

Supplementary Information

Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists

Table of contents

1. Supplementary Note

- 1.1. List of strains
- 1.2. Biological material for Genome sequencing
- 1.3. Genome sequencing and assembly
- 1.4. Genome annotation and sequence analysis
- 1.5. Organismal Phylogeny
- 1.6. Evolution of decay-related gene families
- 1.7. Transcript profiling
- 1.8. MCL/CAFE
- 1.9. CAZyme annotation
- 1.10. Secretome

2. References

3. Supplementary Tables 1 to 11

4. Supplementary Figures 1 to 28

In addition to the listed authors, the Mycorrhizal Genomics Initiative includes:

Marc Buée^{1,2}, Yi Din³, Monique Gardes⁴, Gwen Grelet⁵, Hervé Gryta⁴, Patricia Jargeat⁴, Yaron Sitrit⁶ and Sabine Zimmermann⁷

¹*Inra, Lab of Excellence ARBRE, Unité Mixte de Recherche 1136, 54280 Champenoux, France;*

²*Lorraine University, Lab of Excellence ARBRE, Unité Mixte de Recherche 1136, 54280 Champenoux, France;*

³*Department of Organismic Interactions, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Strasse 10*

⁴*Université de Toulouse, UPS, UMR5174, Laboratoire Evolution et Diversité Biologique, F-31062 Toulouse, France;*

⁵*Landcare Research - Manaaki Whenua, Ecosystems and Global Change team, Gerald Street, PO Box 69040, Lincoln 7640, New Zealand;*

⁶*The Jacob Blaustein Institutes for Desert Research, Bergman Campus, Ben-Gurion University of The Negev, P.O.B. 653 Beer-Sheva 84105, Israel*

⁷*Institut de Biologie Intégrative des Plantes, Biochimie et Physiologie Moléculaire des Plantes, UMR 5004 Agrom-CNRS-INRA-UM2, Campus Agrom-INRA, 34060 Montpellier cedex 1, France;*

1. Supplementary Note

1.1. List of strains (Supplementary Table 1)

Amanita muscaria var. *guessowii* Koide

Agaricales (Basidiomycota)

Common name: Fly Agaric

Abbreviation: Amamu

Collection: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts

Ecology: Ectomycorrhizal fungus

Information: *A. muscaria* (L. per Fr.) Hooker is a filamentous basidiomycete (Agaricales) typically found as mycelia in soils. The morphological species encompasses a complex of eight undescribed biological species. They are geographically widespread symbionts of conifers and hardwoods. We have sequenced a culture of a species named *A. muscaria* var. *guessowii* which grows on the East Coast of North America.

Gymnopus luxurians

Agaricales (Basidiomycota)

Common name: none

Abbreviation: Gymlu

Collection: Forest Product Laboratory, USDA Forest Service, Madison

Ecology: saprotrophic fungus

Information: *Gymnopus luxurians* is a saprotrophic fungus growing on wood chip mulch areas and buried wood, but little is known about the decay chemistry. There is increasing evidence that at least some *Gymnopus* species associate with mycoheterotroph orchids.

Hebeloma cylindrosporum h7

Agaricales (Basidiomycota)

Common name: none

Abbreviation: Hebcy

Collection: Laboratoire d'Ecologie Microbienne, Université Claude Bernard Lyon 1

Ecology: Ectomycorrhizal fungus

Information: *H. cylindrosporum* is frequently found in forest stands developing on sand dunes with very little organic matter along the atlantic or mediterranean coast. In this respect, *H. cylindrosporum* can be qualified as a pioneer species which thrives in newly established forests or in disturbed areas. *H. cylindrosporum* is frequently associated with different pine trees such as *Pinus pinaster*. It can also form mycorrhizas with additional hosts which do not occur in its natural habitats.

Hydnomerulius pinastris

Boletales (Basidiomycota)

Common name: spiny dry rot fungus

Abbreviation: Hydpi

Collection: Forest Product Laboratory, USDA Forest Service, Madison

Ecology: brown-rot wood decayer

Information: *Hydnomerulius pinastri* is the sister group to the Boletineae, which harbors some 800 predominantly ectomycorrhizal species including highly valued edibles, such as porcini mushrooms. The ecological requirements of *H. pinastri* are identical to those of the dry rot fungus *Serpula lacrymans*. The wood of conifers is the substrate of choice and both species attack wooden building structures. However, *H. pinastri* infestations occur rarely and if they do, they are much less aggressive compared to the fatal damage left by *S. lacrymans*.

Hypholoma sublateritium

Agaricales (Basidiomycota)

Common name : Brick Cap

Abbreviation: Hypsu

Collection: Forest Product Laboratory, USDA Forest Service, Madison

Ecology: white-rot wood decayer

Information: *Hypholoma sublateritium* is a widely distributed mushroom and it can be found growing on hardwood trunks in the fall. This wood-decay fungus is receiving increasing attention from commercial growers because it is usually considered edible and is suited for cultivation. Like many other saprotrophic fungi in the Agaricomycotina, *Hypholoma sublateritium* has been scrutinized for pharmacologically active compounds that are of general interest.

Laccaria amethystina LAAM-08-1

Agaricales (Basidiomycota)

Common name: Amethyst Deceiver

Abbreviation: Laam

Collection: Tree-Microbe Interactions Department, INRA-Nancy

Ecology: Ectomycorrhizal fungus

Information: *Laccaria amethystina* is forming symbiotic associations with hardwoods or conifers. It produces deep purple, edible mushrooms, that grow among moss and leaf litter under deciduous as well as coniferous trees. *L. amethystina* diverged from the *L. bicolor* lineage ~20 My ago.

Oidiodendron maius

Incertae sedis (Ascomycota)

Common name: none

Abbreviation: Oidma

Collection: Mycotheca Universitatis Taurinensis, University of Torino

Ecology: ericoid mycorrhiza forming fungus

Information: *O. maius* is an interesting experimental organism being both an endomycorrhizal fungus (with ericaceous plants, e.g. *Vaccinium myrtillus*, *Calluna vulgaris*), and a metal-tolerant fungus.

Paxillus involutus ATCC 200175

Boletales (Basidiomycota)

Common name: brown roll-rim, common roll-rim, or poison pax

Abbreviation: Paxin

Collection: American Type Culture Collection (ATCC) # 200175

Ecology: Ectomycorrhizal fungus

Information: *Paxillus involutus* is one of the most well studied ectomycorrhizal fungi at molecular, physiological, and ecological levels. The fungus forms mycorrhizae with many species of trees, including the genetic model tree species *Populus trichocarpa*.

Paxillus rubicundulus Ve08.2h10

Boletales (Basidiomycota)

Common name : none

Abbreviation: Paxru

Collection: Laboratoire Evolution et Diversité Biologique, CNRS, Université Paul Sabatier-Toulouse

Ecology: Ectomycorrhizal fungus

Information: *Paxillus rubicundulus* is an ectomycorrhizal basidiomycete specifically associated to alders. This species belongs to the Paxillaceae family, in which some members are hygrophilic and highly specialized on alders, such as *P. rubicundulus* or *Gyrodon lividus*, while some other members have a large ecological range and are generalist, such as *P. involutus*.

Piloderma croceum F 1598

Atheliales (Basidiomycota)

Common name : none

Abbreviation: Piler

Collection: Department of Soil Ecology, Helmholtz Centre for Environmental Research, Halle # DSMZ 4824; ATCC MYA-4870.

Ecology: Ectomycorrhizal fungus

Information: *Piloderma croceum* is a broad host range ectomycorrhizal fungus and a common mutualist of both conifer and hardwood species. Typically encountered in boreal and temperate forests, *P. croceum* is one of the most easily spotted ectomycorrhizal fungi, because of the bright yellow color of its mycelium and ectomycorrhizal root tips.

Pisolithus microcarpus 441

Boletales (Basidiomycota)

Synonym: *Polysaccum microcarpum* Cooke & Massee 1887

Common name: none

Abbreviation: Pismi

Collection: Tree-Microbe Interactions Department, INRA-Nancy; American Type Culture Collection # ATCC MYA-4687

Ecology: Ectomycorrhizal fungus

Information: *P. microcarpus* is an ectomycorrhizal basidiomycete forming ectomycorrhizas with angiosperm hosts, especially *Eucalyptus* spp. The ectomycorrhizas are golden yellow with the Hartig net limited to the intercellular spaces of the root epidermis. This gasteromycete fungus has been widely studied owing to its application in eucalypt nurseries and in plant growth promotion worldwide.

Pisolithus tinctorius 270

Boletales (Basidiomycota)

Synonym: *Pisolithus arhizus* (Scop.) Rauschert

Common name: dyemaker's puffball, dog turd fungus, dead man's foot

Abbreviation: Pisti

Collection: Tree-Microbe Interactions Department, INRA-Nancy; American Type Culture Collection # ATCC MYA-4688

Ecology: Ectomycorrhizal fungus

Information: *P. tinctorius* is a cosmopolitan, forming ectomycorrhizas with conifer hosts, mostly in the Northern Hemisphere. The ectomycorrhizas are dark yellow with a pluriseriate Hartig net formed in the intercellular spaces of several root cortex layers. The fungus has been studied owing to its application in forest nurseries and in plant growth promotion worldwide.

Plicaturopsis crispa

Amylocorticiales (Basidiomycota)

Common name: Crimped Gill

Abbreviation: Plicr

Collection: Forest Product Laboratory, USDA Forest Service, Madison

Ecology: white-rot wood decayer

Information: *Plicaturopsis crispa* produces clusters of small, fan-shaped fruiting bodies with wrinkled spore bearing layers that resemble vein-like folds. There is a general notion that this fungus has been expanding its original southern range over the last few decades with temperatures increasing globally. *P. crispa* is an effective decayer in the initial phase of decay colonizing predominantly dead branches of deciduous trees (e.g. *Fagus* and *Betula* species).

Scleroderma citrinum Foug

Boletales (Basidiomycota)

Common name: Common earthball, pigskin poison puffball

Abbreviation: Scld

Collection: Tree-Microbe Interactions Department, INRA-Nancy

Ecology: Ectomycorrhizal fungus

Information: *Scleroderma citrinum* is a very common and widespread ectomycorrhizal gasteromycete species which produces large conspicuous sporocarps, so-called "earthballs", in different forest environments or adjacent to forest areas. The ectomycorrhizal status of *S. citrinum* is proven, but in the *Scleroderma* genus, numerous species appear capable of free-living saprotrophic existence. Unlike most gasteromycetes species, *S. citrinum* can be cultured on synthetic media and its ectomycorrhiza synthesized *in vitro*.

Sebacina vermifera MAFF 305830

Sebacinales (Basidiomycota)

Abbreviation: Sebve

Common name: none

Collection: Institut für Genetik, Biozentrum Köln

Ecology: orchid mycorrhiza forming fungus

Information: The Australian orchid mycorrhizal *Sebacina vermifera* MAFF 305830 belongs to the basidiomycetous order Sebacinales (subgroup B). This order encompasses ubiquitously distributed taxa which are basal in the Agaricomycetes with diverse mycorrhizal abilities, ranging from ectomycorrhizae to ericoid, orchid mycorrhizae and root endophytes. Due to their inconspicuous or even absent basidiomes, this group of fungi has been often overlooked and underestimated in its ecological and potential economic importance. The orchid mycorrhizae represent the most basal group with known mycorrhizal capabilities.

