Gene family expansions and transcriptome signatures uncover fungal adaptations to wood decay

Hayat Hage, Shingo Miyauchi, Máté Virágh, Elodie Drula, Byoungnam Min, Delphine Chaduli, David Navarro, Anne Favel, Manon Norest, Laurence Lesage-meessen, et al.

HAL Id: hal-03226505
https://hal.inrae.fr/hal-03226505v1

Submitted on 14 May 2021

Gene family expansions and transcriptome signatures uncover fungal adaptations to wood decay

Hayat Hage,[1†] Shingo Miyauchi [ID],[1,17†] Máté Virágh,[2]
Elodie Drula,[1,3] Byoungnam Min,[4,5] Delphine Chaduli,[1,6]
David Navarro,[1,6] Anne Favel,[1,6] Manon Norest,[1]
Laurence Lesage-Meessen,[1,6] Balázs Bálint,[2]
Zsolt Merényi,[2] Laura de Eugenio,[7] Emmanuelle Morin,[8]
Angel T. Martínez,[7] Petr Baldrian [ID],[9]
Martina Štursová,[9] María Jesús Martínez,[7]
Cenek Novotny,[9,10] Jon K. Magnuson,[11]
Joey W. Spatafora,[12] Sundy Maurice,[13]
Jasmyn Pangilinan,[4] Willian Andreopoulos,[4]
Kurt LaButti,[4] Hope Hundley,[4] Hyunsoo Na,[4]
Alan Kuo,[4] Kerrie Barry,[4] Anna Lipzen,[4]
Bernard Henrissat,[14] Robert Riley,[4] Steven Ahrendt,[4]
László G. Nagy [ID],[2,15] Igor V. Grigoriev,[4,5,16]
Francis Martin[8] and Marie-Noëlle Rosso [ID][1*]

[1]INRAE, Aix Marseille Univ, UMR1163, Biodiversité et Biotechnologie Fongiques, Marseille, 13009, France.
[2]Synthetic and Systems Biology Unit, Institute of Biochemistry, Biological Research Center, Szeged, 6726, Hungary.
[3]INRAE, USC1408, AFMB, Marseille, 13009, France.
[4]US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720.
[5]Environmental Genomics and Systems Biology, Lawrence Berkeley National Laboratory, Berkeley, CA, 94720.
[6]INRAE, Aix Marseille Univ, CIRM-CF, UMR1163, Marseille, 13009, France.
[7]Centro de Investigaciones Biológicas Margarita Salas, CIB-CSIC, Madrid, 28040, Spain.
[8]Université de Lorraine, INRAE, UMR1136, Interactions Arbres/Microorganismes, Champenoux, 54280, France.
[9]Institute of Microbiology of the Czech Academy of Sciences, Praha 4, 142 20, Czech Republic.
[10]University of Ostrava, Ostrava, 701 03, Czech Republic.
[11]Pacific Northwest National Laboratory, Richland, WA, 99352.
[12]Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, 97331.
[13]Section for Genetics and Evolutionary Biology, University of Oslo, Oslo, 0316, Norway.
[14]Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia.
[15]Department of Plant Anatomy, Institute of Biology, Eötvös Loránd University, Budapest, 1117, Hungary.
[16]Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, CA.
[17]Max Planck Institute for Plant Breeding Research, Department of Plant Microbe Interactions, Köln, Germany.

Summary

Because they comprise some of the most efficient wood-decayers, Polyporales fungi impact carbon cycling in forest environment. Despite continuous discoveries on the enzymatic machinery involved in wood decomposition, the vision on their evolutionary adaptation to wood decay and genome diversity remains incomplete. We combined the genome sequence information from 50 Polyporales species, including 26 newly sequenced genomes and sought for genomic and functional adaptations to wood decay through the analysis of genome composition and transcriptome responses to different carbon sources. The genomes of Polyporales from different phylogenetic clades showed poor conservation in macrosynteny, indicative of genome rearrangements. We observed different gene family expansion/contraction histories for plant cell wall degrading enzymes in core polyporoids and phlebioids and captured expansions for genes involved in signalling and regulation in the lineages of white rotters. Furthermore, we identified conserved cupredoxins, thaumatin-like proteins and lytic polysaccharide monooxygenases with a yet uncharacterized appended module as new candidate players in wood decomposition. Given the current need for enzymatic toolkits dedicated to the transformation of renewable carbon sources, the observed genomic diversity among Polyporales strengthens the relevance of mining

Polyporales biodiversity to understand the molecular mechanisms of wood decay.

Introduction

The Polyporales order, which represents 1.5% of all described fungal species (Kirk *et al*., 2009), belongs to the Agaricomycetes class, a highly diverse and large group of mushroom-forming fungi within the phylum Basidiomycota. Polyporales are widely distributed in boreal, temperate, tropical and subtropical areas. The great majority of Polyporales are efficient wood decay saprotrophs with significant contribution to carbon re-allocation and carbon cycling in forest soils (Boberg *et al*., 2014; Chen *et al*., 2019). Based on decay symptoms, the fungal attack causes to the wood, Polyporales and other wood decay fungi have been classified traditionally into two rotting types: white rotters, which degrade all plant cell wall (PCW) polymers, including lignin, and often leave lighter-coloured cellulose behind, and brown rotters, which degrade the PCW cellulosic and hemicellulosic compounds with partial modification of lignin, resulting in brown cubic wood fragments (Blanchette, 1984; Martínez *et al*., 2005). However, beyond these archetypal situations, a continuum in the ratio of lignin vs. carbohydrate depolymerization during wood colonization has been observed, which highlights diverse strategies for lignocellulose modifications (Schilling *et al*., 2020). Besides saprotrophs, additional nutrition modes have been observed for some species. The pathogens colonize living trees through wounds and cause *wood rot disease*. More occasionally, endophytes in leaves and sapwood of trees are suggested to stand as latent saprotrophs (Sokolski *et al*., 2007; Martin *et al*., 2015). Finally, some have been described as potential temporary parasites of other fungi preliminarily established in the wood (Rayner *et al*., 1987).

The order Polyporales has been subdivided phylogenetically into two lineages including, among others, the core polyporoid, gelatoporia, and antrodia clades on the one hand, and the phlebioid and residual polyporoid clades on the other hand (Justo *et al*., 2017). Most species produce white-rot, the plesiomorphic decay type for Polyporales (Floudas *et al*., 2012), except species from the antrodia lineage and a few genera that convergently transitioned to the brown-rot type (Justo *et al*., 2017).

Each rot type involves a characteristic panel of enzymatic and chemical mechanisms for the degradation of PCW polymers (reviewed in the study by Lundell *et al*., 2014), which is generally reflected by a distinct repertoire of genes coding PCW degrading enzymes (PCWDE) in the genome. Notably, carbohydrate-active enzymes (CAZymes) related to the depolymerization of crystalline cellulose such as lytic polysaccharide monooxygenases (LPMO) or to the depolymerization of lignin (class II peroxidases, POD) and auxiliary enzymes such as multicopper oxidases (MCO) and H_2O_2-generating glucose–methanol–choline-oxidoreductases are generally expanded in the genome of white-rot fungi (Floudas *et al*., 2012; Ruiz-Dueñas *et al*., 2013; Nagy *et al*., 2016; Miyauchi *et al*., 2020), although in some species, the absence of genes for lignin-active POD does not fit the white-rot phenotype (Riley *et al*., 2014). Conversely, brown-rot fungi largely rely on the production of reactive oxygen species (ROS) via the iron-dependent Fenton chemistry for the oxidative attack of wood polymers prior to enzymatic degradation (Martínez *et al*., 2005; Eastwood *et al*., 2011). Accordingly, the transition from white-rot to brown-rot was associated with the loss of key POD enzymes for lignin degradation (Floudas *et al*., 2012) and a change in the temporal expression of oxidative vs. hydrolytic enzymes in the early stages of wood degradation (Zhang *et al*., 2019). Besides PCW degradation, other physiological functions are involved in the ability to grow on wood, such as the intracellular transport and detoxification of toxic products released during the degradation process, as reflected by enrichment in the gene portfolios for Major Facilitator Superfamily transporters, cytochrome P450 and glutathione transferases in the genomes of these fungi (Syed *et al*., 2014; Deroy *et al*., 2015; Qhanya *et al*., 2015; Nagy *et al*., 2016).

Beyond those global characteristics, white-rot and brown-rot fungi each developed a diversity of enzymatic mechanisms to cope with different PCW chemical compositions (Suzuki *et al*., 2012; Deroy *et al*., 2015; Presley and Schilling, 2017). Within Polyporales, white-rot species degrade lignin with different degrees of cellulose preservation (Fernandez-Fueyo *et al*., 2012; Miyauchi *et al*., 2018) and use different sets of enzymes to degrade lignin. For example, within the phlebioid clade, some species have no laccase, an enzyme with ligninolytic activity (Rivera-Hoyos *et al*., 2013), (e.g *Phanerochaete chrysosporium*, *Phanerochaete carnosa*; Suzuki *et al*., 2012), whereas four and five laccase genes have been identified in *Phlebia centrifuga* and *Phlebia radiata* respectively (Kuuskeri *et al*., 2016). Similarly, in the genomes of core polyporoids, the number of lignin-active PODs varies significantly from one species to another. For example, no genes of lignin peroxidase, a subfamily of PODs, were identified in the genome of *Dichomitus squalens*, whereas 10 genes were found in that of *Trametes versicolor* (Ruiz-Dueñas *et al*., 2013).

This diversity in enzymatic systems in Polyporales is of high interest to address the current demand for bio-based chemicals and materials derived from renewable carbon source – the plant biomass – without competition with food or feed production or with the preservation of natural habitats. In this context, the plant tissues to be

transformed are issued from wastes or co-products of agriculture and forestry, which are recalcitrant to degradation. The recent progress in understanding the enzymatic mechanisms used by Polyporales for the oxidative degradation of non-starch polymers (Zhou *et al*., 2015; Berrin *et al*., 2017; Martínez *et al*., 2017; Couturier *et al*., 2018) has stimulated the screening of Polyporales strains for process development and the identification of strains able to use raw biomasses as carbon source (Berrin *et al*., 2012; Zhou *et al*., 2015; Miyauchi *et al*., 2018).

A recent phylogenomic study suggested that high diversification rates occurred in Polyporales (Varga *et al*., 2019), which could have shaped enzymatic mechanisms to adapt diverse ecological niches and substrates, and further strengthens the relevance of the exploration of this taxon for enzymatic diversity. Yet, further studies are required to extend our understanding on the genomic adaptations to wood decay and to assess the functional diversity of the enzymatic systems across Polyporales.

