-regulated genes with corresponding co-secreted proteins of *P. coccineus* CIRM-BRFM 310 growing on four different substrates at day 3 and day 7 in order to see the early transcriptomic and secretomic adaptations. Maltose, which is an easily assimilated source of carbon, was used as a control, and ground wheat straw, pine wood, and aspen wood were used as models for gramineae, softwood, and hardwood biomass, respectively.

Under the conditions imposed in this study, optimal fungal growth was observed at day 3 on lignocellulosic substrates whereas a lag phase of 72 hours was observed on maltose (S1 Fig). A time-course experiment measuring the quantity of maltose consumed by the strain showed the maltose initially added to the basal medium was completely depleted in the culture medium at the point of day 3 (Miyachi et al. in prep). Therefore, the fungal responses for the lignocellulosic substrates at day 3 and 7 were not affected with the residual maltose. The normalized read counts obtained from the RNA-seq of three biological replicates grown under each of the four conditions at each time point were highly correlated (> 0.9 correlation co-

efficient; S2 Fig), and the distributions of the normalized read counts from the replicates were also almost identical (S3 Fig). A total of 26 housekeeping genes showed a stable transcript level between the two time points, suggesting that the physiological state of *P. coccineus* was comparable between all cultivation conditions (S4 Fig). Thus, we concluded that the obtained biological data were consistent and comparable to each other.

We generated fungal omics models with transcriptional changes at two time points. The models constructed in this study were more vast and congruent than those built with a single time point in the previous study [20]. A total of 274,320 points of observations combined from two time points (i.e. 4 conditions x 3 replicates x 2 time points x 11,430 genes with corresponding secretomes) were taken into account to construct omics models of *P. coccineus*. We used a multi-omics profiling platform, Self-organizing map Harboring Informative Nodes with Gene Ontology (SHIN+GO; Fig 1).

The first part of the SHIN+GO platform, Self-organizing map Harboring Informative Nodes (SHIN; the improved version of ABCDEFGT; Miyauchi et al., 2016), was used to merge genome-wide omics data. The SHIN pipeline constructed a single master self-organizing map (SOM) with the normalized read counts of 11,430 genes from the triplicates grown under the four cultivation conditions at day 3 and 7. The genes were sorted into 456 nodes according to their similar transcriptomic patterns. Next, the count of secreted proteins was overlaid onto the master SOM. As a result of this integration of data, SHIN provides nodes made of clustered co-regulated genes (transcriptomes) with corresponding co-secreted proteins (secretomes). The second part of the SHIN+GO platform, Gene Ontology (GO), was newly developed for this study; 1) to measure the frequency of gene functional annotations present in the nodes; and 2) to biologically interpret the outputs of the genome-wide omics models generated. Biological terms with statistically enriched occurrence in a node were used as an indicator of biological functions for the grouped genes and proteins.

All transcriptomic and secretomic topographies are comparable as the positions of the nodes are fixed in the maps (Fig 2A), enabling the instant visual inspection of omics hotspots where biologically intensive activities can be observed (Fig 2 and S5 Fig). The node-wise mean of the normalized transcript read counts reflects the transcription level in response to each condition. The secretomic topography represents the count of secreted proteins detected from the culture medium. Some nodes contained high counts of secreted proteins, indicating hotspots of co-secreted proteins.

Globally, there is a consistency that the transcriptomic topographies showing high transcription contain higher counts of secreted proteins in the corresponding secretomic topographies. It suggests that the high transcription of genes generally leads to the secretion of their proteins (Fig 2). Another observation is that the transcriptomic and secretomic patterns resulting from cultivation on different substrates were slightly different from each other. Particularly, the patterns obtained from maltose differed from those from the lignocellulosic substrates (aspen, pine, wheat straw). Notably, there are overlaps between some of the nodes showing high transcription and intensive secretomic responses to the lignocellulosic substrates (i.e. nodes 1 to 17). We considered such nodes as 'omics hotspots' involved in biologically interesting events at the transcriptomic and secretomic levels. We found that several of these nodes were enriched in CAZyme coding genes (discussed in later sections).

Nodes showing up-regulation and high transcription for the lignocellulosic substrates were determined from the omics models. Nodes that met either of two following criteria were selected for the further analysis; 1) $> 11.7 \log_2$ transcript read counts, which constitute above 90th percentile of the transcription level of the entire gene group; 2) $> 2 \log_2$ transcriptional changes compared to maltose at each time point (Fig 3).