Sphaerobolus stellatus

Phallales (Basidiomycota)

Abbreviation: Sphst

Common name : Cannon ball fungus

Collection: Forest Product Laboratory, USDA Forest Service, Madison

Ecology: white-rot wood decayer

Information: *Sphaerobolus stellatus* grows on woody debris and herbivore dung. *S. stellatus* is a member of the Phallomycetidae, which is a morphologically diverse group that also includes coral fungi, stinkhorns, earthstars, and false truffles. *S. stellatus* produces tiny fruit bodies that violently eject a slimy packet of spores, called a gleba.

Suillus luteus UH-Slu-Lm8-n1

Boletales (Basidiomycota)

Abbreviation: Suilu

Common name : Slippery jack

Collection: Centrum voor Milieukunde, Universiteit Hasselt

Ecology: Ectomycorrhizal fungus

Information: *Suillus luteus* is a cosmopolitan ectomycorrhizal fungus whose natural range of distribution matches the range of distribution of its host plants, the *Pinus* species. It is particularly abundant in young pine forests or planted stands, from the Andes to the boreal forests. The species is a pioneer, which quickly starts sexual reproduction from large edible sporocarps that produce massive quantities of basidiospores, spread by wind and mammals. It forms conspicuous, though relatively few mycorrhizas from which an extensive external mycelium (long distance exploration type) develops into the mineral soil.

Tulasnella calospora AL13/4D

Cantharellales (Basidiomycota)

Abbreviation: Tulca

Common name : pseudobulb, root rot

Collection: Mycotheca Universitatis Taurinensis, University of Torino

Ecology: Orchid mycorrhiza forming fungus

Information: For their development, all orchids rely on the association with symbiotic fungi like *Tulasnella calospora*, that (at least in the early stages) provide the plant with organic carbon. *T. calospora* is the most common mycorrhizal partner of green orchids and belongs to the phylum Basidiomycota (order Cantharellales, family Tulasnellaceae). This fungus is distributed world-wide and normally found in every ecosystem, from tropical to temperate climate zones.

1.2. Biological material for genome sequencing

Free-living vegetative mycelium of sequenced fungi (except *P. involutus*) was grown for about three weeks on cellophane-covered agar (12 g L⁻¹) containing modified Pachlewski P5 medium or low-sugar Pachlewski P20 medium. The edge of the mycelial colonies was then harvested with a razor blade and snap frozen in liquid nitrogen. From 500 mg of finely ground mycelium powder, DNA was extracted in a 50 ml Falcon tube using 17.5 ml of lysis buffer composed of 6.5 ml Buffer A (0.35 M sorbitol, 0.1M Tris-HCl pH 9, 5mM EDTA pH8), 6.5ml Buffer B (0.2 M Tris-HCl pH9, 50mM EDTA pH8, 2M NaCl, 2% CTAB), 2.6 ml Buffer C (5% Sarkosyl (N-lauroylsarcosine sodium salt), 1.75ml 0.1% PVP and 125µl Proteinase K (20mg/ml).

The mycelial suspension was incubated for 30 min at 65°C by frequently mixing the tube. Then, 5.75 ml of 5M potassium acetate was added to the suspension and incubated for 30 min on ice. The tube was then centrifuged for 20 min at 5000g at 4°C in an Eppendorf bench centrifuge and the supernatant transferred to a new 50 ml Falcon tube and one volume of chloroform:isoamylalcohol (24:1) was added. After mixing, the tube was centrifuged for 10 min at 4000g at 4°C. The aqueous phase was transferred to a 50 ml centrifuge tube (Nalgene), 100 µl of RNase A (10 mg/ml) added and incubated for 2 hours at 37°C. Then 1/10 volume of sodium acetate and one volume of isopropanol (RT) were added, the DNA precipitated for 10 min at room temperature (RT) and then centrifuged for 30 min at 10,000g at 4°C. The supernatant was discarded, the pellet washed with 2 ml of 70% (v/v) ethanol and centrifuged for 10 min at 10,000g at 4°C. The supernatant was again discarded, the pellet dried for 5 min at RT and then resuspended in 500 µl of Tris-EDTA (TE) buffer at 65°C. This DNA solution was further purified using Qiagen genomic-tip 500/G columns following the manufacturers instructions. Owing to their high pigment content, *Pisolithus tinctorius* and *P. microcarpus* DNA solutions were further purified by a CHROMA Spin+TE-1000 column (Clontech) to remove the remaining pigments.

P. involutus was maintained on modified Melin-Norkrans (MMN) agar medium containing 2.5 g L⁻¹ glucose¹ and was grown as a free-living vegetative mycelium on the top of cellophane-covered MMN agar Petri dishes for a week at RT and in the dark. At time of harvesting, the mycelium was scraped off the cellophane surface and dropped into liquid nitrogen in a mortar and the frozen mycelium was thoroughly grinded into a powder and left at -80°C until use. For DNA preparation, the DNeasy Plant Maxi kit (Qiagen) was used including the QiaShredder and the on-column RNase treatment according to the manufacturer's instructions. The eluted DNA was then further purified using the Genomic DNA Clean Concentrator kit (Zymo Research) to remove remaining and inhibiting pigments.

1.3. Genome sequencing and assembly

The new genomes and transcriptomes for the fungi reported in this study were sequenced using several sequencing platforms and assembled with tools most appropriate for the datasets obtained from these platforms (Supplementary Table 2) at the U.S. Department of Energy Joint Genome Institute (JGI).

Twelve genomes (*Amanita muscaria*, *Tulasnella calospora*, *Scleroderma citrinum*, *Hydnomerulius pinastris*, *Hypholoma sublateritium*, *Laccaria amethystina*, *Paxillus rubicundulus*, *Piloderma croceum*, *Pisolithus tinctorius*, *Sebacina vermifera*, *Sphaerobolus stellatus*, *Suillus luteus*) were sequenced using Illumina standard paired-end (PE) and long mate-pair (LMP) libraries (Supplementary Table 2). For Illumina

standard libraries, genomic DNA was sheared to 250-300 bp fragments using the Covaris E210. The fragments were treated with end repair, A-tailing, and adaptor ligation, and then sequenced in 2x100 bp or 2x150 bp read formats (Supplementary Table 2). For a majority of LMP libraries, genomic DNA was sheared to its desired insert size (4kb, and 8-10kb) using the Hydroshear®. The ends of the fragments were ligated with biotinylated adapters containing loxP, then circularized via recombination by a Cre excision reaction, and digested using 4-base cutter restriction enzymes followed by self-ligation and inverse PCR. For *A. muscaria*, *L. amethystina*, and *P. rubicundulus*, 4 kb LMP libraries were prepared using 5500 SOLiD Mate-Paired Library Construction Kit (Life Technologies). DNA was sheared with the Covaris g-TUBE™, ligated with biotinylated internal linkers, circularized using intra-molecular hybridization of internal linkers, then nick translated, treated with T7 exonuclease and S1, and ligated with adapters. All LMP libraries were sequenced using 2x100 bp read format on an Illumina platform. Sequenced reads were filtered for artifacts, process contamination, and subsequently assembled with AllPathsLG². For *S. luteus*, LMP reads were simulated from the initial Velvet³ assembly of the Illumina standard library data to improve quality of the final AllPathsLG assembly (Copeland, A. unpublished; ^{4,5}). Library details and run type are listed in Supplementary Table 2. Assembly statistics are summarized in Supplementary Table 3.

Five genomes (*Gymnopus luxurians*, *Oidiodendron maius*, *Paxillus rubicundulus*, *Pisolithus microcarpus*, *Plicaturopsis crispa*) were sequenced using the hybrid approach, which combines 454 (Roche) pyrosequencing with Illumina reads and Sanger sequenced fosmid libraries when available (Supplementary Table 2). For standard libraries, genomic DNA was fragmented by nebulization to an average size of 500-800 bp. 454 LMP were produced using CRE-LoxP protocols as described above. Circularized DNA was randomly sheared using the Covaris E210, ligated with adapters, and amplified via PCR. Sequenced reads produced using 454 Titanium chemistry were quality assessed, trimmed, screened for adapters and contamination. Using Newbler(v.2.5)⁶ 454 reads were assembled with shredded consensus sequences from Velvet assembled Illumina data and Sanger fosmid reads when available. Whenever possible, gaps were closed with gapResolution⁷. The draft assembly of *Hebeloma cylindrosporum* was also produced using the hybrid approach, assembled using AllPathsLG², and further improved using Pacific Biosciences reads from 3kb libraries (v2 chemistry, 31X coverage) to close gaps. Mitochondrial sequences were screened out and assembled separately for all assemblies.

Sequenced transcriptomes were used to assess the completeness of final assemblies and facilitate genome annotations. All transcriptomes except for those of *P. involutus* and *O. maius* were sequenced using the Illumina platform. mRNA was purified from total RNA using oligo-dT beads, chemically fragmented, and reverse transcribed using random hexamers and SuperScript® II transcriptase. The ends of the fragments were treated with end repair and ligated with adapters. For Illumina stranded protocol, the second strand was synthesized using a dNTP/dUTP mix and then removed using AmpErase UNG. RNA-Seq reads were assembled using Rnnotator⁸ after trimming and filtering for low-quality, low-complexity, and adapter sequences and duplications. Assembled contigs > 100bp long with at least 3 mapped reads were corrected for misassemblies, polished, and then clustered into loci to produce final transcripts⁸. *P. involutus* and *O. maius* transcriptomes were sequenced on the 454 (Roche)

pyrosequencing platform. 454 reads were screened for quality and contamination and then assembled using Newbler v.2.5⁶.

1.4. Genome annotation and sequence analysis

The genome assemblies of the 18 fungi were each annotated using the JGI Annotation Pipeline^{9,10}, which 1) detects and masks repeats and transposable elements, 2) predicts genes using a variety of methods, 3) characterizes each conceptually translated protein using a variety of methods, 4) chooses a ‘best’ gene model at each locus to provide a filtered working set, 5) clusters the filtered sets into draft gene families, and 6) creates a JGI Genome Portal with tools for public access and community-driven curation of the annotation¹¹.

The assembly scaffolds were masked by RepeatMasker (<http://www.repeatmasker.org/>) using both the manually curated Repbase library (v. 19.05)¹² and a library of repeats and transposable elements (TEs) individually and specifically constructed *de novo* for each genome as follows. Each custom TE library was created using RepeatScout¹³ to discover repetitive sequences and then filtered by selecting sequences with one of the following properties: similarity to RepBase as assessed by BLASTn¹², presence of a TE-related Pfam domain¹⁴, presence of characteristic TE termini (long terminal repeats, terminal inverted repeats, target site duplications), or presence of > 150 copies in the genome.