In this work, we investigated the genomic basis of evolutionary adaptations related to wood decay within Polyporales. We compared the phylogenomic features and the repertoires of protein-coding genes of 50 fungal species including 26 newly sequenced genomes. The genomic diversity within the order was examined in terms of their genome arrangement and the evolution of gene families. The transcriptomic trends of selected Polyporales species from the core polyporoid and phlebioid clades during degradation of diverse lignocellulosic substrates led to the discovery of conserved gene sets regulated for PCW degradation. Our results unveil some of the mechanisms underlying Polyporales diversification and pinpoint to yet overlooked proteins that could contribute to the ability of Polyporales to degrade recalcitrant PCW polymers.

Results

Polyporales phylogeny and genome features

We sequenced the genome of 26 species, including two species (*Abortiporus biennis* and *Panus rudis*) from the residual polyporoid clade, for which no genome was available, and species from the core polyporoid (16 genomes), antrodia (4 genomes) and phlebioid clades (4 genomes; Supporting Information Tables S1 and S2). We used these 26 genomes along with 24 previously published Polyporales genomes in a phylogenetic reconstruction of Polyporales based on 540 single-copy genes (Supporting Information Table S3). The topology of the maximum likelihood tree confirmed the phylogeny based on three genes previously reported by Justo *et al*. (2017) and the classification into the five main clades; core

polyporoid, antrodia (grouping all brown-rot Polyporales), gelatoporia, phlebioid and residual polyporoid. The tree topology confirmed the placement of the residual polyporoid clade as the sister of the phlebioid clade (Fig. 1A). We estimated the divergence times of the major Polyporales lineages. Molecular clock analyses of a phylogenomic dataset showed that the Polyporales might be around 183 million years old, which is in line with recent molecular clock studies of Agaricomycetes (Varga *et al*., 2019). Trametoid species were estimated at ca. 61 million years, which is consistent with a 45 million years old Trametes fossil from the literature (Knobloch and Kotlaba, 1994).

Overall, our study included 36 white rotters and 10 brown rotters, including *Laetiporus sulphureus*, which has been described both as a brown rotter and a pathogen (Song and Cui, 2017), and four saprotroph species with no attributed decay type (Fig. 1B). The assembly size of the 50 Polyporales genomes ranged from 28 to 66 Mb (mean 39 Mb, standard deviation SD 7) and the transposable element (TE) coverage was positively correlated with genome size (0.37%–16% TE coverage, mean 8%, SD 8, Pearson correlation coefficient = 0.4, *P*-value 0.001). The most abundant TEs were long terminal repeats from the Gypsy and Copia families (Supporting Information Fig. S1).

To assess the conservation of the genome structure across Polyporales, we analysed the genome synteny within and across clades. For this analysis, we selected the 23 Polyporales genomes that had their 10 largest scaffolds covering more than 40% of the genome (Supporting Information Table S4). We quantified macrosynteny by calculating the percentage of hits occurring in the same order on the compared sequences. As expected, the genome macrosynteny was higher between species for which the MRCA had diverged more recently (Fig. 2). Consistent with the estimated time to the MRCA of each clade, the genomes from the core polyporoid clade (94 million years old) showed a higher macrosynteny conservation (mean 13.5% hits in macrosynteny) than those from the antrodia clade (134 million years old; 7.6%) or the phlebioid clade (133 million years old; 7.9%; Supporting Information Fig. S2). It is noteworthy that no significant correlation was observed between TE coverage and synteny loss (Pearson correlation coefficient = −0.12, *P*-value 0.2), suggesting that TE movements were not the major drivers for genome reorganizations.

Core, dispensable and species-specific genes

The protein-coding genes from the 50 Polyporales genomes were clustered in orthologous groups and used to calculate the counts of core, dispensable and species-
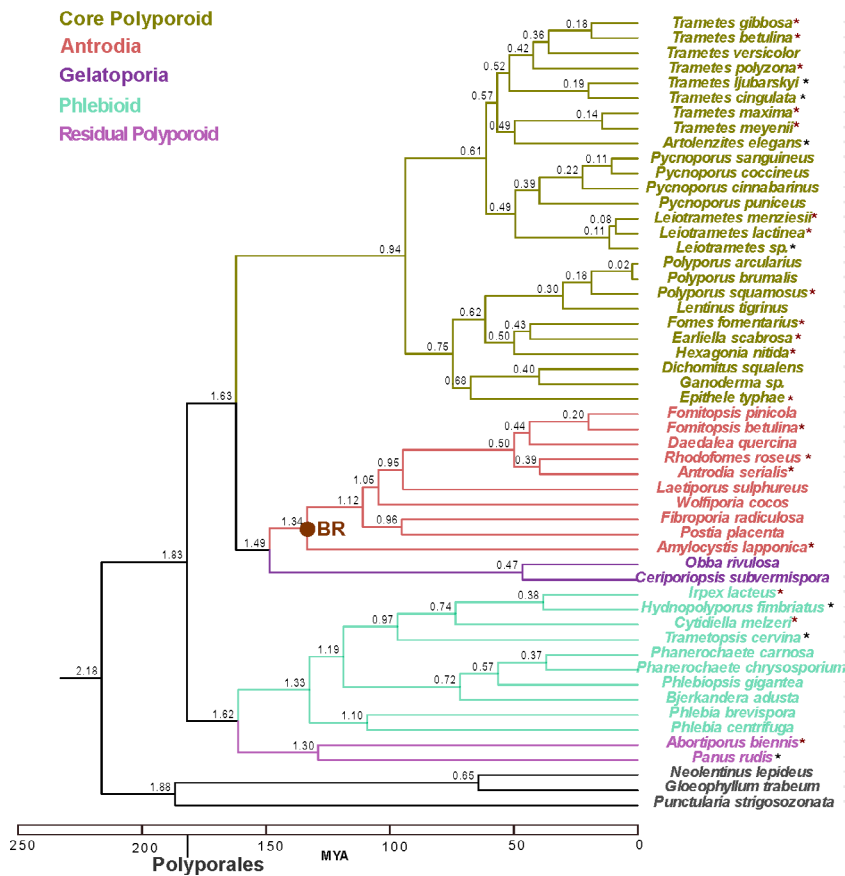
Fig. 1. Phylogenetic relationship and genome features of Polyporales fungi.
A. Time-calibrated species phylogeny. Numbers next to nodes represent mean estimated ages from MCMCTree using six fossil calibrations.
B. Genome size, genome content in transposable elements (TEs) and number of protein-coding genes. The newly sequenced genomes are indicated with asterisks. Genomes sequenced with long-read methods are indicated with red asterisks.
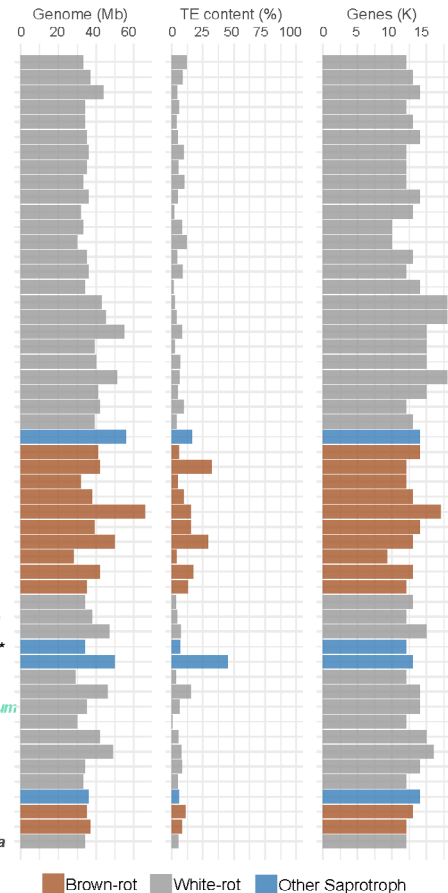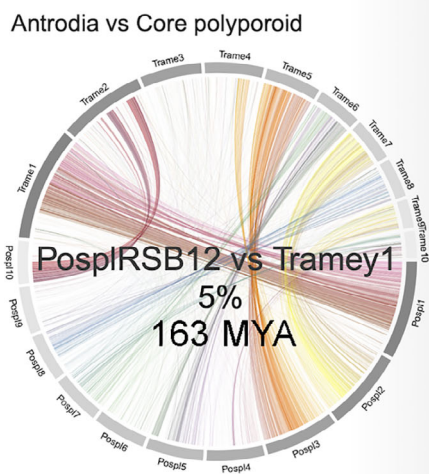
specific genes. The core genes conserved among all genomes consisted of 2169 clusters and its proportion in each species varied from 13% to 26% (mean 18%, SD 2) of the total proteome. The proportion of species-specific genes varied from 11% to 43% (mean 26%, SD 8). From the latter, 97% (mean 98%, SD 5) had no homologues in all annotated fungal genomes in MycoCosm (as of October 2019, 1420 genomes) (Supporting Information Fig. S3; Table S5). Intriguingly, we identified that an average of 81% of these species-unique genes were annotated as small secreted proteins (<300 aa, SSP; Supporting Information Figs. S4 and S5). To test whether the predicted SSP genes could result from defects in gene sequencing and assembly or from defects in gene structural annotations, we looked at transcript evidences in the transcriptomes of six species, *Irpex lacteus*, *Artolenzites elegans*, *Leiotrametes* sp., *Pycnoporus cinnabarinus*, *Pycnoporus coccineus* and *Trametes ljubarskyi* (see below). In each species, in our growth

conditions, we identified transcripts for 87%–94% of the predicted SSP genes, supporting the reliability of our SSP gene predictions.

