Comparative genomics of pyrophilous fungi reveals a link between fire events and developmental genes

Andrei S. Steindorff, Akiko Carver, Sara Calhoun, Krya Stillman, Haowen Liu, Anna Lipzen, Guifen He, Mi Yan, Jasmyn Pangilinan, Kurt LaButti, Vivian Ng, Thomas D. Bruns, Igor V. Grigoriev*

[1]US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA

[2]Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, CA

Corresponding Authors: Thomas D. Bruns <pogon@berkeley.edu> and Igor V. Grigoriev <ivgrigoriev@lbl.gov>

Abstract

Forest fires generate a large amount of carbon that remains resident on the site as dead and partially "pyrolyzed" (i.e., burnt) material that has long residency times and constitutes a significant pool in fire-prone ecosystems. In addition, fire-induced hydrophobic soil layers, caused by condensation of pyrolyzed waxes and lipids, increase post-fire erosion, and can lead to long-term productivity losses. A small set of pyrophilous fungi dominate post-fire soils and are likely to be involved with the degradation of all these compounds, yet almost nothing is currently known about what these fungi do or the metabolic processes they employ**.** In this study, we sequenced and analyzed genomes from fungi isolated after Rim fire near Yosemite National Park in 2013, and showed the enrichment/expansion of CAZymes and families known to be involved in fruiting body initiation when compared to other basidiomycete fungi. We found gene families potentially involved in the degradation of the hydrophobic layer and pyrolyzed organic matter, such

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1462-2920.15273

as hydrophobic surface binding proteins, laccases (AA1_1), xylanases (GH10, GH11), fatty acid desaturases, and tannases. Thus, pyrophilous fungi are important actors to restate the soil's functional capabilities.

Introduction

The frequency of massive, high severity wildfires is increasing in the western U.S. and regions around the world due to fuel accumulations from long-term fire suppression strategies and from the lengthened fire season due to global warming (Westerling *et al.*, 2006). These fires have direct, adverse effects on soil carbon (C) stocks through combustion, but they have indirect, positive effects on soil carbon stocks through the production of pyrolyzed organic matter (PyOM) (González-Pérez *et al.*, 2004). PyOM is chemically heterogeneous material; its composition and structure are determined by both the initial organic matter and the temperature at which it is produced (Whitman *et al.*, 2013; Gul *et al.*, 2015).

Another soil-chemical result of high intensity fires is the production of hydrophobic soils. These soils are thought to result from the deposition of partially pyrolyzed waxes and lipids that are volitalized at high temperatures and condense at slightly lower tempertures in the upper layers of soil (Atanassova *et al.* 2014). The production of hydrophobic soils correlates with higher risk of errosion, and can therefore reduce site productivity for decades (Certini, 2005; Mainwaring *et al.,* 2013). Hydrophobic layers can persist for weeks, months, or even years after a fire.

It seems likely that soil microbes have some role in the degradation of these unique post-fire soil carbon compounds, although this possibility has not yet been explored. While microbes are known to be able to mineralize a small fraction of PyOM, much of the PyOM is highly aromatic and relatively slow to decompose, and thus results in long-term C storage (Debano, 2000; González-Pérez *et al.*, 2004). Nevertheless, the initial post-fire microbial community may determine the rates at which decomposition of PyOM occurs

either by their direct actions or by their capture of more labile carbon compounds that could otherwise prime the decomposition of more recalcitrant forms of PyOM. Microbial decomposition of hydrophobic layers is another possibility that has not yet been explored, but it would be surprising if these energy rich, concentrated layers were not targeted by microbes.

Pyrophilous (fire-loving) fungi are a well-known, ecological guild of fungi that are restricted to post-fire environments and fruit abundantly in months immediately following fire (Petersen 1979, Huges et al 2020). Because of their predicable occurrence after fire, they almost certainly need to be able to interact with post-fire soil chemistry. Here we report on the genomic content of four such pyrophilous that are members of the Agaricomycetes. All four fungi are saprobic (i.e., they live on dead material), yet little is known about what such fungi feed on in these post-fire environments. All of these fungi were found and isolated into culture after the 2013 Rim fire near Yosemite park in California, United States, and they, or closely related species, have been reported from other fires around the world (Bruns *et al.*, 2020).

Although many pyrophilous fungi are members of the Ascomycota, we have focused initially on those that are mushroom-forming fungi (Basidiomycota, Agaricomycetes) because this class of fungi are important wood-degraders involved in global carbon cycling. In the last years, the Agaricomycetes have been of great interest in comparative genomic studies regarding organic matter degradation and as complex multicellular organisms that produce fruiting bodies relevant in agriculture and medicine. The genome sequencing of these fungi showed important components of lignocellulose decomposition and a repertoire of developmental genes involved in fruiting body development in mushroom-forming fungi (Ohm *et al.*, 2010; Sakamoto *et al.*, 2011; Sipos *et al.*, 2017; Almási *et al.*, 2019; Krizsán *et al.*, 2019).

The genomes of *Lyophyllum atratum*, *Coprinellus angulatus*, *Pholiota molesta*, and *Crassisporium funariophilum* were sequenced and reported here for the first time, and we

link the environmental changes caused by fires events to fruiting body development and degradation of pyrolyzed organic matter through a comparative genomics approach. Understanding the role of pyrophilous fungi in post-fire soils is essential, as they are likely to affect the fate of carbon storage in these environments, and may affect the productivity and recovery of burnt soil.

Results

Genome features and phylogeny of pyrophilous genomes

Genomes of pyrophilous species were sequenced using long-read Pacific Biosciences (PacBio) technology, assembled into 41 - 701 scaffolds with length ranging from 37-93 Mbp, and annotated with predicted 13,637 - 25,937 gene models. BUSCO (Seppey *et al.*, 2019) analysis showed 93.9 – 98.7% completeness using the Agaricales database (Table 1).

During the assembly and annotation process, *L. atratum*, *C. angulatus,* and *P. molesta*, secondary scaffolds were detected, suggesting that those genomes are functionally diploid (number of bases in secondary scaffolds > 20%). This is expected in DNA isolated from dikaryon fungi and sequenced using PacBio technology (Kües, 2000; Rhoads and Au, 2015). For all analyses of this work, we moved the secondary scaffolds and their respective annotation into a separate set and used the primary sets for a more accurate comparison. The primary and secondary tracks are available at MycoCosm (https://mycocosm.jgi.doe.gov) (Grigoriev *et al.*, 2014).