Transcript-based gene models were built (1) from the above mentioned *de novo* assembled EST and/or RNA-Seq contigs mapped onto genomic scaffolds using EST_MAP (<http://www.softberry.com/>) and (2) from genome-based assemblies of RNA-Seq reads using COMBEST (Zhou, K unpublished). Protein-based gene models were predicted using GeneWise¹⁵ and FGENESH+¹⁶ seeded by BLASTx alignments of genomic sequence against fungal sequences from the NCBI non-redundant protein set. *Ab initio* gene models were predicted using GeneMark-ES¹⁷ and FGENESH¹⁶, the latter trained on a set of putative full-length transcripts and reliable protein-based models. GeneMark-ES performed unsupervised training through iterative estimation of parameters, starting with a general parameter set and stopping when nucleotide sensitivity and specificity exceeded 97%¹⁷. FGENESH was trained on those protein-based gene models whose protein seeds perfectly aligned with the genome, and then tested on the transcript-based gene models. Each newly computed parameter set’s exon specificity (Sn_e) and sensitivity (Sp_e) on the test models was compared to those of all other parameter sets computed for previous fungal genomes, and the best-performing parameter set was used for the final FGENESH modeling (Sn_e and Sp_e always > 50%). GeneWise models were completed using scaffold data to find start and stop codons. RNA contigs aligned by BLAT¹⁸ to the genome were used by estExt (Grigoriev, IV unpublished) to verify and when possible to extend the predicted gene models when intron-exon boundaries matched between the models and mapped RNA contigs.

All predicted gene models were functionally annotated by the JGI Annotation Pipeline using InterProScan¹⁹, BLASTp alignments against NCBI non-redundant protein set, and hardware-accelerated double-affine Smith-Waterman alignments (<http://www.timelogic.com/>) against highly curated databases such as SwissProt²⁰, KEGG²¹, and Pfam¹⁴. KEGG hits were used to map EC numbers²², and InterPro, KEGG, and SwissProt hits were used to map GO terms²³. In addition, predicted proteins were annotated according to KOG classification²⁴. Protein targeting predictions were made with signalP²⁵ and TMHMM²⁶.

Because the gene-prediction steps usually generated multiple gene models per locus, the JGI Pipeline selected a single representative gene model for each locus using a heuristic approach based on protein similarity, and RNA coverage, leading to a filtered set of gene models for downstream analysis (see Supplementary Table 4). Protein similarity was calculated from the above-described BLASTp alignments, considering only those with bit score (s_b) > 50 and coverage ($0 \leq \text{cov} \leq 1$) > 25% of the length of both subject and query. RNA coverage was calculated from the above-described RNA contig BLAT alignments, considering only those with positive average correlation coefficient (CC) between a gene model and overlapping BLAT alignments (CC = +1 for complete agreement and CC = -1 for complete disagreement). Gene models that did not fulfill both protein and RNA criteria were assigned a score $S = 0$. Each remaining gene model was assigned a score $S = s_b \times (\text{cov}_s \times \text{cov}_q + \text{CC})$. For a given locus, the model with the highest score was selected, and all other models that had greater than 5% overlap with the selected model were excluded from the filtered model set.

All of the filtered model set proteins in a genome were aligned by BLASTp to each other and the alignment scores were used as a distance metric for clustering by MCL (<http://www.micans.org/mcl/>) with inflation parameter=2 into a first draft of candidate multigene families for each genome. In addition, segmental duplications were selected as duplicated genome fragments with minimum of three genes in each fragment with at least of 50% of genes between fragments being mutual best Blastp hits to each other (Supplementary Table 4).

Genome assemblies and annotations for the organisms used in this study are available via the JGI fungal genome portal MycoCosm²⁴ (<http://jgi.doe.gov/fungi>; Supplementary Tables 2 & 3). In addition, the newly sequenced genome assemblies and annotations have been deposited to GenBank under the following accessions/BioProjects: *Amanita muscaria* Koide BX008: JMDV000000000/ PRJNA207684; *Gymnopus luxurians* FD-317 M1: JJNP000000000/ PRJNA68535; *Hebeloma cylindrosporum* h7: JMDQ000000000/ PRJNA207849; *Hydnomerulius pinastri* MD 312: JMSK000000000/ PRJNA207871; *Hypholoma sublateritium* FD-334 SS-4: JMSJ000000000/ PRJNA70685; *Laccaria amethystina* LaAM-08-1: JMSL000000000/ PRJNA196025; *Oidiodendron maius* Zn: JMDP000000000/ PRJNA74727; *Paxillus involutus* ATCC 200175: JOMD000000000/ PRJNA60449; *Paxillus rubicundulus* Ve08.2h10: JMDR000000000/ PRJNA243391; *Piloderma croceum* F 1598: JMDN000000000/ PRJNA61203; *Pisolithus microcarpus* 441: JMDM000000000/ PRJNA60815; *Pisolithus tinctorius* Marx 270: JMDO000000000/ PRJNA207840; *Plicaturopsis crispa* FD-325 SS-3: JOMB000000000/ PRJNA207847; *Scleroderma citrinum* Foug A: JMDU000000000/ PRJNA207859; *Sebacina vermifera* MAFF 305830: JMDS000000000/ PRJNA207844; *Sphaerobolus stellatus* SS14: JOMA000000000/ PRJNA207858; *Suillus luteus* UH-Slu-Lm8-n1: JMSM000000000/ PRJNA242126; *Tulasnella calospora* MUT 4182: JMDT000000000/ PRJNA20784).

1.5. Organismal Phylogeny

Phylogenomic analyses. Dataset assembly. We performed phylogenomic analyses using 49 genomes (Supplementary Table 1). We identified gene families with only one gene per species by clustering protein sequences using the Markov clustering algorithm, with an inflation parameter of 2.0 (cluster 1973), resulting in 1809 single-copy gene families. 617 of the single-copy gene families were represented in at least 15 species. For these

gene families, we performed multiple sequence alignment and Maximum Likelihood (ML) tree estimation. Multiple sequence alignment was performed in PRANK v.1.11.130²⁸ with default settings, one round of alignment improvement and invoking the option to output XML formatted alignment files in addition to fasta. To assess the effects of alignment assumptions on the resulting phylogenies, we also used MAFFT 6.864b²⁹ to align protein sequences. Gene trees were estimated in RAxML 7.2.8³⁰ using the standard algorithm and the PROTGAMMAWAG model of sequence evolution. We filtered gene families for potentially non-orthologous proteins following dos Reis *et al.*³¹. Unusually long terminal branches (accounting for >60% of the total tree length) were checked manually.

In five gene families a single gene accounted for more than 60% of sum of branch lengths in the gene tree. These gene families were omitted from further analyses. The resulting 612 gene trees were screened for topological conflict by computing weighted Robinson-Foulds distances as implemented in hashRF 6.0.1³². We computed weighted RF-distances for each pair of trees, which takes branch length information into account. Robinson-Foulds distances were uniform across gene trees suggesting congruent phylogenetic signal, with the exception of one gene family, which showed a significantly greater RF distance against all other trees and was omitted from subsequent analyses. A flowchart illustrating data collection and filtering steps is presented in Supplementary Figure 1 and the list of gene families used for phylogenetic analyses has been deposited in Dryad.

For each gene alignment, we excluded ambiguously aligned sites using two strategies. The first strategy used GBlocks 0.91³³ with default settings applied to the alignments generated by MAFFT 6.864b2. The second strategy (developed for this project) used site posterior probabilities calculated by PRANK and stored in the xml-files. These posterior probabilities represent the probability that the alignment site is correctly aligned and take into account insertion and deletion events and the guide tree connecting the sequences (in contrast, GBlocks considers only gap content). For each alignment site, the posterior probabilities were summed over all residues and their mean calculated; only sites for which the mean of the posterior probabilities exceeded a user-set threshold were retained. These analyses were performed using a custom perl script (FilterPostProb), which is available from the authors upon request and is deposited in Dryad. We used two different posterior probability thresholds, 0.95 and 1.0, to exclude unreliable columns from the alignments. A threshold of 1.0 means that only alignment columns for which all residues have a posterior probability of 1 are retained for phylogenetic analyses. This represents a very strict criterion, yet it has several advantages over GBlocks. First, missing sequences are not counted as gaps, so alignment accuracy is evaluated only in the context of the species that have the sequence (in contrast, with default settings GBlocks would delete the entire gapped region of the alignment). Second, since PRANK posterior probabilities take into account the insertion-deletion process, high-reliability regions with indels will be retained, whereas they would be removed by GBlocks.

Following the exclusion of unreliable alignment sites, we concatenated single-gene alignments into a supermatrix. We excluded alignments that had less than 50 amino acid residues left after the exclusion of unreliable sites, because such short alignments may be insufficient for accurate estimation of model parameters in partitioned analyses. We recorded the starting and end positions of each single- gene alignment during

concatenation, in order to use that information in partitioning the dataset for model-based phylogenetic analyses.

The strategy described above resulted in three phylogenomic datasets: (1) PRANK alignments with 1.0 exclusion threshold (PR1.0 dataset); (2) PRANK alignments with 0.95 exclusion threshold (PR0.95 dataset); and (3) MAFFT alignments curated by Gblocks (MAFFT-Gbl). The PR1.0 dataset contained 19,567 aligned sites representing 149 loci, whereas due to the lower stringency of site exclusion, the PR0.95 dataset comprised 114,814 aligned sites from 542 loci. The MAFFT-Gbl dataset contained 34,323 aligned sites from 259 loci. On average, each species had 74.4% of all the loci (450 out of the 611 loci, with the lowest value (36.7%) in the yeast *Pichia stipitis*, and the highest value (94.05%) in *Jaapia argillacea*. The datasets and partition tables have been deposited in Dryad.

Phylogenetic analyses. We performed ML and Bayesian phylogenetic inference for each of the phylogenomic datasets. During initial runs the datasets were unpartitioned, but in subsequent analyses the datasets were partitioned according to single-gene alignments and the model parameters were unlinked between partitions. For each dataset, the tree topology obtained in these initial runs was identical to the one inferred in the partitioned analyses. Nonetheless, partitioned models have been shown to outperform unpartitioned models^{34,35}, especially in phylogenomic studies including protein sequences evolving at a potentially wide range of evolutionary rates^{36,37}. Throughout the analyses, we used the WAG model of protein evolution with a gamma distribution (4 discrete categories) to account for rate heterogeneity within single loci. The WAG model was implemented for each locus separately, and the model parameters unlinked between partitions.

For all three datasets, we performed ML bootstrapping using the PTHREADS version of RaxML 7.2.83. For all runs we ran 1000 thorough bootstrap replicates, using the rapid hill climbing algorithm and a partitioned model. Bootstrapped trees were summarized and mapped to the ML tree using the SumTrees script of the Dendropy package³⁸.

Bayesian phylogenetic analyses were performed in PhyloBayes 3.3, with the CAT mixture model of protein evolution³⁹. We ran three replicates with one chain per replicate and a chain length of 100,000 cycles. Bayesian phylogenetic analyses were also attempted using MrBayes 3.2.1, but they failed to converge to a reasonable posterior distribution, despite multiple parameter settings and replicated runs, so those results were not used.