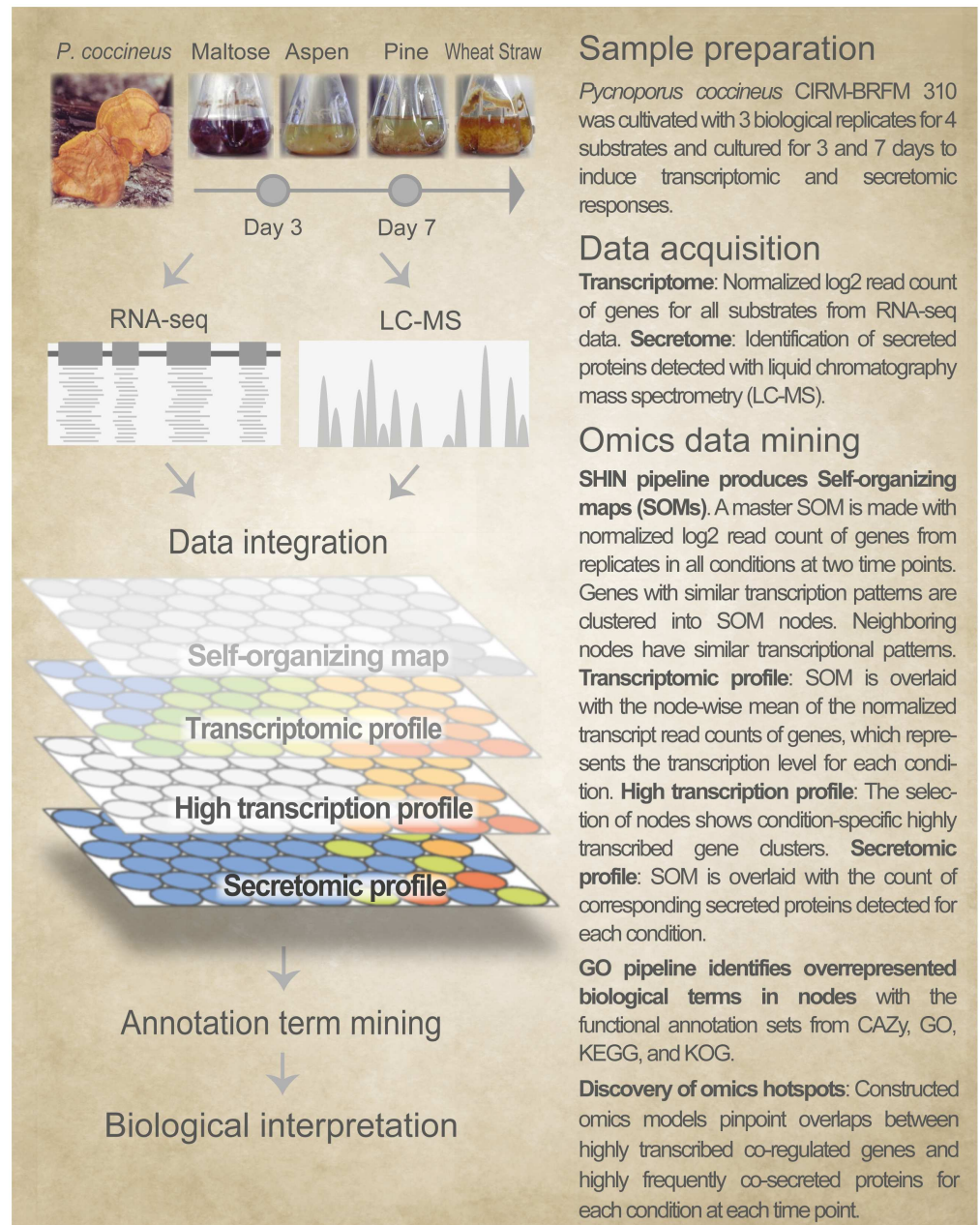


Fig 1. Overview of the genome-wide integrative omics profiling of *Pycnoporus coccineus* CIRM-BRFM 310 at two time points. The SHIN+GO platform; 1) integrated the fungal transcriptome from RNA-seq data and the secretome from liquid chromatography mass spectrometry; and 2) assisted the biological interpretation of the outputs of the omics models with functional gene annotations.

<https://doi.org/10.1371/journal.pone.0175528.g001>

In the transcriptome at day 3, a majority of genes in the nodes showed a similar response to the three lignocellulosic substrates (i.e. 32 nodes showed high transcription or high up-regulation on each of aspen, pine, and wheat straw; Fig 3). There were eleven nodes showing substrate-specific responses to either only pine (ten nodes) or aspen (one node). However, there were no nodes specific to wheat straw. At day 7, a higher number of genes in the nodes were highly transcribed or highly up-regulated on a single substrate compared to day 3

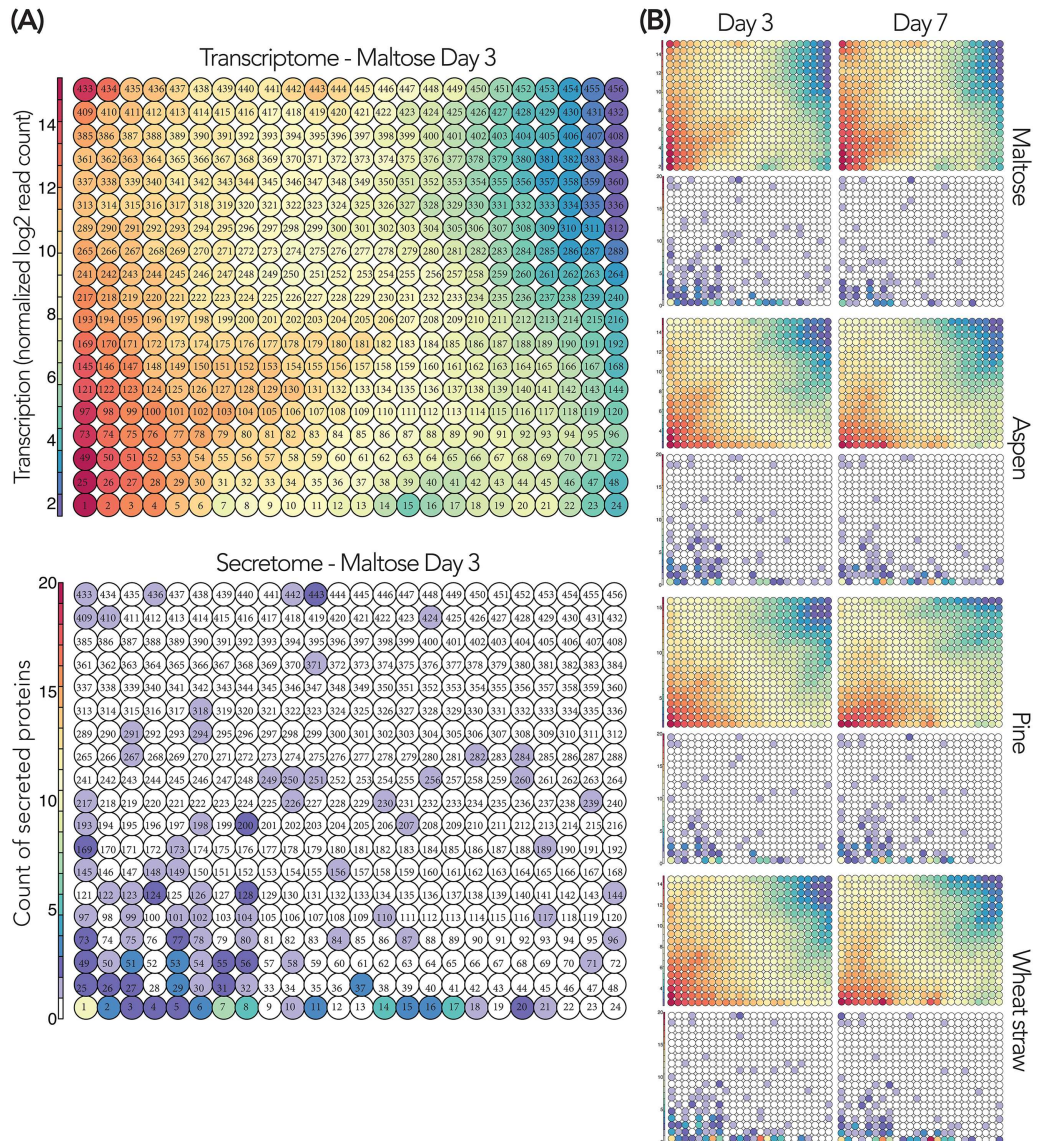


Fig 2. Genome-wide integrative omics models of *P. coccineus* in response to the substrates at day 3 and 7. Transcriptomic topography: Mean transcription levels per node for each cultivation condition. Secretomic topography: The total count of secreted proteins per node indicates secretion hotspots. (A): Magnified version of the topographies. The node identification is labeled (i.e. 1 to 456). (B): Transcriptomic and secretomic topographies from the four substrates. An animated version of transcriptomic topographies is available (S1 File).

<https://doi.org/10.1371/journal.pone.0175528.g002>

(i.e. nineteen, four, and five nodes for pine, aspen, and wheat straw respectively). The observation indicated that common molecular functions were triggered by the exposure to the lignocellulosic substrates at the early stage (day 3). Then, the fungus adjusted more specific responses to the specific substrate over time at the later stage (day 7). Remarkably, the number of pine-specific nodes was increased by nine from day 3 (nodes 32, *33, *56, *57, 71, 78, 79, *126, 150, 175) to day 7 (*2, *3, *4, 10, *33, 34, 42, *56, *57, *58, 62, 79, *80, 81, *102, 103, 104, 105, *125). Four and ten of these nodes at day 3 and day 7 respectively contained statistically enriched gene annotations (i.e. the nodes with asterisks above and discussed in later sections).