We reconstructed the species phylogeny from 1,134 single-copy orthologs (1,604,504 amino acid characters) using 29 Agaricomycetes taxa: 25 Agaricales, three Polyporales, and one Boletales (Fig. 1A). With high bootstrap support (> 90%), the tree shows a topology expected of an Agaricomycetes phylogenetic tree (Almási *et al.*, 2019) and a broad distribution of pyrophilous species across the tree, with *C. funariophilum*

placed outside, but near the family Strophariaceae. *P. molesta* and *C. funariophilum* show similar genomic metrics when compared with other related white-rot fungi. However, *C. angulatus,* on the Psathyrellaceae clade, shows a slightly larger genome size and a higher number of gene models when compared with the other 29 fungi (Figure 1B). Finally, *L. atratum*, the only species from the Lyophyllaceae family, follows a similar pattern of other white-rot fungi (Figure 1B).

As seen in Figure 1C, the core gene set has a low variation (mean ± s.d. = 4,392 ± 369 genes) among the 29 genomes in this study. On the other hand, the number of unique genes varies greatly, ranging from 410 genes (*O. olearius*) to 7,294 (*C. micaceus*) and correlates with the proteome size (R^2 = 0.78). Regarding these unique genes (unassigned to clusters), pyrophilous genomes show numbers (3,393 ± 2,052) on par with other Agaricales genomes (3,490 ± 1,753). Enrichment analysis of PFAM domains on pyrophilous unique genes ($P < 0.05$, Fisher exact test Benjamini Hochberg adjusted P values, abbreviated as FET) showed eight domains enriched, seven of them in *C. angulatus* (Supplemental file 1) which have the largest genome in our dataset. Among these enriched PFAM domains, there are two cytosol aminopeptidases (PF00883 and PF02789), two ribonucleotide reductase (PF02867 and PF00317), two mannosyltransferases (PF16192, PF02366), a MIR domain (PF02815), and a carboxyl transferase (PF02626). Interestingly, 23 domains are more frequent (frequency > 50%, Supplemental file 1) in the unique dataset than in the rest of the genome with a high diversity of functions reflecting the metabolic repertoire present in pyrophilous fungi. To infer functions enriched in pyrophilous fungi, we analyzed all PFAM domain counts across 29 fungal Agaricomycetes genomes, which revealed 48 significantly overrepresented domains and 21 underrepresented (FET, $P \leq 0.05$) (Supplemental file 2).

Comparing the functional annotation normalized by proteome size (Fig. 1D), there is a trend of basal nodes of the tree showing higher counts, except for the small secreted proteins (with signal peptide and < 300 amino acids, abbreviated as SSP), which follows

the opposite trend. Among different broad classes of annotation used in this study, the pyrophilous genomes do not show any clear distinctions. The only feature shared by them is the high number of small secreted proteins (SSP), but this expansion is not exclusive to pyrophilous fungi. SSPs are associated with several eco-physiological features of mushroom-forming fungi linked to developmental stages (Almási *et al.*, 2019; Krizsán *et al.*, 2019). We found two SSP domains – Hydrophobin (PF01185) and Hydrophobic surface binding protein (PF12296) – being enriched (FET, $P \leq 0.05$) and having a strong negative correlation with proteome size in pyrophilous ($R^2 = -0.971$ and $R^2 = -0.861$, respectively) when compared with all genomes ($R^2 = 0.354$ and $R^2 = 0.147$, respectively) (Supplemental files 2 and 3). Other domains, like cutinase (PF01083) and alpha-L-arabinofuranosidase (PF6964) follow a similar pattern (Supplemental file 2).

Few CAZymes are correlated with pyrophilous lifestyle

To access the enrichment of CAZymes between different trophic modes, we performed a Fisher-exact test with FDR correction ($P \leq 0.05$) followed by contrast analysis (Felsenstein, 1985) and phylo-PCA (Revell, 2009) (Figure 2A, 2B, and 2C respectively). In pyrophilous genomes, six CAZy families are enriched: endo-α-1,4-polygalactosaminidase (GH114), gluco/chitooligosaccharide oxidase (AA7_dist), acetyl xylan esterase/cutinase (CE5), α-mannosyltransferase (GT39), lytic polysaccharide monooxygenases (LPMO/AA9), and cellulose-binding module (CMB1). There is a clear imbalance of gene counts for most families, except for AA9 and CBM1, which are enriched in all pyrophilous genomes. The families GH114 and GT39 are enriched mainly in *C. angulatus*, family AA7_dist in *P. molesta* and *L. atratum*, and CE5, mainly in *C. funariophilum* and *L. atratum*. In addition, we used the contrast program (Felsenstein, 1985) to correlate gene counts with nutritional mode (Fig. 2B). We highlighted families with correlation value > |0.4| using this threshold. We found eight families enriched in pyrophilous genomes: all six described earlier, distantly related to plant expansins

(EXPN), laccase (AA1_1), and 4-O-methyl-glucuronoyl methylesterase (CE15). To visualize the enriched CAZy families in different trophic modes, we applied phylogenetically informed principal component analysis (phylo-PCA). Since pyrophilous fungi have other trophic modes (Fig. 1A), they clustered together with similar white-rot and soil/litter decomposers, and no clear separation was found (Fig. 2C).

Enriched clusters involved in the degradation of the hydrophobic zone and fruiting body development

Since the fungi isolated from post-fire soil are polyphyletic and show different trophic modes, the prediction of gene families related to this environment is not straightforward. In order to access orthologous clusters (or gene families) for these 29 fungi proteomes, the program Orthofinder v 2.3.0 with default parameters (Emms and Kelly, 2015) was used to generate the count table used as input for Wilcoxon rank-sum test and contrast analysis for pyrophilous and non-pyrophilous fungi (Fig 3). Comparing the distribution of protein count within each cluster of pyrophilous and other nutritional modes, 94 clusters were found statistically significant (rank-sums p-value ≤ 0.05) and contrast correlation ≥ 0.4 (Fig. 2A and 2B, respectively). Among the clusters enriched on pyrophilous genomes, 70 showed at least one PFAM domain or presence of functional annotation (e.g., signal peptide), and 14 different clusters of SSP were enriched.

These families show a high diversity of functions and some are associated with fungal cell remodeling like α-mannosyltransferase (GT39), endo-1,3-β-glucosidase (GH17), chitinase (GH18), and exo-β-1,3-glucanase (GH55); transcription factors and transcriptional regulators: TFIIF, fungal Zn(2)-Cys(6), homeodomain, protein kinases, F-box proteins; and 14 small secreted protein (SSP) clusters. Comparing this results with developmentally regulated genes in other Agaricales genomes (Muraguchi *et al.*, 2015; Almási *et al.*, 2019; Krizsán *et al.*, 2019), we found that 67% of the enriched clusters are involved in at least one developmental stage (black circles in Fig. 3B and Fig. 3D). To

confirm the specificity of these clusters to the pyrophilous lifestyle, we used phylo-PCA, showing a clear separation of pyrophilous genomes on the first two principal components (Figure 3C).