Results. Altogether 611 single-gene alignments passed the initial filtering criteria. After concatenation, the proportion of missing data ranged from 12% in the PR1.0 dataset, to 14% in the MAFFT-Gbl dataset, and 30% in the PR0.95 dataset. The trees obtained from the different ML and Bayesian phylogenetic analyses were largely congruent with each other and with results of previous multigene⁴⁰ and phylogenomic analyses⁴¹ (Supplementary Fig. 2). Virtually all nodes were supported by >70% bootstrap or 0.95 posterior probability values. Weak and conflicting placement of taxa in Basidiomycota was limited to two areas of the phylogeny, the placement of the Ustilaginomycotina/Pucciniomycotina and the Auriculariales/Phallomycetidae. The placement of the *Jaapia-Gloeophyllum-Punctularia* clade received moderate to weak

support in some analyses, although its placement is consistent across the analyses and with previous genome-based phylogenies⁴¹.

The placement of the Ustilaginomycotina and Pucciniomycotina along the backbone of the Basidiomycota phylogeny has been poorly supported in multiple studies^{42–44,45,46}. Some of our analyses united the Ustilaginomycotina and Pucciniomycotina as a weakly supported clade (PR1.0-ML and PR0.95-ML, 47% and 78%, respectively), but other analyses placed the Ustilaginomycotina as the sister group of the Agaricomycotina. Statistical support for the latter placement is stronger (Bayesian posterior probability of 0.98 for PR1.0, PR0.95 and MAFFT-Gbl, whereas the ML bootstrap support for the MAFFT-Gbl dataset was 54%), but the relationships among the three subphyla of Basidiomycota cannot be considered to be resolved. It is possible that these splits represent a hard polytomy that cannot be appropriately handled in phylogenetic software assuming bifurcating evolution^{47,48}.

Another problematic area of the tree concerns the placement of the Phallomycetidae relative to the Auriculariales, represented here by *Sphaerobolus stellatus* and *Auricularia delicata*, respectively. In two of the six analyses (MAFFT-Gbl-ML, PR1.0-ML) these taxa were inferred as monophyletic, although with weak to moderate support (50% and 86% bootstrap). The rest of the analyses agree in their placement as separate clades along the backbone of the Agaricomycetes, with *Sphaerobolus stellatus* being the sister group to the rest of the Agaricomycetes (excluding the Auriculariales, Cantharellales and Sebaciniales), whereas *Auricularia* being the sister to Phallomycetidae and the rest of the Agaricomycetes. This conformation received strong support (1.0 for PR0.95-Bayes, PR1.0-Bayes, MAFFT-Gbl-Bayes and 96% for PR0.95-ML) and agrees well with previous studies^{43,45,49}.

Molecular clock analyses

Methods. To obtain divergence times for the genome-based tree, we used the penalized likelihood algorithm as implemented in the program r8s⁵⁰, using the ML tree obtained with the PR0.95 dataset. To obtain an optimal smoothing parameter, we first performed a cross-validation analysis, in which we tested smoothing parameter values across 6 orders of magnitudes, starting from 0.01. The POWELL optimization algorithm was used during all analyses. To calibrate nodes of the phylogeny we used three fungal fossils, a suilloid ectomycorrhiza associated with pine roots from the middle Eocene (50 MYR⁵¹), *Archaeomarasmium legettii* from the mid-Cretaceous (94–90 MYR⁵²), *Paleopyrenomycites devonicus* from the early Devonian (ca. 400 MYR⁵³). The suilloid ectomycorrhiza from the middle Eocene is the only known fossilized ectomycorrhizal structure and can be placed in the Suillineae of the Boletales because of its tuberculate morphology and pinaceous host. We calibrated the Suillinae/Paxillinae-Sclerodermatinae split comprising the taxa *Suillus luteus*, *Hydnomerulius pinastri*, *Paxillus rubicundulus*, *P. involutus*, *Scleroderma citrinum*, *Pisolithus tinctorius* and *P. microcarpus*. *Archaeomarasmium legettii* is the oldest known fossil resembling the morphology of modern Agaricales, and can be placed in the marasmioid clade of this order. Hence, we used this fossil to calibrate the origin of the clade formed by *Gymnopus luxurians* and *Schizophyllum commune* on the species tree. *Paleopyrenomycites* represents an ascomycete with a well-developed perithecium, which resembles similar structures found in extant Dothideomycetes. We used this fossil to calibrate the split between *Tuber melanosporum* and the clade formed by *Cryphonectria parasitica*, *Trichoderma reesei*, *Oidiodendron maius*, *Stagonospora*

nodorum, and *Aspergillus nidulans*. After determining the optimal smoothing parameter value, we inferred node ages using the three calibrations mentioned above. We specified age intervals for each of the three fossils as follows: *Archaeomarasmius* minimum age 70 MA, maximum age 110 MA; Suilloid ectomycorrhiza minimum age 40 MA, maximum age 60 MA; and for *Paleopyrenomyces* a minimum age of 360 MA, and a maximum age of 440 MA was set. The analysis was run with the optimal smoothing parameter under the penalized likelihood method. Age estimates were summarized by the 'chrono_description' command on r8s. A Bayesian MCMC approach using BEAST⁵⁴ was also attempted, but the analyses failed to converge to reasonable posterior distributions, so we discarded those results.

Results. The fossils that we used for calibration have been used in numerous prior molecular clock analyses, including that of Floudas *et al.*⁴¹, which used the same three fossils in a very similar manner. Therefore the node age estimates presented here are not wholly independent of those obtained in other studies. Nonetheless, the results presented here (see Supplementary Table 6) are similar to those of Eastwood *et al.*⁵⁵. For example, in our r8s analysis the age estimates for Dikarya, Basidiomycota and Agaricomycetidae are 635, 500 and 125 Ma, while the average ages for these nodes estimated by Floudas *et al.* using BEAST with three fossil calibrations are 662, 521 and 149 Ma (with large 95% highest posterior density intervals). To estimate the temporal range during which the evolution of ECM is most plausible, we considered the ages of the major and potentially oldest ECM hosts, namely Pinaceae (which are exclusively ECM) and Rosids (which include such major ECM groups as Fagales and Myrtaceae, among others). The oldest Pinaceae fossil is *Eathiestrobus mackenziei* from the upper Jurassic⁵⁶, while molecular clock analyses⁵⁷ suggest that the most recent common ancestor of Pinaceae existed in the mid-Jurassic, which is much earlier than the estimated radiation of Rosids ca. 91-108 MA⁵⁸. Therefore, we considered all nodes in the fungal phylogeny younger than ca. 170 Ma to be plausibly ECM. In the Agaricomycotina, the minimum age of ECM roughly corresponds to that of the most recent common ancestor of the clade containing Polyporales, Corticiales, Gloeophyllales, Jaapiales, Russulales and Agaricomycetidae.

1.6. Evolution of decay-related gene families

Methods. To reconstruct the evolution of saprotrophic capabilities, we performed gene tree/species tree reconciliation, focusing on gene families encoding enzymes thought to be involved in plant cell wall (PCW) degradation. We aimed to analyze gene families with diverse biochemical functions (including oxidoreductases, hydrolases and esterases) that target diverse macromolecules, including lignin, cellulose, hemicellulose, and pectin, as well as secondary products of the PCW degradation such as cellobiose and xylobiose. We caution that many large gene families, and in particular CAZymes, are polyspecific and may encompass functions other than PCW degradation⁵⁹. We compiled data from 19 CAZY and nine oxidoreductase gene families from the 49 genomes (Table1, Supplementary Table 9). Of these, we selected 11 CAZY and 5 oxidoreductase gene families for phylogenetic analyses and gene tree/species tree reconciliation analyses. The assembled datasets include glycoside hydrolases (GH) of the families 3, 5, 6,7, 10, 12, 28, 43, and 61 (lytic polysaccharide monooxygenases), carbohydrate esterases (CE) of the families 1, and 16, Class II peroxidases (POD),

multicopper oxidases (MCO), Dye decolorizing peroxidases (DyP), heme-thiolate peroxidases (HTP), and copper radical oxidases (CRO).

We considered POD, MCO, HTP, and DyP to be related, or potentially related, to lignin degradation. While PODs seem to be strictly involved in the oxidative degradation of lignin^{60,61}, MCOs, HTPs and DyPs appear to participate in diverse processes, such as pigment production and defense mechanisms⁶² and oxidation of lignin residues, humic substances and xenobiotic compounds⁶³. CROs represent only one of the gene families involved in hydrogen peroxide generation during wood degradation, assisting the function of peroxidases^{64,65}. We also targeted enzymes involved in degradation of crystalline and amorphous cellulose. We included cellobiohydrolases and endoglucanases of the families GH6, GH7, GH5 (subfamily 5) and GH12⁶⁶⁻⁷¹, and LPMO genes, which are involved in the oxidative attack of crystalline cellulose^{72,73}. In addition, we investigated hemicellulose-acting enzymes including GH10 xylanases, GH5 mannanases (subfamily 7), GH28/GH43 pectin-acting enzymes⁷⁴ and carbohydrate esterases of the families 1 and 16. The latter enzymes disrupt the esteric bonds of hemicellulose⁷⁵.

We performed preliminary alignments for each gene family using MAFFT v.7⁷⁶ (<http://mafft.cbrc.jp/alignment/software/>) with the FFT-NS-i or E-INS-i strategies. We divided large, heterogeneous datasets to facilitate phylogenetic analyses (GH3, GH28, GH5, CE1, CE16, CRO, MCO) following previous subclassification studies, when possible. We checked the preliminary alignments manually and removed very short fragments or potential pseudogenes. When necessary, we replaced models that represented fragments of the same locus with a complete model or replaced models of the gene catalog with other better available models found on the genome browser. We performed more thorough alignments of the final datasets using PRANK1 (<http://www.ebi.ac.uk/goldman-srv/webprank/>) with the default settings, and we manually corrected and removed ambiguous regions of the resulting alignments using MacClade v. 4.08⁷⁷.

To generate the input tree files for gene tree/species tree reconciliation, we performed ML analyses using RAxML 7.2.8 with the PROTGAMMAWAG model of evolution. Bootstrap support from 1000 replicates was mapped onto branches of the ML tree using the SumTrees script³⁸. We performed gene tree/species tree reconciliation analyses using Notung 2.6⁷⁸ with an edge weight threshold of 90 (RAxML bootstrap support) in the gene trees. For the species tree, we used the topology obtained with the PR0.95 dataset. Unreconciled gene trees generated with RAxML are presented in Supplementary Figs. 4-11 and organismal phylogenies indicating changes in gene copy number based on Notung analyses are presented in Supplementary Figs. 12-21. Gene copy numbers at internal nodes estimated with Notung are presented in Supplementary Table 9. Alignments and RAxML trees for each gene family have been deposited in Dryad.

Overview of results. The 49 fungal genomes include 5167 protein models in the 16 targeted gene families. We excluded 83 models due to their low quality (short fragments, potential pseudogenes) or potential origin via horizontal gene transfer from bacteria (based on BLAST searches and preliminary phylogenetic analyses). Twenty-four of the excluded models belong to *S. stellatus*, while the rest of the excluded sequences belong to 27 genomes, ranging from 1 to 7 models per genome (Supplementary Tables 1 & 8). We also excluded 36 models from families GH5 and

GH28 that represent small groups of genes with scattered representation across the 49 genomes (described below).

