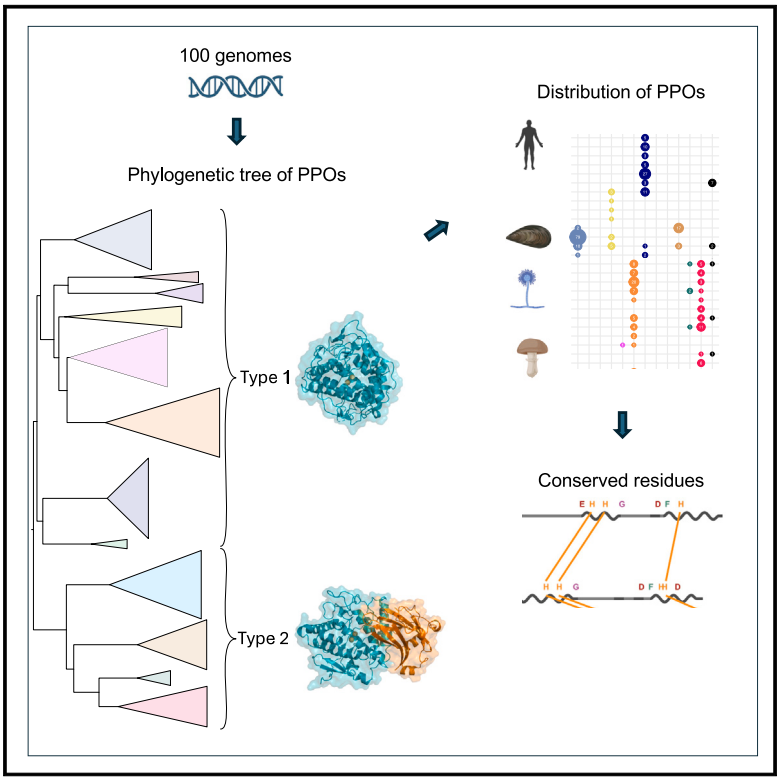
iScience

Article

Classification of polyphenol oxidases shows ancient gene duplication leading to two distinct enzyme types

Graphical abstract



Authors

Ida K.S. Meitil, Caio de O.G. Silva, Anders Gorm Pedersen, Jane W. Agger

Correspondence

jaag@dtu.dk

In brief

Biochemistry; Molecular biology; Evolutionary biology; Phylogenetics

Highlights

- A comprehensive classification of polyphenol oxidases (PPOs) reveals 12 PPO groups
- An ancient gene duplication has led to two distinct PPO types
- Most type 2 PPOs have a shielding domain and a thioether bond in the active site
- The genomic distribution of PPOs is provided for a wide range of species

Meitil et al., 2025, iScience *28*, 111771
February 21, 2025
https://doi.org/10.1016/j.isci.2025.111771

CellPress

iScience

CellPress
OPEN ACCESS

Article

Classification of polyphenol oxidases shows ancient gene duplication leading to two distinct enzyme types

Ida K.S. Meitil,[1] Caio de O.G. Silva,[1] Anders Gorm Pedersen,[2] and Jane W. Agger[1,3,*]
[1]Department of Biotechnology and Biomedicine, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[2]Department of Health Technology, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark
[3]Lead contact
*Correspondence: jaag@dtu.dk
https://doi.org/10.1016/j.isci.2025.111771

SUMMARY

Polyphenol oxidases (PPOs) are coupled binuclear copper proteins that catalyze the oxidation of phenols. New functions of PPOs are continuously being discovered, latest with several fungal *o*-methoxy phenolases, which are active on lignin-derived compounds. Here, we perform a comprehensive phylogenetic analysis of PPOs from a wide taxonomic origin and define 12 PPO groups. We find that a deep gene duplication has led to two distinct PPO types. Type 1 includes PPOs from chordates and molluscs, as well as the fungal *o*-methoxy phenolases. Type 2 includes plant PPOs, molluscan hemocyanins, and fungal tyrosinases. Most of the type 2 proteins have a C-terminal shielding domain and a thioether bond in the copper-binding site. We also find that most ascomycetes contain high numbers of the PPO type 1 that includes the *o*-methoxy phenolases, which may indicate a role in the lignin conversion strategy of these fungi.

INTRODUCTION

Polyphenol oxidases (PPOs) are a large group of enzymes found across all domains of life[1] where they perform a wide range of functions, which are still not fully understood. PPOs are mostly associated with the biosynthesis of melanins, which have several biological roles such as protection from UV radiation and oxidative stress (due to its free radical scavenging properties), skin wound healing in animals, virulence in pathogenic microorganisms, defense against pathogens, and enhancement of microbial resistance to environmental damage.[2,3] Beyond melanin synthesis, PPOs are involved in the synthesis of specialized bioactive metabolites in plants and microorganisms.[4–6] Bacterial PPOs from wetland environments are also regarded as key enzymes in the "latch mechanism" hypothesis, impacting carbon storage in these ecosystems.[7]

PPOs belong to a larger group of proteins called coupled binuclear copper (CBC) proteins (Figure 1A). CBC proteins are defined by having a single type-3 copper binding site, containing two copper ions (Cu_A and Cu_B) each coordinated by three histidine residues in a four-helix bundle. Besides PPOs, the CBC protein class includes hemocyanins, which are proteins without enzymatic activity that transport oxygen in arthropods and molluscs[8] (Figure 1A). Most PPOs, as well as molluscan hemocyanins, contain the Pfam domain PF00264 (common central domain of tyrosinase). Arthropods, on the other hand, contain hemocyanins and prophenoloxidases,[9] which contain the distantly related Pfam domain PF00372 (hemocyanin, copper containing domain), which will not be the focus of this study.

Other examples of important and well-studied copper-containing enzymes are laccases and ascorbate oxidases, which both have four copper atoms in their active sites, two of which are coordinated in a type-3 site, similar to PPOs,[10] and both are prevalent in plants.[11] Single copper-enzymes like lytic polysaccharide monooxygenases (LPMOs)[12] also receive much attention from research due to their biological significance and industrial potential.

PPOs are classically divided into two functional groups; those that possess both mono- and diphenolase activity (tyrosinases, EC 1.14.18.1), and those that possess only diphenolase activity (catechol oxidases, EC 1.10.3.1) (Figure 1A). Monophenolase activity refers to the monooxygenation of a monophenol, generating an *o*-diphenol, while diphenolase activity refers to the two-electron oxidation of a diphenol, generating an *o*-quinone (Figure 1B). Tyrosinases are named so because their canonical substrate is L-tyrosine which they convert into L-dopaquinone in the synthesis of melanin. Catechol oxidases, on the other hand, are named so because they are active on L-DOPA (catechol) but not on L-tyrosine. This nomenclature can, however, be misleading, because some PPOs can be inactive on tyrosine, but still possess monophenolase activity toward other monophenols. Additionally, insufficient characterization can lead to wrong naming.

Several other functions have been discovered, which do not fit in any of the above two groups. The aurone synthase performs monooxygenation and oxidation of chalcones in the anabolic pathway of aurone in plants[4,13] (Figure 1A). These enzymes perform monooxygenation, but are not active on tyrosine. Another

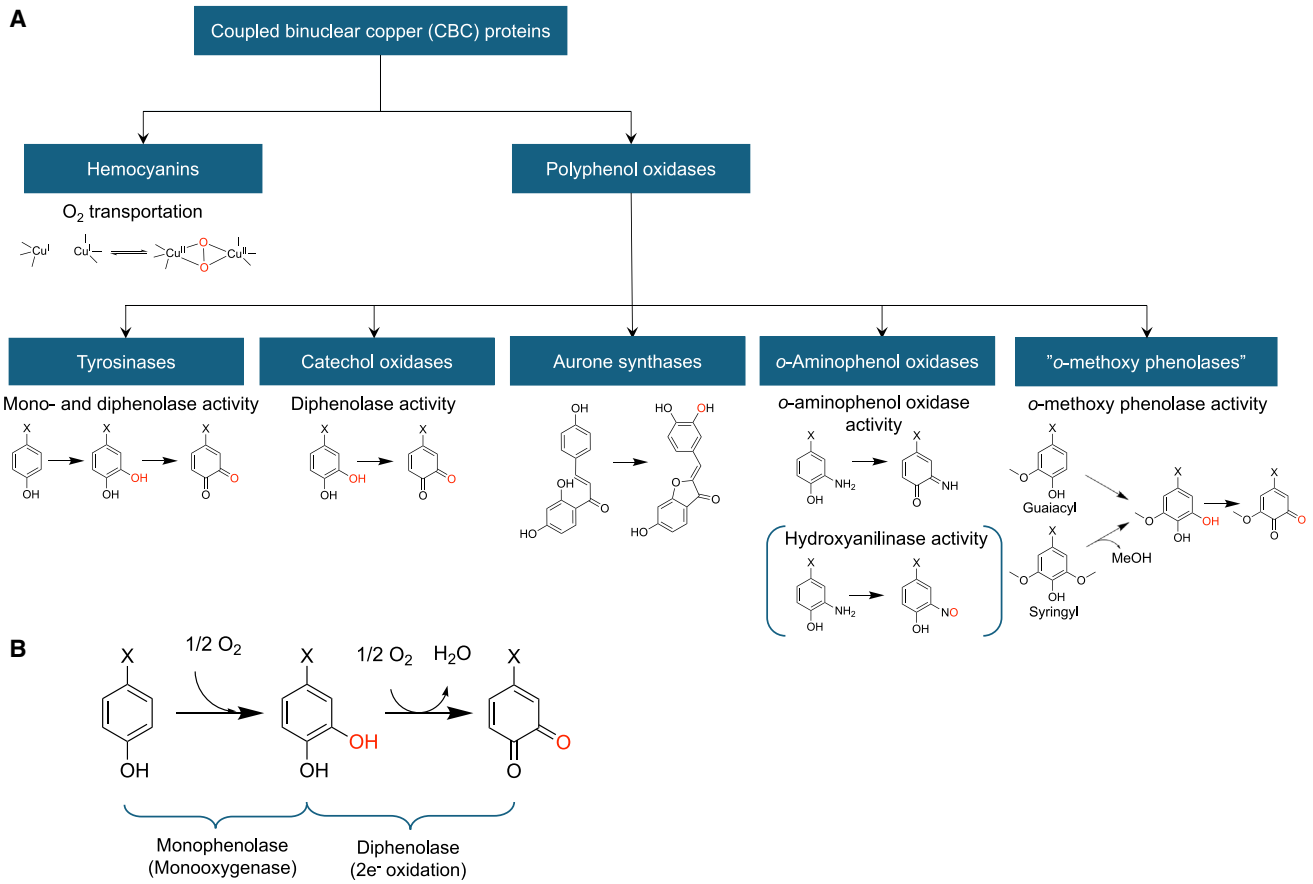


Figure 1. Functions of CBC proteins

(A) Functional groups of CBC proteins.