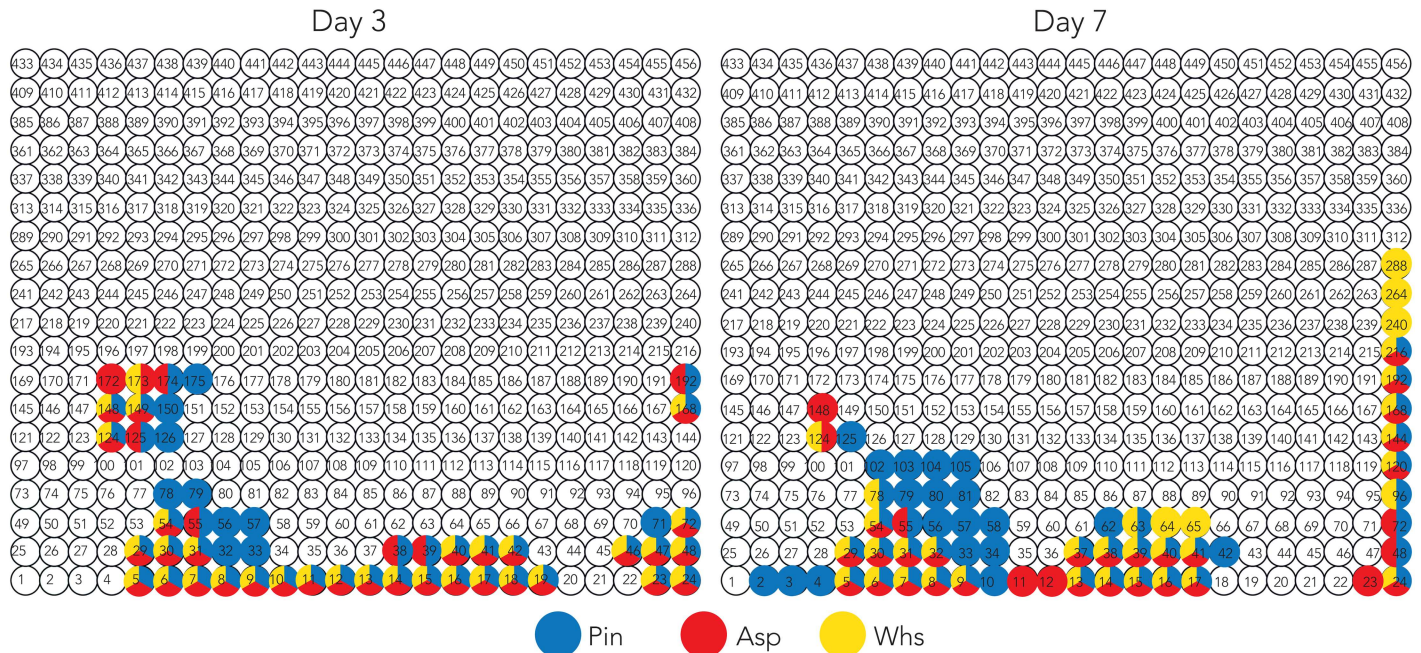


Fig 3. Transcriptomic changes of the lignocellulosic substrate-specific nodes from day 3 to day 7. The specific transcription patterns for the lignocellulosic substrates were extracted from Fig 2. The highlighted nodes met either of two criteria; 1) > 11.7 log2 read counts; or 2) > 2 log2 fold changes on aspen (Asp), pine (Pin), and wheat straw (Whs) in comparison with maltose at each time point.

<https://doi.org/10.1371/journal.pone.0175528.g003>

Transcriptomic responses to the lignocellulosic substrates

We attempted to interpret the fungal adaptations to the lignocellulosic substrates at the transcriptomic level using the enriched gene annotations present in the nodes. The frequency of gene annotations in the 456 nodes of the omics models revealed that five nodes (1, 7, 8, 14, 15) contained a significant number of genes coding for CAZymes (S1 Table and S4 File). Among them, four nodes (7, 8, 14, 15) showed high transcription during the growth on the lignocellulosic substrates compared to maltose (Fig 3 and Table 1). These nodes contained the highest number of frequently secreted proteins (secretomic topographies in Fig 2), suggesting that the corresponding CAZymes were contributing to the decomposition of the lignocellulosic substrates. Nodes in proximity share similar transcriptomic patterns due to the nature of SOM used for the construction of the omics models [20]. Hence, the regulations of genes coding for

Table 1. Nodes with highly enriched CAZyme coding genes up-regulated on the lignocellulosic substrates.

Node	CAZymes
7	AA2, AA3_3, AA8-AA3_1, AA9, AA9-CBM1(2), CBM1-CE16, GH1, CBM1-GH10, CBM1-GH5_7, CBM1-GH6, GH7, GH51, GH16, GH30, GH35, GH74-CBM1, GH79, PL8_4
8	AA3_2(2), AA5_1, CE4, GH3, GH28, GH55, GH76, GH92
14	CBM1-CE1(2), CE8, CBM1-CE15, CE16, GH3, CBM1-GH10, CBM1-GH5_5, GH12(2), GH131-CBM1, GH28, GH32, GH43(2), GH45, GH53, GH78, GH115
15	AA2, AA9(5), CE16, GH28(2), CBM1-GH5_5, CBM1, GH7(2)

Different types of CAZymes present in node 7, 8, 14, and 15. AA: Auxiliary activity enzymes, GH: Glycoside hydrolases, CE: Carbohydrate esterases, CBM1: Carbohydrate binding module 1 (CBM1) [8]. The number of gene copies is indicated in brackets.

<https://doi.org/10.1371/journal.pone.0175528.t001>

CAZymes in nodes 7 and 8 on the one hand, 14 and 15 on the other hand can be considered as co-regulated groups.

Each of four nodes contained a set of CAZymes active on the four types of lignocellulose polymers (i.e. cellulose, hemicelluloses, pectin and lignin), suggesting that the fungus simultaneously recruited a variety of enzymes and executed the orchestrated depolymerization (Table 1). There were eight LPMO (AA9) coding genes in nodes 7 and 15 that were up-regulated in response to lignocellulosic substrates at the two time points (Table 1). Other highly up-regulated genes were two predicted glucose-methanol-choline (GMC) oxidoreductases (family AA3_2; protein ID: 1463000, 1465734) and one glyoxal oxidase (AA5_1; 1480943) in node 8, which could give hydrogen peroxide to AA2 in nodes 7 (1468611) and 15 (1430659) for their enzymatic activities (Table 1). There were three LPMOs in node 7 (AA9; 1374028, 1382161, 1428145) along with the presence of one cellobiose dehydrogenase (AA8-AA3_1; 1401955) and two GMC oxidoreductases (AA3_2; 1463000, 1465734) in node 7 and 8, suggesting co-regulation for the oxidative cleavage of cellulose polymers [12]. Also, there were genes encoding hemicellulose-active enzymes such as glycosyl hydrolases (GH) and carbohydrate esterases (CE). Some of the GHs are for the degradation of xylan. GH10 acts on the main chain of xylan, while GH43 and 51 are arabinofuranosidases for the arabinose substitution of xylan. GH115 is a glucuronidase for the glucuronoyl substitution of woody xylan. Meanwhile, some of the CEs are targeting acetyl xylans (CE4, CE16) and 4-O-Methyl glucuronoyl side chains (CE15). In addition, nodes 7 and 14 were highly enriched with CBM1 associated CAZymes (Table 1 and S2 File; adjusted $p < 0.01$). CBM1 has an affinity to crystalline cellulose and directs CBM1 associated enzymes to potentiate cellulolytic activities on insoluble substrates [23,24].