The phylogenetic clades within Polyporales have different portfolios of secreted CAZymes

We assessed the diversity of the gene repertoires for enzymes involved in lignocellulose degradation across Polyporales, by comparing the repertoires of genes coding for secreted CAZymes in the 50 genomes. Phylogeny had a higher impact than rot type on the repertoires of secreted CAZymes in Polyporales (38% and 7% of the variance respectively, PERMANOVA; Mesny, 2020). As expected, we observed that the gene repertoires of brown rotters (from the antrodia clade) were distinct from those of white rotters (Fig. 3A, Supporting Information Fig. S6), mainly due to significantly reduced numbers of
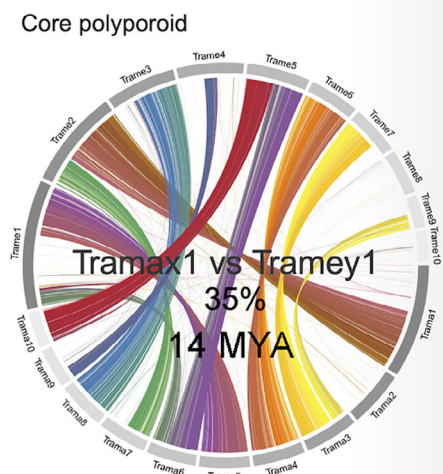
Fig. 2. Comparison of synteny conservation between genomes from different Polyporales clades: core polyporoid (*Trametes meyenii* vs. *Trametes maxima*), phlebioid (*Phanerochaete chrysosporium* vs. *Phlebia brevispora*) and antrodia (*Postia placenta vs. Wolfiporia cocos*). The percentage of hits occurring in the same order on the compared block of sequences between two genomes and the estimated time to the most recent common ancestor between the two species are shown inside each circos plot.

A

B
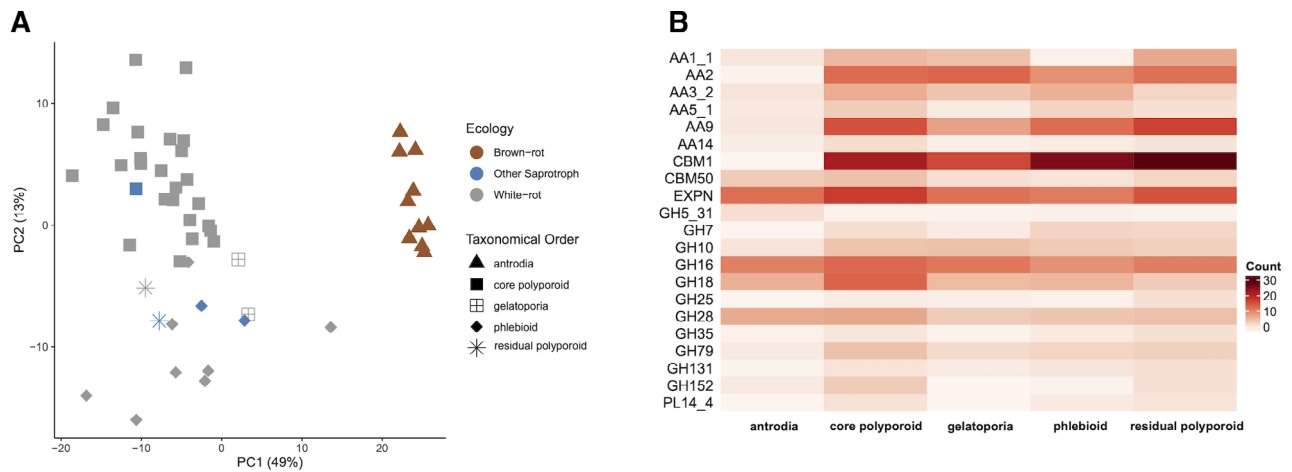
Fig. 3. Comparison of the counts of secreted CAZyme genes in 50 Polyporales species.
A. PCA of the counts of genes coding for secreted CAZymes in 50 Polyporales species.
B. Heatmap showing the differences between Polyporales clades in counts of the secreted CAZymes contributing the most to the distribution of the species in the PCA plot.

PCWDE in antrodia genomes, notably laccases (AA1_1), POD (AA2), carbohydrate binding module 1 (CBM1), AA9 LPMOs and expansins (Fig. 3B; Supporting Information Fig. S7; Table S6). Among white rotters, the phlebioids had reduced numbers of laccase genes (Kruskal–Wallis test, *P* value <0.05), whereas core polyporoids were enriched for GH18 chitinases and for GH79 β-glucuronosidases, which target 4-*O*-methyl-glucuronic acids from plant hemicelluloses (Kuroyama *et al*., 2001). Interestingly, the gelatoporia (white-rotters), a sister clade of antrodia (brown rotters), had similar gene counts as core polyporoids or phlebioids for PODs, CBM1 modules, AA5_1 and PL14_4, but resembled antrodia by low numbers of laccase, AA9 LPMO and expansin genes. Although the two gelatoporia species analysed here (*Obba rivulosa* and *Ceriporiopsis subvermispora*) are white rot fungi (Blanchette *et al*., 1997; Miettinen *et al*., 2016), this similarity with antrodia repertoires suggested gene losses had occurred in a common ancestor of the two clades (see below).

CAZyme genes underwent different gain/loss trajectories in Polyporales lineages

We further investigated the gene family expansion/contraction events that took place on eight early nodes of the Polyporales, partly because we wanted to avoid the effects of large, species-specific expansions that can stem from genome assembly artefacts and because early events are more likely to generate conserved genetic signatures of wood decay capabilities of the whole order. We found that extensive CAZyme gene loss took place before diversification of antrodia and gelatoporia, which was congruent with reduced numbers of AA9 and
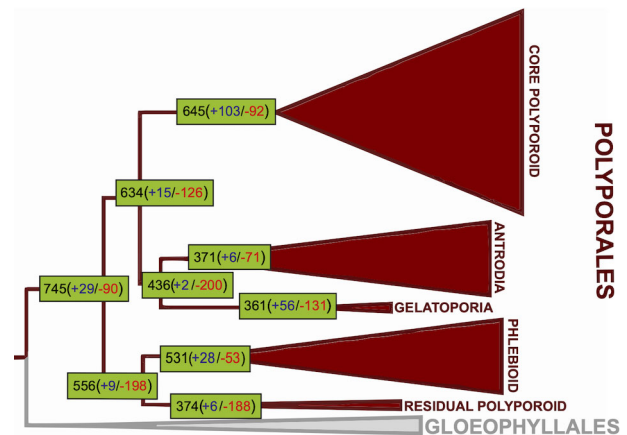
Fig. 4. Numbers of genes coding for CAZymes in the most recent common ancestors of the five Polyporales clades.

expansins in white rot gelatoporia (Fig. 4). Surprisingly, the MRCA of phlebioids had a smaller reconstructed ancestral CAZyme repertoire (531 genes) than that of the core polyporoid clade (645 genes). This is driven by a large difference in the number of gene duplications inferred in the MRCA of the core polyporoid vs that of the phlebioid clade: 103 duplications vs nine (in comparison, only six in the antrodia clade). We find that several oxidoreductases, CBM1-s and expansins contribute to this pattern. For example, the laccase, POD and GMC-oxidoreductase families showed four, three and 12 duplications in the most recent common ancestor (MRCA) of the core polyporoid clade, respectively, but 0, 0 and 5 duplications in the MRCA of the phlebioid clade, respectively. Similarly, CBM1-containing genes and expansins showed 17 and 13 duplications in the former while only three and two in the latter clade. LPMO-s

(AA9) showed a more moderate difference (three duplications in the MRCA of the core polyporoid vs one in that of the phlebioid). These differences in gene family expansions in the MRCA of Polyporoid clades have resulted in reduced numbers of laccase genes (AA1_1) in extant phlebioid genomes (Supporting Information Fig. S7). However, in other CAZyme families, our reconstruction suggests that successive gene loss events took place in core polyporoids whereas gene gains were predominant in phlebioids. For example, in core polyporoids, gene gains in general peroxidase (GP) were followed by extensive gene loss in several lineages (Supporting Information Fig. S8).

We further inspected the diversity of POD gene portfolios among Polyporales. We classified the AA2 POD genes based on the presence of known conserved residues into the three POD types active on lignin; MnP, VP and LiP (Ruiz-Dueñas *et al*., 2013). Our study on 10 antrodia genomes confirmed previous observations on the presence of non-ligninolytic GP and the absence of MnP, LiP and VP in these brown rotters. Conversely, the white-rot species showed diverse combinations of the POD types. From the 40 genomes analysed, 15 contained the three types of lignin-active peroxidases, 19 contained combinations of MnP and VP or MnP and LiP and six contained MnP only (Fig. 5). None of the genomes contained the LiP type only. In addition, we noticed rare occurrences of atypical MnP with two acidic amino acids (QED and EQE) instead of three at the Mn(II) binding site, which were not previously described in Agaricomycetes (Ruiz-Dueñas *et al*., 2020).

Gene tree-species tree reconciliation studies suggested that the diversity in POD portfolios among Polyporales resulted from different evolutionary histories in the different lineages, and from recent gene gain events for LiP genes in the phlebioids *Phanerochaete chrysosporium* and *Bjerkandera adusta*, for MnP in gelatoporia (*Ceriporiopsis subvermispora*) and residual polyporoids (*Abortiporus biennis)*, and for VP in core polyporoids and in the residual polyporoid *Panus rudis* (Supporting Information Fig. S8).

Several gene families underwent intensive expansions in early Polyporales

We searched for gene families whose expansion/contraction trajectories had potential functional consequences upon major Polyporales clades, with no *a priori* knowledge on these functions. To answer this question, we searched which gene families expanded most significantly in basal nodes of the Polyporales and converted gene duplication statistics to duplication rates; this allowed us to rank gene families by their duplication rate in early Polyporales clades. The predicted functions for

each cluster were deduced from InterPro domains (Supporting Information Table S7). The 10 most significantly expanding clusters at the Polyporales MRCA had predicted functions related to signalling and protein–protein interactions. Cluster12444, which contains proteins with an EF-Hand, calcium binding domain (IPR018247), had the highest gain rate and showed expansions in seven of the eight examined early Polyporales nodes suggesting that it is involved in key wood-decay processes. Both a survey of published RNASeq data (Miyauchi *et al*., 2018, 2020; Wu *et al*., 2018) and the data generated in this study showed that genes from this cluster were differentially expressed in seven out of eight inspected Polyporales species in response to diverse ligno(cellulosic) substrates (Supporting Information Table S8). Similarly, we found evidence for differential expression during growth on lignocellulosic substrates, for 8 out of 10 most significantly expanding clusters, which contained protein kinase domains (Supporting Information Table S8).