Of the 5048 retained sequences, 4323 belong to Agaricomycotina, while 725 sequences belong to *U. maydis*, *M. laricis-populina*, *P. blakesleeanus*, *B. dendrobatitis* and the ascomycete genomes. Excluding *S. stellatus*, the largest overall number of sequences in the targeted families belongs to *G. marginata* (205 sequences), followed by *A. delicata* (200 sequences), while the smallest number belongs to *B. dendrobatitis* (8 sequences).

Sphaerobolus stellatus, which represents the first genome of the Phallomycetidae, exhibited a high degree of fragmented models, which made it difficult to confidently estimate gene contents. The *S. stellatus* genome annotation includes 62 POD genes, 29 laccases sensu stricto, 31 DyP, 151 HTP genes and one GLX, as well as members of all examined families of CAZs, with LPMO presenting the largest number of sequences (37 copies; Supplementary Table 7). Although though the precise the number of gene copies in *S. stellatus* is uncertain, its gene repertoire strongly suggests that it is capable of producing a white rot. Lignin modification and degradation Class II peroxidases (POD). POD genes are found in both Ascomycota and Basidiomycota, indicating that they were present in the ancestor of Dikarya, but are absent from Ustilaginomycotina and Pucciniomycotina, suggesting that they were lost repeatedly in these clades. Species of Agaricomycotina possess from 0 to 26 copies of POD genes, with an average of four copies per genome (Supplementary Figs. 4, 11, Supplementary Table 7). The largest numbers of POD genes are found in white rot species (7-26 copies), except *S. commune*, which has no POD genes and does not appear to degrade lignin⁷⁹. The soil saprotrophs possess fewer POD genes, ranging from 0 for *A. thiersii* to 5 for *G. luxurians*. The latter species is associated with wood chips and leaf litter. In contrast, mycorrhizal and brown rot species have few if any POD genes. Among the ECM species, *P. croceum* and *L. bicolor* possess one POD gene each, while *H. cylindrosporum* possesses 3 copies. The other mycorrhizal species have no POD genes (Supplementary Table 7). Not all POD genes exhibit high redox potential, which is necessary for lignin degradation. To assess the ligninolytic potential of PODs in the newly sequenced genomes, we searched for Mn-binding site residues (characteristic of MnPs) and the long range electron transfer (LRET) tryptophan (characteristic of LiPs) and subclassified the PODs following the terminology developed in previous studies^{41, 55, 60} (Supplementary Table 10).

Gymnopus luxurians contains both short and long/extra long types of MnP (Supplementary Fig. 4). *Hypholoma sublateritium* contains only atypical MnP genes, nested within two different clades (Supplementary Table 10, Supplementary Fig. 4). *Galerina marginata* has 23 POD genes, including 16 atypical MnPs, one atypical versatile peroxidase (VP), one generic peroxidase (GP), one gene that could be either a GP or an atypical MnP (a gap in the sequence did not allow precise classification, Supplementary Table 10), and four genes that appear to code for lignin peroxidase (LiP) type enzymes (the translated proteins have the LRET tryptophan, but they each have only one of the three amino acids necessary for the formation of the Mn-binding site). The putative LiPs genes of *G. marginata* cluster together with the VP genes of the species, similar to the clustering of a VP gene from *T. versicolor* (ID: Trave1_43289) with the LiP gene in Polyporales (Supplementary Fig. 4). These would be the first LiP genes to be found outside of the Polyporales, but biochemical confirmation is needed to resolve their function.

In contrast to the diverse set of POD genes found in *H. sublateralitium*, *G. marginata* and *G. luxurians*, the ECM *P. croceum* (Atheliales) contains only a generic peroxidase similar to that of *L. bicolor*, which has a basal placement in the POD phylogeny (Supplementary Fig. 4). *Hebeloma cylindrosporum* is the only ECM species to code for 3 PODs, which are atypical MnPs (Supplementary Table 10) nested within a strongly supported clade among other atypical MnP sequences of *H. sublateralitium* and *G. marginata* (Supplementary Fig. 4).

Gene tree/species tree reconciliation results suggest that the first duplication event for POD genes in Basidiomycota occurred prior to the separation of Auriculariales (node 18, Supplementary Fig. 12), with additional expansions later (from 2 genes at node 18, to 7 genes at node 27), suggesting the establishment of the white rot mechanism during the diversification of Agaricomycetes. Parallel reductions in PODs have occurred in each of the lineages leading to the mycorrhizal Agaricomycetes. However, in two cases, the Amanitaceae and Boletaceae, the loss of all copies of PODs appears to have preceded the origin of the ecomycorrhizal lifestyle (Supplementary Fig. 12).

Multicopper oxidases. We separated the MCOs into 5 groups following Hoegger et al.⁸⁰. Two of the recognized groups include putative ascorbate oxidases and fungal pigment MCO sequences. Both groups represent only a small fraction of the MCO dataset (14 and 17 sequences respectively); they have only a scattered representation in Agaricomycotina and were not included in the reconciliation analysis. The ECM *P. croceum* encodes two putative ascorbate oxidases (*S. commune* and *T. mesenterica* are the only other two Agaricomycotina species that have ascorbate oxidases), while 5 copies of fungal pigment MCOs are found in *B. botryosum*, *T. calospora*, *S. commune* and *A. delicata*.

The other three groups of MCO genes have a wide representation across Agaricomycotina. Fet3 is involved in iron homeostasis⁸¹. Most species of Agaricomycotina have one Fet3 gene copy per genome (Supplementary Table 7). However, *S. commune* and *C. cinerea* lack Fet3 sequences, while few species have multiple copies, including the ectomycorrhizal *L. bicolor* and *P. croceum*. The conservation of Fet3 copy number is also supported in the reconciliation results, which suggest that 1 to 2 copies of Fet3 were present throughout the diversification of Agaricomycotina.

In contrast to Fet3 genes, laccase/ferroxidase genes have a more scattered distribution across Agaricomycotina and their abundance is not obviously correlated with any particular nutritional strategy (Supplementary Table 7). Most Agaricales genomes lack laccase/ferroxidase genes, but they are present in most of the Boletales. The largest number of laccase/ferroxidase genes is found for *P. croceum* (6 copies), but the rest of the mycorrhizal species have one or zero copies (Supplementary Table 7). Reconciliation results suggest an early expansion of these genes in the evolution of basidiomycetes and Agaricomycotina, with subsequent reductions during establishment of the Agaricomycotina orders (Supplementary Table 5).

Laccases *sensu stricto*, the most abundant group of MCO genes, are widespread in Dikarya, but their numbers are variable between species. The largest number of laccases is found in *O. maius* (15 copies) for ascomycetes and for *A. muscaria* (18 copies) for basidiomycetes, but there are no copies of laccases found for Ustilaginomycotina and Pucciniomycotina in this dataset (Supplementary Table 7, Supplementary Fig. 12). The

Agaricomycotina genomes have on average 8 copies of laccase genes, but there are no laccase genes found for Tremelomycetes, Dacrymycetales, Cantharellales and *P. chrysosporium*. In contrast to the POD genes, laccase genes are widespread in mycorrhizal genomes and most, except *H. cylindrosporum*, and *T. calospora*, encode average or above average numbers of laccases (Supplementary Table 7, Supplementary Fig. 12).

Gene tree/species tree reconciliation analyses suggest that the expansions of laccase genes began after the divergence of Cantharellales (Supplementary Fig. 12). Throughout the early diversification of Agaricomycotina the laccase genes appear to have undergone small changes in numbers, but there were many lineage-specific expansions in terminal groups of Agaricomycotina, including expansions in ECM Agaricomycetidae lineages, such as *Amanita muscaria*, *Laccaria bicolor*, *Piloderma crocea*, and the ECM Boletales (Supplementary Figs. 6, 12).

Dye-decolorizing peroxidases (DyP) and heme-thiolate peroxidases (HTP). HTP genes are present in all Agaricomycotina sampled, but one third of the species lack DyP genes (Supplementary Fig. 13, Supplementary Table 7). When present, DyP genes usually occur in 1-5 copies, but *G. luxurians* and *A. delicata* have 13 and 11 copies, respectively. HTP genes are more numerous, with an average of 7 copies per species. Several species show expansions in HTP genes, including soil saprotrophs (*A. bisporus*, *C. cinerea*), and wood decayers (*G. marginata*, *H. sublateralitium*, *G. luxurians* and *A. delicata*). *Piloderma croceum* is the only ECM species that has more than the average number of HTP gene copies (11 copies).

HTP and DyP genes both increased in number early in the diversification of Agaricomycotina (node 16 to node 18 for HTP genes and node 13 to 18 for DyP genes, Supplementary Fig. 13). Subsequently, the results suggest contractions of DyP genes at the base of Amylocorticiales-Boletales, Polyporales, and at the base of Gloeophyllales, Corticiales and Jaapiales. Similarly, reductions are suggested for HTP genes in lineages leading to the Boletales and Polyporales. However, HTP genes have been maintained in the ancestral nodes in Agaricales, where several species, enriched in HTP genes, are nested. Both gene families show variable patterns of gene gains or reductions in symbiotic lineages (Supplementary Fig. 13).

Copper radical oxidases (CRO). CRO genes can be divided into the groups CRO1, CRO2, CRO3, CRO4, CRO5, CRO6 and glyoxal oxidases^{41,82} (GLX). The role of CRO proteins is not completely understood, but they may participate in hydrogen peroxide production during wood colonization from *P. chrysosporium*^{64,83}. Among the CROs, CRO 3-5 sequences form a phylogenetic group and include N- terminal repeats of the WSC domain, which is of unknown function. Most species possess one CRO 3-5 gene, but some have lost these genes. In addition, a few species have multiple copies of CRO 3-5 genes, including *S. luteus* and *L. bicolor* (Supplementary Table 7). Reconciliation results support the maintenance of low numbers of these genes during the diversification of Agaricomycotina (Supplementary Table 9).

CRO 1, 2, 6 and GLX genes are widespread and more variable in copy numbers in basidiomycetes in comparison to CRO 3-5 genes (Supplementary Table 7). The reconciliation results also suggest a more complex history of gene duplications and losses for CRO 1, 2, 6, and GLX (Supplementary Fig. 14). Five copies of these genes are reconstructed early during the early diversification of Basidiomycota (nodes 10-13), with expansions during the diversification of Agaricomycetes (nodes 14 to 18), and

contractions in several clades, including the Boletales-Atheliales-Amylocorticiales clade, Gloeophyllales-Jaapiiales clade, and the Hymenochaetales (Supplementary Fig. 14).