(B) Mono- and diphenolase activity.

group is the *o*-aminophenol oxidases, which convert *o*-aminophenols into *o*-quinone imines by a two-electron oxidation (EC 1.10.3.4) (Figure 1A). The *o*-aminophenol oxidases also possess "normal" mono- and diphenolase activity. In addition to *o*-aminophenol oxidase activity, hydroxyanilinase activity (C-nitrosation) (Figure 1A) has been reported for two characterized bacterial enzymes, GriF from *Streptomyces griseus*[14] and NspF from *Streptomyces murayamaensis*.[5,15] The hydroxyanilinase activity can be considered a special case of monooxygenation.

In addition, a set of fungal PPOs have been discovered in recent years, which oxidize *ortho*-methoxylated phenols, suggesting potential roles in the valorization of lignin-derived compounds.[16–20] These enzymes funnel lignin-derived guaiacyl (G)- and syringyl (S)-type monophenols into the same product (methoxy-*ortho*-diphenols) (Figure 1A), which suggests that they may be part of the strategy of these organisms to transform lignin-derived compounds into common intermediates before further uptake and metabolization. They have also been shown to boost LPMO-driven degradation of lignocellulose by supplying LPMOs with electrons from the oxidation of G-units.[16,21] These PPOs are not active on tyrosine and do not fit into any of the other three functional groups, thus we refer to them here as "*o*-methoxy phenolases" (Figure 1A). In addition, the reaction on S-type monophenols is different as it is an oxidative demethoxylation (Figure 1A). All of the *o*-methoxy phenolases described so far belong to a group of fungal PPOs termed short PPOs, which differ from the canonical fungal long PPOs (tyrosinases) by lacking the C-terminal shielding domain (PF18132).[22]

In addition to the aforementioned functions, some reports show unusual functionalities for PPOs, such as dopachrome tautomerase for the human tyrosinase related protein 2 (TYRP2), which contains zinc ions in the active site,[23,24] and proteolytic activity for two PPOs, *Md*PPO2 from the "Golden Delicious" apple (*Malus domestica*) and *Ab*PPO4 from button mushroom (*A. bisporus*).[25]

Some PPOs are produced with a C-terminal shielding domain, which blocks the active site. This includes plant catechol oxidases, fungal tyrosinases, and some bacterial tyrosinases, such as *Vs*Tyr from *Verrucomicromium spinosum*.[26] This domain presumably protects against the production of highly reactive quinonoid products inside the cell.[27] The C-terminal domain is cleaved off by a proteolytic enzyme to induce activity.[27]

It is evident that there is a large functional diversity in the PPOs, and that grouping these enzymes based on function is difficult. Furthermore, it is poorly understood how the differences in function relate to sequence and structure. Previous

phylogenetic studies have focused mainly on animal and plant CBC proteins[1,28] with fungal PPOs being underrepresented. Moreover, with the many newly discovered activities, a phylogenetic analysis showing how the different characterized enzymes are related, is needed.

Here, we perform a comprehensive phylogenetic analysis of PPOs from all kingdoms, and define 12 types of PPOs. We investigate how these PPO types relate to activity and substrate preference and how they are distributed in different genomes. We further give an overview of the conserved residues in each type. Finally, we perform a detailed analysis of the fungal short PPOs, including their differences in substrate preference, and their presence in fungi with different lignin-degrading profiles.

RESULTS

Phylogeny of PPOs

An analysis of all the NCBI reference genomes shows that the number of PPO genes varies considerably between different organisms (Figure S1). Some bacterial and archaeal genomes contain PPO genes, but the large majority do not (92% for bacteria and 99% for archaea). Most land plants (Magnoliopsida) contain high numbers of PPOs (up to 22), except for plants of the Brassicaceae family (including the model organism *Arabidopsis thaliana*), which interestingly do not possess any PPOs. In the fungal classes, the ascomycetes such as sordariomycetes and dothideomycetes usually contain high numbers of PPOs, while several other fungal classes do not possess any PPOs, such as saccharomycetes (budding yeast), mucoromycetes, and tremellomycetes. Most animals possess proteins with the PF00264 domain, except for insects and malacostracans (e.g., crabs, lobsters, and shrimp) which, as mentioned previously, contain hemocyanins and prophenoloxidases with the distantly related Pfam domain PF00372.[9] The highest numbers of PPO genes occur in bivalves, such as 81 PPO genes in blue mussel. The mammalian genomes contain between two and nine PPO genes.

We collected a representative set of PPO sequences by selecting one reference proteome from each taxonomic class in which PPOs were present, excluding the classes where there were no proteomes with BUSCO completeness score[29] greater than 80%. This resulted in 100 selected proteomes (Table S1). We then retrieved all the proteins matching the PPO Pfam domain (PF00264, common central domain of tyrosinase) in these genomes (749 proteins). For proteins with several instances of the domain, all the instances of the domain were included, resulting in 856 sequences (Table S2). We then constructed a Bayesian phylogenetic tree of the selected domains, excluding low scoring hits (Figure 2, detailed Figure S2). The tree was rooted using minimum variance rooting.[30] Based on the tree structure, we defined 12 PPO types (**a**-**l** in Figure 2).

Type **a** consists of sequences mainly from molluscs, but also a few sequences from brachiopods, nematodes, and annelids. Type **b** consists of sequences from Cnidaria, type **c** consists of sequences from oomycotes, and types **d** and **e** consist mainly of sequences from zoopagomycotes. Most of the sequences from types **a**-**e** contain signal peptides, and there is to our knowledge no experimentally characterized proteins of these types.

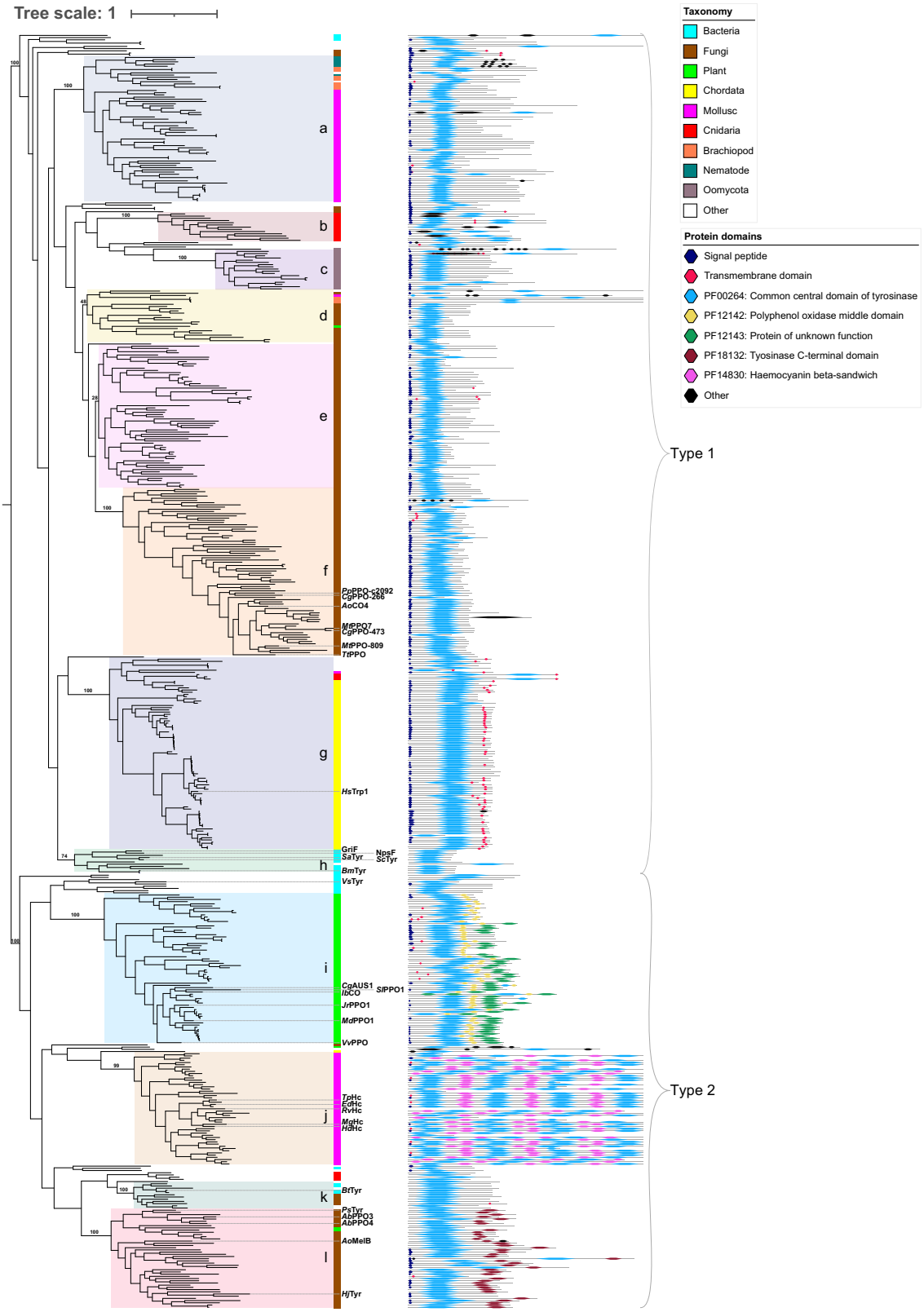
Type **f** encompasses the so-called fungal short PPOs, named so because they do not possess a C-terminal shielding domain as the fungal long PPOs. All the characterized proteins of this type are *o*-methoxy phenolases and include *Ao*CO4 from *Aspergillus oryzae* (pdb: 4J3P),[18,22,31] *Mt*PPO7,[16] *Tt*PPO (pdb: 6Z1S),[17,32] and MtPPO-809[20] from *Myceliophthora thermophila* (syn. *Thermothelomyces thermophilus*), *Pp*PPO-c2092[19] from *Parascedosporium putredinis*, and *Cg*PPO-473 and *Cg*PPO-266 from *Chaetomium globusum*. All characterized members are active on lignin-derived G units, and some are additionally active on lignin-derived S units. The **f** type is present in both ascomycetes, basidiomycetes, and chytridiomycetes, and most of the sequences have a predicted signal peptide.

Type **g** consists of sequences from chordates (vertebrates, tunicates, and cephalochordates). Most of these proteins contain a C-terminal transmembrane domain and a signal peptide. There is one structurally characterized protein of this type; the human tyrosinase-related protein 1 (*Hs*Trp1, pdb:5M8L).[33] Type **h** is a small clade of bacterial sequences that form a sister clade to type **g**. This clade includes the structurally characterized tyrosinases *Sc*Tyr (from *Streptomyces castaneoglobisporus*, pdb:1WX2),[34] *Bm*Tyr (from *Priestia megaterium*, pdb:3NM8),[35] as well as the two *o*-aminophenol oxidases with hydroxyanilinase activity, GriF from *S. griseus*[14] and NspF from *S. murayamaensis*.[5,15] GriF and NspF form a small clade within the **h** clade.