To perform an in-depth analysis of these developmentally regulated genes, we used proteinortho v6 tool (Lechner *et al.*, 2011) using *-synteny* option to get as close as 1 to 1 gene orthologs with all four pyrophilous genomes and three Agaricales genomes with transcriptome data available (Krizsán *et al.*, 2019) (Supplemental file 4). We decided to use Agaricales genomes since all pyrophilous fungi analyzed here belong to this order. As seen in Fig. 4, about half of the gene space found differentially expressed in *C. cinerea*, *A. ostoyae*, and *S. commune* on developmental stages are shared with at least one pyrophilous genome. Among those genes, we can see that most of them are shared with all four pyrophilous genomes (UpSet diagram on Fig. 4A and 4B), and on average, 17.7% and 19.8%, are expanded in pyrophilous genomes on the conditions fruiting body initiation and fruiting body development (Krizsán *et al.*, 2019), respectively.

The nodes of the phylogenetic tree on Fig 1A show the number of gene families expanded and contracted (family-wide $P < 0.01$) using CAFE program (De Bie *et al.*, 2006). Among the pyrophilous fungi, only *C. angulatus* and *P. molesta* showed a higher proportion of expansions/contractions (3.62 and 1.77, respectively). Looking at the rapidly evolving families (Viterbi $P<0.01$) (Supplemental file 5), all pyrophilous genomes show a proportion of expansions/contractions > 1, totaling 408 expanded and 100 contracted families. In addition, those CAFE expanded families are not only related to fruiting body development but potentially involved with the degradation of the hydrophobic layer – condensed waxes and lipids formed after the fire (Keiluweit *et al.*, 2010; Bruns *et al.*, 2020) – and pyrolyzed biomass that is enriched in aromatics. Genes such as a laccase (AA1_1), xylanases (GH10, GH11), a fatty acid desaturase and a tannase; and detoxification: DNA repair protein XPG_1, hsp20, p450, among others (Supplemental file 5) are examples.

Discussion

In this study, for the first time we sequenced, annotated, and analyzed four genomes from basidiomycetes/Agaricales fungi isolated from post-fire soil, and compared with 25 other fungi with different known nutritional modes in basidiomycetes (Riley *et al.*, 2014; Almási *et al.*, 2019). For this study, we considered pyrophilous as a nutritional mode. However, they show a similar arsenal of enzymes to degrade cellulose when compared with white-rot and soil/litter decomposer fungi (supplemental file 6). The purpose of this separation was to highlight that they are restricted to fruiting on post-fire soil and to explore whether the altered soil chemistry of post-fire soils correlates with changes in gene content. Their position in the phylogenetic tree follows what we expected based on the taxonomy and pyrophilous fungi are broadly distributed across different families rather than being members of a single clade; however, it is essential to note that other ascomycetes fungi are often found in post-fire soil (El-Abyad and Webster, 1968; Bruns *et al.*, 2020) (e.g., *Pyronema spp*.), but we restricted the analysis to basidiomycetes.

We pinpointed 94 orthologous gene families enriched in pyrophilous genomes, 48 overrepresented PFAM domains, and another 508 rapidly evolving families in at least one pyrophilous fungi. Among these gene families and domains, we see several cell-wall modifying enzymes, various small secreted proteins, kinases, and transcription factors (TF). Agaricomycetes-specific phylostrata of developmentally regulated genes are particularly enriched in F-box genes, TFs, and kinases, which indicates an increased rate of origin for these genes in mushroom-forming fungi (Krizsán *et al.*, 2019). We can see this pattern even more pronounced in pyrophilous fungi when compared with other Agaricomycetes (Fig. 3, supplemental file 2). This suggests that developmental processes in pyrophilous fungi can be even more complex than is expected from a regular mushroom-forming fungus.

Hydrophobins are a particularly important class of SSP, playing a role in development and morphogenesis in the majority of the filamentous fungi, influencing spore properties, and fitness-related traits (Bayry *et al.*, 2012; Cai *et al.*, 2020). Interestingly, there is a slight enrichment of this domain on pyrophilous genomes (PF01185, P = 0.045 – Supplemental file 2), and a negative correlation of hydrophobin domain counts compared to total proteome size (pyrophilous R^2 = -0.94; all genomes R^2 = 0.35), suggesting the retention of these domains during its evolutionary history (Supplemental file 3). A similar pattern was found for hydrophobic surface binding protein (PF12296), cutinase (PF01083) and alpha-L-arabinofuranosidase (PF06964). The latter two are CAZymes that might be involved in the enzymatic arsenal to degrade pyrolyzed organic matter.

Carbohydrate-active enzymes is an essential class of proteins in fungi since it is involved in the nutrient acquisition through degradation of biomass and cell wall transformation, among other specific functions (Lombard *et al.*, 2014). Our analysis results reveal that CAZy families involved with the degradation of plant biomass (a mix of with PyOM in the context of this study), and cell wall remodeling were enriched in pyrophilous genomes. Some examples are CAZymes active in the fungal cell wall glucan and chitins, like GH17, GH18, GH55, GH71, and GH88, all developmentally regulated (Krizsán *et al.*, 2019). The degradation of PyOM and the hydrophobic layer are crucial to recovering the soil properties. Enzymes belonging to auxiliary activities (AA) such as laccases (AA1), peroxidases (AA2), and lytic polysaccharide monooxygenases (AA9) are candidates to perform this degradation. Laccases and peroxidases are well known to be active in phenolic compounds and highly recalcitrant environmental pollutants (Novotný *et al.*, 2004; Viswanath *et al.*, 2014). And finally, enzymes active on cellulose/hemicellulose such as GH1, CBM1 containing protein, GH114, and GH10/GH11 were also enriched/expanded on pyrophilous fungi suggesting a role of these families processing the pyrolyzed biomass. Interestingly, classic cellulases families like GH5,

GH6, GH7, GH12, and GH45 were not enriched on pyrophilous fungi, which might be related to the low accessibility to cellulose due to the carbonization of plant biomass on post-fire environments.

Soil is an excellent insulator, and as a result, there is a steep drop in temperature with depth during a fire (Massman, 2012; Bruns *et al.*, 2020). Peak soil temperature is predictable and drops as the log of soil depth. That means that within the span of only 15 cm, temperatures can change from over 500 °C at the surface to pre-fire ambient temperatures below. Pyrophilous fungi are likely to survive as spores or sclerotia at soil depths with intermediate temperatures that kill most of their competitors and allow them to rapidly colonize an open, post-fire niche (Bruns *et al.*, 2020). All four of our sampled pyrophilous fungi were common fruiters after the 2013 Rim fire in California, and they are common throughout the Northern Hemisphere in prescribed fires and wildfire events (Bruns *et al.*, 2020). Interestingly, a significant number of gene families involved in fruiting body development were found enriched on these genomes (Figs 3 and 4).