The overall pattern of expansions and contractions for CRO 1, 2, 6 and GLX genes is not obviously associated with any specific nutritional strategy in Agaricomycotina, but an examination of the phylogenetic distribution of genes within the 4 categories indicates that each has variable retention in Agaricomycotina. CRO 1 and 2 sequences form a strongly supported clade in the phylogenetic analysis (Supplementary Fig. 5), but CRO 1 genes appear to be paraphyletic, with the CRO 2 sequences forming a strongly supported clade nested among them. CRO 2 genes appear to be conserved, being found in *U. maydis* and all the Agaricomycotina genomes sampled here, except for *Dacryopinax*. Thus, they may be critical in function. This is in agreement with previous findings on the role of the CRO2 protein (Ustma_2411) for filamentous growth and pathogenicity in *U. maydis*⁸⁴. The mycorrhizal genomes have maintained from 1 to 3 CRO 2 genes each. The CRO 1 genes are also relatively conserved, but have been lost in both of the ECM *Paxillus* species (Supplementary Fig. 5).

GLX and CRO 6 genes form a strongly supported clade, including one main group of CRO 6 genes and two smaller clades of putative CRO 6 sequences, which collectively form a paraphyletic group within which the strongly supported GLX sequences are nested (Supplementary Fig. 5). CRO 6 sequences are in general conserved across the Agaricomycotina, but some mycorrhizal species have lost those genes, including both *Pisolithus* species, *S. citrinum*, *L. bicolor* and *T. calospora*. GLX genes are maintained mostly in white rot genomes, as shown previously³⁶, but not in mycorrhizal or brown rot species. Here, we show that some soil saprotrophs maintain GLX (Supplementary Fig. 5, Supplementary Table 7). The oldest Agaricomycotina lineage to contain GLX genes is the Auriculariales (*A. delicata*), suggesting that the GLX genes emerged early in the diversification of Agaricomycetes, roughly coinciding with the expansion of POD genes. The absence of GLX in Sebaciniales and Cantharellales may be a plesiomorphic feature. In the more recently evolved Agaricomycetes, the loss of GLX genes observed for ECM and brown rot species, and their maintenance in white rot species and soil saprotrophs suggests a possible connection of GLX to lignin degradation. The GLX of *P. chrysosporium* has been associated with hydrogen peroxide production, potentially provided for the function of class II peroxidases^{64,65}. However, presence of GLX genes is not always coupled with presence of ligninolytic POD genes. For example, *A. thiersii* possesses GLX genes, but no POD genes. Conversely, *H. cylindrosporum*, *H. annosum* and *F. mediterranea* contain ligninolytic PODs, but lack GLXs (Supplementary Table 7).

CAZYmes

GH6, *GH7*, *LPMO*. We focused on GH6 and GH7, which include cellobiohydrolases and endoglucanases, and LPMO, which are lytic polysaccharide monooxygenases, because they have been shown to participate in the degradation of crystalline cellulose^{67, 69, 72, 73, 85}. The three gene families have homologs across Ascomycota and Basidiomycota, but are absent from *Phycomyces blakesleeanus* (Mucoromycotina) and *Batrachochytrium dendrobatidis* (Chytridiomycota) (Supplementary Table 7). The lowest overall number of genes is found for GH6 (Fig missing), while LPMO is the most diverse (Supplementary Fig. 7). On average across the Agaricomycotina, there is

one copy of GH6, three copies of GH7 and 13 copies of LPMO per genome (Supplementary Table 7).

The presence or absence of GH6 and GH7 copies within genomes appears to be correlated (Supplementary Fig. 15). Both gene families are present in soil saprotrophs and white rot species, as well as the orchid mycorrhizal *S. vermifera*, *T. calospora* and the endophytic *P. indica*, but they are absent from the genomes of the ectomycorrhizal Agaricomycetidae and most brown rot fungi. Exceptions to this pattern occur in the ECM *H. cylindrosporum* and *P. croceum*, which have each maintained one GH7 gene, and the brown rot Boletales, including *S. lacrymans*, which has one GH6 gene, and *C. puteana* and *H. pinastri*, which have both GH6 and GH7 genes (Supplementary Table 7).

LPMO genes occur in species with diverse nutritional strategies, but their number per genome is highly variable. *Tremella mesenterica* and *Dacryopinax* are the only Agaricomycotina that lack LPMO, while the largest number of LPMO genes is found in *C. cinerea* (Supplementary Fig. 16, Supplementary Table 7). The number of LPMO genes is near or above average for white rot species, soil saprotrophs, the orchid mycorrhizal *S. vermifera* and *T. calospora* and the endophyte *P. indica* (from nine in *P. crispa* to 35 in *C. cinerea*), around or below the average for brown rot genomes (0 copies for *Dacryopinax* to 15 copies for *H. pinastri*), and below average for the ectomycorrhizal Agaricomycetidae (one for *P. croceum* to 11 for *L. bicolor*, Supplementary Fig. 16).

Reconciliation analyses suggest that the common ancestor of Agaricomycotina possessed 4 genes for each of the gene families GH6, GH7 and LPMO (Supplementary Figs. 15, 16, Supplementary Table 5). The number of GH6 genes was conserved during the diversification of Agaricomycotina, especially between nodes 12 to 28. In contrast, the GH7 and LPMO families appear to have diversified early in the evolution of Agaricomycetes, with LPMO genes increasing from 4 to 27 copies (node 13 to node 14, Supplementary Fig. 15). In addition, reconciliation analyses suggest maintenance or lineage specific expansions for the three gene families in the Cantharellales and Sebaciniales (except for GH6 genes in *P. indica*), with *Tulasnella calospora* possessing the largest number of GH7 and GH6 genes (27 and seven copies respectively) among the genomes sampled here. Species in these orders also possess among the largest numbers of LPMO genes, surpassed only by *Coprinopsis cinerea* (Supplementary Fig. 15).

Reconciliation analyses suggest parallel gene losses for GH6, GH7 and LPMO in lineages leading to ectomycorrhizal species in Agaricomycetidae (Supplementary Fig. 15, 16). The ten ECM Agaricomycetidae species sampled here all lack GH6 genes, and eight lack GH7 genes. Even though *H. cylindrosporum* possesses one GH7 gene, the reconciliation suggest that there were three losses of GH7 genes in the lineage leading to the species (Supplementary Fig. 15). In contrast to GH6 and GH7, all ECM Agaricomycetidae have at least one LPMO gene. Nonetheless, reconciliation analyses also suggest gene losses in lineages leading to ECM Agaricomycetidae (Supplementary Fig. 16). Phylogenetic analysis of LPMO genes suggests that gene retention for ectomycorrhizal Agaricomycetidae is not uniformly distributed. Despite the multiple origins of the ectomycorrhizal lifestyle in the Agaricomycetidae, most of the maintained LPMO genes for the species sampled here are nested in one moderately supported clade (RAxML bootstrap = 70%; Supplementary Fig. 7). *Hebeloma cylindrosporum* is the

only ECM species of Agaricomycetidae that maintains two LPMO copies from outside of this clade, while *L. bicolor* has only one. This pattern of differential retention supports the view that there may be more than one function distributed across the subclades of this diverse gene family^{82, 86, 87}.

Glycoside hydrolase families 10, 28, 43, 3 and 12. The preliminary alignment of GH28 proteins revealed diverse groups of sequences, which we treated as separate datasets. One of these datasets includes only five models (three from *A. delicata*, and one each from *H. cylindrosporum* and *O. maius*) and was not analyzed further. Most species of Agaricomycotina maintain copies for both GH28 and GH43 gene families, but the number of copies per genome is variable (Supplementary Fig. 10a, b). GH28 sequences range from zero (*A. muscaria* and *T. mesenterica*) to 19 copies per genome (*G. luxurians*), while GH43 sequences range from zero (*S. citrinum*, both *Paxillus* species, *L. bicolor*, *T. mesenterica*) to 28 copies (*A. delicata*). On average, there are seven GH28 and five GH43 sequences across Agaricomycotina (Supplementary Fig. 20, Supplementary Table 7). Saprotrophs maintain higher numbers of sequences for both families than symbiotic species. Among Agaricomycotina, GH43 or GH28 sequences are absent only from the genomes of the mycoparasite *T. mesenterica* and the ECM species *A. muscaria*, *S. citrinum*, both *Paxillus* species, and *L. bicolor* (Supplementary Fig. 20).

GH10 xylanases are widespread in Agaricomycotina, which have on average four GH10 copies per species (Supplementary Table 7). The largest number of GH10 genes is found in *T. calospora*, followed by *J. argillacea* and *G. marginata*. All saprotrophic Agaricomycotina and the symbiotic Cantharellales and Sebaciales possess GH10 genes, but there is variation in the number of genes harbored in these genomes, with above average numbers of copies usually found in white rot species, soil saprotrophs, and the symbiotic Cantharellales/Sebaciales, and below average numbers in brown rot saprotrophs (Supplementary Fig. 17). Only three ECM species in Agaricomycetidae possess GH10 genes (Supplementary Fig. 17). Reconciliation analyses suggest an increase in number of GH10 genes early in the diversification of Agaricomycetes (node 13 is estimated to have had 3 copies while node 24 had 8 copies), with parallel reductions in ECM Agaricomycetidae (Supplementary Fig. 17).

GH3 xylosidase gene sequences are broadly distributed across the Agaricomycotina, with as few as two (*L. bicolor*) to as many as 14 (*P. strigosozonata*) copies per species, with average of nine genes per species (Supplementary Table 7, Supplementary Fig. 19). Mycorrhizal genomes tend to have below or around the average number of GH3 sequences in comparison to saprotrophs, but exceptions can be found for *P. croceum* and *S. luteus* (with eleven and ten copies, respectively; Supplementary Fig. 18).

Reconciliation analyses suggest a high number of GH3 genes across many ancestral nodes of Agaricomycotina, with subsequent lineage specific reductions (Supplementary Fig. 19). Transitions to the ECM lifestyle are accompanied by reductions in GH3 gene copy numbers in the lineages leading to Sclerodermatineae, the *Paxillus* species, *H. cylindrosporum*, and *L. bicolor*.

GH12 endoglucanase genes appear to be relatively conserved across Agaricomycotina, with an average of two copies per species (Supplementary Fig. 19, Supplementary Table 7). A few saprotrophic genomes have increased number of GH12 genes, with up to eight copies in *J. argillacea*, while some ECM species have completely lost GH12 genes, including all *Pisolithus* and *Paxillus* species, *A. muscaria*,

as the orchid mycorrhizal *T. calospora*, as has the mycoparasite *T. mesenterica*. The reconciliation analysis suggests also that GH12 was conserved in gene copies family during the diversification of Agaricomycotina, with most backbone nodes reconstructed as having 2 genes for the family (Supplementary Fig. 19).

Carbohydrate esterases (CE). Carbohydrate esterases are a diverse group of enzymes. We analyzed the carbohydrate esterases families CE1 and CE16 as exemplars of esterases participating in plant cell wall degradation. Acetyl-xylan, cinnamoyl, feruloyl esterase and carboxymethylesterase activities are reported for the CE1 enzymes, while acetyl-xylan esterases are the known enzymatic function for the CE16 enzymes⁸⁸⁻⁹⁰.