Type **i** consists of sequences from plants only, including several characterized catechol oxidases with elucidated crystal structure; *Md*PPO1 from apple (*Malus domestica*, pdb:6ELS),[36] *Vv*PPO from grape (*Vitis vinifera*, pdb:2P3X),[37] *Ib*CO from sweet potato (*Ipomoea batatas*, pdb:1BT3),[38] *Jr*PPO1 from walnut (*Juglans regia*, pdb:5CE9),[39] *Sl*PPO1 from tomato (*Solanum lycopersicum*, pdb: 6HQI),[40] and the aurone synthase *Cg*AUS from *Coreopsis grandiflora* (pdb: 4Z0Y).[4] In all the sequences of type **i**, the active domain (PF00264) is followed by a PF12142 domain (polyphenol oxidase middle domain) and most of the sequences additionally contain the C-terminal shielding domain PF12143. Some proteins of type **i** harbor a signal peptide, and a few harbor a transmembrane domain close to the N-terminus.

Type **j** consists of sequences from molluscs, including the characterized hemocyanins *Ed*Hc from octopus (*Enteroctopus dofleini*, pdb: 1JS8),[41] *Rv*Hc from the sea snail *Rapana venosa* (pdb: 1LNL),[42] *Tp*Hc from Japanese flying squid (*Todarodes pacificus*, pdb: 6R83),[43] *Hd* from the sea snail *Haliotis diversicolor* (pdb: 3J32),[44] and *Mg*Hc from giant keyhole limpet (*Megathura crenulata*, pdb: 3QJO).[45] These proteins have a "beads-on-a-string" structure with 7–8 repetitive PF00264 domains interspersed with the shielding domain PF14830 (hemocyanin beta-sandwich). These proteins gather in decamers or multidecamers, forming some of the largest known proteins.[46]

Type **k** is a small group containing sequences from various taxonomic origins, including bacteria and fungi. There is one characterized protein of type **k**, *Bt*Tyr from *Burkholderia thailandensis* (pdb:5ZRD).[47] Finally, type **l** consists of the so-called fungal long PPOs, which harbor the C-terminal shielding domain PF18132. This type consists of sequences from ascomycetes and basidiomycetes including the structurally resolved *Ab*PPO3 (pdb:2Y9X)[48]) and *Ab*PPO4 (pdb: 5M6B)[49] from button mushroom (*A. bisporus*),



(legend on next page)

and *Ao*MelB from *Aspergillus oryzae* (pdb:3W6Q).[27] All of the characterized PPOs of this type exhibit canonical tyrosinase activity. Approximately half of the sequences of type **l** contain a signal peptide, indicating different cell compartmentalization and hence different functional roles. One characterized bacterial tyrosinase, *Vs*Tyr (8BBQ) from *Verrucomicromium spinosum*[26] falls outside of the defined types, in a small sister clade to the plant PPOs (type **i**).

Distribution of PPO types in genomes

The presence of multiple PPO genes in many species shows that gene duplication events have occurred. Moreover, the presence of sequences from multiple taxonomic groups (e.g., fungi) in distinct, distant regions of the phylogenetic tree suggests that at least one ancient gene duplication event occurred. A gene duplication in the last universal common ancestor (LUCA) will appear as a split in the gene tree, creating two subtrees. Each subtree retains the topology of the Tree of Life. On the other hand, gene loss event will result in sections of the Tree of Life being absent in the tree.[50] The phylogenetic tree of the PPOs shows a split at the root with sequences from fungi, molluscs, and cnidarians on both sides of the tree, indicating a deep gene duplication event (Figure 2). Since it can be difficult to see the exact position of different species in the PPO tree, we display in Figure 3 the number of each PPO type in each species alongside a phylogenetic tree illustrating the evolutionary relationship between the species.

The chordate genomes almost exclusively contain **g** type PPOs, except for a few **d** type and unclassified PPOs. The number of PPO genes varies from 3 in *Gallus gallus* (red junglefowl) to 27 in *Callorhinchus milli* (Australian ghostshark).

The molluscan genomes contain several different PPO types. All three investigated molluscan genomes contain the **a** type, and *Mytilus edulis* (blue mussel) contains 79 PPOs of this type. Apart from the **a** type PPOs, *Octopus vulgaris* (common octopus) and *M. edulis* also contain **d** type PPOs, and *Candidula unifasciata* (land snail) and *O. vulgaris* also contain the **j** type PPOs (hemocyanins). Furthermore, *O. vulgaris* contains one **g** type PPO and two unclassified PPOs. PPOs in molluscs are known to contribute to shell formation and pigmentation.[52] In mussels, PPOs furthermore have a role in the production of adhesive proteins.[53,54] This is a post-translational modification, where PPOs hydroxylate tyrosine residues in the foot-proteins into DOPA, and in some cases they oxidize it further into dopaquinone.[53] These adhesive proteins are important for the mussel to maintain attachment to their solid support in the sea bed (rocks and stones).[55] The exceptionally high number of **a** type PPOs in mussels may provide the mussels with a wide array of PPOs exhibiting different substrate specificities and functions (monophenolase or diphenolase activity),[53] but it could also be linked to increased gene expression.

The ascomycetes and basidiomycetes contain mainly **f** and **l** type PPOs, while the zoopagomycetes contain mainly **d** and **e** types (detailed further). The two genomes of oomycetes contain high numbers of PPOs (14 and 18 type **c** PPOs). In plants, most of the streptophytes (land plants and green algae except chlorophytes) contain high numbers of the **i** type PPOs; 34 for the liverwort *Marchantia polymorpha* and 14 for poplar (*Populus trichocarpa*). The PPOs of land plants have been extensively reviewed elsewhere, finding that monocots typically contain two to eight PPO genes, and that gymnosperms presumably possess PPOs, but that they are not widely expressed.[28] The green algae *Klebsormidium nitens* has only 2 **i** type PPOs (Figure 2). The examined chlorophytes (green algae) have 1 PPO gene each, of the **l** type and unclassified, which are very different in sequence from the **i** type PPOs (Figure 2). Others have previously noted that the chlorophyte PPOs are not homologous to the streptophyte PPOs and have suggested that they have been acquired via horizontal gene transfer from bacteria.[56] The examined rhodophytes (red algae) contain 1-3 PPO genes of types **k**, as well as unclassified PPOs. All of the examined bacterial genomes contain one PPO gene from either the **b**, **h**, or **k** type or unclassified PPOs.

Based on the topology of the protein tree and the distribution of PPOs across various genomes, we propose that the root of the PPO tree corresponds to a deep gene duplication event, leading to the emergence of two distinct types (Figure 2). Type 2 is distinguished by the presence of C-terminal shielding domains in most of the proteins (in molluscan hemocyanins, these domains alternate with tyrosinase domains). If this hypothesis is correct, we would expect the C-terminal domains in these groups to be homologous. To test this, we used the COMER software[57] to compare the plant C-terminal domain (PF14830) against the entire Pfam database. The search produced three significant hits: the search domain itself, the hemocyanin domain (PF14830; E-value = 6.9E-5), and the fungal domain (PF18132; E-value = 3.7E-4). These results support the idea that the C-terminal shielding domains evolved from a common ancestral gene.

Both type 1 and type 2 PPOs are present in bacteria, animals, and fungi, indicating that the gene duplication occurred in the last universal common ancestor. However, it is notable that very few bacteria possess PPOs, and even fewer archaea. Therefore, it is also possible that the PPO gene originated in an early eukaryote and was later transferred horizontally to some bacteria and archaea.

Differences in conserved residues between the PPO types

We investigated the conserved residues in all of the PPO sequences as well as in the individual families. There are common features in the sequences between all of the types, particularly in

Figure 2. Phylogenetic tree of PPO sequences

The tree contains protein sequences matching PF00264 from one NCBI reference genome from each taxonomic class (E-value<1e-20). For proteins with several instances of the domain, all the domains were included. The phylogenetic tree is constructed with MrBayes using only the PF00264 domain, and the tree is rooted using minimum variance rooting. Characterized proteins (with priority to proteins having an elucidated crystal structure) are shown on the tree with their respective protein names. The domain architecture of the proteins is shown on the right with a cut-off at 1500 residues. The 12 PPO types are marked **a**-**l**. The taxonomic origin of the proteins is marked by color-coding according to the legend in the figure. Posterior probabilities are shown as support values on the nodes for the major clades. A detailed tree with accession numbers and species is shown in Figure S2.

(legend on next page)

the two copper binding sites (Figures 4A and S3). The Cu_A binding site shows the most conservation with a fully conserved Phe, Arg, and Glu residue besides the three copper-coordinating histidines. Yet, the distance between the 1st and the 2nd His residue varies considerably. In types **a**-**f** (PPOs from molluscs, cnidaria, oomycetes, and zoopagomycetes, and the fungal short PPOs), the distance is only 7–9 residues, whereas it is 16–24 residues in types **g**-**l** (PPOs from chordates, bacteria and plants, molluscan hemocyanins, and fungal long PPOs).

One feature that is common between all the structures of the type 2 proteins is that they contain a thioether bond, between a Cys residue and the 2nd His residue of the Cu_A binding motif. This peculiar bond is known to stabilize the 2nd His residue, which is located in a flexible loop.[58] The location of the Cys residue involved in the bond, however, varies. In types **j**, **k**, and **l** (mollusc hemocyanins, bacterial PPOs, and fungal long PPOs), the Cys residue is located two positions before the 2nd His Cu_A (Figures 4A and S3), while in the plant PPOs (type **i**), it is located four residues after the 1st His of Cu_A and is only 76% conserved and therefore not depicted in Figure 4A.

The Cu_B binding site shows less conservation with only the copper-coordinating histidines being fully conserved. All types except for **b**, additionally have a conserved Phe located four residues before the 3rd His of Cu_B, and most types have a conserved Asp located four residues after the 3rd His of Cu_B. Many of the types (**c**-**f** and **l**) also contain a seventh conserved His residue located immediately before the third His of Cu_B, which has been shown to assist in Cu_B-coordination.[59]

A residue that is known to be important for activity is the so-called waterkeeper residue, a Glu located 3–4 residues before the 1st Cu_B-coordinating His (Glu-256 in *Ab*PPO3, Figure 4B). This residue has been shown essential for monophenolase activity in several PPOs and is proposed to activate a water molecule, which deprotonates the monophenolic substrate.[59] The Glu residue is conserved in all types except for **b**, **c**, **e**, and **f**. In the fungal short PPOs, the residue at this position varies, but is often Gln or Glu (Gln-273 in *Ao*CO4, Figure 4B). Most of the fungal short sequences have an insertion between this residue and the 1st Cu_B-coordinating His, making the distance between them 10 residues.

Another position that has been proposed to be important to distinguish PPO activity, is the 1st activity controller, located directly after the 1st His of CuB.[60] The 1st activity controller is a conserved Asn in type **j** (molluscan hemocyanin) and **g** (chordate PPOs). An Asn at this position is known to allow for tyrosinase activity.[61] The fungal long PPOs, mostly contain Asn or Asp in this position (Asn-260 in *Ab*PPO3, Figure 4B), which is proposed to control tyrosinase activity.[60] The fungal short PPOs, on the other hand, mostly have Gly or Ala at this position (Gly-285 in *Ao*CO4).