Expansion and contraction signatures of white-rot and brown-rot Polyporales

To refine our search for potentially decay-associated gene families, we constructed a model with the prerequisite that a gene family showed the following attributes: (i) expansion in early white-rot-associated nodes of the Polyporales, (ii) losses in the brown-rot antrodia clade and (iii) the family is conserved in >50% of white-rot species. We detected 338 gene families that fit this model (Supporting Information Table S9). Somewhat surprisingly, the most expanding families in the white-rot Polyporales included only a few PCWDEs, AA9 LPMOs (three clusters), a GH5 cluster (sub-family GH5_9 exo-β-1,3-glucanases), a GH7 cellulase, a GH30 endo-1,4-β-xylanase, a GMC-oxidoreductase cluster, as well as PODs. The lack of expansion for other PCWDE is consistent with the inference that white-rot originated long before the Polyporales (Floudas *et al*., 2012), so probably relevant PCWDE expansions also happened earlier. On the other hand, a preponderance of peptidase (13 clusters), cytochrome p450s (13 clusters) and transporter (13 clusters) clusters was obvious among the families which are most highly duplicated in early Polyporales but are lost in brown rotters (Supporting Information Table S9). Furthermore, several groups of regulator families, including seven clusters of F-box domain proteins, eight of transcription factors, two clusters of BTB/POZ proteins, four clusters of RING/MYND type zinc finger proteins (both of which can indirectly regulate transcription), and six of protein kinases were found. Based on their expansion/contraction patterns, we speculate that these families may be functionally linked to wood decay;
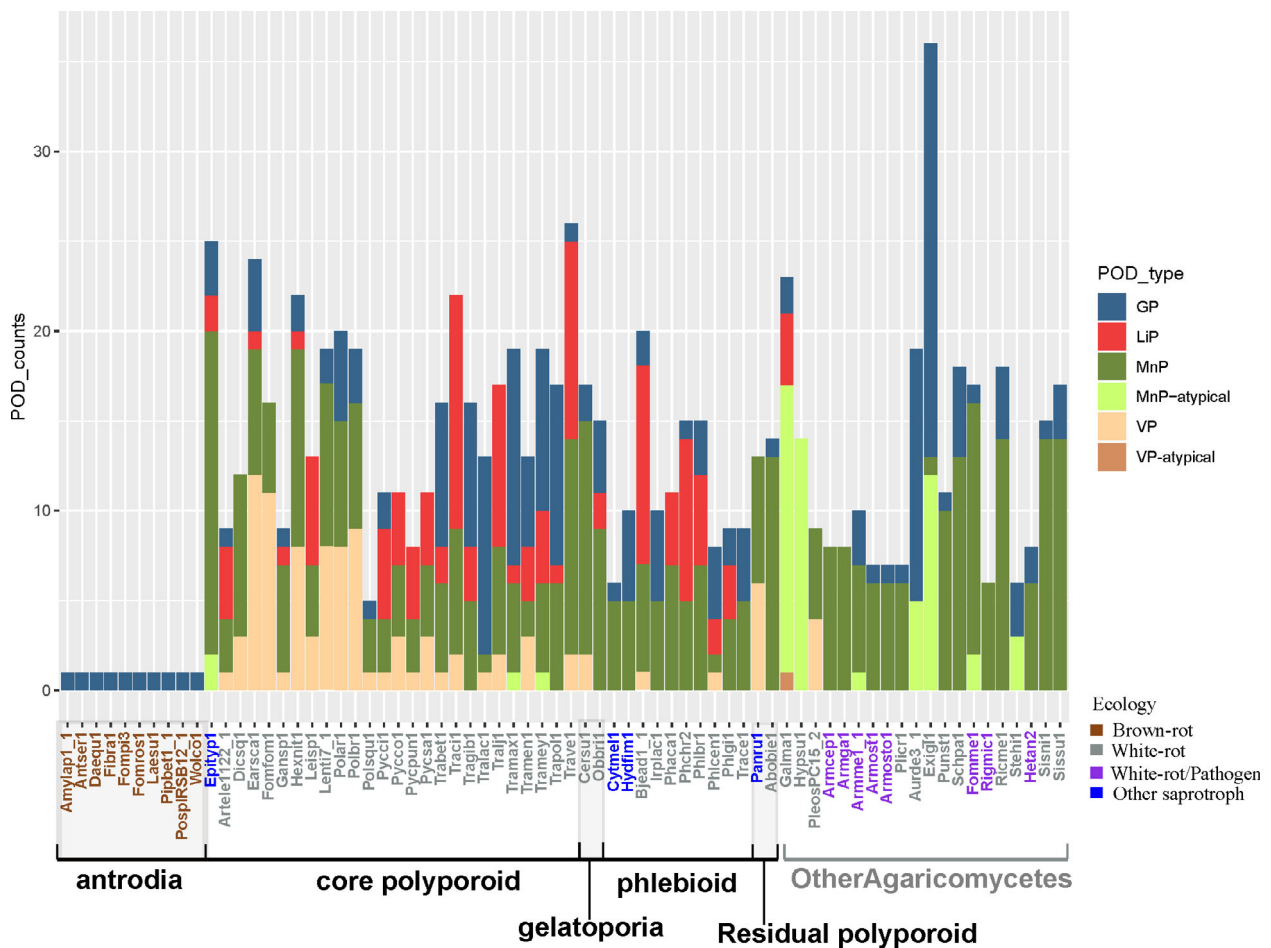
Fig. 5. Count of class II peroxidases (PODs) in Polyporales genomes and in Agaricomycetes species from other taxonomic orders.

they comprise worthy targets for detailed functional follow-up experiments.

Transcriptomics show conserved signatures for white-rot decay across Polyporales

To further assess the functional adaptation of Polyporales for wood decay activities, we compared the transcriptomic responses to three lignocellulosic substrates representing gramineae (wheat-straw), softwood (pine) or hardwood (aspen) using the following species, *Artolenzites elegans*, *Leiotrametes* sp., *Pycnoporus cinnabarinus*, *Pycnoporus coccineus*, *Trametes ljubarskyi* and *Irpex lacteus*. We used the differential expression level of the genes coding predicted secreted proteins and CAZymes during growth on each substrate (i.e. differences in transcription levels compared to the common maltose control for fair cross-comparisons). In total, we retrieved 5536 genes from the six species and examined the transcription trends by grouping similarly regulated genes using Self-Organizing Map (SOMs; see Methods;

Supporting Information Table S10). SOM clusters the genes according to the transcription profiles into nodes and locates nodes with similar patterns in the vicinity of each other on a map. The trained SOM containing 195 nodes was sliced into each of the substrate conditions. Such maps (i.e. Tatami maps) gave an overview of the transcriptomic patterns showing the mean normalized log2 fold change in each node (Fig. 6A). We observed that the genes up-regulated on the substrates were clustered in the bottom left area of the map. The overall sets of genes up-regulated in response to pine and aspen were similar, and the set of genes up-regulated in response to wheat straw globally overlapped those responsive to cellulose and to woody substrates. Looking at the distribution of CAZyme genes on the map for an overview of CAZyme expression regulations, we observed that almost all genes coding for CBM1-associated CAZymes were localized in nodes with upregulation (Fig. 6B). We also noticed enrichment for glycoside hydrolase and carbohydrate esterase genes in the up-regulated fraction of the transcriptome. On the contrary, expansin and polysaccharide
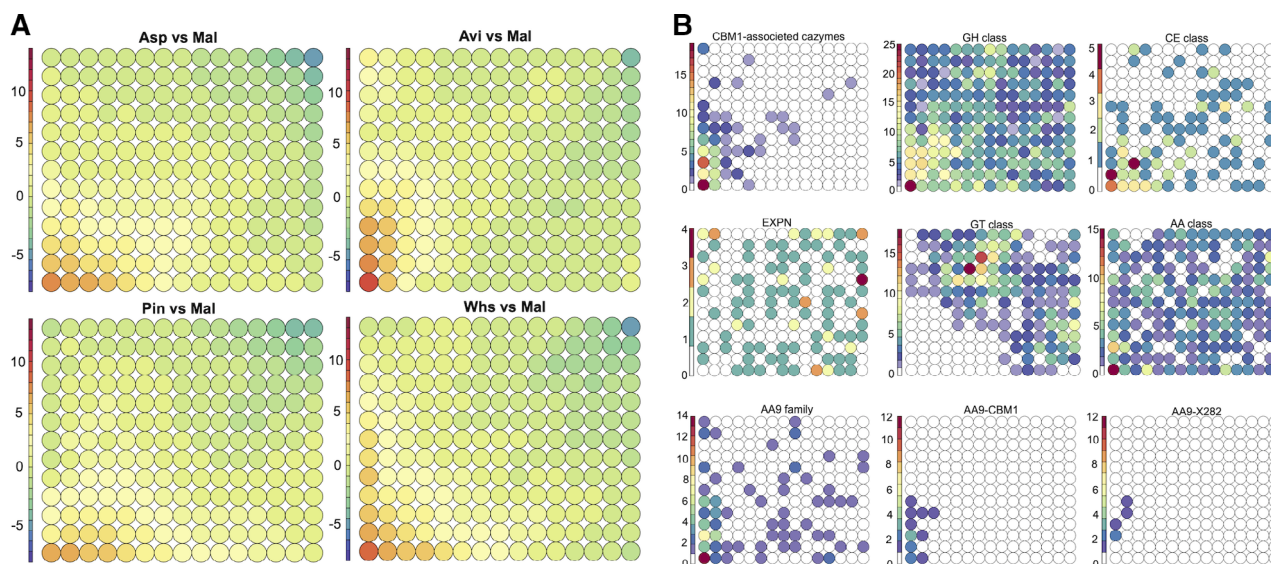
Fig. 6. Overview of the transcription regulation of genes coding for CAZymes and predicted secreted proteins from six Polyporales species. The Tatami maps show, for each node, the mean log2 differential transcription of genes at day 3 on cellulose, aspen, pine, or wheat straw in comparison to maltose (A) and the counts for selected CAZyme genes in each node (B).

lyase genes (Supporting Information Fig. S9) were uniformly localized on the map and were not preferentially clustered in up-regulated nodes. Interestingly, the genes coding for glycosyltransferases were located outside the up-regulated nodes, indicating these genes were not differentially regulated by the substrates. Globally, the AA genes were evenly distributed over the map, except for AA9 LPMOs. In particular, AA9 LPMOs with a conserved C-terminal module (called X282; Lenfant *et al*., 2017) were all clustered with cellulose responsive genes. For each fungus, one single AA9-X282 gene was identified, which was strongly up-regulated (28–500-fold change in transcript read counts) after 3 day-growth on cellulose (Supporting Information Table S11).