Environmental conditions play a crucial role in the decision of whether a fruiting body will be formed. The optimal environmental for fruiting body development is often induced after drastically altering the environmental circumstances, and there is no universal set of conditions that leads to fructification in all fungi (Kües and Liu, 2000). Some environmental changes known to trigger this process are drop in temperature, nutrient depletion, shifts in light condition, pH, humidity, salinity, and CO_2 concentration. It seems counterintuitive, but fire events alter all these environmental parameters, and they can match the required triggers for fruiting body initiation, though not necessarily in the short term. Drop in temperature, and CO_2 concentration are good examples: not surprisingly, fire causes a rise in temperature and CO_2 levels, which cools down after it extinguishes. Therefore, these two artificial changes can work as a trigger; for instance, *A. bisporus* requires at least a drop of 5 °C to start mushroom formation (Kües and Liu, 2000).

Another component is the formation of black carbon, commonly known as charcoal (González-Pérez *et al.*, 2004), which can influence mushroom formation as well. Experiments with charcoal suggest that it induces fruiting body development by eliminating inhibitory compounds present in the substrate (De Groot *et al.*, 1998). It is known that pH influences the fruiting body initiation, requiring neutral to slightly acidic pH (6.5 - 7) to trigger this process (Kües and Liu, 2000). The ashes formed burning increases soil pH (Certini, 2005), and considering that Yosemite park is acidic with pH ranging from 4.5 – 6.5 (USDA, 2007), this rise in pH could be another factor inducing mushroom formation. The onset of fruiting body development correlates with the nutritional exhaustion of the growth substrates (Kües and Liu, 2000). The C/N ratios of soil after burning are usually lower than in the original soils, a phenomenon frequently cited in several types of post-fire soils (Almendros and Leal, 1990; González-Pérez *et al.*, 2004), and nutrient depletion is an essential trigger for fruiting body initiation (Sakamoto, 2018; Almási *et al.*, 2019).

It is important to note that pyrophilous fungi fruit are found only in burned habitats and are abundant in the first weeks, months, or years after fire (Bruns *et al.*, 2020, Huges *et al* 2020, Petersen 1970), not necessarily right after the event. At first glance, the causal relationship between fire events and fruiting body development seems circumstantial. However, in this study, we sequenced and analyzed genomes from fungi isolated after Rim fire at Yosemite Park, and showed the enrichment/expansion of families known to be involved in fruiting body initiation when compared to other basidiomycete fungi. We found gene families potentially involved in the degradation of the hydrophobic layer. This is important because the hydrophobic layer can persist for several years and dramatically increases post-fire soil erosion, resulting in losses in productivity and long-term C storage (Mainwaring *et al.*, 2013). Thus, pyrophilous fungi may be important actors to restate the soil's functional capabilities.

Methods

Isolation of fungi on post-fire soil and nucleotide extraction

All four fungi were isolated from tissue explants from inner pileus tissue of young mushrooms. Mushroom caps were ripped open to expose inner, uncontaminated tissue, and small pieces of these tissues were cut out with sterile scalpels, dipped into 30% H_2O_2 for 10 to 15 seconds, and rinsed in sterile water before being pushed into 2% malt extract agar (MA) containing 50 mg/L Streptomycin sulfate and 50 mg/L chloramphenicol. After mycelium grew out into the agar, explants were transferred to 2% MA plates lacking antibiotics and maintained on this media.

All samples were grown in Petri dishes on cellophane over cornmeal yeast malt agar media at 25°C for 3-14 days, depending on the growth rate, and were allowed to experience light and dark cycles. The samples were ground by mortar and pestle in liquid nitrogen and frozen at -80° C.

RNA extraction was performed using the “Fleming method” (Sessitsch et al. 2002; Fleming, Yao, and Sayler 1998) modified by using as a lysis buffer a solution of 50% dH_2O and 50% Buffer Buffer RLT from the Qiagen RNeasy Mini Kit (Qiagen, Germantown, MD, USA) and precipitated by lithium chloride. DNA was extracted using the Qiagen Monarch Genomic DNA Purification Kit (Qiagen, Germantown, MD, USA). DNA and RNA samples were quantified by Qubit fluorospectrometer (Thermo Fisher Scientific, Waltham, MA USA), and the quality of samples was assessed by 260:230 and 260:280 ratios using Nanodrop spectrophotometers (Thermo Fisher Scientific, Waltham, MA USA). DNA and RNA were visualized via gel electrophoresis to confirm integrity. DNA fragments were size-selected using AmPure XP bead cleanup kit (Beckman Coulter, Indianapolis, IN, USA) or Blue Pippin (Sage Science, Inc., Beverly, MA, USA).

Genome sequencing and assembly

The genomes in this study (*Lyophyllum atratum* CBS 144462, *Coprinellus angulatus* CBS 144469, *Pholiota molesta* CBS 144467, and *Crassipodium funariophilum* CBS 144457) were sequenced using PacBio AMPure Bead Size Selection. Unamplified libraries were generated using the Pacific Biosciences standard template preparation protocol for creating >10kb libraries. Five µg of gDNA (10 µg for Blue Pippin protocol) was used to generate each library, and the DNA was sheared using Covaris g-Tubes™ to generate sheared fragments of >10kb in length. The sheared DNA fragments were then prepared using Pacific Biosciences SMRTbell template preparation kit, where the fragments were treated with DNA damage repair, had their ends repaired so that they were blunt-ended, and 5' phosphorylated. Pacific Biosciences hairpin adapters were then ligated to the fragments to create the SMRTbell template for sequencing. The SMRTbell templates were then purified using exonuclease treatments and size-selected using AMPure PB beads. For Blue Pippin protocol, the SMRTbell templates were size-selected using the Sage Science BluePippin instrument with a 4kb or 7kb lower cutoff depending on DNA quality. In both protocols, PacBio Sequencing primer was then annealed to the SMRTbell template library, and sequencing polymerase was bound to them using Sequel Binding kit 2.0. The prepared SMRTbell template libraries were then sequenced on a Pacific Biosystem's Sequel sequencer using v3 sequencing primer, 1M v2 SMRT cells, and Version 2.0 (2.1 for Blue Pippin protocol) sequencing chemistry with 6- and 10-hour sequencing movie run times.

The genomes were assembled with Falcon version 1.8.8 (Chin *et al.*, 2016) to generate an initial assembly. *L. atratum* CBS 144462 passed through an extra step to remove contaminant data using kmer matching with bbtools bbduk.sh [k=25 mm=f mkf=0.05] (https://sourceforge.net/projects/bbtools/). The mitochondria were assembled separately from the Falcon pre-assembled reads (preads) using an in-house tool (assemblemito.sh), used to filter the preads, and polished with Arrow version SMRTLink v5.0.1.9578 (https://github.com/PacificBiosciences/GenomicConsensus). A secondary

Falcon assembly was generated using the mitochondria-filtered preads, improved with finisherSC version 2.1 (Lam *et al.*), and polished with Arrow version SMRTLink v5.0.1.9578.