We also observe that type **g** and **h** (chordate and bacterial PPOs, respectively) share a conserved motif in a helix before the Cu_B-binding site; $FRNX_2EG$. Finally, we observe that all PPO types have a conserved Arg located between the two copper binding sites in an area with low conservation. This Arg is located slightly outside of the active site (Figures 4B and 4C), and its role has to our knowledge not been examined.

Distribution of PPOs in fungal genomes

Next, we examined the occurrence of the different PPO types in a larger selection of fungal genomes. Instead of one proteome per taxonomic class, we selected one proteome for each fungal taxonomic order (66 proteomes, Table S3). We additionally added a number of fungal species that are known to grow on living plants or decaying plant biomass (Table S5). From these proteomes, we retrieved the proteins matching the PF00264 Pfam domain (769 proteins, Tables S4 and S6). In order to determine the types of the PPO sequences, we used phylogenetic placement. This method places new sequences on the phylogenetic tree (Figure 2), rather than building a new phylogenetic tree.

Almost all the included ascomycetes contain both short and long PPOs (type **l** and **f**), and many of them contain high numbers of the **f** type PPOs (Figure 5). For example, the two plant pathogens *Alternaria alternata* and *Colletotrichum gloeosporioides* contain 20 and 30 short PPOs, respectively.

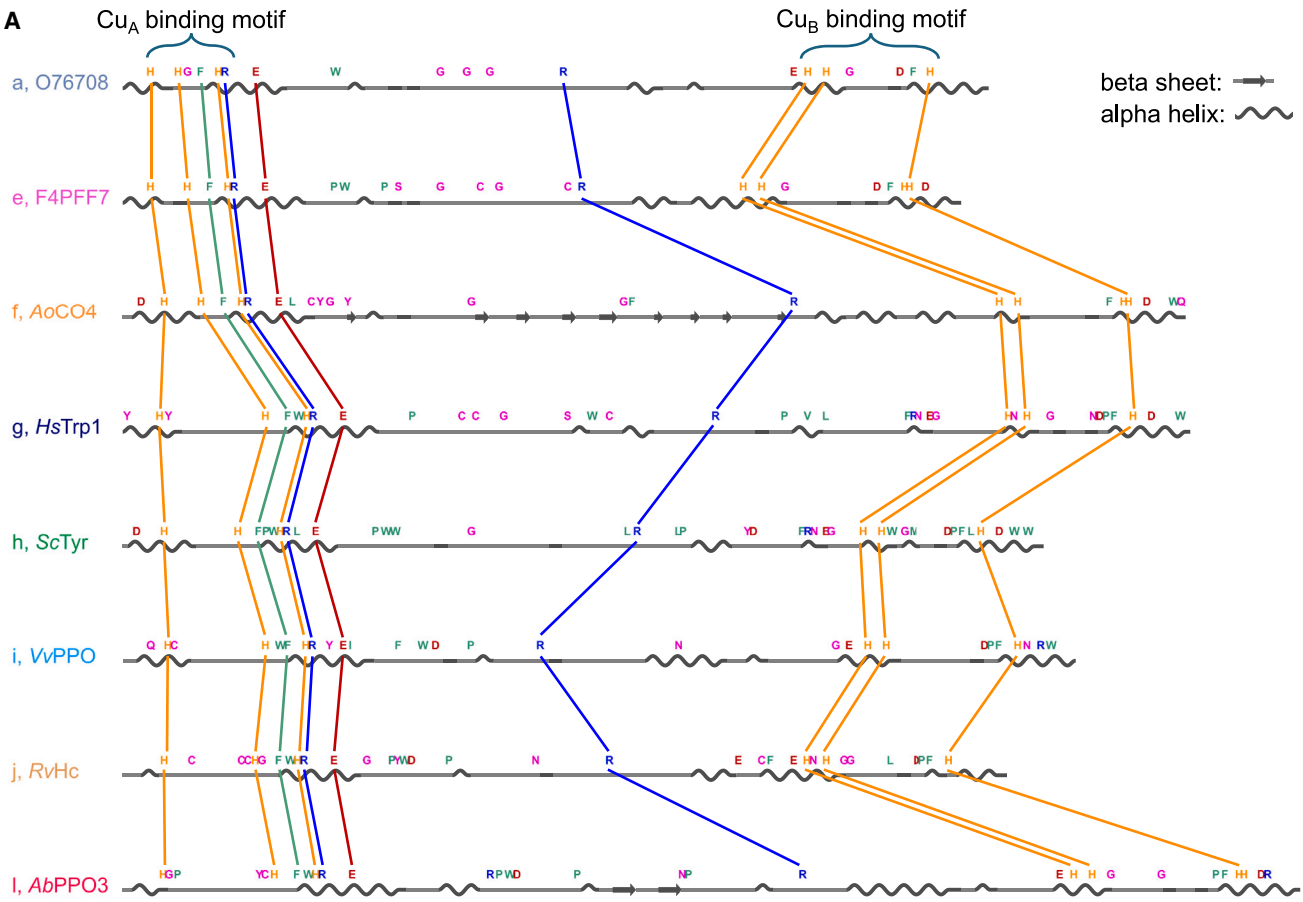
In the basidiomycetes, on the other hand, several species do not have any short PPOs, e.g., the lignin-degrading fungi *A. bisporus* (button mushroom)[62] and *Phanerochaete carnosa*.[63] However, others have high numbers of short PPOs, e.g., the plant pathogen *Rhizoctonia solani* which has 26 short PPOs, and the good lignin degrader *Coprinopsis cinerea*[64] which has 18 short PPOs. There are no characterized fungal short PPOs from basidiomycetes and hence it is still not known whether they are *o*-methoxy phenolases like the characterized ones from ascomycetes.

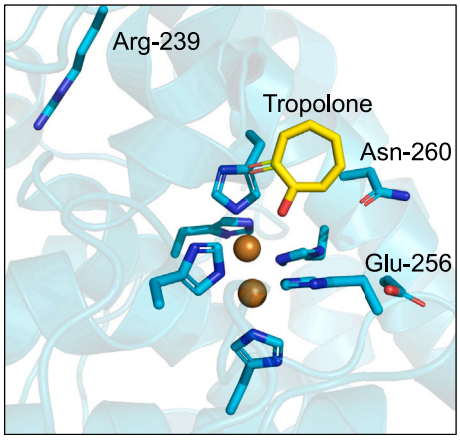
While almost all agaricomycetes contain long PPOs, the rest of the examined basidiomycetes (exobasidiomycetes, pucciniomycetes, ustilaginomycetes, and wallemiomycetes) do not contain any long PPOs (Figure 5). One of these, corn smut (*Ustilago maydis*), has been shown to have a different melanin synthesis pathway, which does not involve L-DOPA.[65] This is congruent with its lack of long PPOs.

The examined chytridiomycetes contain **f** type (fungal short PPOs), **e** type, and unclassified PPOs, but no **l** type PPOs (fungal long) (Figure 5). The mucoromycetes contain only **l** type PPOs (fungal long PPOs). The zoopagomycetes contain very high numbers of PPOs (up to 48), of the **d** and **e** type (Figure 5). These fungi are obligate parasites of other zygomycete fungi and of microscopic soil animals such as amebae,[66] and it is tempting to speculate that their high numbers of PPOs are related to this lifestyle. While the **d** and **e** type PPOs are similar in sequence to the fungal short PPOs (type **f**), their function is still unknown, as there are no characterized PPOs of this type.

Figure 3. Distribution of PPO types in genomes

On the left, the time tree is displayed showing the evolutionary relationship between the different species in the tree.[51] On the right, the number of genes of each type of PPO is shown (protein type-naming according to Figure 2). Unclassified PPOs are shown as "u". The size of the circles is proportional to the number of PPOs of each type. Proteins with several instances of the PF00264 domain are counted as one.





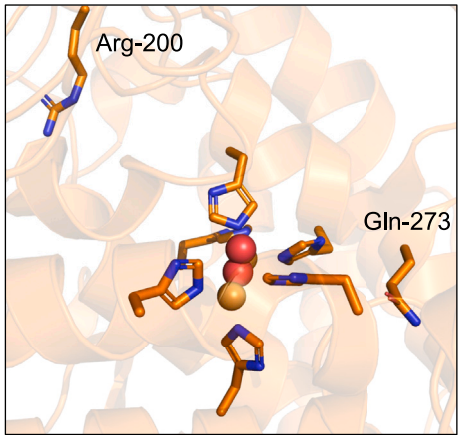


Figure 4. Conserved residues in the PPO types

(A) Conserved residues in selected PPO types are shown on representative sequences. Only the PF00264 domain is shown. The secondary structure of the sequences was retrieved from the crystal structures for types **f**, **g**, **h**, **i**, **j**, and **l** and from AlphaFold models for types **a** and **e**. Lines are shown between residues that are conserved across all the PPO types. An aligned version including all PPO types and residue numbers is shown in Figure S3.

(B) Structure of a fungal long PPO (*Ab*PPO3, pdb: 2Y9X) with the inhibitor tropolone bound showing the conserved histidines, Arg, waterkeeper and 1st activity controller.

(C) Structure of a fungal short PPO (*Ao*CO4, pdb: 4J3P) showing the conserved histidines, Arg and "waterkeeper". The 1st activity controller is a Gly and is therefore not shown. The copper and oxygen atoms are shown in brown and red respectively.