For a closer inspection of the conserved responses across the six species, the genes up-regulated on the tested substrates were retrieved from the nodes with the mean fold change in normalized transcript read counts >4 after growth on cellulose, aspen, pine, or wheat straw in comparison to maltose. We first analysed the expression of CAZyme genes. Out of the 63 CAZyme families analysed, 11 were up-regulated in the six species in response to the four substrates (Fig. 7A; Supporting Information Fig. S10). This set of CAZyme hence represented a core response to (ligno)cellulose and contained the expected GH1 β-glucosidases, GH5_5 β-1,4-endoglucanasse, GH6 and GH7 cellobiohydrolases, GH131 glucanases, GH10 xylanases, CE1 and CE16 carbohydrate esterases, GH28 polygalacturonases and numerous oxidative AA9 LPMOs active on crystalline

cellulose. Cellobiose dehydrogenases (AA8-AA3_1 modular enzymes), β-mannosidases (GH5_7) and β-glucuronidases that target glycosylated proteins (GH79) were found up-regulated in five out of the six species. We then examined the CAZyme genes specifically regulated in response to complex lignin and hemicellulose containing substrates, not to cellulose. We observed no conservation in the specific responses to complex substrates across the tested Polyporales (Supporting Information Fig. S11). These results suggested that cellulose was the main inducer of the shared response to the substrates within Polyporales. However, we cannot rule out the possibility that the absence of conservation in CAZyme response to complex substrates could be due to the analysis of a single time-point in our study, with the observation window being too narrow to see common changes.

Using a similar approach with no *a priori* on the functions of the differentially regulated genes, we looked at the genes with the same InterPro description that were up-regulated in at least five species. We found SSP genes with a thaumatin domain (IPR001938) were up-regulated in all species, except in *Pycnoporus cinnabarinus* (Fig. 7B; Supporting Information Table S12). In plants, thaumatins and thaumatin-like proteins are involved in plant defence against pathogens. However, the newly identified proteins had no similarity with known proteins from the SwissProt database and had four predicted disulfide bridges, instead of eight in plant thaumatins, suggesting they might have different functions (Supporting Information Table S13). We also found genes coding predicted secreted proteins with a
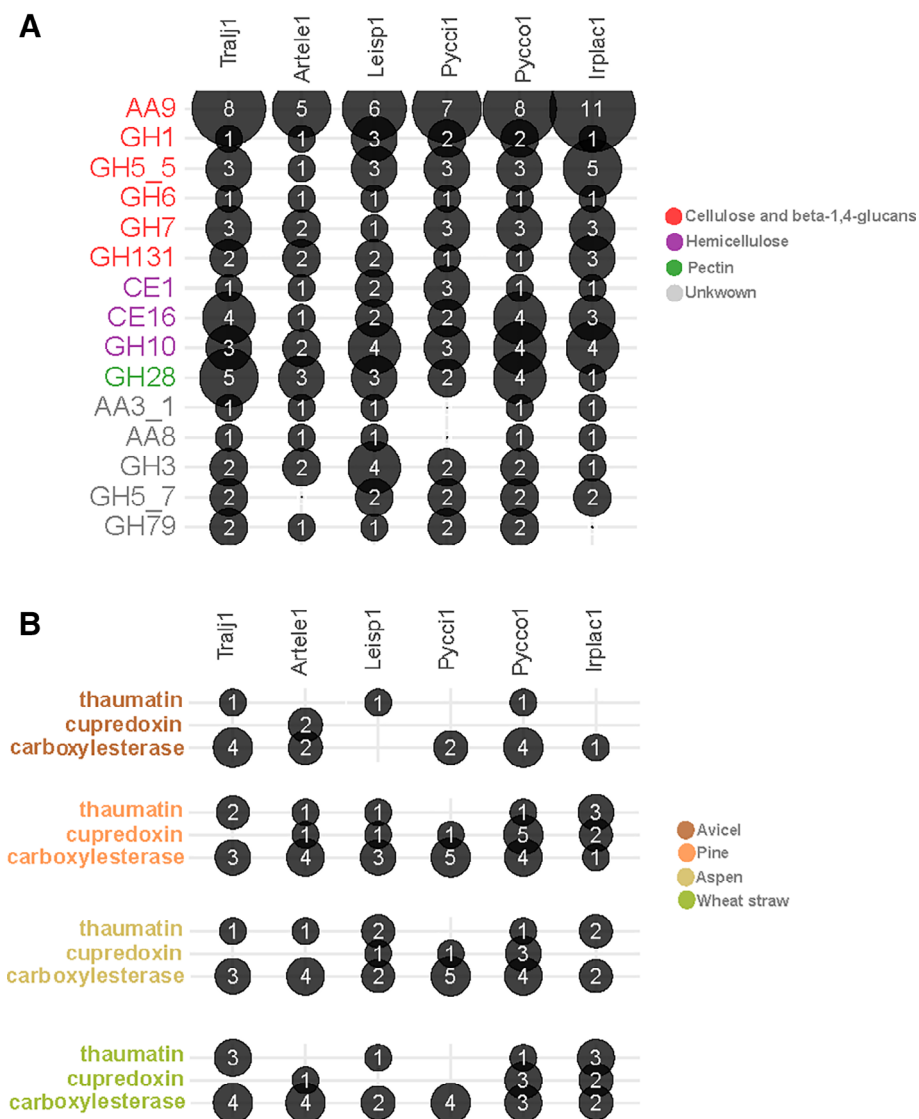
A

B

Fig. 7. Count of CAZyme genes commonly up-regulated in response to cellulose, aspen, pine and wheat straw in at least five of the species.
A. The colours indicate the substrates targeted by each CAZy family.
B. Count of Carboxylesterase, cupredoxin and thaumatin genes up-regulated in response to cellulose (Avicel), pine, aspen and wheat straw.

'Carboxylesterase, type_B' domain (IPR002018) were commonly up-regulated on all the tested substrates except in *Leiotrametes* sp. grown on cellulose (Fig. 7B). Using BLASTP search on Fungi proteins in the SwissProt database, we identified these proteins as candidate lipases (Supporting Information Table S13). Finally, we identified genes coding for predicted 'cupredoxin' proteins (IPR008972) were up-regulated in response to pine in all species except *T. ljubarskyi*. We identified the copper protein ARB_05732-1 from the ascomycete *Arthroderma benhamiae* as the closest homologue among Fungi proteins (Supporting Information Table S13). Similar to ARB_05732-1, the candidate cupredoxin proteins we identified carried a predicted signal peptide for secretion and a probable glycosylphosphatidylinositol (GPI)-anchor domain.

Discussion

We found Polyporales might be around 183 million years old and have undergone intensive diversification, which is consistent with previous findings (Varga *et al*., 2019). First, we observed poor synteny conservation within Polyporales clades, and the loss of synteny was not related to TE coverage. Further studies may be required to elucidate possible mechanisms for genome rearrangements. Second, on average 26% of the protein-coding genes were species-specific, among which on average 81% were coding for predicted SSPs. The abundance of species-specific SSPs in Polyporales genomes is in line with previous observations on the genomes of saprotroph fungi including white-rot and brown-rot fungi, which encode numerous SSPs (Pellegrin *et al*., 2015;

Kim *et al*., 2016), and suggests they might have a role in the determination of specific traits. For example, SSPs could be used as weapons or decoys in intra- and inter-specific cell-to-cell communication in soil, wood, or the rhizosphere. They could also be used as regulators of quorum sensing (Feldman *et al*., 2020). Our transcriptomic analysis on six Polyporales species revealed that about half of the SSPs with transcript evidence in each genome had no predicted function (55.6%, SD 4.4). And 8.7%–32.7% (mean 22.9%, SD 9) of those were highly transcribed or up-regulated after 3 days of growth on lignocellulosic substrates. Among the up-regulated SSPs with an InterPro domain, proteins with a thaumatin domain were up-regulated on pine and aspen in five out of the six species analysed. Yet, no predicted function was identified for these SSPs by homology search. SSPs with a GPI membrane anchoring domain and a single cupredoxin (IPR008972) domain were induced in response to the lignocellulosic substrates in all six species analysed. The cupredoxin fold allows copper-mediated electron transfer with the reduction of molecular oxygen to water. In eukaryotes, toxic copper (II) ions are reduced to copper (I) by monodomain cupredoxins before being transported inside the cell via the high-affinity copper transporters of the CTR family (Arnesano *et al*., 2002). Further studies are required to elucidate whether SSPs harbouring a cupredoxin domain are involved in copper homeostasis at the fungal cell wall, as recently suggested for LPMO-like copper containing proteins with a GPI-anchor domain located at the fungal cell wall of ectomycorrhizal and pathogenic fungi (Garcia-Santamarina *et al*., 2020; Labourel *et al*., 2020).

In a search for idiosyncrasies related to wood decay, we analysed gene family expansion/contraction events inferred in early ancestors of the Polyporales. Surprisingly, we observed only few expansions for PCWDE in early Polyporales. White-rot fungi most likely evolved in the Carboniferous in the most recent common ancestor of the Auriculariales and higher Agaricomycetes (Floudas *et al*., 2012; Nagy *et al*., 2016), which predates the Polyporales MRCA by at least 100–150 million years (this study, Varga *et al*., 2019). We explain the lack of clear PCWDE expansions in early Polyporales by the already diverse lignocellulosic toolkit present before the origin of the Polyporales (see, e.g. Nagy *et al*., 2016), which may have provided early members of the Polyporales with the necessary tools to degrade plant cell walls. A remarkable exception of this pattern are class-II peroxidases (POD), that existed before, but further evolved for nonphenolic lignin degradation in the Polyporales (Ayuso-Fernández *et al*., 2018), which coincided with duplication events also captured by our analyses. Among POD, MnP were identified in all white-rot Polyporales genomes, consistent with MnP being the first lignin-active POD type, which

appeared in the common ancestor of white rotters (Floudas *et al*., 2012; Nagy *et al*., 2016). VP and LiP were present each in 69% of the white-rot Polyporales genomes but had scarce occurrences in other white-rot Agaricomycetes we analysed. Conversely, atypical MnP, with unconventional catalytic residues, were rare in Polyporales genomes (maximum two gene copies identified in three Polyporales species), whereas they are abundant in fungi from the Agaricomycetes orders (Ruiz-Dueñas *et al*., 2020). Whether these atypical MnP are involved in the degradation of lignin or lignin-derived polymers, such as the humic acids from forest litter is still to be determined.

Interestingly, our analysis of gene loss/gain within Polyporales showed different trajectories in the evolution of PCWDE gene families in the MRCA of the core polyporoid and phlebioid clade. It follows that functional differences in the wood-decay modes of white rotters from these two clades might also exist, in particular with regard to AA families, expansins and CBM1-containing proteins, but these will require further study.