Transcriptome sequencing and assembly

For transcriptomes, stranded cDNA libraries were generated using the Illumina Truseq Stranded RNA LT kit. mRNA was purified from 1 ug of total RNA using magnetic beads containing poly-T oligos. mRNA was fragmented and reversed transcribed using random hexamers and SSII (Invitrogen) followed by second-strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and eight cycles of PCR. The prepared library was quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified library was then multiplexed with other libraries, and the pool of libraries was then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit v4, and Illumina's cBot instrument to generate a clustered flow cell for sequencing. Sequencing of the flow cell was performed on the Illumina HiSeq 2500 sequencer using HiSeq TruSeq SBS sequencing kits, v4, following a 2x150 bp indexed run recipe. For *Lyophyllum atratum* CBS 144462, the pool of libraries was prepared for sequencing on the Illumina NovaSeq 6000 sequencing platform using NovaSeq XP v1 reagent kits, S4 flow cell, following a 2x150 bp indexed run recipe.

Raw fastq file reads were filtered and trimmed using the JGI QC pipeline resulting in the filtered fastq file (*.filter-RNA.fastq.gz files). Using BBDuk (https://sourceforge.net/projects/bbmap/), raw reads were evaluated for artifact sequence by kmer matching (kmer=25), allowing one mismatch and the detected artifact was trimmed from the 3' end of the reads. RNA spike-in reads, PhiX reads and reads containing any Ns were removed. Quality trimming was performed using the Phred

trimming method set at Q6. Finally, following trimming, reads under the length threshold were removed (minimum length 25 bases or 1/3 of the original read length - whichever is longer). Filtered fastq files were used as input for de novo assembly of RNA contigs. Reads were assembled into consensus sequences using Trinity (ver. 2.3.2) (Grabherr *et al.*, 2011). Trinity partitions the sequence data into many individuals de Bruijn graphs, each representing the transcriptional complexity at a given gene or locus, and then processes each graph independently to extract full-length splicing isoforms and to tease apart transcripts derived from paralogous genes. Trinity was run with the --normalize_reads (In-silico normalization routine) and --jaccard_clip (Minimizing fusion transcripts derived from gene dense genomes) options.

Genome annotation

All genomes were annotated using the JGI Annotation Pipeline (Grigoriev *et al.*, 2014), which combines several gene predictions and annotation methods with transcriptomics data and integrates the annotated genomes into MycoCosm (https://mycocosm.jgi.doe.gov), a Web-based fungal resource for comparative analysis. Completeness of genome annotation was accessed using BUSCO v4.0.6 (Seppey *et al.*, 2019) using the agaricales_odb10 database.

Phylogenetic analysis

Single-copy orthologs were identified in MCL clusters of the 29 Agaricomycetes, including brown rotters, ectomycorrhizal, saprotrophs/litter decomposers/organic matter degraders, white rotters and fungi of uncertain ecology. Protein sequences in the clusters were aligned using MAFFT (Katoh *et al.*, 2019). Ambiguously aligned regions were removed using the -automated1 tag of Trimal (Capella-Gutiérrez *et al.*, 2009). Maximum-likelihood (ML) inference was performed in iqtree v 1.6.9, using 5000 ultrafast bootstrap and partition model. We considered each gene as a partition (ModelFinder to select the

best model) and used the partition-resampling strategy with –sampling option (Nguyen *et al.*).

Identification of orthologous clusters

Orthogroups were identified using OrthoFinder v.2.3.7 (Emms and Kelly, 2015). Two analyses were performed, one to delimit orthogroups across the 29 Agaricomycetes species and the second using proteinortho v6.0.10 (Lechner *et al.*, 2011) with -synteny option to find a near 1 to 1 orthologous shared by pyrophilous genomes and other Agaricales with transcriptome during fruiting body development: *A. ostoaye*, *C. cinerea*, and *S. commune* (Krizsán *et al.*, 2019).

Functional annotation and principal component analysis

Functional annotation was done using INTERPROSCAN v.5.35-74.0. CAZymes were annotated using the CAZy annotation pipeline (Lombard *et al.*, 2014). Gene count heatmaps were created using the 'heatmap.2' function of gplots-R using z-score normalization. Hierarchical clustering with Euclidean distance and averaged-linkage clustering was carried out on copy-number using the 'hclust' function in R. Small secreted proteins (SSPs) were defined as proteins with < 300 amino acids, having a signal peptide, an extracellular localization and no transmembrane domain using TMHMM 2.0 (Krogh *et al.*, 2001). Proteins < 300 amino acids long were subjected to signal peptide prediction through signalP5.0 (Almagro Armenteros *et al.*, 2019) with the 'eukaryotic' option.

Phylogenetic PCA was performed on CAZyme copy numbers using the phyl.pca (Revell, 2009) function from PHYTOOLS. A copy number matrix normalized by proteome size (Table S1) and the ML species tree, were used as input. Independent contrasts were calculated under the Brownian motion model and the parameter mode="cov".

Enrichment and CAFE analysis

The enrichment of PFAM domains was tested using Fisher's exact test and corrected for multiple testing by the Benjamini-Hochberg method in R. To identify orthologous clusters signal in pyrophilous fungi; we used Wilcoxon-test method in R. For both cases, $P \leq 0.05$ was considered significant. To access gene family expansions and contraction analysis, we used CAFE v 4.2.1 (De Bie *et al.*, 2006), a tool that uses a stochastic birth and death process to model the evolution of gene family sizes over a phylogeny. We used the auto-estimation of λ values based on our dataset and a Viterbi $P \leq 0.01$ to be considered a rapidly evolving family (Supplemental file 5).

Data availability

Genome assemblies and annotations are available from available at MycoCosm (https://mycocosm.jgi.doe.gov) (Grigoriev *et al.*, 2014) and has been deposited at DDBJ/ENA/GenBank under the following accessions *L. atratum* CBS 144462 (JACGCN000000000)*, C. angulatus* CBS 144469 (JACGCI000000000)*, C. funariophilum* CBS 144457 (JACGCL000000000), and *P. molesta* CBS 144467 (JACGCK000000000).