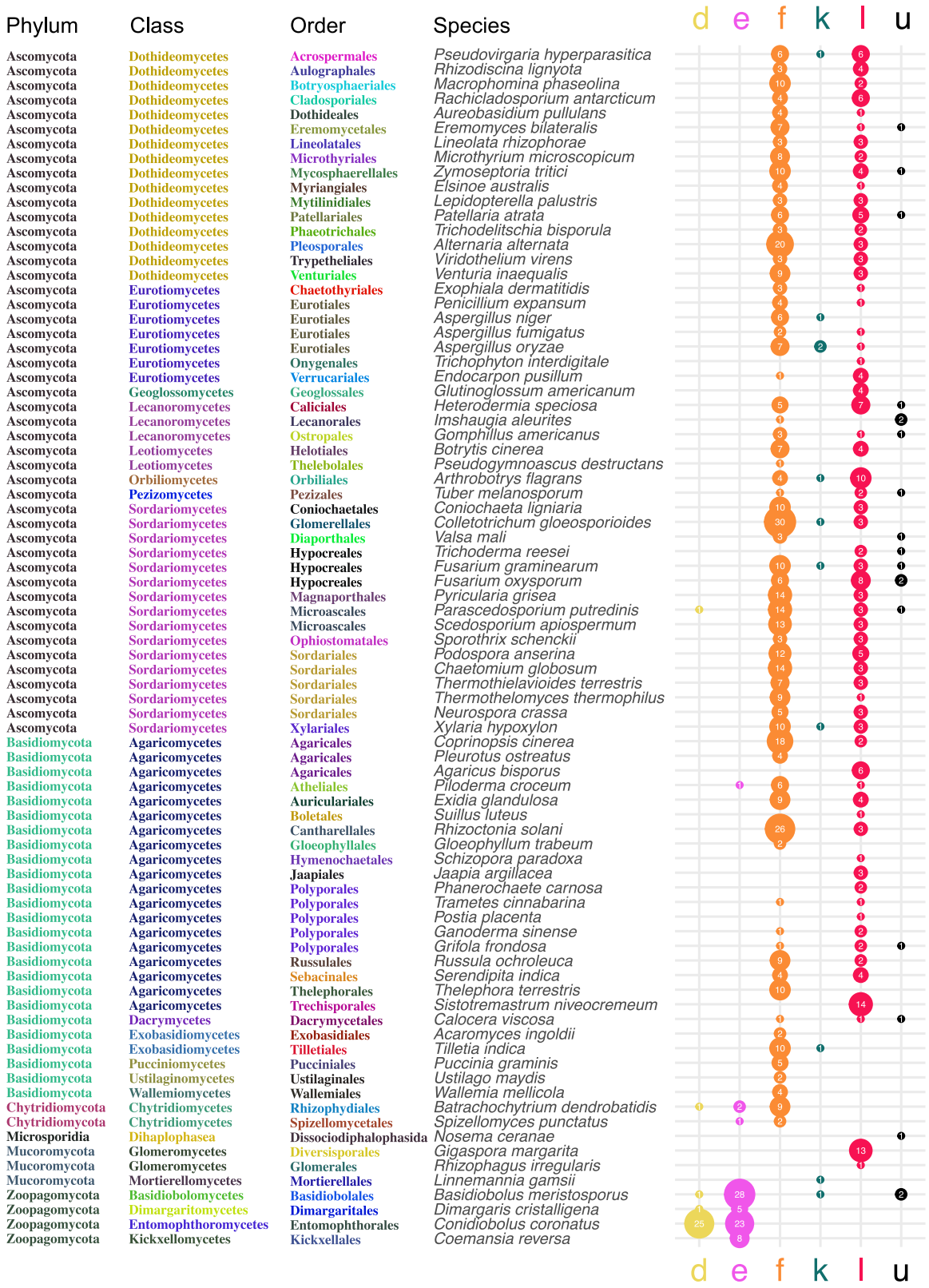


Figure 5. Distribution of PPO types in fungal genomes

The number of the different PPO types (Figure 2) in the selected fungal genomes is shown on the right. Phylum, class and order is shown on the left.

Specificity of fungal PPOs on lignin-derived monophenols

Next, we investigated how the differences in activity of the *o*-methoxy phenolases (short fungal PPOs) relate to differences in sequence. As mentioned previously, the *o*-methoxy phenolases are active on G-type monophenols, which have a methoxy group in the ortho-position relative to the hydroxyl group (Figure 1A). Additionally, activity on S-type monophenols have been shown for three PPOs. The reaction on S-type monophenols is different, because the carbon attacked by the PPO is protected by a methoxy group, and the *ortho*-methoxy group is released as methanol[20] (Figure 1A).

To investigate this difference in specificity, we constructed a Bayesian phylogenetic tree of the **f** type PPOs from the extended fungal dataset (Table S4) (Figure 6). The tree shows that the S-type active *Cg*PP O-473 and *Mt*PPO7 are located very closely in the tree. However, the third S-type active PPO, *Cg*PPO -266, is located in another branch close to *Pp*PPO-c2092, which is not active on S-type monophenols. The sequence identity between *Cg*PPO-266 and *Pp*PPO-c2092 is 39.9% over 416 residues. Thus, the tree does not show a clear division of the PPOs that are active on both G- and S-type monophenols and those that are active only on G-type monophenols. This suggests that the ability to oxidize S-type monophenols has evolved multiple times. The difference in substrate preference may be related to subtle differences in the substrate binding pocket, and increasing the number of characterized PPOs of type **f** together with structural data may help to understand these differences.

DISCUSSION

We have presented a comprehensive phylogenetic analysis of PPOs, showing 12 PPO types (**a**-**l**). The topology of the phylogenetic tree indicates that the root of the tree corresponds to a deep gene duplication, which has likely occurred in LUCA. This duplication has led to two overall PPO types; type 1, which includes sequences from chordates, molluscs, cnidarians, and fungi, including *o*-methoxy phenolases (types **a**-**h**), and type 2 which includes PPOs from plants, molluscan hemocyanins, and fungal tyrosinases (types **i**-**l**). The type 2 proteins are characterized by most proteins having a thioether bond in the copper-binding site and a C-terminal shielding domain. The shielding domains from plants, hemocyanins, and fungi show distant similarity, indicating that they have evolved from the same ancestral gene.

There is, however, some uncertainty regarding the rooting of the PPO tree. The tree was rooted using minimum variance rooting, which does not always result in the correct rooting. However, the distinct differences between the two resulting gene families—specifically the C-terminal shielding domain and the thioether bond—further support this rooting. We also attempted outgroup rooting using the Pfam module of arthropod hemocyanins (PF00372) as an outgroup, as it is distantly related to the PPO Pfam module (PF00264). However, the outgroup did not provide a definitive root, likely due to the low similarity between the two Pfam modules.

Since both type 1 and type 2 PPOs are present in both bacteria, fungi, and animals, the gene duplication has likely occurred in LUCA. However, it cannot be excluded that it has occurred in an early eukaryote with subsequent horizontal gene transfer to bacteria. Especially, since very few bacteria contain PPOs (Figure S1).

The classification system presented here can be used to indicate the function of new sequences, as the PPO types often preserve the overall functions. Specifically, the *o*-methoxy phenolases are all classified as type **f**, the fungal tyrosinases as type **l**, the plant catechol oxidases as type **i**, and the *o*-aminophenol oxidases with hydroxyanilinase (C-nitrosation) activity as type **h**. However, more functionally characterized PPOs are needed, before we can make more precise functional predictions.

While some areas of the phylogenetic tree contain many characterized enzymes, others remain largely unexplored. Our analysis can help guide the selection of targets for enzyme characterization, which may reveal new and important functionalities.

It is evident that the functional diversity of PPOs has not yet been fully explored. The high abundance of PPOs in certain organisms suggests different functional roles. These differences may be associated with substrate specificity, reaction speed, and pH preference, as demonstrated in a recent report for the six PPOs of button mushroom (*A. bisporus*).[60]

Our analysis of fungal genomes showed that the fungal short PPOs are more prevalent in ascomycetes than in basidiomycetes. All currently characterized short PPOs exhibit *o*-methoxy phenolase activity, suggesting that these enzymes may play a role in the lignin utilization, especially in ascomycetes, likely through the oxidation of low molecular weight lignin-derived phenolic compounds. These enzymes convert both guaiacyl and syringyl lignin-derived units into *o*-methoxy-catechols,[20] and this may be part of a strategy to transform lignin-derived compounds into common intermediates before further uptake and use.[67] Notably, the formation of catechols is a prerequisite for ring-opening reactions by dioxygenases and incorporation of lignin derivatives into central carbon metabolism. We also see some examples of basidiomycetes with high numbers of short PPOs. However, no short PPOs from basidiomycetes have been characterized, and they are located in a different region of the phylogenetic tree than the characterized short PPOs (Figure 6), and thus their functions are still unknown.

The characterized *o*-methoxy phenolases exhibit different substrate specificities; some are active on S-type monophenols, performing oxidative demethoxylation, while others are only active on G-type monophenols.[20] In our phylogenetic tree of the short PPOs, the S-type active PPOs did not cluster separately from those that are not active on S-type monophenols. Therefore, the features that confer S-type activity are still not known, and may involve subtle differences that enable the accommodation and stabilization of the extra *o*-methoxy group present in S-units. To understand these differences, further characterization of *o*-methoxy phenolases, along with structural data, is needed.

Limitations of the study

We did not include the Pfam domain PF00372 (hemocyanin, copper containing domain) in our analysis. This domain covers mainly the arthropod hemocyanins and prophenoloxidases. Although this Pfam domain is related to the PPO Pfam

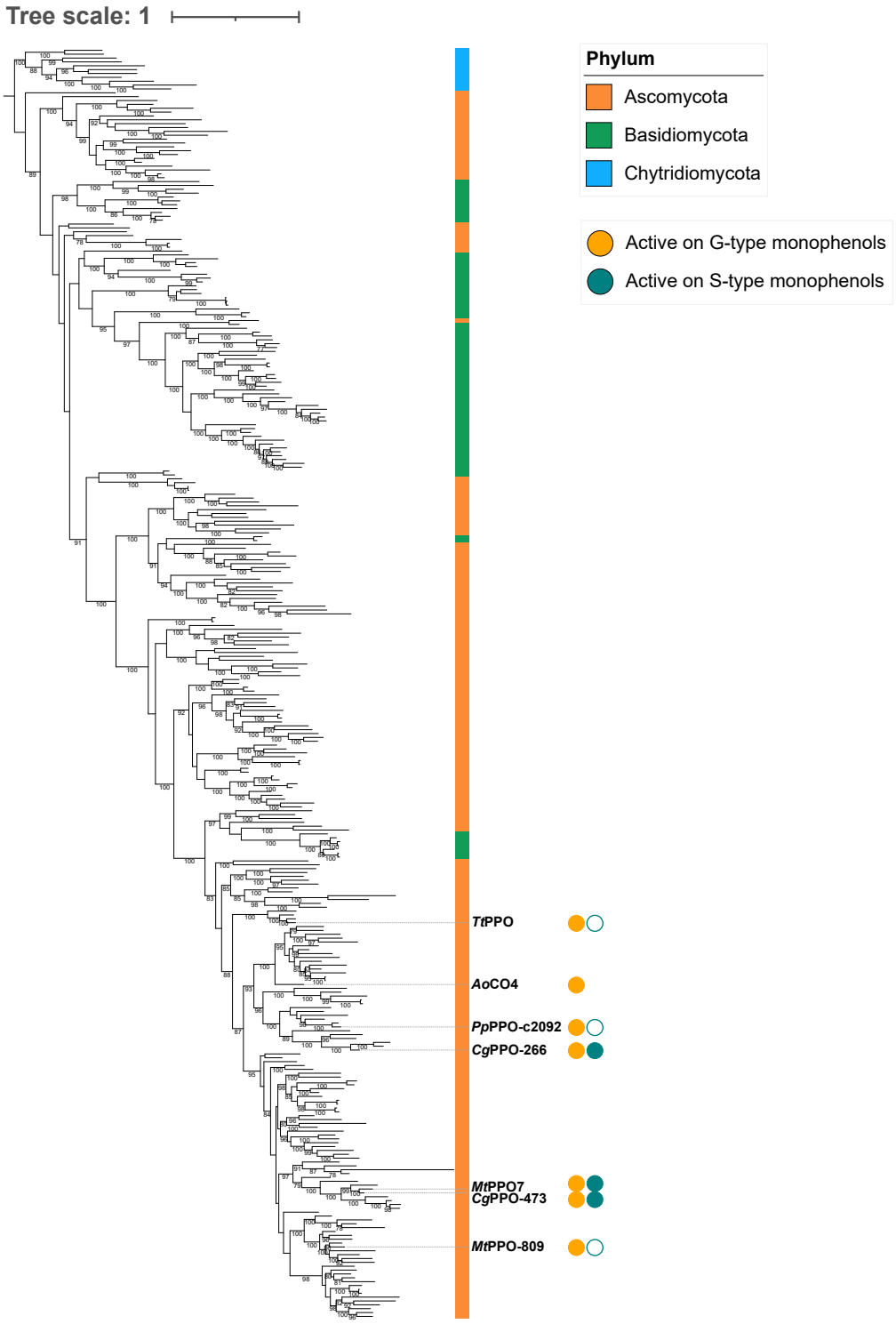


Figure 6. Phylogenetic tree of fungal short PPOs (type f)

One genome was selected from each taxonomic order of fungi, and the proteins matching PF00264 were retrieved. A Bayesian phylogenetic tree of the **f** type PPOs (fungal short PPOs) was constructed. The characterized proteins are shown on the left along with their substrate specificity. An empty circle is shown for PPOs that were tested on the given substrate and were not active on the substrate, and no circle is shown for PPOs that have not been tested on the given substrate. Posterior probabilities are shown as support values on the node, with only values > 50% shown.