We searched for transcriptomic evidence for shared mechanisms used by species from the core polyporoid and phlebioid clade as indicators of conserved toolkits of efficient wood decay. By comparing the transcriptomes of six species during growth on four different ligno (cellulosic) substrates, we found that genes coding for AA9 LPMOs appended to a CBM1 carbohydrate binding module were up-regulated in the six tested species. Our results are in line with a major role for AA9 LPMOs in cellulose degradation by Polyporales fungi (Berrin *et al*., 2017) and the previous observation that the CBM1 module can enhance targeting of the LPMO catalytic module to the substrate (Bennati-Granier *et al*., 2015; Chalak *et al*., 2019). Interestingly, we observed the conserved upregulation of genes coding for AA9 LPMOs with an atypical ~30 amino acid module fused to the C-terminal of the catalytic domain. The X282 module is to date only found in saprotroph Agaricomycetes and absent from the genome sequences of mycorrhizal Agaricomycetes (Supporting Information Fig. S12), and each of the six Polyporales genomes analysed here contained one single AA9-X282 gene. In the case of the phlebioid *Irpex lacteus*, the enzyme also carried a CBM1 module. Further studies are required to understand the role of the X282 extension in the activity or specificity of the enzyme and its contribution to wood decay.

We also identified in the six analysed species several lipase genes (triacylglycerol acylhydrolases, EC 3.1.1.3) strongly up-regulated during growth on lignocellulosic substrates. The secreted lipases could catalyse the hydrolysis of long chain triacylglycerol substrates (>C8) present in the wood extractives or at the fungal cell wall (Kadimaliev *et al*., 2006) or could contribute to the

degradation of suberin, a highly recalcitrant polyester embedded in the secondary cell walls in the phellem of tree barks and in the endodermis of roots (Martins *et al*., 2014). Besides, wood lipid degradation generates free unsaturated fatty acids further converted by MnP-mediated peroxidation into lipid hydroperoxides, which could be involved in lignin and xenobiotic degradation (Gutiérrez *et al*., 2002).

Finally, we identified gene families with different expansion/contraction in early white-rot vs. brown-rot Polyporales. We identified gene expansions for peptidases, cytochrome p450s and transporters in early white-rot Polyporales. These findings support a role for these proteins in wood decay and are congruent with the previously observed upregulation of these genes and the secretion of peptidases during growth on lignocellulosic substrates (e.g. this study; Hori *et al*., 2014; Korripally *et al*., 2015; Miyauchi *et al*., 2017, 2018; Moudy and Stites, 2017). In early Polyporales, we also observed gene expansion for several groups of regulator families, including F-box domain proteins, transcription factors, BTB/POZ proteins, RING/MYND type zinc finger proteins, LLR domain proteins and protein kinases. Although at the moment, we do not have evidence for the involvement of these regulator families in wood decay, these gene duplication/loss patterns in the Polyporales is remarkable and lend support to the hypothesis that, while PCWDEs were in place in the ancestors of the Polyporales, the diversification of some regulatory families followed only after the origin of the order. This hypothesis is consistent with the generally low conservation of regulators of lignocellulose degradation (see e.g. Mattila *et al*., 2020). Nevertheless, more detailed experimental follow-up are necessary to establish potential direct links with wood decay.

Altogether, our results highlight Polyporales genomes as remarkable resources to unravel the enzymatic mechanisms of lignocellulose breakdown. Because they have evolved a diversity of enzymatic mechanisms to overcome the PCW polymer barrier, Polyporales are promising resources to inspire enzymatic toolkits aimed at ecofriendly processes to access PCW polysaccharides and to facilitate the extraction and bioconversion of plant molecules (Arantes *et al*., 2012).

Materials and methods

Sequencing, assembly, and annotation of 26 new Polyporales genomes

Genomic libraries were created using either 'Illumina Regular Fragment, 300 bp', 'Illumina Regular Long Mate Pair, 4kb, CLRS', 'PacBio Low Input 10kb', or 'PacBio > 10kb' protocols (Tables S1 and S2; see Supporting

Information). PacBio SMRTbell templates were sequenced using the Pacific Biosciences RSII, or SEQUEL sequencers with the following parameters: Version C4 chemistry, 4 h run times (RSII); Version 2.0 chemistry, 6 h and 10 h run times (SEQUEL). Libraries were assembled with Falcon (Maeder, 2018), finished with finisherSC (Lam *et al*., 2015), and polished with Arrow (Maeder, 2019) or Quiver (Chin *et al*., 2013). The Illumina genomic libraries were sequenced with Illumina HiSeq2500 or HiSeq2000 using HiSeq TruSeq SBS sequencing kits, v4 (2×150 indexed run for 'Regular Fragment' genomic libraries; 2×100 indexed run for 'Regular LMP' genomic libraries). Illumina genomic libraries were assembled with AllPathsLG (Gnerre *et al*., 2011) (Supporting Information Table S2). For RNA sequencing, the mycelia were ground in liquid nitrogen using a Freezer/Mill Cryogenic Grinder (SPEX Sample Prep, United Kingdom). Total RNA was extracted from 100 mg ground tissue using TRIZOL (Ambion), precipitated with isopropanol, resuspended in water and treated with RNase-Free DNase I (QIAGEN). Total RNA was precipitated again with LiCl and resuspended in DEPC-treated water as described in the study by Miyauchi *et al*. (2020). RNA Sequencing was performed using an Illumina HiSeq 2000 or HiSeq 2500 instrument. All genomes were annotated using the JGI annotation pipeline (Grigoriev *et al*., 2014). CAZymes and auxiliary activity enzymes (AA) were annotated as in the study by Lombard *et al*. (2014). Gene models from PODs were further classified based on the presence of known conserved residues i.e. manganese peroxidase (MnP), with a Mn(II)-oxidation site formed by three conserved acidic residues, lignin peroxidase (LiP) with an exposed catalytic tryptophan, and versatile peroxidase (VP) with both the Mn(II) oxidation site found in MnP and the catalytic tryptophan found in LiP (Floudas *et al*., 2012; Ruiz-Dueñas *et al*., 2013).

Reconstruction of the duplication/loss history of gene families

Whole proteomes of the 50 Polyporales species, and three from the orders Gloeophyllum and Corticiales (Supporting Information Table S3) were clustered to identify single-copy orthogroups. After removal of poorly aligned regions, the trimmed alignments of single-copy clusters were concatenated into a supermatrix to build a maximum likelihood tree that was used as input to MCMCTree (Puttick, 2019) for molecular clock dating analyses (see Supporting Information). To investigate the evolutionary history of each gene family, protein alignments of each cluster containing at least four proteins were first trimmed using trimAl 1.4 (Capella-Gutiérrez *et al*., 2009) with a parameter –gt 0.2. Then we inferred a

maximum likelihood phylogeny for each gene cluster and calculated Shimodaira-Hasegawa-like (SH-like) branch support values using the PTHREADS version of RAxML 8.2.12 under the PROTGAMMAWAG model. Next, we midpoint-rooted and reconciled the rooted gene trees with the species tree using Notung 2.9 (Chen *et al*., 2000) with an edge-weight threshold of 0.95. Gene trees were then further processed by cutting them at basal duplication nodes where possible, which resulted in a final number of 36 153 clusters containing at least four proteins. Together with the clusters containing less than four proteins, we reconstructed the duplication/loss history of 288 253 protein clusters across the species tree using the COMPARE pipeline (Nagy *et al*., 2014, 2016). Duplication/loss rates were computed by dividing the number of inferred duplication/loss events by the length of the respective branch of the species tree the event occurred on. Gene families were functionally characterized based on InterPro annotations. Graphical maps of gene duplication/loss histories were generated using custom scripts in R.

Protein conservation across Polyporales

We downloaded the protein sequences from the JGI MycoCosm database, for the 50 Polyporales species. A reciprocal protein BLAST was performed using BLAST +2.7.1. Orthologous genes were then clustered using FASTORTHO (Wattam *et al*., 2014) with 50% identity and 50% length coverage thresholds. The predicted proteins shared by all 50 genomes were counted as core proteins. The ones shared by at least two genomes were counted as dispensable. The sequences of species-specific proteins were further used to search for similarities in all available fungal proteomes by running BLAST on the MycoCosm (50% identity and 50% coverage).

Identification of transposable elements

We identified transposable elements (TEs) using a previously described method (Payen *et al*., 2016). Briefly, REPEATSCOUT 1.0.5 (Price *et al*., 2005) with default parameters was used to generate *de novo* predictions of repeat elements in the unmasked genomes. Repeated sequences found more than 10 times in the genomes were annotated using TBLASTX (Altschul *et al*., 1990) against the fungal reference sequences from REPBASE version 22.08 (http://www.girinst.org/server/RepBase/index.php). The coverage of TEs in the genomes, including unknown TEs, was estimated by REPEATMASKER open 4.0.6 (http://www.repeatmasker.org). The results were integrated and visualized using the Transposon Identification Nominative Genome Overview custom script in R (TINGO; Morin *et al*., 2019).

Genome macrosynteny

The analysis of genome macrosynteny across Polyporales was done on 23 selected genomes that had their 10 largest scaffolds covering more than 40% of their genomes. Pair-wise comparisons and identification of syntenic blocks were done on the 10 largest scaffolds using the R package DECIPHER (Wright, 2015), with default parameters and without masking repeat sequences. The macrosynteny between two species was visualized with the R package Circlize (Gu *et al*., 2014). Data manipulation, integration, and visualization were coordinated using a set of custom R scripts Syntenic Governance Overview (SynGO, available on request).

Transcriptomics of six Polyporales species

Artolenzites elegans BRFM 1663, *Leiotrametes* sp. BRFM 1775, *Pycnoporus cinnabarinus* BRFM 137, *Pycnoporus coccineus* BRFM 310, *Trametes ljubarskyi* BRFM 1659 and *Irpex lacteus* CCBAS Fr. 238617/93 were grown for 3 days in liquid culture medium supplemented with cellulose Avicel PH 101 (Fluka) (15 g.L^{-1}), ground and sifted wheat straw fragments <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 tremuloides* (15 g.L^{-1}) (see Supporting Information). For control, each strain was grown in the same medium supplemented with maltose (20 g.L^{-1}). Each culture was done in triplicate. RNA sequencing was performed as described above. The sequence reads were trimmed for quality and aligned to the reference genome to generate the raw gene counts (see Supporting Information). Raw gene counts were used to evaluate the level of correlation between biological replicates using Pearson's correlation. DESeq2 (version 1.2.10; Love *et al*., 2014) was subsequently used to determine which genes were differentially expressed between pairs of conditions. The parameters used to identify differentially expressed genes were Bonferroni-adjusted *P*-value < 0.05 and Benjamini–Hochberg-adjusted *P*-value < 0.05. The transcript log2 fold changes were transformed into the range zero to one using unity-based normalization (Aksoy and Haralick, 2001; Supporting Information Fig. S13). The normalized values were used to train the Self-organizing map (SOM) with 975 000 training iterations (195 map units × 5000 times), using the visual multi-omics pipeline Self-organizing map Harbouring Informative Nodes with Gene Ontology (SHIN + GO; Miyauchi *et al*., 2017, 2020). The Tatami maps generated from SHIN + GO represented clustered genes from the six strains with similar transcription profiles.