Species abbreviations

Hypholoma sublateritium FD-334 SS-4 – **Hypsu**; *Pholiota molesta* CBS 144467 – **Phohi**; *Hebeloma cylindrosporum* h7 – **Hebcy**; *Galerina marginata* CBS 339.88 – **Galma**; *Crassisporium funariophilum* CBS 144457 – **Crafu**; *Laccaria amethystina* LaAM-08-1 – **Lacam**; *Laccaria bicolor* – **Lacbi**; *Crucibulum laeve* CBS 166.37 – **Crula**; *Coprinellus angulatus* CBS 144469 – **Copan**; *Coprinellus micaceus* FP101781 – **Copmi**; *Coprinopsis cinerea* – **Copci**; *Coprinopsis marcescibilis* CBS121175 – **Copma**; *Agaricus bisporus* var. *bisporus* H97 – **Agabi**; *Amanita thiersii* Skay4041 – **Amath**; *Amanita muscaria* Koide BX008 – **Amamu**; *Lyophyllum atratum* CBS 144462 – **Lyoat**; *Pluteus cervinus* NL-1719 – **Pluce**; *Armillaria gallica* 21-2 – **Armga**; *Armillaria solidipes* 28-4 – **Armos**;

Cylindrobasidium torrendii FP15055 – **Cylto**; *Gymnopus luxurians* FD-317 M1 – **Gymlu**; *Omphalotus olearius* – **Ompol**; *Fistulina hepatica* ATCC 64428 – **Fishe**; *Auriculariopsis ampla* NL-1724 – **Auram**; *Pleurotus ostreatus* PC15 – **Pleos**; *Coniophora puteana* RWD-64-598 SS2 – **Conpu**; *Fomitopsis pinicola* FP-58527 SS1 – **Fompi**; *Postia placenta* MAD-698-R-SB12 – **Pospl**; *Phanerochaete chrysosporium* RP-78 – **Phchr**.

Acknowledgments. The work by the US Department of Energy (DOE) Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the US DOE under Contract DE-AC02-05CH11231. The work was funded by the Department of Energy grants DE-SC0016365 and DE- SC0020351.

The authors declare no conflict of interest.

References

Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., et al. (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**: 420–423.

Almási, É., Sahu, N., Krizsán, K., Bálint, B., Kovács, G.M., Kiss, B., et al. (2019) Comparative genomics reveals unique wood-decay strategies and fruiting body development in the Schizophyllaceae. *New Phytol* **224**: 902–915.

Almendros, G. and Leal, J.A. (1990) An evaluation of some oxidative degradation methods of humic substances applied to carbohydrate-derived humic-like polymers. *J Soil Sci* **41**: 51–59.

Atanassova, I.D., S.H. Doerr, and G.L. Mills (2014) Hot-Water-Soluble Organic Compounds Related to Hydrophobicity in Sandy Soils. Soil Carbon, ed. A.E. Hartemink and K. McSweeney: 137-146.

Bayry, J., Aimanianda, V., Guijarro, J.I., Sunde, M., and Latgé, J.-P. (2012) Hydrophobins—Unique Fungal Proteins. *PLoS Pathog* **8**: e1002700.

De Bie, T., Cristianini, N., Demuth, J.P., and Hahn, M.W. (2006) CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22**: 1269–1271.

Bruns, T.D., Chung, J.A., Carver, A.A., and Glassman, S.I. (2020) A simple pyrocosm for studying soil microbial response to fire reveals a rapid, massive response by Pyronema species. *PLoS One* **15**: e0222691.

Cai, F., Gao, R., Zhao, Z., Ding, M., Jiang, S., Yagtu, C., et al. (2020) Evolutionary compromises in fungal fitness: hydrophobins can hinder the adverse dispersal of conidiospores and challenge their survival. *ISME J* 1–15.

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

Certini, G. (2005) Effects of fire on properties of forest soils: A review. *Oecologia* **143**: 1–10.

Chin, C.S., Peluso, P., Sedlazeck, F.J., Nattestad, M., Concepcion, G.T., Clum, A., et al. (2016) Phased diploid genome assembly with single-molecule real-time sequencing. *Nat Methods* **13**: 1050–1054.

Debano, L.F. (2000) The role of fire and soil heating on water repellency in wildland environments: A review. In *Journal of Hydrology*. Elsevier Science B.V., pp. 195–206.

El-Abyad, M.S.H. and Webster, J. (1968) Studies on Pyrophilous Discomycetes. *Trans Br Mycol Soc* **51**: 353–367.

Emms, D.M. and Kelly, S. (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* **16**: 157.

Felsenstein, J. (1985) Phylogenies and the comparative method. *Am Nat* **125**: 1–15.

González-Pérez, J.A., González-Vila, F.J., Almendros, G., and Knicker, H. (2004) The effect of fire on soil organic matter - A review. *Environ Int* **30**: 855–870.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**: 644–652.

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.

De Groot, P.W.J., Visser, J., Van Griensven, L.J.L.D., and Schaap, P.J. (1998) Biochemical and molecular aspects of growth and fruiting of the edible mushroom Agaricus bisporus. *Mycol Res* **102**: 1297–1308.

Gul, S., Whalen, J.K., Thomas, B.W., Sachdeva, V., and Deng, H. (2015) Physico-chemical properties and microbial responses in biochar-amended soils: Mechanisms and future directions. *"Agriculture, Ecosyst Environ* **206**: 46–59.

Hughes, KW, Matheny PB, Miller AN, Petersen RH, Iturriaga TM, Johnson KD, Methven AS, Raudabaugh, DB, Swenie RA, Bruns, TD. 2020. Pyrophilous fungi detected after wildfires in the Great Smoky Mountains National Park expands known species ranges and biodiversity estimates. **Mycologia** (in press)

Katoh, K., Rozewicki, J., and Yamada, K.D. (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. **20**: 1160–1166.

Keiluweit, M., Nico, P.S., Johnson, M., and Kleber, M. (2010) Dynamic molecular structure of plant biomass-derived black carbon (biochar). *Environ Sci Technol* **44**: 1247–1253.

Krizsán, K., Almási, É., Merényi, Z., Sahu, N., Virágh, M., Kószó, T., et al. (2019) Transcriptomic atlas of mushroom development reveals conserved genes behind complex multicellularity in fungi. *Proc Natl Acad Sci* **116**: 7409–7418.

Krogh, A., Larsson, B., Von Heijne, G., and Sonnhammer, E.L.L. (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J Mol Biol* **305**: 567–580.

Kües, U. (2000) Life History and Developmental Processes in the Basidiomycete Coprinus cinereus. *Microbiol Mol Biol Rev* **64**: 316–353.

Kües, U. and Liu, Y. (2000) Fruiting body production in basidiomycetes. *Appl Microbiol Biotechnol* **54**: 141–152.

Lam, K.-K., Labutti, K., Khalak, A., and Tse, D. FinisherSC: a repeat-aware tool for upgrading de novo assembly using long reads.

Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P.F., and Prohaska, S.J. (2011) Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics* **12**: 124.

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**: 490–495.

Mainwaring, K., Hallin, I.L., Douglas, P., Doerr, S.H., and Morley, C.P. (2013) The role of naturally occurring organic compounds in causing soil water repellency. *Eur J Soil Sci* **64**: 667–680.

Massman, W.J. (2012) Modeling soil heating and moisture transport under extreme conditions: Forest fires and slash pile burns. *Water Resour Res* **48**:.