PF00264, they differ substantially in sequence, and it did not make sense to include them in the same phylogenetic tree.

RESOURCE AVAILABILITY

Lead contact

Requests for further information and resources should be directed to and will be fulfilled by the lead contact, Jane W. Agger (jaag@dtu.dk).

Materials availability

This study did not generate new materials.

Data and code availability

- This article does not report any original data.
- This article does not report any original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We thank Albert Þór Þórhallsson for his assistance in setting up MrBayes on a high performance cluster. This work is funded by Novo Nordisk Foundation via grant awards NNF22OC0074634 and NNF23OC0086414.

AUTHOR CONTRIBUTIONS

I.K.S.M. performed data acquisition, sequence analysis, and interpretation; C.d.O.G.S. supervised and interpreted data; A.G.P. supervised and interpreted data; J.W.A. supervised and interpreted data. All co-authors contributed to the conceptualization of the study. The manuscript was written by I.K.S.M. with the help of all co-authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
  - PPO sequence dataset
  - Phylogenetic tree
  - Phylogenetic placement
  - Genome distribution plot
  - Similarity of C-terminal domains
  - Conserved residues
  - Extended fungal PPO sequence dataset
  - Genome distribution plot for extended fungal dataset
  - Detailed phylogenetic tree of fungal short PPOs

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2025.111771.

Received: October 3, 2024
Revised: November 18, 2024
Accepted: January 6, 2025
Published: January 10, 2025

REFERENCES

1. Aguilera, F., McDougall, C., and Degnan, B.M. (2013). Origin, evolution and classification of type-3 copper proteins: lineage-specific gene expansions and losses across the Metazoa. BMC Evol. Biol. *13*, 96. https://doi.org/10.1186/1471-2148-13-96.
2. Nosanchuk, J.D., and Casadevall, A. (2006). Impact of Melanin on Microbial Virulence and Clinical Resistance to Antimicrobial Compounds. Antimicrob. Agents Chemother. *50*, 3519–3528. https://doi.org/10.1128/AAC.00545-06.
3. Brenner, M., and Hearing, V.J. (2008). The Protective Role of Melanin Against UV Damage in Human Skin. Photochem. Photobiol. *84*, 539–549. https://doi.org/10.1111/j.1751-1097.2007.00226.x.
4. Molitor, C., Mauracher, S.G., and Rompel, A. (2016). Aurone synthase is a catechol oxidase with hydroxylase activity and provides insights into the mechanism of plant polyphenol oxidases. Proc. Natl. Acad. Sci. *113*, E1806–E1815. https://doi.org/10.1073/pnas.1523575113.
5. Noguchi, A., Kitamura, T., Onaka, H., Horinouchi, S., and Ohnishi, Y. (2010). A copper-containing oxidase catalyzes C-nitrosation in nitrosobenzamide biosynthesis. Nat. Chem. Biol. *6*, 641–643. https://doi.org/10.1038/nchembio.418.
6. Sullivan, M.L. (2014). Beyond brown: polyphenol oxidases as enzymes of plant specialized metabolism. Front. Plant Sci. *5*, 783. https://doi.org/10.3389/fpls.2014.00783.
7. Panis, F., and Rompel, A. (2022). The Novel Role of Tyrosinase Enzymes in the Storage of Globally Significant Amounts of Carbon in Wetland Ecosystems. Environ. Sci. Technol. *56*, 11952–11968. https://doi.org/10.1021/acs.est.2c03770.
8. Decker, H., Hellmann, N., Jaenicke, E., Lieb, B., Meissner, U., and Markl, J. (2007). Minireview: Recent progress in hemocyanin research. Integr. Comp. Biol. *47*, 631–644. https://doi.org/10.1093/icb/icm063.
9. Lu, A., Zhang, Q., Zhang, J., Yang, B., Wu, K., Xie, W., Luan, Y.-X., and Ling, E. (2014). Insect prophenoloxidase: the view beyond immunity. Front. Physiol. *5*, 252. https://doi.org/10.3389/fphys.2014.00252.
10. Solomon, E.I., Heppner, D.E., Johnston, E.M., Ginsbach, J.W., Cirera, J., Qayyum, M., Kieber-Emmons, M.T., Kjaergaard, C.H., Hadt, R.G., and Tian, L. (2014). Copper Active Sites in Biology. Chem. Rev. *114*, 3659–3853. https://doi.org/10.1021/cr400327t.
11. Mydy, L.S., Chigumba, D.N., and Kersten, R.D. (2021). Plant Copper Metalloenzymes As Prospects for New Metabolism Involving Aromatic Compounds. Front. Plant Sci. *12*, 692108. https://doi.org/10.3389/fpls.2021.692108.
12. Eijsink, V.G.H., Petrovic, D., Forsberg, Z., Mekasha, S., Røhr, Å.K., Várnai, A., Bissaro, B., and Vaaje-Kolstad, G. (2019). On the functional characterization of lytic polysaccharide monooxygenases (LPMOs). Biotechnol. Biofuels *12*, 58. https://doi.org/10.1186/s13068-019-1392-0.
13. Furudate, H., Manabe, M., Oshikiri, H., Matsushita, A., Watanabe, B., Waki, T., Nakayama, T., Kubo, H., and Takanashi, K. (2023). A Polyphenol Oxidase Catalyzes Aurone Synthesis in Marchantia polymorpha. Plant Cell Physiol. *64*, 637–645. https://doi.org/10.1093/pcp/pcad024.
14. Suzuki, H., Furusho, Y., Higashi, T., Ohnishi, Y., and Horinouchi, S. (2006). A Novel o-Aminophenol Oxidase Responsible for Formation of the Phenoxazinone Chromophore of Grixazone. J. Biol. Chem. *281*, 824–833. https://doi.org/10.1074/jbc.M505806200.
15. Ginsbach, J.W., Kieber-Emmons, M.T., Nomoto, R., Noguchi, A., Ohnishi, Y., and Solomon, E.I. (2012). Structure/function correlations among coupled binuclear copper proteins through spectroscopic and reactivity studies of NspF. Proc. Natl. Acad. Sci. *109*, 10793–10797. https://doi.org/10.1073/pnas.1208718109.
16. Frommhagen, M., Mutte, S.K., Westphal, A.H., Koetsier, M.J., Hinz, S.W.A., Visser, J., Vincken, J.-P., Weijers, D., Van Berkel, W.J.H., Gruppen, H., and Kabel, M.A. (2017). Boosting LPMO-driven lignocellulose

degradation by polyphenol oxidase-activated lignin building blocks. Biotechnol. Biofuels *10*, 121. https://doi.org/10.1186/s13068-017-0810-4.