We visualized the transcription intensity of CAZyme-coding genes from nodes 7, 8, 14, and 15 in order to capture the specific regulation patterns at two time points (Fig 4). Globally, the transcript level of the genes for plant cell wall-active enzymes went up from day 3 to day 7.

At day 3, the most highly up-regulated CAZymes were the lignin-active peroxidases in nodes 7 and 15 (AA2-Mnp/Lip; protein ID: 1468611, 1431101), which were less-strongly up-regulated at day 7, suggesting that their regulations were finely controlled. The tight regulation could possibly expose cellulose and hemicellulose polymers from lignin while minimizing oxidative damage to the fungus. This phenomenon is in accordance with the dynamics of MnP or LiP gene transcription in *P. carnosus* and the detection of the enzymes in the secretomes of *C. subvermisporea* [25–27]. At day 7, almost all of highly up-regulated CAZymes (transcription induction factor > 200) were from the cultivation on the wheat straw substrate (Fig 4). There were 18 CAZymes and the majority had a CBM1 attached including four AA9, AA8-AA3_1, two GH10, GH5_5, GH5_7, GH6, GH7, GH12, GH45, GH131, CE1, CE15, and CBM1. The cellobiose dehydrogenase-coding gene (CDH; AA8-AA3_1; protein ID:1401955) co-regulated with three LPMOs (AA9) grouped in node 7 (Table 1) showed strong up-regulation on the three lignocellulosic substrates (e.g. transcription induction factor 351 on the wheat straw substrate at day 7; Fig 4), suggesting that the CDH might promote the activity of these LPMOs via electron transfer [28]. LPMO coding genes were seemingly co-regulated with hydrolytic enzymes active on beta-1,4-glycans such as GH3, GH5, GH6, GH7, GH45, and GH131. Since LPMOs display oxidative cleavage of crystalline cellulose [28], the combinations of these CAZymes indicated the simultaneous induction of hydrolytic and oxidative activities (Fig 4). Some LPMOs were highly up-regulated as the transcription induction factors were 115 at day 3 and 412 at day 7 on the wheat straw substrate (AA9; 1374028, 1417214; Fig 4 and S2 Table). LPMOs in different fungal species had similar transcriptional regulations during lignocellulosic degradation. The white-rot fungi *Phlebia radiata* and *Phlebiopsis gigantea* showed up-regulation of the genes during the growth on spruce and loblolly pine wood respectively [29,30].

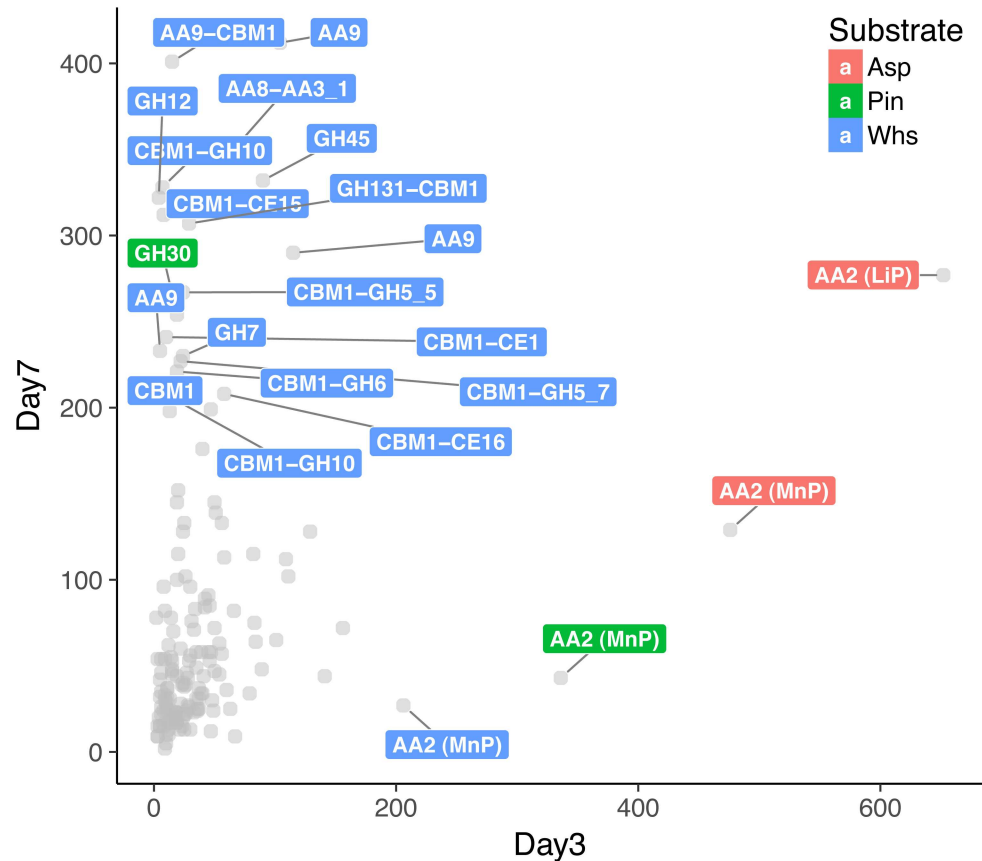


Fig 4. Transcription intensity of CAZyme coding genes in nodes 7, 8, 14, and 15 at day 3 and day 7. The X and Y axes represent the values of transcription induction factor (TIF). TIF values > 200 are labeled at each time point. TIF were estimated by squaring log2 fold change values of the transcript read counts on lignocellulosic substrates compared to the control condition with maltose. Detailed information is provided (S2 Table).

<https://doi.org/10.1371/journal.pone.0175528.g004>

Also, secreted LPMOs were found in *Phanerochaete chrysosporium* and *Ceriporiopsis subvermispota* [26,27,31,32].

Incidentally, expansin (EXPN) showed unique transcriptional patterns for the lignocellulosic substrates in comparison to maltose (S4 File). While four expansin coding genes showed the moderate up-regulation only on the pine substrates at day 3 and 7 (EXPN; 1451679, 1467568, 1470236, 316489), other ten genes were mostly down-regulated on each of the lignocellulosic substrates at two time points (EXPN; 1440495, 1439344, 1463757, 1477520, 1424314, 1422224, 1463980, 1386069, 1373750, 148891). There might have been some transcriptional regulations induced by the pine substrate. This peculiar phenomenon remains to be elucidated.