Matheny, P.B., Moreau, P.A., Vizzini, A., Harrower, E., De Haan, A., Contu, M., and Curti, M. (2015) Crassisporium and Romagnesiella: Two new genera of dark-spored Agaricales. *Syst Biodivers* **13**: 28–41.

Matheny, P.B., Swenie, R.A., Miller, A.N., Petersen, R.H., and Hughes, K.W. (2018) Revision of pyrophilous taxa of *Pholiota* described from North America reveals four species— *P. brunnescens, P. castanea, P. highlandensis* , and *P. molesta*. *Mycologia* **110**: 997–1016.

Muraguchi, H., Umezawa, K., Niikura, M., Yoshida, M., Kozaki, T., Ishii, K., et al. (2015) Strand-Specific RNA-Seq Analyses of Fruiting Body Development in Coprinopsis cinerea. *PLoS One* **10**: e0141586.

Nguyen, L.-T., Schmidt, H.A., Von Haeseler, A., and Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies.

Novotný, Č., Svobodová, K., Erbanová, P., Cajthaml, T., Kasinath, A., Lang, E., and Šašek, V. (2004) Ligninolytic fungi in bioremediation: Extracellular enzyme production and degradation rate. In *Soil Biology and Biochemistry*. Pergamon, pp. 1545–1551.

Ohm, R.A., De Jong, J.F., Lugones, L.G., Aerts, A., Kothe, E., Stajich, J.E., et al. (2010) Genome sequence of the model mushroom Schizophyllum commune. *Nat Biotechnol* **28**: 957–963.

Petersen, P.M. 1970. Danish Fireplace Fungi. Dansk Botanisk Arkiv, 27(3): 1-97.

Revell, L.J. (2009) Size-correction and principal components for interspecific comparative studies. *Evolution (N Y)* **63**: 3258–3268.

Rhoads, A. and Au, K.F. (2015) PacBio Sequencing and Its Applications. *Genomics, Proteomics Bioinforma* **13**: 278–289.

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 U S A* **111**: 9923–9928.

Sakamoto, Y. (2018) Influences of environmental factors on fruiting body induction, development and maturation in mushroom-forming fungi. *Fungal Biol Rev* **32**: 236–248.

Sakamoto, Y., Nakade, K., and Konno, N. (2011) Endo-β-1,3-Glucanase GLU1, from the fruiting body of Lentinula edodes, belongs to a new glycoside hydrolase family. *Appl Environ Microbiol* **77**: 8350–8354.

Seppey, M., Manni, M., and Zdobnov, E.M. (2019) BUSCO: Assessing genome assembly and annotation completeness. In *Methods in Molecular Biology*. Humana Press Inc., pp. 227–245.

Sipos, G., Prasanna, A.N., Walter, M.C., O'Connor, E., Bálint, B., Krizsán, K., et al. (2017) Genome expansion and lineage-specific genetic innovations in the forest pathogenic fungi Armillaria. *Nat Ecol Evol* **1**: 1931–1941.

Smith, A. and Hesler, L.R. (1968) THE NORTH AMERICAN SPECIES OF PHOLIOTA, New York.

Viswanath, B., Rajesh, B., Janardhan, A., Kumar, A.P., and Narasimha, G. (2014) Fungal Laccases and Their Applications in Bioremediation.

Westerling, A.L., Hidalgo, H.G., Cayan, D.R., and Swetnam, T.W. (2006) Warming and earlier spring increase Western U.S. forest wildfire activity. *Science (80-)* **313**: 940–943.

Whitman, T., Hanley, K., Enders, A., and Lehmann, J. (2013) Predicting pyrogenic organic matter mineralization from its initial properties and implications for carbon management. *Org Geochem* **64**: 76–83.

Table 1. Summary statistics of pyrophilous genomes. BUSCO represents the percentage of complete BUSCO models using the Agaricales dataset.

Species	Assembly size (Mbp)	# contigs	Coverage	N50	L50 (Mbp)	# genes	BUSCO
Lyophyllum atratum CBS 144462	64.90	171	30.2X	15	1.21	15,046	97.7%
Coprinellus angulatus CBS 144469	93.57	517	63.55X	56	0.48	25,937	96.6%
Pholiota molesta CBS 144467	66.04	701	65.79X	65	0.25	17,825	93.9%
Crassisporium funariophilum CBS 144457	37.12	41	120.54X	7	2.16	13,637	98.7%

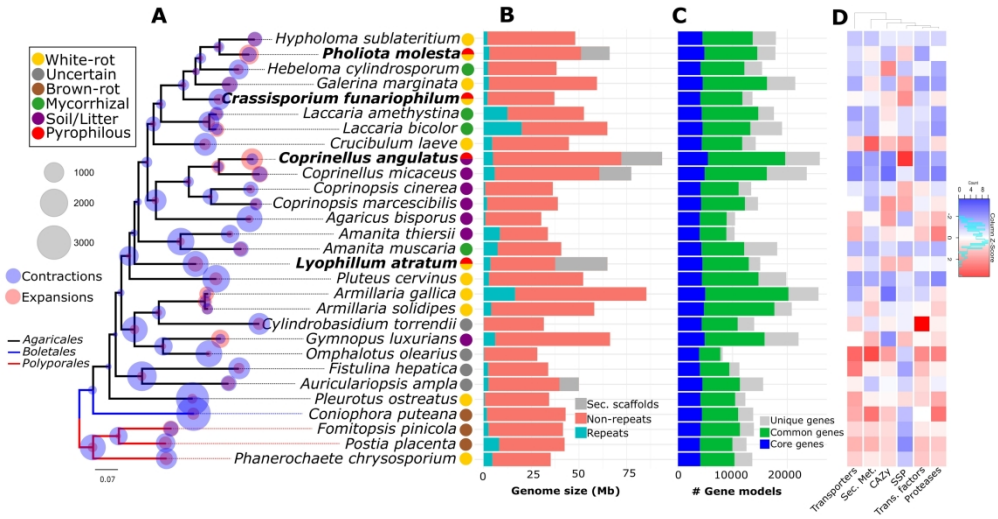

Figure 1. Genome features of 29 Agaricomycetes genomes. A) Maximum likelihood tree showing the phylogenetic relationship based on 1,134 single-copy orthologs of 29 agaricomycetes fungi and their respective nutritional mode. Pyrophilous species are represented in bold. All support values are > 90%. Circles in the nodes represent all family-wide gene expansions (red) and contractions (blue) using CAFE. B) Genome size in Mb, showing the distribution repeats and non-repeats content. C) Gene model counts of each genome (only primary alleles are shown) divided by core genes (present in all genomes), common genes (present in two or more genomes), and unique genes (exclusively found on that genome). D) Column z-score heatmap of functional annotations normalized by total proteome size.