17. Nikolaivits, E., Dimarogona, M., Karagiannaki, I., Chalima, A., Fishman, A., and Topakas, E. (2018). Versatile Fungal Polyphenol Oxidase with Chlorophenol Bioremediation Potential: Characterization and Protein Engineering. Appl. Environ. Microbiol. *84*, e1628-18. https://doi.org/10.1128/AEM.01628-18.
18. Penttinen, L., Rutanen, C., Jänis, J., Rouvinen, J., and Hakulinen, N. (2018). Unraveling Substrate Specificity and Catalytic Promiscuity of *Aspergillus oryzae* Catechol Oxidase. Chembiochem *19*, 2348–2352. https://doi.org/10.1002/cbic.201800387.
19. Oates, N.C., Abood, A., Schirmacher, A.M., Alessi, A.M., Bird, S.M., Bennett, J.P., Leadbeater, D.R., Li, Y., Dowle, A.A., Liu, S., et al. (2021). A multi-omics approach to lignocellulolytic enzyme discovery reveals a new ligninase activity from *Parascedosporium putredinis* NO1. Proc. Natl. Acad. Sci. *118*, e2008888118. https://doi.org/10.1073/pnas.2008888118.
20. Silva, C.D.O.G., Sun, P., Barrett, K., Sanders, M.G., Van Berkel, W.J., Kabel, M.A., Meyer, A.S., and Wittrup Agger, J. (2024). Polyphenol Oxidase Activity on Guaiacyl and Syringyl Lignin Units. Angew. Chem. Int. Ed. *63*, e202409324. https://doi.org/10.1002/anie.202409324.
21. De Oliveira Gorgulho Silva, C., Vuillemin, M., Kabel, M.A., Van Berkel, W.J.H., Meyer, A.S., and Agger, J.W. (2023). Polyphenol Oxidase Products Are Priming Agents for LPMO Peroxygenase Activity. ChemSusChem *16*, e202300559. https://doi.org/10.1002/cssc.202300559.
22. Gasparetti, C., Faccio, G., Arvas, M., Buchert, J., Saloheimo, M., and Kruus, K. (2010). Discovery of a new tyrosinase-like enzyme family lacking a C-terminally processed domain: production and characterization of an Aspergillus oryzae catechol oxidase. Appl. Microbiol. Biotechnol. *86*, 213–226. https://doi.org/10.1007/s00253-009-2258-3.
23. Furumura, M., Solano, F., Matsunaga, N., Sakai, C., Spritz, R.A., and Hearing, V.J. (1998). Metal Ligand-Binding Specificities of the Tyrosinase-Related Proteins. BBRC (Biochem. Biophys. Res. Commun.) *242*, 579–585. https://doi.org/10.1006/bbrc.1997.8007.
24. Tsukamoto, K., Jackson, I.J., Urabe, K., Montague, P.M., and Hearing, V.J. (1992). A second tyrosinase-related protein, TRP-2, is a melanogenic enzyme termed DOPAchrome tautomerase. EMBO J. *11*, 519–526. https://doi.org/10.1002/j.1460-2075.1992.tb05082.x.
25. Biundo, A., Braunschmid, V., Pretzler, M., Kampatsikas, I., Darnhofer, B., Birner-Gruenberger, R., Rompel, A., Ribitsch, D., and Guebitz, G.M. (2020). Polyphenol oxidases exhibit promiscuous proteolytic activity. Commun. Chem. *3*, 62. https://doi.org/10.1038/s42004-020-0305-2.
26. Fekry, M., Dave, K.K., Badgujar, D., Hamnevik, E., Aurelius, O., Dobritzsch, D., and Danielson, U.H. (2023). The Crystal Structure of Tyrosinase from Verrucomicrobium spinosum Reveals It to Be an Atypical Bacterial Tyrosinase. Biomol *13*, 1360. https://doi.org/10.3390/biom13091360.
27. Fujieda, N., Yabuta, S., Ikeda, T., Oyama, T., Muraki, N., Kurisu, G., and Itoh, S. (2013). Crystal Structures of Copper-depleted and Copper-bound Fungal Pro-tyrosinase. J. Biol. Chem. *288*, 22128–22140. https://doi.org/10.1074/jbc.M113.477612.
28. Tran, L.T., Taylor, J.S., and Constabel, C.P. (2012). The polyphenol oxidase gene family in land plants: Lineage-specific duplication and expansion. BMC Genom. *13*, 395. https://doi.org/10.1186/1471-2164-13-395.
29. Waterhouse, R.M., Seppey, M., Simão, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E.V., and Zdobnov, E.M. (2018). BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. Mol. Evol. *35*, 543–548. https://doi.org/10.1093/molbev/msx319.
30. Mai, U., Sayyari, E., and Mirarab, S. (2017). Minimum variance rooting of phylogenetic trees and implications for species tree reconstruction. PLoS One *12*, e0182238. https://doi.org/10.1371/journal.pone.0182238.
31. Hakulinen, N., Gasparetti, C., Kaljunen, H., Kruus, K., and Rouvinen, J. (2013). The crystal structure of an extracellular catechol oxidase from the ascomycete fungus Aspergillus oryzae. JBIC *18*, 917–929. https://doi.org/10.1007/s00775-013-1038-9.
32. Nikolaivits, E., Valmas, A., Dedes, G., Topakas, E., and Dimarogona, M. (2021). Considerations Regarding Activity Determinants of Fungal Polyphenol Oxidases Based on Mutational and Structural Studies. Appl. Environ. Microbiol. *87*, e00396-21. https://doi.org/10.1128/AEM.00396-21.
33. Lai, X., Wichers, H.J., Soler-Lopez, M., and Dijkstra, B.W. (2017). Structure of Human Tyrosinase Related Protein 1 Reveals a Binuclear Zinc Active Site Important for Melanogenesis. Angew. Chem. Int. Ed. *56*, 9812–9815. https://doi.org/10.1002/anie.201704616.
34. Matoba, Y., Kumagai, T., Yamamoto, A., Yoshitsu, H., and Sugiyama, M. (2006). Crystallographic Evidence That the Dinuclear Copper Center of Tyrosinase Is Flexible during Catalysis. J. Biol. Chem. *281*, 8981–8990. https://doi.org/10.1074/jbc.M509785200.
35. Sendovski, M., Kanteev, M., Ben-Yosef, V.S., Adir, N., and Fishman, A. (2011). First Structures of an Active Bacterial Tyrosinase Reveal Copper Plasticity. J. Mol. Biol. *405*, 227–237. https://doi.org/10.1016/j.jmb.2010.10.048.
36. Kampatsikas, I., Bijelic, A., Pretzler, M., and Rompel, A. (2019). A Peptide-Induced Self-Cleavage Reaction Initiates the Activation of Tyrosinase. Angew. Chem. Int. Ed. *58*, 7475–7479. https://doi.org/10.1002/anie.201901332.
37. Virador, V.M., Reyes Grajeda, J.P., Blanco-Labra, A., Mendiola-Olaya, E., Smith, G.M., Moreno, A., and Whitaker, J.R. (2010). Cloning, Sequencing, Purification, and Crystal Structure of Grenache (*Vitis vinifera*) Polyphenol Oxidase. J. Agric. Food Chem. *58*, 1189–1201. https://doi.org/10.1021/jf902939q.
38. Klabunde, T., Eicken, C., Sacchettini, J.C., and Krebs, B. (1998). Crystal structure of a plant catechol oxidase containing a dicopper center. Nat. Struct. Biol. *5*, 1084–1090. https://doi.org/10.1038/4193.
39. Bijelic, A., Pretzler, M., Molitor, C., Zekiri, F., and Rompel, A. (2015). The Structure of a Plant Tyrosinase from Walnut Leaves Reveals the Importance of "Substrate-Guiding Residues" for Enzymatic Specificity. Angew. Chem. Int. Ed. *54*, 14677–14680. https://doi.org/10.1002/anie.201506994.
40. Kampatsikas, I., Bijelic, A., and Rompel, A. (2019). Biochemical and structural characterization of tomato polyphenol oxidases provide novel insights into their substrate specificity. Sci. Rep. *9*, 4022. https://doi.org/10.1038/s41598-019-39687-0.
41. Cuff, M.E., Miller, K.I., Van Holde, K.E., and Hendrickson, W.A. (1998). Crystal structure of a functional unit from Octopus hemocyanin. J. Mol. Biol. *278*, 855–870. https://doi.org/10.1006/jmbi.1998.1647.
42. Perbandt, M., Guthöhrlein, E.W., Rypniewski, W., Idakieva, K., Stoeva, S., Voelter, W., Genov, N., and Betzel, C. (2003). The Structure of a Functional Unit from the Wall of a Gastropod Hemocyanin Offers a Possible Mechanism for Cooperativity. Biochem. *42*, 6341–6346. https://doi.org/10.1021/bi020672x.
43. Tanaka, Y., Kato, S., Stabrin, M., Raunser, S., Matsui, T., and Gatsogiannis, C. (2019). Cryo-EM reveals the asymmetric assembly of squid hemocyanin. IUCrJ *6*, 426–437. https://doi.org/10.1107/S205225251900321X.
44. Zhang, Q., Dai, X., Cong, Y., Zhang, J., Chen, D.-H., Dougherty, M.T., Wang, J., Ludtke, S.J., Schmid, M.F., and Chiu, W. (2013). Cryo-EM Structure of a Molluscan Hemocyanin Suggests Its Allosteric Mechanism. Structure *21*, 604–613. https://doi.org/10.1016/j.str.2013.02.018.
45. Jaenicke, E., Büchler, K., Decker, H., Markl, J., and Schröder, G.F. (2011). The refined structure of functional unit h of keyhole limpet hemocyanin (KLH1-h) reveals disulfide bridges. IUBMB Life *63*, 183–187. https://doi.org/10.1002/iub.435.
46. Kato, S., Matsui, T., Gatsogiannis, C., and Tanaka, Y. (2018). Molluscan hemocyanin: structure, evolution, and physiology. Biophys. Rev. *10*, 191–202. https://doi.org/10.1007/s12551-017-0349-4.

47. Son, H.F., Lee, S.-H., Lee, S.H., Kim, H., Hong, H., Lee, U.-J., Lee, P.-G., Kim, B.-G., and Kim, K.-J. (2018). Structural Basis for Highly Efficient Production of Catechol Derivatives at Acidic pH by Tyrosinase from *Burkholderia thailandensis*. ACS Catal. *8*, 10375–10382. https://doi.org/10.1021/acscatal.8b02635.

48. Ismaya, W.T., Rozeboom, H.J., Weijn, A., Mes, J.J., Fusetti, F., Wichers, H.J., and Dijkstra, B.W. (2011). Crystal Structure of *Agaricus bisporus* Mushroom Tyrosinase: Identity of the Tetramer Subunits and Interaction with Tropolone. Biochemistry *50*, 5477–5486. https://doi.org/10.1021/bi200395t.

49. Pretzler, M., Bijelic, A., and Rompel, A. (2017). Heterologous expression and characterization of functional mushroom tyrosinase (AbPPO4). Sci. Rep. *7*, 1810. https://doi.org/10.1038/s41598-017-01813-1.

50. Koonin, E.V. (2005). Orthologs, Paralogs, and Evolutionary Genomics. Annu. Rev. Genet. *39*, 309–338. https://doi.org/10.1146/annurev.genet.39.073003.114725.

51. Kumar, S., Suleski, M., Craig, J.M., Kasprowicz, A.E., Sanderford, M., Li, M., Stecher, G., and Hedges, S.B. (2022). TimeTree 5: An Expanded Resource for Species Divergence Times. Mol. Evol. *39*, msac174. https://doi.org/10.1093/molbev/msac174.

52. Aguilera, F., McDougall, C., and Degnan, B.M. (2014). Evolution of the tyrosinase gene family in bivalve molluscs: Independent expansion of the mantle gene repertoire. Acta Biomater. *10*, 3855–3865. https://doi.org/10.1016/j.actbio.2014.03.031.

53. Duthoo, E., Delroisse, J., Maldonado, B., Sinot, F., Mascolo, C., Wattiez, R., Lopez, P.J., De Weerdt, C.V., Harrington, M.J., and Flammang, P. (2024). Abundance, diversity and evolution of tyrosinase enzymes involved in the adhesive systems of mussels and tubeworms. Preprint at bioRxiv. https://doi.org/10.1101/2024.07.05.602216.

54. Suci, P.A., and Geesey, G.G. (2000). Influence of Sodium Periodate and Tyrosinase on Binding of Alginate to Adlayers of Mytilus edulis Foot Protein 1. J. Colloid Interface Sci. *230*, 340–348. https://doi.org/10.1006/jcis.2000.7120.

55. Silverman, H.G., and Roberto, F.F. (2007). Understanding Marine Mussel Adhesion. Mar. Biotechnol. *9*, 661–681. https://doi.org/10.1007/s10126-007-9053-x.

56. Liao, J., Wei, X., Tao, K., Deng, G., Shu, J., Qiao, Q., Chen, G., Wei, Z., Fan, M., Saud, S., et al. (2023). Phenoloxidases: catechol oxidase - the temporary employer and laccase - the rising star of vascular plants. Hortic. Res. *10*, uhad102. https://doi.org/10.1093/hr/uhad102.

57. Margelevičius, M. (2019). Estimating statistical significance of local protein profile-profile alignments. BMC Bioinf. *20*, 419. https://doi.org/10.1186/s12859-019-2913-3.

58. Kanteev, M., Goldfeder, M., and Fishman, A. (2015). Structure-function correlations in tyrosinases. Protein Sci. *24*, 1360–1369. https://doi.org/10.1002/pro.2734.

59. Kampatsikas, I., and Rompel, A. (2021). Similar but Still Different: Which Amino Acid Residues Are Responsible for Varying Activities in Type-III Copper Enzymes? Chembiochem *22*, 1161–1175. https://doi.org/10.1002/cbic.202000647.

60. Pretzler, M., and Rompel, A. (2024). Mushroom Tyrosinase: Six Isoenzymes Catalyzing Distinct Reactions. Chembiochem *25*, e202400050. https://doi.org/10.1002/cbic.202400050.

61. Solem, E., Tuczek, F., and Decker, H. (2016). Tyrosinase versus Catechol Oxidase: One Asparagine Makes the Difference. Angew. Chem. Int. Ed. *55*, 2884–2888. https://doi.org/10.1002/anie.201508534.