Secretomic responses to the lignocellulosic substrates

We examined secretomic responses to the lignocellulosic substrates by comparing secreted proteins detected at day 3 and day 7. The CAZyme-enriched nodes (7, 8, 14, 15) were part of six nodes that showed the highest number of secreted proteins (> 9 protein count) among the total of 49 nodes containing secreted proteins (Table 2). A greater number of proteins was detected at day 7 compared to day 3, suggesting the accumulation of secreted enzymes in the culture medium over time.

Table 2. Secretomic responses to the lignocellulosic substrates at day 3 and day 7.

Node	Day 3 only	Day 3 and Day 7	Day 7 only
7		AA8-AA3_1, AA9-CBM1, AA9, CBM1-CE16, CBM1-GH10, CBM1-GH5_7, GH79, Peptidase S53	AA9-CBM1
8	Ceramidases	GH28	AA3_2, AA5_1
14		Aldose_1-epimerase, CBM1-CE1 (2), CBM1-CE15, CBM1-GH10, CBM1-GH5_5, CE16, CE8, GH12 (2), GH131-CBM1, GH43, GH45	GH53, GH78, GH115
15		AA9 (4), Carboxylesterase_and_related_proteins (#1468641), CBM1-GH5_5, GH7 (2), SSP (#1437297), Unknown (# 1439153)	AA2 (LiP), CE16, Unknown (#1446065)

Secreted proteins were detected under the cultivations on the lignocellulosic substrates. Unknown: Proteins with unknown function. SSP: Small secreted proteins which size is smaller than 300 aa. #: Protein ID. The number of gene copies is indicated in brackets. Details of the selected six nodes are provided (S3 File).

<https://doi.org/10.1371/journal.pone.0175528.t002>

Notably, the oxidative enzymes (AA2-LiP; protein ID: 1431101) for lignin degradation and the partner enzymes glyoxal oxidase (AA5_1, 1480943) and GMC oxido-reductases (AA3_2, 1463000, 1465734) for hydrogen peroxide production were detected only at day 7, suggesting that the secretion of such enzymes at day 3 could have been limited by post-transcriptional regulations. We evaluated the correlation between the transcriptome and secretome in response to each substrate at each time point. The two omics patterns were moderately correlated (the correlation co-efficient ranging from 0.30 to 0.44; $p < 0.001$; S3 Table). The moderate correlation could be due to the narrow window of our observation points (at day 3 and 7). Also, difficulties in the detection of proteins in the liquid fraction of culture media could have arisen from; 1) delays in mRNA translation; 2) fungal proteins bound to the substrates; or 3) a lower amount of proteins secreted than the detection limit.

Metabolic adaptations to the lignocellulosic substrates

We investigated fungal metabolic adaptations during the growth on the lignocellulosic substrates at day 3 and day 7 using the annotation databases Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and EuKaryotic Orthologous Groups (KOG) [33–35]. The functional annotations found in all 456 nodes are provided (S4 File).

A total of 150 out of 456 nodes contained statistically enriched functional annotations (adjusted $p < 0.01$; S1 Table). There were twenty nodes showing either up-regulation/high transcription on two or three lignocellulosic substrates, and also up-regulation/high transcription specific to the pine substrate (S4 and S5 Tables).

The terms relating to putative hydrolases active on carbohydrates were enriched in nodes 7, 14, and 15 (S4 Table). Putative dehydrogenases, oxidoreductases or peroxidases were also enriched in nodes 7, 8, 19, 22, and 23. These enriched terms confirmed that the fungus executed hydrolytic and oxidative processes for the breakdown of the lignocellulosic substrates, which was discussed earlier in the transcriptomic and secretomic responses (Tables 1 and 2). The up-regulation of the genes was accompanied by predicted transporters in nodes 23 and 39 potentially involved in carbohydrate assimilation (S4 Table). The terms for predicted cytochrome P450s were enriched in nodes 9, 40, and 63, suggesting that the fungus adjusted to exposure to toxic compounds released from the lignocellulosic substrates during the degradation process (S4 Table). The findings suggest that the white-rot fungus *P. coccineus* might have used a common set of P450 genes for the fungal adaptation to diverse lignocellulosic substrates. The detoxification response of *P. coccineus* to lignocellulosic substrates somewhat differs from that of the brown-rot fungus *Postia placenta*, which showed a higher transcription of P450 genes and specific P450 protein secretion during the growth on pine (i.e. *Pinus strobes*)

in comparison to aspen (i.e. *Populus gigantea*) [32]. The different culture conditions and compositions in wood extractives might have contributed to the variations of the fungal responses.

Nodes 2 and 3 were enriched with the terms for putative chitinases and putative 1,3-beta-glucan synthases respectively (S5 Table). Comparing the enriched terms with the CAZy annotations confirmed the predicted functions for one chitinase (protein ID: 1433997, CAZyme: GH18), with a chitin synthase (1467191, GT2) in node 2 and two b-1,3-glucan synthases (1436152 and 1479981, GT48) in node 3. These terms indicate that there might have been remodeling of the fungal cell wall during the growth on the pine substrate. In addition, four nodes (4, 33, 56, 58) had enriched terms for putatively involved in signal transduction such as protein kinases, which might have contributed to the signaling cascade driving the fungal response to the pine substrate (S5 Table).

Conclusions

We examined transcriptomic and secretomic models of *P. coccineus* CIRM-BRFM 310 at two time points in response to three different types of lignocellulosic substrates representing gramineae, softwood and hardwood. The transcriptomic approach unveiled the dynamic changes of genome-wide transcription levels while the secretomic approach identified sets of enzymes involved in the synergistic depolymerization of the lignocellulosic substrates. Co-regulated genes and corresponding co-secreted proteins were determined and their biological roles were extrapolated using the enriched CAZymes and other functional annotations. Although our observations were limited to two time points, the fungus seemed to execute an almost simultaneous recruitment of various enzymes targeting cellulose, hemicelluloses, lignin and pectin, rather than a step-by-step degradation of the plant polymers. A composite of hydrolases, esterases and oxido-reductases active on plant cell walls might have been simultaneously secreted to achieve lignocellulose decomposition. Genes coding for enzymes associated with a carbohydrate binding module (CBM1) were strongly up-regulated and their proteins were secreted in response to the lignocellulosic substrates. Secreting abundant sets of CBM1-associated CAZymes could improve the efficiency of the degradation of plant cell wall polymers as CBM1 enhances the binding of CAZymes to cellulose. The plant cell wall-degrading CAZymes exhibited a global increase at the transcription level from day 3 to day 7, except for lignin active peroxidases (AA2), which displayed a unique early regulation at day 3. A large number of genes coding for LPMOs (AA9) appeared in the omics hotspots of our models, indicating a close involvement of this enzyme family for adaptive responses of *P. coccineus* to the complex substrates. Our integrative omics strategies enabled us to combine layers of biological information for better understandings of fungal machineries for decomposition of plant cell walls. The SHIN+GO platform is versatile and can be applied for comparative transcriptomics of different strains or species. Our findings may shed some light on the design of customized enzyme cocktails for the conversion of diverse and highly recalcitrant biomass resources.