On average, two CE1 genes are found per species of Agaricomycotina (Supplementary Fig. 21, Supplementary Table 7), but most saprotrophs have more gene copies. In contrast, nine of the ten ECM Agaricomycetidae species lack CE1 genes (*P. croceum* has two copies). Mycorrhizal species of Cantharellales and Sebaciniales have CE1 numbers similar to those of saprotrophs. *Amanita muscaria* appears to have no fungal CE1 sequences, but it has four genes annotated as CE1 that are very distinct from other fungal sequences and are potentially of bacterial origin (Supplementary Table 4).

CE16 genes are widely distributed across Agaricomycotina, with an average of seven copies per species; only *T. mesenterica* lacks CE16 genes (Supplementary Fig. 21, Supplementary Table 7). Some mycorrhizal species have lower than average numbers of CE16 genes, but there are also examples of mycorrhizal species with above average numbers of CE16 genes, such as *S. vermifera*, *T. calospora*, *P. croceum* and *P. rubicundulus*. Reconciliation analyses suggest that both CE1 and CE16 expanded early in the diversification of the Agaricomycetes. Expansions are reconstructed between nodes 14 and 16 for CE1 genes, while expansions are supported for the CE16 genes from node 13 to node 14, and from node 16 to node 19 (Supplementary Fig. 21). In addition, reconciliation analyses suggest that losses of CE1 genes occurred in multiple lineages leading to most of the ECM Agaricomycetidae.

Glycoside hydrolase family 5. The preliminary data set of GH5 included 1078 protein models. We followed the Aspeborg et al.⁷¹ classification to divide this very diverse dataset into smaller groups of proteins, putatively belonging to GH5 subfamilies 4, 5, 7, 9, 11, 12, 15, 16, 22, 23, 30, 49, and 50. In addition, we recovered two groups of sequences of uncertain classification, which we named groups A and B. Several subfamilies of GH5 genes had patchy distributions across the genomes studied, and were not subjected to phylogenetic and reconciliation analyses, including GH5_4 (with six proteins, including three in Cantharellales and three in ascomycetes), GH5_23 (two proteins in *A. nidulans*), GH5_11 and GH5_16 (ten proteins total, including seven in ascomycetes and one each in *A. delicata*, *C. cinerea* and *U. maydis*), models Dacsp1_119053 and Tremel1_61952 (which returned weak blastp hits to non-fungal eukaryotes only). Subfamilies 7 and 30, 49 and 50 were analyzed in groups because of their overall phylogenetic relationship⁷¹.

GH5 subfamilies GH5_12 (β -glucosylceramidases and β -glucosidases), GH5_49, GH5_50 (uncharacterized), GH5_22 (potential endo- β -1,4-glucanases), and GH5_15 (fungal cell wall related enzymes, β -1,6-glucanases) are relatively conserved in copy number, with one to three copies per genome, but a few species have lost the genes for those subfamilies (Supplementary Table 7). The only exceptions to this pattern are *G. marginata*, which has eleven GH5_22 genes, and *A. delicata* which has six genes for

subfamily GH5_15. Reconciliation analyses suggest conserved numbers for these subfamilies throughout the diversification of fungi, suggesting a critical role of these proteins for fungal biology (Supplementary Table 5).

In contrast, subgroups GH5_A and GH5_B, which we could not assign with certainty to any of the groups recognized by Aspeborg et al.⁷¹, show a scattered distribution across the genomes sampled here (Supplementary Table 7). Most of the species have no homologs in these groups; only ten species have genes for GH5_A, while eighteen species have GH5_B. GH5_A is represented by one or two genes in most species (with five and nine genes for *Dacryopinax* sp. and *M. larici populina*, respectively), and GH5_B is represented by one to three genes for some of the saprotrophs. The reconciliation results suggest also rather stable numbers for these genes during the diversification of fungi (Supplementary Table 9). GH5 subfamily GH5_9 (fungal cell-wall modifying enzymes) shows a different pattern of gene copy numbers in comparison to the subfamilies described above. All genomes included here contain GH5_9 genes, except *B. dendrobatitis*, but the number of genes per species ranges from three (for the Boletales) to 15 for *A. delicata* (Supplementary Table 9).

GH5 subfamilies GH5_5, GH5_7 and GH5_30 include endo- β -1,4-glucanases (GH5_5), endo- β -1,4-mannanases (GH5_7), and uncharacterized proteins (GH5_30) that are closely related to GH5_7. GH5_5 genes are widespread in Dikarya and the representation of genes for the Agaricomycotina species varies from zero to eighteen (*T. calospora*) copies (Supplementary Fig. 8, Supplementary Table 7), with generally higher abundance in saprotrophic species as well as mycorrhizal species in Cantharellales and Sebaciniales. In contrast, the ECM Agaricomycetidae have either lost the GH5_5 genes (*A. muscaria* and the Sclerodermatineae) or have a reduced number of homologs (one or two copies), which are scattered on different subclades of the GH5_5 phylogeny (Supplementary Fig. 8). Reconciliation analyses suggest losses of GH5_5 genes in lineages leading to several ectomycorrhizal species, including the Sclerodermatineae, the *Paxillus* species, *S. luteus*, *H. cylindrosporum*, *L. bicolor* and *A. muscaria* (Supplementary Fig. 8).

The combined dataset for GH5_7 and GH5_30 proteins indicates that proteins of these subfamilies are widespread across Agaricomycotina, while no copies are found for basal lineages in Basidiomycota and Ascomycota, as well as *P. blakeeslanus* and *B. dendrobatitis* (Supplementary Fig. 9, Supplementary Table 7). An examination of each subfamily separately (Supplementary Fig. 9) shows that GH5_30 has no sequences from ascomycetes, Polyporales, *P. strigosozonata* or *S. commune*. On the other hand, every basidiomycete mycorrhizal species has at least one GH5_30 sequence. GH5_7 includes sequences from both Ascomycetes and Basidiomycetes (Supplementary Fig. 9). GH5_7 sequences are found in Cantharellales and Sebaciniales, but many ectomycorrhizal Agaricomycetidae lack these genes, except *Paxillus* species and *S. luteus*. The reconciliation analysis for GH5_7 and GH5_30 suggests that their number increased early in the diversification of Agaricomycotina (node 12 to node 13, Supplementary Fig. 17), while additional expansions are suggested later (node 13 to node 28).

Proteins of subfamilies GH5_5 and GH5_7 are frequently modular including a CBM1 domain. The mapping of CBM1 domain for both families (Supplementary Figs. 8, 9) indicates that CBM1 domains can be found more frequently on proteins of white rot and soil saprotrophs, but are also frequent in the genomes of the orchid mycorrhizal species *S. vermifera* and *T. calospora*. In contrast, ECM genomes in Agaricomycetidae

not only have fewer genes for both subfamilies, but also the translated proteins frequently lack a CBM1, as has been shown before for brown rot species⁴¹. In our dataset, the only ECM species that possess a GH5_5 protein with a CBM1 domain are *L. bicolor* and *H. cylindrosporium* (Supplementary Fig. 9).

Discussion. Independent losses of genes related to saprotrophy in ECM Agaricomycetidae. Analyses of the first ECM basidiomycete genome, *L. bicolor*, suggested that transitions from saprotrophy to ECM symbiosis are accompanied by multiple losses of gene families involved in plant cell wall (PCW) decay^{41,91}. The present study increases the number of sequenced ECM Agaricomycetidae genomes by nine species, allowing us to address the generality of results obtained with *L. bicolor*. Focusing on 16 gene families related to saprotrophy, we found that on average, ECM species have 62 gene copies, whereas saprotrophic species have 120 copies (Supplementary Table 7), which supports the view that evolution of ECM is associated with a reduced capacity for saprotrophy. However, some species deviate from this pattern, such as the brown rot *Serpula lacrymans*, which has 63 gene copies for these 16 gene families, and the ECM *P. croceum*, which has 91 gene copies.

Gene losses in ECM lineages are not uniformly distributed across gene families encoding PCW degrading enzymes. Rather, losses are most pronounced in genes involved in the depolymerization of lignin and crystalline cellulose, which are the most recalcitrant macromolecules in the PCW. Compared to saprotrophs, ECM species have reduced numbers of copper dependent polysaccharide monooxygenases (LPMO). Within Agaricomycetidae, all ten ECM genomes lack GH6, and only two possess GH7 genes (one copy each in *P. croceum* and *H. cylindrosporium*). In contrast, the orchid mycorrhizal Cantharellales and Sebaciniales contain two to seven GH6 genes and one to 27 GH7 genes (Supplementary Fig. 15). Class II peroxidases (PODs) are also reduced in the ECM Agaricomycetidae (and are absent from the sampled Cantharellales and Sebaciniales, discussed below), with *H. cylindrosporium* being the only ECM species that still possesses PODs that appear to have ligninolytic capability, albeit in a reduced number compared to the ancestor of Hymenogastraceae (Supplementary Figs. 4, 12). Thus, reduction or loss of PODs and reduced dependence on crystalline cellulose as a carbon source appear to be unifying features of ECM fungi. Nonetheless, some ECM species may still be able to degrade lignin or lignin residues found in the soil, which is suggested by the POD genes found in *H. cylindrosporium* and other ECM species, such as *Cortinarius* and *Hygrophorus*⁹². In addition, other oxidative enzymes have been suggested to participate in degradation of lignin, lignin residues, and humic substances, such as laccases *sensu stricto*⁵⁷, HTPs, and DyPs⁹³⁻⁹⁵. All ECM fungi have maintained HTP and laccase genes, and many have DyP genes as well (Supplementary Figs. 12, 13). The maintenance of these genes suggests that they play important roles, but not all may be related to PCW degradation. For example, some laccase *sensu stricto* genes of *A. bisporus* are upregulated during growth on compost, possibly reflecting a role in degradation of humic materials, but others appear to be related to fruiting body development^{95,96}.

A complex picture emerges for the CAZymes related to the degradation of less recalcitrant macromolecules, such as amorphous cellulose, hemicellulose and pectin. Each ECM lineage has a different repertoire of the CAZY gene families studied here. An example is provided by the putative endoglucanases of families GH12, GH5_5, and GH5_22; the *Pisolithus* species lack genes for all three gene families, while *S. citrinum*

and *A. muscaria* have each retained one out of the three families, the *Paxillus* species and *L. bicolor* have retained two of the three families, and the rest of ECM species have retained copies of all three gene families (Supplementary Figs. 18, 19). Similar examples are found in xylanases (GH10), endomannanases (GH5_7), and the hemicellulose/pectin related GH43 enzymes (Table S7). In contrast, CE1 is almost uniformly absent from ECM genomes (*P. croceum* is the exception), while CE16, GH3, and GH28 are almost uniformly present in ECM genomes (except for *A. muscaria*, which has no GH28 genes). The maintenance of GH28 genes supports the suggested importance of GH28 for the ECM lifestyle⁸. Overall, the variation in repertoires of CAZymes and decay-related oxidoreductases across ECM species suggests that a diversity of saprotrophic abilities exists in these symbiotic lineages.