62. Duran, K., Kohlstedt, M., Van Erven, G., Klostermann, C.E., America, A.H.P., Bakx, E., Baars, J.J.P., Gorissen, A., De Visser, R., De Vries, R.P., et al. (2024). From ^{13}C-lignin to ^{13}C-mycelium: *Agaricus bisporus* uses polymeric lignin as a carbon source. Sci. Adv. *10*, eadl3419. https://doi.org/10.1126/sciadv.adl3419.

63. Mahajan, S., Jeremic, D., Goacher, R.E., and Master, E.R. (2012). Mode of coniferous wood decay by the white rot fungus Phanerochaete carnosa as elucidated by FTIR and ToF-SIMS. Appl. Microbiol. Biotechnol. *94*, 1303–1311. https://doi.org/10.1007/s00253-011-3830-1.

64. Kontro, J., Lyra, C., Koponen, M., Kuuskeri, J., Kähkönen, M.A., Wallenius, J., Wan, X., Sipilä, J., Mäkelä, M.R., Nousiainen, P., and Hildén, K. (2021). Production of Recombinant Laccase From Coprinopsis cinerea and Its Effect in Mediator Promoted Lignin Oxidation at Neutral pH. Front. Bioeng. Biotechnol. *9*, 767139. https://doi.org/10.3389/fbioe.2021.767139.

65. Reyes-Fernández, E.Z., Shi, Y.-M., Grün, P., Bode, H.B., and Bölker, M. (2021). An Unconventional Melanin Biosynthesis Pathway in Ustilago maydis. Appl. Environ. Microbiol. *87*, e01510–e01520. https://doi.org/10.1128/AEM.01510-20.

66. Corsaro, D., Köhsler, M., Wylezich, C., Venditti, D., Walochnik, J., and Michel, R. (2018). New insights from molecular phylogenetics of amoebophagous fungi (Zoopagomycota, Zoopagales). Parasitol. Res. *117*, 157–167. https://doi.org/10.1007/s00436-017-5685-6.

67. Werner, A.Z., Cordell, W.T., Lahive, C.W., Klein, B.C., Singer, C.A., Tan, E.C.D., Ingraham, M.A., Ramirez, K.J., Kim, D.H., Pedersen, J.N., et al. (2023). Lignin conversion to -ketoadipic acid by *Pseudomonas putida* via metabolic engineering and bioprocess development. Sci. Adv. *9*, eadj0053. https://doi.org/10.1126/sciadv.adj0053.

68. Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

69. Pedersen, A.G. (2023). Seqconverter: A Command-Line Program for Reading, Writing, Analyzing, and Manipulating Sequence Files (Zenodo). https://doi.org/10.5281/zenodo.10411474 Version 3.0.0.

70. Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. Syst. Biol. *61*, 539–542. https://doi.org/10.1093/sysbio/sys029.

71. Pedersen, A.G. (2023). Sumt: A Command-Line Program for Computing Consensus Trees and Other Phylogenetic Tree Summaries (Zenodo). https://doi.org/10.5281/zenodo.10148693 Version 3.8.1.

72. Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. *49*, W293–W296. https://doi.org/10.1093/nar/gkab301.

73. Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: genome-scale protein function classification. Bioinformatics *30*, 1236–1240. https://doi.org/10.1093/bioinformatics/btu031.

74. Barbera, P., Kozlov, A.M., Czech, L., Morel, B., Darriba, D., Flouri, T., and Stamatakis, A. (2019). EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. Syst. Biol. *68*, 365–369. https://doi.org/10.1093/sysbio/syy054.

75. Czech, L., Barbera, P., and Stamatakis, A. (2020). Genesis and Gappa: Processing, Analyzing and Visualizing Phylogenetic (Placement) Data. Bioinformatics *36*, 3263–3265. https://doi.org/10.1093/bioinformatics/btaa070.

76. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis (New York: Springer-Verlag).

77. Joosten, R.P., Te Beek, T.A.H., Krieger, E., Hekkelman, M.L., Hooft, R.W.W., Schneider, R., Sander, C., and Vriend, G. (2011). A series of PDB related databases for everyday needs. Nucleic Acids Res. *39*, D411–D419. https://doi.org/10.1093/nar/gkq1105.

78. Schrödinger, L. (2020). The PyMOL Molecular Graphics System, Version 3.0 (Schrödinger, LLC).

79. Ayres, D.L., Darling, A., Zwickl, D.J., Beerli, P., Holder, M.T., Lewis, P.O., Huelsenbeck, J.P., Ronquist, F., Swofford, D.L., Cummings, M.P., et al. (2012). BEAGLE: An Application Programming Interface and High-Performance Computing Library for Statistical Phylogenetics. Syst. Biol. *61*, 170–173. https://doi.org/10.1093/sysbio/syr100.

80. Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. *37*, 420–423. https://doi.org/10.1038/s41587-019-0036-z.

81. Käll, L., Krogh, A., and Sonnhammer, E.L.L. (2004). A Combined Transmembrane Topology and Signal Peptide Prediction Method. J. Mol. Biol. *338*, 1027–1036. https://doi.org/10.1016/j.jmb.2004.03.016.

82. R Core Team (2023). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

83. Posit team (2023). RStudio: Integrated Development Environment for R (Boston, MA: Posit Software, PBC).

84. Shindyalov, I.N., and Bourne, P.E. (1998). Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. Protein Eng. *11*, 739–747. https://doi.org/10.1093/protein/11.9.739.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
MAFFT v7.508	Katoh and Standley[68]	https://mafft.cbrc.jp/alignment/server/index.html
seqconverter v3.0.0	Pedersen[69]	https://github.com/agormp/seqconverter
MrBayes v3.2	Ronquist et al.[70]	https://nbisweden.github.io/MrBayes/
sumt v3.8.1	Pedersen[71]	https://github.com/agormp/sumt
iTOL	Letunic[72]	https://itol.embl.de/
InterProScan	Jones[73]	https://www.ebi.ac.uk/interpro
EPA-NG v0.3.8	Barbera[74]	https://github.com/pierrebarbera/epa-ng
gappa v0.8.0	Czech[75]	https://github.com/lczech/gappa
ggplot2 3.5.1	Wickham[76]	https://ggplot2.tidyverse.org/
TimeTree 5	Kumar[51]	https://timetree.org/
COMER	Margelevičius[57]	https://bioinformatics.lt/comer/search/input/
DSSP	Joosten[77]	https://biopython.org/docs/1.75/api/Bio.PDB.DSSP.html
PyMOL	Schrödinger[78]	https://www.pymol.org/

METHOD DETAILS

PPO sequence dataset

A file listing all the Uniprot reference proteomes was downloaded from Uniprot on 2024-02-28, and a file listing all the Uniprot reference proteomes containing proteins matching Pfam PF00264 was downloaded from InterProScan[73] on 2024-01-22. One proteome was selected from each of the taxonomic classes represented in this data set, based on BUSCO scores.[29] If no proteome was available with BUSCO completeness score greater than 80%, the taxonomic class was omitted. Since no classes are defined for Oomycota, one protome from each order of Oomycota was selected. In total 100 proteomes were retrieved (Table S1).

All proteins matching PF00264 were downloaded from InterProScan[73] along with their E-values and proteome IDs. The proteins from the selected proteomes were retrieved. If a protein contained several PF00264 domains, all of the domains were included in the dataset. One protein from *Octopus vulgaris*, which contained 14 PF00264 domains, was discarded, as there were no similar proteins in Uniprot. In total 856 sequences were retrieved (Table S2).

Phylogenetic tree

The selected PF00264 domains with an E-value lower than 1e-20 were used for the phylogenetic tree. An alignment was constructed of the selected PF00264 domains using MAFFT v7.508 with the L-INS-i strategy.[68] Columns with more than 95% gaps were removed using seqconverter v3.0.0.[69]

A Bayesian phylogenetic tree was generated with MrBayes v3.2[70] using metropolus-coupled Markov chain Monte Carlo with the BEAGLE library.[79] Two runs with 3 chains were performed for 210,000,000 generations with gamma distributed rates and without a fixed substitution model, reaching an average standard deviation of split frequencies of 0.015. A consensus tree with all compatible groups was generated. Minimum variance rooting was performed using sumt v3.8.1.[71] The tree was visualized using iTOL.[72] Protein domain annotations were downloaded from InterProScan, including Pfam, SignalP[80] and Phobius[81] annotations.

Phylogenetic placement

To place new sequences on the phylogenetic tree, a maximum clade credibility tree was created using sumt v3.8.1.[71] EPA-NG v0.3.8 was used to perform phylogenetic placement of the new sequences on to the maximum clade credibility tree,[74] and a grafted tree was built using the "graft" command of gappa v0.8.0.[75]

Genome distribution plot

Sequences that were filtered out due to high E-values were placed on the tree using phylogenetic placement (see above). The members of each type in the tree were retrieved and the number of PPOs of the different types was plotted for each proteome in the dataset using R,[82] Rstudio,[83] and the ggplot2 v3.5.1 package.[76] A species tree showing the evolutionary relationship between the species was constructed using TimeTree 5.[51] Species that were not available in TimeTree were omitted.

Similarity of C-terminal domains

To investigate the similarity between the C-terminal shielding domains, we searched with the full alignment of the PF12143 Pfam domain downloaded from InterProScan on 2024-09-30 against the Pfam database (37.0) using COMER.[57]

Conserved residues

The conserved residues for each PPO type were determined based on alignments of the proteins of each type generated with MAFFT v7.508 using the L-INS-i strategy. Protein secondary structure was calculated from the PDB files using the DSSP module.[77]

Protein structures were visualized using PyMOL,[78] and structural alignments were performed using the CEalign algorithm[84] in pymol.

Extended fungal PPO sequence dataset

For the extended fungal PPO dataset, one proteome was selected from each fungal taxonomic order. If no proteome was available with BUSCO completeness score greater than 80%, the taxonomic order was omitted. In total 66 proteomes were retrieved (Table S3). From the selected proteomes, all the domains matching the PF00264 domain were retrieved. Additionally, a number of genomes of fungi that are known to grow on plant biomass were included (Table S5), and the PF00264 domains were likewise retrieved from these proteomes.

Genome distribution plot for extended fungal dataset

All the sequences from the extended fungal dataset (Table S4) as well as the lignin-degraders (Table S6) were placed on the phylogenetic tree using phylogenetic placement (see above). The members of each group in the tree were retrieved and the number of PPOs of the different types was plotted for each proteome.

Detailed phylogenetic tree of fungal short PPOs

All sequences in the extended fungal dataset (Table S4) with an E-value lower than 1e-20 were placed on the overall phylogenetic tree (Figure 2) using phylogenetic placement (see above). The sequences that located in the group of the fungal short PPOs (type **f**) were used to build a detailed phylogenetic tree of this group.

A Bayesian phylogenetic tree was generated with MrBayes v3.2[70] using metropolus-coupled Markov chain Monte Carlo with the BEAGLE library.[79] Two runs with 3 chains were performed for 60,000,000 generations with gamma distributed rates and without a fixed substitution model, reaching an average standard deviation of split frequencies of 0.011. A consensus tree with all compatible groups was generated. The tree was rooted on the branch that connects the chytridiomycetes sequences to the rest of the tree, as observed in the phylogenetic tree of PPOs from all taxonomic classes (Figure 2).