Materials and methods

Fungal strains and cultures

P. coccineus CIRM-BRFM 310 was obtained from the International Center of Microbial Resources (CIRM-CF; https://www6.inra.fr/cirm_eng). 5mm discs of mycelium grown on malt agar were used to inoculate Roux flasks containing 200 ml of medium (five discs per flask) and incubated at 30°C. After 15 days, the mycelium was ground (ultraturax 10000 rpm, 60 s) in 100 ml of purified water (MilliQ, Millipore). Five milliliters of this suspension were used for the inoculation of each 250-ml Erlenmeyer flask containing 100 ml of basal medium: diammonium tartrate (1.84 g.L⁻¹), Yeast Extract (2.5 g.L⁻¹), Maltose (2.5 g.L⁻¹), KH₂PO₄

(0.2 g.L⁻¹), CaCl₂·2H₂O (1.32·10⁻² g.L⁻¹), MgSO₄·7H₂O (0.5 g.L⁻¹), FeSO₄·7H₂O (0.07 g.L⁻¹), ZnSO₄·7H₂O (7.77·10⁻³ g.L⁻¹), MnSO₄·H₂O (3.63·10⁻³ g.L⁻¹), CuSO₄·5H₂O (7.2·10⁻⁴ g.L⁻¹) and thiamine (250·10⁻³ g.L⁻¹). The test conditions were created by adding one of the following additional components: maltose (20 g.L⁻¹ final), ground and sifted wheat straw < 2 mm (15 g.L⁻¹), ground and sifted *Pinus halepensis* wood fragments < 2 mm (15 g.L⁻¹) or 1 mm Wiley-milled *Populus grandidentata* (15 g.L⁻¹) kindly provided by Dan Cullen (Forest Product Laboratory, USDA, Madison, WI). Incubation was carried out at 30°C in a rotary shaker (Infors AG) at 120 rpm. Assays were performed in triplicate.

RNA preparation and RNA-sequencing

Total RNA was extracted from 3 day and 7 day-old cultures as described [16]. In short, mycelia were harvested by filtering through Miracloth (Calbiochem), squeeze dried, snap frozen in liquid nitrogen and stored at -80°C until use. Frozen fungal pellets were ground using the SamplePrep 6770 FreezerMill (Spex). RNA was extracted from one hundred milligram of ground mycelium in 1 mL TRIZOL (Ambion). RNAs were precipitated with isopropanol (Sigma-Aldrich), treated with DNase I (QIAGEN) and resuspended in 25 µL RNase Free water. RNA purity and integrity were analyzed on NanoDrop Spectrophotometer and Agilent 2100 BioAnalyzer. For RNASeq, cDNA libraries were prepared using the TrueSeq RNA-Seq Sample Prep Kit V2 (Illumina Inc., San Diego, CA), and submitted to sequencing using Illumina 2x75 bp technology (Beckman Coulter Genomics). The RNA-seq data are available on NCBI's Gene Expression Omnibus and accession number GSE94878 [36].

Protein extraction and detection

Secreted proteins were collected from the same cultures, dia-filtered and identified by ESI-MS/MS as described in Couturier et al. (2015). Briefly, short SDS-PAGE runs were performed, allowing 10 µg of proteins to migrate on a 0.5 cm length. Each one-dimensional electrophoresis lane was cut into two slices of gel and protein identification was performed using PAPPISO "Plate-forme d'Analyse Protéomique de Paris Sud-Ouest" platform facilities. In-gel digestion was carried out according to a standard trypsinolysis protocol. Online analysis of peptides was performed with a Q-exactive mass spectrometer (Thermo Fisher Scientific), using a nano-electrospray ion source. Protein identification was performed by querying MS/MS data against the genome *P. coccineus* BRFM 310 v1.0, together with an in-house contaminant database, using the X!Tandem software (X!Tandem Cyclone, Jouy en Josas, France). All peptides matched with an E-value lower than 0.05 were parsed with X!Tandem pipeline software. Proteins identified with at least two unique peptides and a log (E-value) lower than -2.6 were validated. The secretomic information is provided (S5 File).

Genome and gene annotations

Functional annotations of *P. coccineus* with GO, KEGG, KOG, SignalP were obtained from MycoCosm, JGI (<http://genome.jgi.doe.gov/Pycco1/Pycco1.home.html>). CAZy annotations were obtained from the CAZy group led by Bernard Henrissat (AFMB, Aix Marseille University).

Data preparation and manipulations

QCed reads from each library were aligned to the genome *P. coccineus* BRFM 310 v1.0 using TopHat [37] with only unique mapping allowed. R was used for data manipulations using our customized scripts [38]. Raw read counts obtained from HTSeq were used for the calculation

using DESeq2 [39]. A total of 11,430 genes having more than five reads were selected for the analysis. The normalized log₂ transformed read count of genes showed similar mean and distribution patterns. The log₂ fold difference of the gene expression induced by cultivation on aspen, wheat straw, pine, compared to maltose, was calculated at each time point. The statistically significant genes were selected based on adjusted p value < 0.05 (FDR and Bonferroni correction). Normalized read counts of the genes were produced with the function, counts in DESeq2, which were subsequently log₂ transformed [39].

Transcript abundance of housekeeping genes

Housekeeping genes coding for chitin synthases and NADH dehydrogenases were chosen for comparison of transcription levels between day 3 and 7. Normalized log₂ read counts were compared for a total of 26 genes between the combined triplicates from each cultivation condition.

Construction of integrated omics models

Two-time-point transcriptome and secretome models were constructed using the SHIN pipeline (i.e. the ABCDEFGT workflow with improved capacity) [20]. A SOM was trained with the normalised read count of all conditions using Rsomoclu [40] on the computer clusters at Mesocentre (<https://equipex-mesocentre.univ-amu.fr/en/welcome-mesocentre-aix-marseille-universite/>). The matrix of 24 x 19 (456) was used with four neighbouring nodes. The resolution of 25 genes per node was used, which was empirically determined. The epoch of 1000 times more than the map size was applied, i.e. 456,000 = 456 (map size) multiplied by 1000. The initial radius for SOM calculation was determined using neighbour distance function in R Kohonen package [41]. The following outputs were visualized; 1) transcriptomic models of all biological replicates in four cultivation conditions at two time points; and 2) combined transcriptomic and secretomic models in four cultivation conditions at two time points.

Determination of time-course secretomic profiles of *P. coccineus*

Nodes containing high numbers of secreted proteins were determined from the 3- and 7-day cultivation on the aspen, pine, and wheat straw substrates. Protein IDs detected in the cultivations on maltose were removed as a background. There were 49 nodes containing secreted proteins for the lignocellulosic substrates. Then, a total of 6 nodes were selected based on the total occurrence of proteins detected above 90th percentile of the protein counts in 49 nodes.

Correlation of transcriptome and secretome

Spearman's rank correlation was calculated for; 1) genome-wide transcriptome and secretome; and 2) selected genes for only secreted proteins. The correlation coefficients were calculated with the mean transcription level of all nodes from the normalised log₂ reads (transcriptome) and the frequency of proteins secreted in each condition at each time point (secretome).