Comparison of the predicted secretomes

The secreted proteins were predicted using a custom pipeline described in the study by Pellegrin *et al*. (2015) (Supporting Information). Predicted secreted proteins smaller than 300 amino acids were considered as small secreted proteins (SSPs). CAZymes were annotated as in the study by Lombard *et al*. (2014). The impacts of phylogeny and lifestyle on secreted CAZyme counts were assessed using the function adonis2 from the Vegan R package (Oksanen *et al*., 2012) after converting the Poly-porales species tree into evolutionary distance using R package *ape* (Paradis *et al*., 2004) and CAZyme counts into distance matrix using Vegdist function. RVAideMemoire package (Maxim, 2018) was used to check significant difference among Polyporales clades using a pair-wise PERMANOVA (Mesny, 2020).

Data availability

The assemblies and annotations of 26 newly sequenced genomes are available at the US Department of Energy Joint Genome Institute (JGI) fungal genome portal.

MycoCosm (https://mycocosm.jgi.doe.gov; Grigoriev *et al*., 2014) and in the DDBJ/EMBL/GenBank repository (https://www.ncbi.nlm.nih.gov; Table S1). The RNASeq data generated in the current study are available at Gene Expression Omnibus repository under accession nos. GSE82427, GSE82419, GSE156901.

Acknowledgements

Author contributions

L.L.M., M.N.R., and F.M. conceived the work. I.V.G. and K.B. coordinated JGI projects. A.F., D.C., D.N., S.G., S.M., L.L.M., M.N., A.T.M., L.d.E., M.J.M., C.N., P.B., M.S., S.M., J.M., and J.W.S. produced the biological material. J.P., W.A., K.L.B., H.H., H.N., A.K., K.B., A.L., R.R., S.A., B.H., E.D., and S.M. generated and analysed the data. L.G.N., M.V., Z.M., and B.B. performed the phylogenomics analyses. S.M. did the SOM analysis. H.H. and S.M. did the comparative genomics analysis. H.H., S.M., L.G.N., and M.N.R. wrote the manuscript with the help of all co-authors.

References

Aksoy, S., and Haralick, R.M. (2001) Feature normalization and likelihood-based similarity measures for image retrieval. *Pattern Recogn Lett* **22**: 563–582.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.

Arantes, V., Jellison, J., and Goodell, B. (2012) Peculiarities of brown-rot fungi and biochemical Fenton reaction with regard to their potential as a model for bioprocessing biomass. *Appl Microbiol Biotechnol* **94**: 323–338.

Arnesano, F., Banci, L., Bertini, I., and Thompsett, A.R. (2002) Solution structure of CopC: a cupredoxin-like protein involved in copper homeostasis. *Structure* **10**: 1337–1347.

Ayuso-Fernández, I., Ruiz-Dueñas, F.J., and Martínez, A.T. (2018) Evolutionary convergence in lignin-degrading enzymes. *Proc Natl Acad Sci* **115**: 6428–6433.

Bennati-Granier, C., Garajova, S., Champion, C., Grisel, S., Haon, M., Zhou, S., *et al* (2015) Substrate specificity and regioselectivity of fungal AA9 lytic polysaccharide mono-oxygenases secreted by *Podospora anserina*. *Biotechnol Biofuels* **8**: 90.

Berrin, J.-G., Navarro, D., Couturier, M., Olivé, C., Grisel, S., Haon, M., *et al*. (2012) Exploring the natural fungal biodiversity of tropical and temperate forests toward improvement of biomass conversion. *Appl Environ Microbiol* **78**: 6483–6490.

Berrin, J.-G., Rosso, M.-N., and Abou, H.M. (2017) Fungal secretomics to probe the biological functions of lytic polysaccharide monooxygenases. *Carbohydr Res* **448**: 155–160.

Blanchette, R.A. (1984) Screening wood decayed by white rot fungi for preferential lignin degradation. *Appl Environ Microbiol* **48**: 647–653.

Blanchette, R.A., Krueger, E.W., Haight, J.E., Akhtar, M., and Akin, D.E. (1997) Cell wall alterations in loblolly pine wood decayed by the white-rot fungus, *Ceriporiopsis subvermispora*. *J Biotechnol* **53**: 203–213.

Boberg, J.B., Finlay, R.D., Stenlid, J., Ekblad, A., and Lindahl, B.D. (2014) Nitrogen and carbon reallocation in fungal mycelia during decomposition of boreal forest litter (RA Wilson, Ed.). *PLoS ONE* **9**: e92897.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics (Oxford, England)* **25**: 1972–1973.

Chalak, A., Villares, A., Moreau, C., Haon, M., Grisel, S., d'Orlando, A., *et al* (2019) Influence of the carbohydrate-

binding module on the activity of a fungal AA9 lytic polysaccharide monooxygenase on cellulosic substrates. *Biotechnol Biofuels* **12**: 206.

Chen, K., Durand, D., and Farach-Colton, M. (2000) NOTUNG: a program for dating gene duplications and optimizing gene family trees. *J Comput Biol* **7**: 429–447.

Chen, J., Heikkinen, J., Hobbie, E.A., Rinne-Garmston, K.T., Penttilä, R., and Mäkipää, R. (2019) Strategies of carbon and nitrogen acquisition by saprotrophic and ectomycorrhizal fungi in Finnish boreal Picea abies-dominated forests. *Fungal Biol* **123**: 456–464.

Chin, C.-S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., et al. (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* **10**: 563–569.

Couturier, M., Ladevèze, S., Sulzenbacher, G., Ciano, L., Fanuel, M., Moreau, C., et al. (2018) Lytic xylan oxidases from wood-decay fungi unlock biomass degradation. *Nat Chem Biol* **14**: 306–310.

Deroy, A., Saiag, F., Kebbi-Benkeder, Z., Touahri, N., Hecker, A., Morel-Rouhier, M., et al. (2015) The GSTome reflects the chemical environment of white-rot fungi. *PLoS one* **10**: e0137083.

Eastwood, D.C., Floudas, D., Binder, M., Majcherczyk, A., Schneider, P., Aerts, A., et al. (2011) The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science (New York, N.Y.)* **333**: 762–765.

Feldman, D., Yarden, O., and Hadar, Y. (2020) Seeking the roles for fungal small-secreted proteins in affecting saprophytic lifestyles. *Front Microbiol* **11**: 455.

Fernandez-Fueyo, E., Ruiz-Dueñas, F.J., Ferreira, P., Floudas, D., Hibbett, D.S., Canessa, P., et al. (2012) Comparative genomics of *Ceriporiopsis subvermispora* and *Phanerochaete chrysosporium* provide insight into selective ligninolysis. *Proc Natl Acad Sci U S A* **109**: 5458–5463.

Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R. A., Henrissat, B., et al. (2012) The paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* **336**: 1715–1719.

Garcia-Santamarina, S., Probst, C., Festa, R.A., Ding, C., Smith, A.D., Conklin, S.E., et al. (2020) A lytic polysaccharide monooxygenase-like protein functions in fungal copper import and meningitis. *Nat Chem Biol* **16**: 337–344.

Gnerre, S., MacCallum, I., Przybylski, D., Ribeiro, F.J., Burton, J.N., Walker, B.J., et al. (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci U S A* **108**: 1513–1518.

Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., et al. (2014) MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* **42**: D699–D704.

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014) Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**: 2811–2812.

Gutiérrez, A., del Río, J.C., Martínez-Íñigo, M.J., Martínez, M.J., and Martínez, Á.T. (2002) Production of new unsaturated lipids during wood decay by ligninolytic basidiomycetes. *Appl Environ Microbiol* **68**: 1344–1350.

Hori, C., Gaskell, J., Igarashi, K., Kersten, P., Mozuch, M., Samejima, M., and Cullen, D. (2014) 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* **80**: 2062–2070.

Justo, A., Miettinen, O., Floudas, D., Ortiz-Santana, B., Sjökvist, E., Lindner, D., et al. (2017) A revised family-level classification of the Polyporales (Basidiomycota). *Fungal Biol* **121**: 798–824.

Kadimaliev, D.A., Nadezhina, O.S., Atykyan, N.A., Revin, V. V., and Samuilov, V.D. (2006) Interrelation between the composition of lipids and their peroxidation products and the secretion of ligninolytic enzymes during growth of *Lentinus (Panus) tigrinus*. *Microbiology* **75**: 563–567.

Kim, K.-T., Jeon, J., Choi, J., Cheong, K., Song, H., Choi, G., et al. (2016) Kingdom-wide analysis of fungal small secreted proteins (SSPs) reveals their potential role in host association. *Front Plant Sci* **7**: 186.

Kirk, P., Cannon, P., Minter, D., and Stalpers, J. (2009) Dictionary of the fungi. *Mycol Res* **113**: 908–910.

Knobloch, E., and Kotlaba, F. (1994) *Trametites eocenicus*, a new fossil polypore from the Bohemian Eocene. *Czech Mycol* **47**: 207–214.

Korripally, P., Hunt, C.G., Houtman, C.J., Jones, D.C., Kitin, P.J., Cullen, D., and Hammel, K.E. (2015) Regulation of gene expression during the onset of ligninolytic oxidation by *Phanerochaete chrysosporium* on spruce wood. *Appl Environ Microbiol* **81**: 7802–7812.

Kuroyama, H., Tsutsui, N., Hashimoto, Y., and Tsumuraya, Y. (2001) Purification and characterization of a beta-glucuronidase from *Aspergillus niger*. *Carbohydr Res* **333**: 27–39.

Kuuskeri, J., Häkkinen, M., Laine, P., Smolander, O.-P., Tamene, F., Miettinen, S., et al. (2016) 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* **9**: 192.