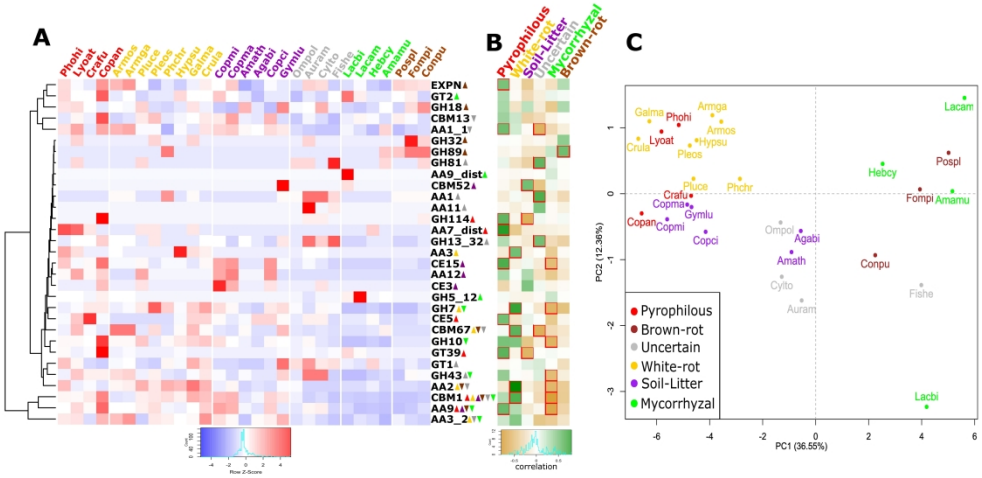

Figure 2. CAZy families enriched on pyrophilous genomes (FDR < 0.05). A) Heat-map (blue-red) of CAZy counts (z-score normalization). The arrows on each CAZy row represents whether that family is overrepresented (up-arrow) or underrepresented (down-arrow). B) Phylogenetic Principal Component Analysis of CAZy families enriched at least one nutritional mode. C) The heatmap (brown-green) shows the correlation of each cluster with trophic mode using the program contrast from phylip package (Felsenstein, 1985). Red squares represent correlation values ≥ |0.4|.

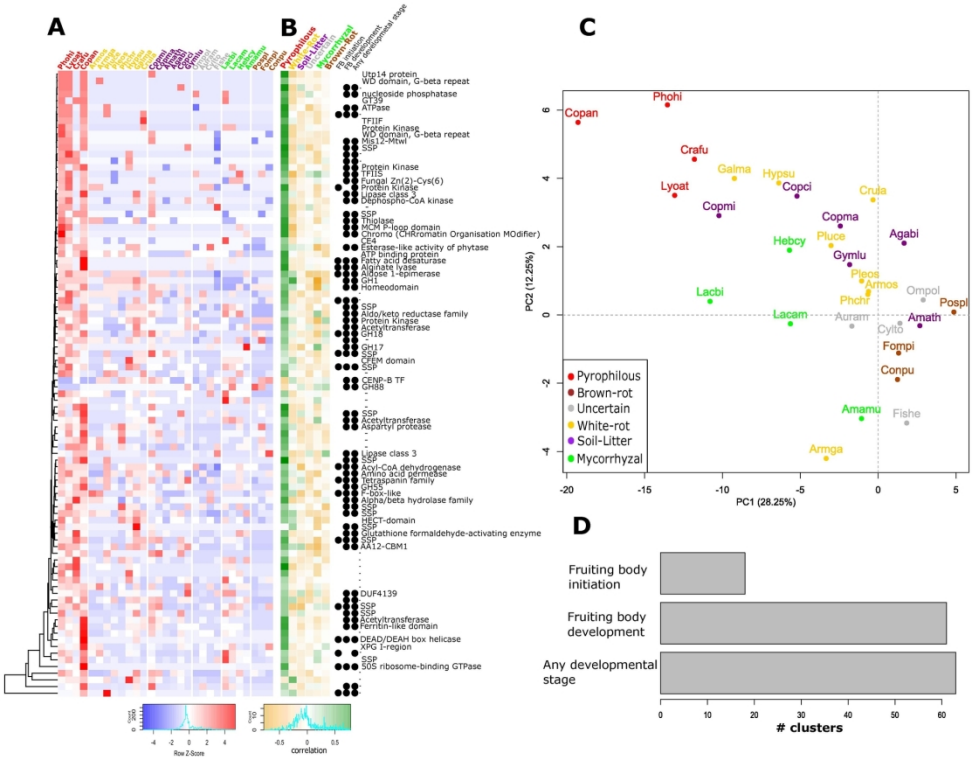

Figure 3. Orthologous clusters enriched (p-value of Wilcoxon test P ≤ 0.05) on pyrophilous fungi. A) The heatmap (blue-red) shows z-score of gene copy number on each cluster detected using Orthofinder counts. B) The heatmap (yellow-green) shows the correlation of each cluster with trophic mode using the program contrast from phylip package (ref). Genes involved in developmental stages are shown as a black circle. C) Phylogenetic PCA of the same clusters showing a separation of pyrophilous species on both principal components. D) The number of clusters enriched and involved in developmental stages by the condition.

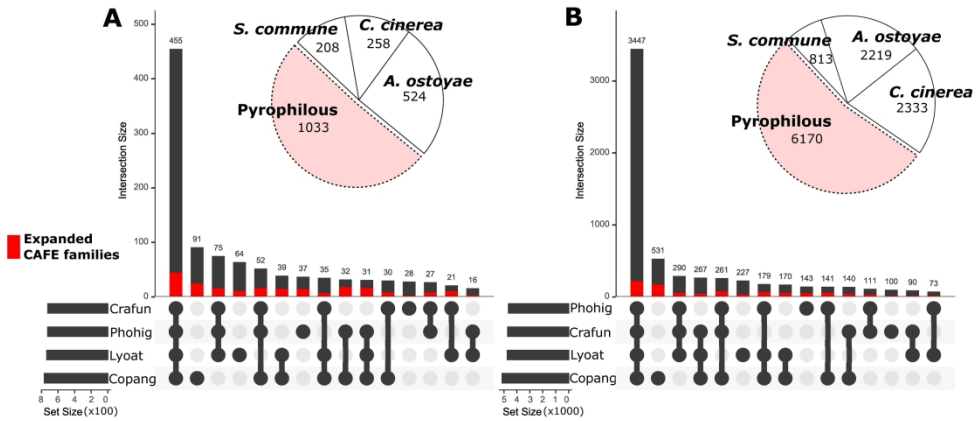

Figure 4. Clusters involved in different developmental stages (Supplemental file 5 from Krisan et al., 2019) in Agaricales genomes and shared with at least one pyrophilous genome (light red on the pie chart). Each upset plot shows the distribution of respective pyrophilous slice. CAFE expanded clusters are shown in red for each intersection. A) Clusters that showed over four-fold upregulation at fruiting body initiation. B) Clusters that show over four-fold expression dynamics (up- or downregulation) across the range of fruiting body stages.