Term-mining and enrichment analysis

The GO pipeline was used to count the frequency of gene annotations per nodes in the integrated omics models generated from the SHIN pipeline. It was performed using the individual annotation data sets CAZy database [7,8], The Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and EuKaryotic Orthologous Groups (KOG) [33–35]. The appearance of terms was counted using tm package [42]. The genes with the same annotation appearing more than twice in a single node were selected for enrichment analysis. P values of

enriched gene annotations per node were calculated using a function `phyper` in R stats package for the hypergeometric test [43]. The following parameters were considered to estimate the overrepresentation (enrichment) of biological terms in nodes; 1) the number of specific terms of interest in nodes; 2) the number of all terms present in the nodes; 3) the number of specific terms of interest in the genome; 4) the number of all terms present in the genome excluding the specific terms of interest. P values were adjusted using Benjamini-Hochberg (FDR). Annotations with adjusted p values of $FDR < 0.01$ were considered to be statistically significant.

Supporting information

S1 Fig. Growth of *P. coccineus* CIRM-BRFM310 on the substrates.

(PDF)

S2 Fig. Correlation of normalized read counts from the three biological replicates in four conditions at two time points.

(PDF)

S3 Fig. Box and density plot of the normalized log₂ read counts from three biological replicates in four cultivation conditions at two time points.

(PDF)

S4 Fig. The normalized log₂ transformed read count of housekeeping genes.

(PDF)

S5 Fig. Transcriptomic profiles of three biological replicates in four cultivation conditions at two time points.

(PDF)

S1 Table. Statistically significant enrichments in gene annotations per node.

(PDF)

S2 Table. Transcription induction factors for the up-regulated genes.

(PDF)

S3 Table. Spearman's rank correlation of transcriptome and secretome per node for day 3 and 7.

(PDF)

S4 Table. Nodes containing genes up-regulated on at least two lignocellulosic substrates and statistically enriched gene annotations.

(PDF)

S5 Table. Nodes containing genes specifically up-regulated on pine and statistically enriched gene annotations.

(PDF)

S1 File. Animated transcriptomic topographies. The animation was made based on the transcriptomic topographies made with mean transcription levels per node for each cultivation condition (Fig 2).

(ZIP)

S2 File. Enriched CBM1 attached CAZymes in nodes 7 and 14. The output of the hypergeometric test (adjusted $p < 0.05$).

(CSV)

S3 File. Selected six nodes containing the secreted proteins detected. The table contains protein ID, node ID, secretomic observations at day 3 and 7, high transcription, high log₂ fold difference, averaged log₂ normalized read counts of replicates, log₂ transcriptional changes against maltose, CAZyme and other functional annotations.
(XLSX)

S4 File. All 456 nodes with detailed information. The large excel spread sheet contains protein ID, node ID, averaged normalized log₂ transcript read counts of replicates, log₂ transcriptional changes against maltose, specific high transcription, secreted proteins, and corresponding gene annotations (CAZyme, KOG, KEGG, GO).
(XLSX)

S5 File. Summary of secretomic information. The excel spread sheet contains protein ID and the number of spectra for all cultivation conditions.
(XLSX)

Acknowledgments

We would like to express our gratitude to the Mesocentre at Aix-Marseille University for the computing facilities, the INRA MIGALE bioinformatics platform (<http://migale.jouy.inra.fr>) for providing computational resources, Pedro Coutinho from Glycogenomics group at AFMB, Polytech Marseille, France for assisting RNA-sequencing.

Author Contributions

Conceptualization: MNR.

Data curation: SM MNR.

Formal analysis: SM DN DC.

Funding acquisition: MNR.

Investigation: MNR DN SG DC.

Methodology: SM MNR.

Project administration: MNR.

Resources: SM MNR DN SG DC.

Software: SM.

Supervision: MNR JGB.

Validation: SM MNR.

Visualization: SM.

Writing – original draft: SM MNR.

Writing – review & editing: SM MNR JGB.

References

1. Scarlat N, Dallemand J-F, Monforti-Ferrario F, Nita V. The role of biomass and bioenergy in a future bioeconomy: Policies and facts. *Environ Dev.* 2015; 15: 3–34.

2. Kawaguchi H, Hasunuma T, Ogino C, Kondo A. Bioprocessing of bio-based chemicals produced from lignocellulosic feedstocks. *Curr Opin Biotechnol.* 2016; 42: 30–39. <https://doi.org/10.1016/j.copbio.2016.02.031> PMID: 26970511
3. Beckham GT, Johnson CW, Karp EM, Salvachúa D, Vardon DR. Opportunities and challenges in biological lignin valorization. *Curr Opin Biotechnol.* 2016; 42: 40–53. <https://doi.org/10.1016/j.copbio.2016.02.030> PMID: 26974563
4. Blanchette R. Delignification by wood-decay fungi. *Annu Rev Phytopathol.* 1991; 29: 381–398.
5. Floudas D, Bnder M, Riley R, Barry K, Blanchette RA, Henrissat B, et al. 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336:1715–1719. <https://doi.org/10.1126/science.1221748> PMID: 22745431
6. Nagy LG, Riley R, Bergmann PJ, Krizsán K, Martin FM, Grigoriev I V, et al. Genetic Bases of Fungal White Rot Wood Decay Predicted by Phylogenomic Analysis of Correlated Gene-Phenotype Evolution. *Mol Biol Evol.* 2017; 34: 35–44. <https://doi.org/10.1093/molbev/msw238> PMID: 27834665
7. 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(1):41. <https://doi.org/10.1186/1754-6834-6-41> PMID: 23514094
8. 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. <https://doi.org/10.1093/nar/gkt1178> PMID: 24270786
9. Baldrian P, Valášková V. Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiol Rev.* 2008; 32: 501–521. <https://doi.org/10.1111/j.1574-6976.2008.00106.x> PMID: 18371173
10. Langston JA, Shaghasi T, Abbate E, Xu F, Vlasenko E, Sweeney MD. Oxidoreductive Cellulose Depolymerization by the Enzymes Cellobiose Dehydrogenase and Glycoside Hydrolase 61. *Appl Environ Microbiol.* 2011; 77: 7007–7015. <https://doi.org/10.1128/AEM.05815-11> PMID: 21821740
11. Phillips CM, Beeson WT, Cate JH, Marletta MA. Cellobiose Dehydrogenase and a Copper-Dependent Polysaccharide Monooxygenase Potentiate Cellulose Degradation by *Neurospora crassa*. *ACS Chem Biol.* 2011; 6: 1399–1406. <https://doi.org/10.1021/cb200351y> PMID: 22004347
12. Garajova S, Mathieu Y, Beccia MR, Bennati-Granier C, Biaso F, Fanuel M, et al. Single-domain flavoenzymes trigger lytic polysaccharide monooxygenases for oxidative degradation of cellulose. *Sci Rep.* Nature Publishing Group; 2016; 6: 28276.
13. Kersten P, Cullen D. Copper radical oxidases and related extracellular oxidoreductases of wood-decay Agaricomycetes. *Fungal Genet Biol.* 2014; 72: 124–130. <https://doi.org/10.1016/j.fgb.2014.05.011> PMID: 24915038
14. Amore A, Giacobbe S, Faraco V. Regulation of cellulase and hemicellulase gene expression in fungi. *Curr Genomics.* 2013; 14: 230–49. <https://doi.org/10.2174/1389202911314040002> PMID: 24294104
15. Fernández-Fueyo E, Ruiz-Dueñas FJ, López-Lucendo MF, Pérez-Boada M, Rencoret J, Gutiérrez A, et al. A secretomic view of woody and nonwoody lignocellulose degradation by *Pleurotus ostreatus*. *Biotechnol Biofuels.* 2016; 9: 49. <https://doi.org/10.1186/s13068-016-0462-9> PMID: 26933449
16. Couturier M, Navarro D, Chevret D, Henrissat B, Piumi F, Ruiz-Dueñas FJ, et al. Enhanced degradation of softwood versus hardwood by the white-rot fungus *Pycnoporus coccineus*. *Biotechnol Biofuels.* 2015; 8: 216. <https://doi.org/10.1186/s13068-015-0407-8> PMID: 26692083
17. Korripally P, Hunt CG, Houtman CJ, Jones DC, Kitin PJ, Cullen D, et al. Regulation of Gene Expression during the Onset of Lignolytic Oxidation by *Phanerochaete chrysosporium* on Spruce Wood. *Appl Environ Microbiol.* 2015; 81: 7802–7812. <https://doi.org/10.1128/AEM.02064-15> PMID: 26341198
18. Gupta R, Mehta G, Khalsa YP, Kuhad RC. Fungal delignification of lignocellulosic biomass improves the saccharification of celluloses. *Biodegradation.* 2011; 22: 797–804. <https://doi.org/10.1007/s10532-010-9404-6> PMID: 20711746
19. Lomascolo A, Uzan-Boukhris E, Herpoël-Gimbert I, Sigoillot J-C, Lesage-Meessen L. Peculiarities of *Pycnoporus* species for applications in biotechnology. *Appl Microbiol Biotechnol.* 2011; 92: 1129–49. <https://doi.org/10.1007/s00253-011-3596-5> PMID: 22038244
20. Miyauchi S, Navarro D, Grigoriev I V, Lipzen A, Riley R, Chevret D, et al. Visual Comparative Omics of Fungi for Plant Biomass Deconstruction. *Front Microbiol.* 2016; 7: 1335. <https://doi.org/10.3389/fmicb.2016.01335> PMID: 27605927
21. Kohonen T. Self-organized formation of topologically correct feature maps. *Biol Cybern.* 1982; 43: 59–69.
22. Perlack RD, Stokes BJ, Eaton LM, Turnhollow AF. US Billion-ton update. *Biomass Supply for a Bioenergy and Bioproducts Industry.* 2011; 1–227.