Labourel, A., Frandsen, K.E.H., Zhang, F., Brouilly, N., Grisel, S., Haon, M., et al. (2020) A fungal family of lytic polysaccharide monooxygenase-like copper proteins. *Nat Chem Biol* **16**: 345–350.

Lam, K.-K., LaButti, K., Khalak, A., and Tse, D. (2015) FinisherSC: a repeat-aware tool for upgrading de novo assembly using long reads. *Bioinformatics* **31**: 3207–3209.

Lenfant, N., Hainaut, M., Terrapon, N., Drula, E., Lombard, V., and Henrissat, B. (2017) A bioinformatics analysis of 3400 lytic polysaccharide oxidases from family AA9. *Carbohydr Res* **448**: 166–174.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**: D490–D495.

Love, M.I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

Lundell, T.K., Mäkelä, M.R., de Vries, R.P., and Hildén, K.S. (2014) Chapter eleven - genomics, lifestyles and future prospects of wood-decay and litter-decomposing

Basidiomycota. In *Advances in Botanical Research: Fungi*, Vol. **70**, Martin, F.M. (ed). London: Academic Press, pp. 329–370.

Martin, R., Gazis, R., Skaltsas, D., Chaverri, P., and Hibbett, D. (2015) Unexpected diversity of basidiomycetous endophytes in sapwood and leaves of Hevea. *Mycologia* **107**: 284–297.

Martínez, A.T., Ruiz-Dueñas, F.J., Camarero, S., Serrano, A., Linde, D., Lund, H., *et al*. (2017) Oxidoreductases on their way to industrial biotransformations. *Biotechnol Adv* **35**: 815–831.

Martínez, A.T., Speranza, M., Ruiz-Dueñas, F.J., Ferreira, P., Camarero, S., Guillén, F., *et al*. (2005) Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. *Int Microbiol* **8**: 195–204.

Martins, I., Hartmann, D.O., Alves, P.C., Martins, C., Garcia, H., Leclercq, C.C., *et al*. (2014) Elucidating how the saprophytic fungus *Aspergillus nidulans* uses the plant polyester suberin as carbon source. *BMC Genomics* **15**: 613.

Maeder H. (2018) FALCON: experimental PacBio diploid assembler. [WWW document] URL https://github.com/PacificBiosciences/FALCON.

Maeder H. (2019) *PacBio® variant and consensus caller* [WWW document]. URL https://github.com/PacificBiosciences/GenomicConsensus

Mattila, H.K., Mäkinen, M., and Lundell, T. (2020) Hypoxia is regulating enzymatic wood decomposition and intracellular carbohydrate metabolism in filamentous white rot fungus. *Biotechnol Biofuels* **13**: 26.

Maxim H. (2018) RVAideMemoire: testing and plotting procedures for biostatistics. R Package Version 0.9–69.

Mesny F. (2020) Detecting the effect of biological categories on genomes composition [WWW document]. URL https://github.com/fantin-mesny/Effect-Of-Biological-Categories-On-Genomes-Composition

Miettinen, O., Riley, R., Barry, K., Cullen, D., de Vries, R.P., Hainaut, M., *et al*. (2016) Draft genome sequence of the white-rot fungus *Obba rivulosa* 3A-2. *Genome Announc* **4**: e00976–e00916.

Miyauchi, S., Hage, H., Drula, E., Lesage-Meessen, L., Berrin, J.-G., Navarro, D., *et al*. (2020) Conserved white-rot enzymatic mechanism for wood decay in the Basidiomycota genus *Pycnoporus*. *DNA Res* **27**: dsaa011.

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

Miyauchi, S., Rancon, A., Drula, E., Hage, H., Chaduli, D., Favel, A., *et al*. (2018) Integrative visual omics of the white-rot fungus *Polyporus brumalis* exposes the biotechnological potential of its oxidative enzymes for delignifying raw plant biomass. *Biotechnol Biofuels* **11**: 201.

Morin, E., Miyauchi, S., San Clemente, H., Chen, E.C.H., Pelin, A., de la Providencia, I., *et al*. (2019) Comparative genomics of *Rhizophagus irregularis*, *R. cerebriforme*, *R. diaphanus* and *Gigaspora rosea* highlights specific genetic features in Glomeromycotina. *New Phytol* **222**: 1584–1598.

Moudy, M., and Stites, W. (2017) Investigation of methionine sulfoxide formation as a regulator of proteolysis. *Biophys J* **112**: 202a.

Nagy, L.G., Ohm, R.A., Kovács, G.M., Floudas, D., Riley, R., Gácser, A., *et al*. (2014) Latent homology and convergent regulatory evolution underlies the repeated emergence of yeasts. *Nat Commun* **5**: 4471.

Nagy, L.G., Riley, R., Bergmann, P.J., Krizsán, K., Martin, F.M., Grigoriev, I.V., *et al*. (2016) Genetic bases of fungal white rot wood decay predicted by phylogenomic analysis of correlated gene-phenotype evolution. *Mol Biol Evol* **34**: 35–44.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin P, O'Hara RB, et al. (2012) Vegan: Community Ecology Package. R package version 2.0-2.

Paradis, E., Claude, J., and Strimmer, K. (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**: 289–290.

Payen, T., Murat, C., and Martin, F. (2016) Reconstructing the evolutionary history of gypsy retrotransposons in the Périgord black truffle (*Tuber melanosporum* Vittad.). *Mycorrhiza* **26**: 553–563.

Pellegrin, C., Morin, E., Martin, F.M., and Veneault-Fourrey, C. (2015) Comparative analysis of secretomes from ectomycorrhizal fungi with an emphasis on small-secreted proteins. *Front Microbiol* **6**: 1278.

Presley, G.N., and Schilling, J.S. (2017) Distinct growth and secretome strategies for two taxonomically divergent brown rot fungi. *Appl Environ Microbiol* **83**: e02987–e02916.

Price, A.L., Jones, N.C., and Pevzner, P.A. (2005) De novo identification of repeat families in large genomes. *Bioinformatics* **21**: i351–i358.

Puttick, M.N. (2019) MCMCtreeR: functions to prepare MCMCtree analyses and visualize posterior ages on trees. *Bioinformatics* **35**: 5321–5322.

Qhanya, L.B., Matowane, G., Chen, W., Sun, Y., Letsimo, E.M., Parvez, M., *et al*. (2015) Genome-wide annotation and comparative analysis of cytochrome P450 monooxygenases in Basidiomycete biotrophic plant pathogens. *PLOS One* **10**: e0142100.

Rayner, A.D.M., Boddy, L., and Dowson, C.G. (1987) Temporary parasitism of *Coriolus* spp. by *Lenzites betulina*: a strategy for domain capture in wood decay fungi. *FEMS Microbiol Lett* **45**: 53–58.

Riley, R., Salamov, A.A., Brown, D.W., Nagy, L.G., Floudas, D., Held, B.W., *et al*. (2014) Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. *Proc Natl Acad Sci* **111**: 9923–9928.

Rivera-Hoyos, C.M., Morales-Álvarez, E.D., Poutou-Piñales, R.A., Pedroza-Rodríguez, A.M., RodrÍguez-Vázquez, R., and Delgado-Boada, J.M. (2013) Fungal laccases. *Fungal Biol Rev* **27**: 67–82.

Ruiz-Dueñas, F.J., Lundell, T., Floudas, D., Nagy, L.G., Barrasa, J.M., Hibbett, D.S., and Martínez, A.T. (2013) Lignin-degrading peroxidases in Polyporales: an evolutionary survey based on 10 sequenced genomes. *Mycologia* **105**: 1428–1444.

Ruiz-Dueñas, F.J., Barrasa, J.M., Sánchez-García, M., Camarero, S., Miyauchi, S., Serrano, A., *et al*. (2020)

Genomic analysis enlightens Agaricales lifestyle evolution and increasing peroxidase diversity. *Mol Biol Evol*.

Schilling, J.S., Kaffenberger, J.T., Held, B.W., Ortiz, R., and Blanchette, R.A. (2020) Using wood rot phenotypes to illuminate the "gray" among decomposer fungi. *Front Microbiol* **11**: 1288.

Sokolski, S., Bernier-Cardou, M., Piché, Y., and Bérubé, J.A. (2007) Black spruce (*Picea mariana*) foliage hosts numerous and potentially endemic fungal endophytes. *Can J For Res* **37**: 1737–1747.

Song, J., and Cui, B.-K. (2017) Phylogeny, divergence time and historical biogeography of *Laetiporus* (Basidiomycota, Polyporales). *BMC Evol Biol* **17**: 102.

Suzuki, H., MacDonald, J., Syed, K., Salamov, A., Hori, C., Aerts, A., *et al*. (2012) Comparative genomics of the white-rot fungi, *Phanerochaete carnosa* and *P. chrysosporium*, to elucidate the genetic basis of the distinct wood types they colonize. *BMC Genomics* **13**: 444.

Syed, K., Shale, K., Pagadala, N.S., and Tuszynski, J. (2014) Systematic identification and evolutionary analysis of catalytically versatile cytochrome p450 monooxygenase families enriched in model basidiomycete fungi. *PLOS ONE* **9**: e86683.

Varga, T., Krizsán, K., Földi, C., Dima, B., Sánchez-García, M., Sánchez-Ramírez, S., *et al*. (2019) Megaphylogeny resolves global patterns of mushroom evolution. *Nat Ecol Evol* **3**: 668–678.

Wattam, A.R., Abraham, D., Dalay, O., Disz, T.L., Driscoll, T., Gabbard, J.L., *et al*. (2014) PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res* **42**: D581–D591.

Wright, E.S. (2015) DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics* **16**: 322.

Wu, B., Gaskell, J., Held, B.W., Toapanta, C., Vuong, T., Ahrendt, S., *et al*. (2018) Substrate-specific differential gene expression and RNA editing in the brown rot fungus *Fomitopsis pinicola*. *Appl Environ Microbiol* **84**: e00991–e00918.

Zhang, J., Silverstein, K.A.T., Castaño, J.D., Figueroa, M., and Schilling, J.S. (2019) Gene regulation shifts shed light on fungal adaption in plant biomass decomposers. *mBio* **10**: e02176–e02119.

Zhou, S., Raouche, S., Grisel, S., Navarro, D., Sigoillot, J.-C., and Herpoël-Gimbert, I. (2015) Solid-state fermentation in multi-well plates to assess pretreatment efficiency of rot fungi on lignocellulose biomass. *J Microbial Biotechnol* **8**: 940–949.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Supporting Information.
Appendix S2. Tables.