23. Lehtio J, Sugiyama J, Gustavsson M, Fransson L, Linder M, Teeri TT. The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proc Natl Acad Sci*. 2003; 100: 484–489. <https://doi.org/10.1073/pnas.212651999> PMID: 12522267
24. Fong M, Berrin J-G, Paës G. Investigation of the binding properties of a multi-modular GH45 cellulase using bioinspired model assemblies. *Biotechnol Biofuels*. 2016; 9: 12. <https://doi.org/10.1186/s13068-016-0428-y> PMID: 26788125
25. Macdonald J, Master ER. Time-dependent profiles of transcripts encoding lignocellulose-modifying enzymes of the white rot fungus *Phanerochaete carnos*a grown on multiple wood substrates. *Appl Environ Microbiol*. 2012; 78: 1596–600. <https://doi.org/10.1128/AEM.06511-11> PMID: 22210217
26. Hori C, Gaskell J, Igarashi K, Kersten P, Mozuch M, Samejima M, et al. 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;
27. Zhu N, Liu J, Yang J, Lin Y, Yang Y, Ji L, et al. 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. <https://doi.org/10.1186/s13068-016-0461-x> PMID: 26900401
28. Bey M, Zhou S, Poidevin L, Henrissat B, Coutinho PM, Berrin J-G, et al. Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monoxygenases (family GH61) from *Podospora anserina*. *Appl Environ Microbiol*. American Society for Microbiology; 2013; 79: 488–96.
29. Hori C, Ishida T, Igarashi K, Samejima M, Suzuki H, Master E, et al. Analysis of the *Phlebiopsis gigantea* Genome, Transcriptome and Secretome Provides Insight into Its Pioneer Colonization Strategies of Wood. *PLoS Genet*. 2014; 10: e1004759. <https://doi.org/10.1371/journal.pgen.1004759> PMID: 25474575
30. Kuuskeri J, Häkkinen M, Laine P, Smolander O-P, Tamene F, Miettinen S, et al. 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. <https://doi.org/10.1186/s13068-016-0608-9> PMID: 27602055
31. Gaskell J, Marty A, Mozuch M, Kersten PJ, Splinter BonDurant S, Sabat G, et al. Influence of *Populus* genotype on gene expression by the wood decay fungus *Phanerochaete chrysosporium*. *Appl Environ Microbiol*. 2014; 80: 5828–35. <https://doi.org/10.1128/AEM.01604-14> PMID: 25015893
32. Vanden Wymelenberg A, Gaskell J, Mozuch M, Splinter BonDurant S, Sabat G, Ralph J, et al. Significant Alteration of Gene Expression in Wood Decay Fungi *Postia placenta* and *Phanerochaete chrysosporium* by Plant Species. *Appl Environ Microbiol*. 2011; 77: 4499–4507. <https://doi.org/10.1128/AEM.00508-11> PMID: 21551287
33. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*. 1999. pp. 29–34. PMID: 9847135
34. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin E V, et al. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics*. 2003; 4: 41. <https://doi.org/10.1186/1471-2105-4-41> PMID: 12969510
35. Blake JA, Christie KR, Dolan ME, Drabkin HJ, Hill DP, Ni L, et al. Gene ontology consortium: Going forward. *Nucleic Acids Res*. 2015; 43: D1049–D1056. <https://doi.org/10.1093/nar/gku1179> PMID: 25428369
36. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002; 30: 207–10 PMID: 11752295
37. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013; 14: R36. <https://doi.org/10.1186/gb-2013-14-4-r36> PMID: 23618408
38. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna; 2013. <http://www.r-project.org/>
39. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15: 550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
40. Wittek P, Gao SC, Lim IS, Zhao L. Somoclu: An Efficient Parallel Library for Self-Organizing Maps. 2013; <http://arxiv.org/abs/1305.1422>
41. Wehrens R, Buydens LMC. Self-and Super-organizing Maps in R: The kohonen Package. *JSS J Stat Softw*. 2007; 21: <http://www.jstatsoft.org/>
42. Feinerer I, Hornik K, Meyer D. Text Mining Infrastructure in R. *J Stat Softw*. 2008; 25: 1–54.
43. Norman L. Johnson, Adrienne W. Kemp SK. Univariate Discrete Distributions. Computational Statistics & Data Analysis. North-Holland; 1994.