Origin(s) of ECM and retention of brown rot in Boletales. Previous analyses of Boletales suggested that ECM arose at least twice in the order from a paraphyletic group of brown rot ancestors⁹⁷. The brown rot *H. pinastri* was nested among ECM lineages, suggesting a possible reversal from ECM to saprotrophy⁹⁷. Results of phylogenomic analyses confirm the placement of *H. pinastri* among ECM species in the “core Boletales” (Sclerodermatineae, Suillineae, and Boletineae) (Supplementary Fig. 2). To assess whether *H. pinastri* represents a reversal from ECM to brown rot, we examined phylogenies of gene families encoding PCW-degrading enzymes. Twelve gene families (GH6, GH7, GH5_5, GH5_22, GH12, GH10, GH43, CE1, GH5_A, GH5_B, DyP, Lac/Fet3) are reconstructed as being present in the ancestor of the core Boletales (Supplementary Fig. 22). Parallel losses of these gene families are reconstructed in some or all of the lineages leading to the ECM members of the core Boletales, but *H. pinastri* retains members of all twelve gene families (Supplementary Fig. 22, Supplementary Table 7). Moreover, *H. pinastri* possesses fifteen copies of LPMO, while the ECM species of the core Boletales possess two to five copies each. The LPMO phylogeny indicates that the LPMO genes of *H. pinastri* represent an ancestral diversity of gene lineages, rather than a recent expansion (Supplementary Figs. 7, 16, 23). Similar patterns are evident in the GH28, GH5_5 and GH3_B gene families (Supplementary Figs. 8, 9, 10, 11). Taken together, these results suggest that the ancestor of the core Boletales had substantial saprotrophic ability, which has been retained in *H. pinastri*. Nevertheless, this interpretation does not rule out the possibility that the ancestor of the core Boletales could have been capable of forming ECM association. Among extant ECM species, some have considerable arsenals of decay enzymes. For example, *P. croceum* (Atheliales) has 91 gene copies in the 16 decay-related gene families mentioned previously, compared to 89 gene copies in *P. crispa*, which is a closely related white rot species in the Amylocorticiales. (The Atheliales-Amylocorticiales clade is the sister group of the Boletales.) Gene tree/species tree reconciliation analyses suggest a progressive loss of decay-related genes in multiple lineages of ECM Boletales, perhaps indicating that saprotrophic abilities decline gradually after establishment of ECM symbioses. Consistent with this view, a recent study suggests that ECM Boletales could still have an active Fenton reaction⁹⁸, which might facilitate losses of ECM associations.

Diversification of PCW degrading enzymes in the early evolution of Agaricomycetes. Previous studies have suggested that many PCW degrading enzymes, and particularly POD genes, diversified early in the evolution of Agaricomycetes¹⁹. Here, we expanded the sampling of genomes for early-diverging lineages in Agaricomycetes to include

Cantharellales and Sebaciales. Results of Notung analyses suggest that expansions occurred around node 13 (the common ancestor of Dacrymycetes and Agaricomycetes) and node 14 (the ancestor of Agaricomycetes) in multiple gene families encoding PCW degrading enzymes, including those encoding enzymes related to pectin, hemicellulose and amorphous cellulose degradation (GH28, GH43, GH3, GH5_5, and GH5_7_30), as well as Lac/Fet3 genes, and gene families related to crystalline cellulose degradation (GH7, LPMO) (Supplementary Figs. 12-21). Other genes families that show expansions around these nodes include xylanases (GH10), esterases (CE16), copper radical oxidases (CRO 1, 2, 6 and GLX) and DyP genes. Later, around node 16, expansions are suggested for laccases *sensu stricto* and esterases (CE1). POD and HTP genes are suggested to have started diversifying around node 18, prior to the divergence of Auriculariales. The common ancestor of Dikarya (node 3) is reconstructed as having 22 GH3 genes, 16 GH28 genes, and 26 GH43 genes, suggesting that diversification in these gene families began before the origin of Agaricomycetes, and continued early in the evolution of Agaricomycetes (Supplementary Fig. 19, 20). Overall, results of Notung analyses suggest that plant cell-wall degradation machinery diversified gradually during the early evolution of Agaricomycotina, starting with hemicellulose/pectin degrading enzymes, followed by expansions of crystalline cellulose degrading enzymes, and eventually by the diversification of the major lignin degrading enzymes (POD) along with HTPs.

Symbiosis in the Cantharellales and Sebaciales, in contrast to ECM Agaricomycetidae, is not associated with reduced numbers of saprotrophic genes. The orchid mycorrhizal and root endophytic *Sebacina vermifera* and *P. indica* (Sebaciales) and the orchid symbiont *T. calospora* (Cantharellales) have rich repertoires of CAZymes related to PCW degradation, including GH6, GH7 and LPMO, which attack crystalline cellulose, endoglucanases of the family GH5_5, and gene families related to hemicellulose degradation, such as CE1 and CE16, GH10, GH43 and mannanases of the family GH5_7. Decay-related oxidoreductases present a more heterogeneous picture; PODs and GLXs are absent and laccases *sensu stricto* are only found in the Sebaciales, but both Sebaciales and Cantharellales have HTPs, DyPs and other CROs (Supplementary Table 7). Gene tree/species tree reconciliation analyses suggest that the symbiotic Sebaciales and Cantharellales sampled here have maintained large arsenals of CAZymes, which might have been inherited from saprotrophic ancestors, but these lineages may have originated prior to the evolution of ligninolytic PODs or the GLX genes. The presence of numerous CAZs in symbiotic Sebaciales and Cantharellales further demonstrates that transitions from saprotrophy toward symbiosis are not uniformly associated with losses of PCW degrading enzymes, and blurs the distinction between decayer and mycorrhizal lifestyles. A good example in Cantharellales is provided by *B. botryosum* and *T. calospora*. *Botryobasidium* species have been associated with white rot⁹⁹, but a comparison of *B. botryosum* and *T. calospora* indicates only minor differences in copy numbers for the gene families examined here. It is therefore possible that *B. botryosum* is capable for forming mycorrhizal associations, and that *T. calospora* can obtain carbon nutrition as a saprotroph.

1.7. Transcript profiling Biological material for RNA-Seq

Amanita muscaria–*Populus tremula x tremuloides* ectomycorrhizal interaction. *A. muscaria* ([L.] ex Fr.) Hooker strain MEII was isolated from fruiting bodies collected under Norway spruce obtained in the Schönbuch forest near Tübingen, Germany. Stock cultures were grown on MMN-agar at 8°C according to ¹⁰⁰ and were used to inoculate 300 ml flasks containing 80 ml of modified MMN ¹⁰¹ medium. Liquid cultures were grown at 20°C on a rotary shaker (120 rpm) and were used as inoculum for ectomycorrhiza formation in Petri dishes according to ¹⁰². *Populus tremula x tremuloides* (clone T89) ¹⁰³ cuttings were rooted in MS-medium (Sigma, St. Louis, MI, USA) under sterile conditions for about 2 month. Rooted plants were transferred under axenic conditions onto the agar surface of Petri dishes containing sugar-free modified MMN medium with 300 µM ammonium as sole nitrogen source such that the root system was inside and the shoot outside of the Petri dish. Inoculated and non-inoculated control plants were grown in parallel in small plastic greenhouses for additional six weeks at 18°C, 12 h day/night periods with 150 µE x m⁻² x s⁻¹ illumination. Fine roots of non-inoculated plants, and fully developed ectomycorrhizas were isolated around noontime, frozen in liquid nitrogen, and stored at –80°C.

Total RNA was isolated according to ¹⁰⁴. In brief, frozen material was ground to a fine powder in a mortar with pestle under liquid nitrogen and resuspended in five volumes of 4 M guanidine isothiocyanate, 96 mM β-mercaptoethanol, 25 mM sodium acetate, pH 6.0. Insoluble material was removed by centrifugation (20 min, 20,000 g, 20°C), and the supernatant was mixed with one volume of 5.7 M CsCl, 25 mM sodium acetate, pH 6.0 in an ultracentrifuge tube. The RNA was sedimented by centrifugation (20 h, 200,000 g, 20°C) and the pellet was resuspended in 200 µl DEPC-treated H₂O. After transfer into an Eppendorf tube and the addition of 20 µl 3 M sodium acetate, 500 µl ethanol, the RNA was precipitated by centrifugation (40 min, 14,000 g, 20°C), air dried and resuspended in 50 µl DEPC-treated H₂O. Aliquots were frozen in liquid nitrogen and stored at -80 °C.

Tulasnella calospora–*Serapias vomeracea* orchid mycorrhiza. The fungal isolate is *T. calospora* strain AL13, deposited at the Mycotheca Universitatis Taurinensis collection (MUT4182) at the Department of Life Sciences and Systems Biology, University of Turin, Italy. This strain was isolated from the roots of *Anacamptis laxiflora* collected in a meadow in northern Italy (*Festuco-Brometalia*, at 410–450 m a.s.l.). The characteristics of the site and the identification of this fungal isolate are described in more detail in ¹⁰⁵. The fungus was grown as a free living culture for 14 days at 24°C on Oatmeal-agar medium (3% oat flakes, 1.5% agar); ten plugs of 0.5 cm of diameter were then aseptically collected from the actively growing margin of the colony, and inoculated in 100 mL Erlenmeyer flasks containing 50 mL of liquid Oatmeal modified medium (0.75% oat flakes). Flasks were then incubated on an orbital shaker at 150 rpm for 14 days at 24°C, and three separate RNA extractions were performed. Symbiotic germination of *Serapias vomeracea* seeds was performed in plastic petri plates (9 cm in diameter, 1.5 cm in height) containing Oatmeal-agar medium (0.3% milled oat, 2.0% agar) inoculated with *T. calospora* following the protocol described in ¹⁰⁶. Protocorms were incubated at 20°C in darkness and collected 30 days after inoculation, immediately frozen in liquid nitrogen and stored at -80°C before RNA extraction.

Total RNA was extracted from 200 mg aliquots of *T. calospora* free living mycelia grown in Oatmeal modified liquid medium. Mycelium was mechanically ground in liquid nitrogen and RNA was extracted in Tris-HCl extraction buffer (100 mM Tris-HCl pH 8, 100 mM NaCl, 20 mM Na-EDTA, 0.1% PVP, 1% Na-laurylsarcosine sodium salt dissolved in DEPC-treated deionised water), mixed 1:1 with phenol (pH 4.5-5; Roti-Phenol, Roth A980). A phenol:chloroform:isoamyl alcohol (25:24:1, v/v/v) and a chloroform extraction followed, with 5 min centrifugation at 14000 rpm and 4°C after each extraction step. RNA was precipitated with isopropyl alcohol at -80°C for 30 min, followed by 30 min centrifugation at 14,000 rpm and 4°C. The pellet was then re-suspended in DEPC treated water: 6M LiCl solution (1:1, v/v) and precipitated overnight at 4°C. After 30 min centrifugation at 14,000 rpm and 4°C, the RNA was rinsed with 70% ethanol, centrifuged for 5 min at 14,000 rpm and 4°C, air dried on ice, re-suspended in DEPC-treated water and quantified using Nanodrop 1000 (Thermo Scientific, USA) and Qubit 2.0 (Life Technologies, Italy). RNA integrity was checked using a Bioanalyzer 2100 (Agilent Technologies, Italy).

Total RNA was extracted from mycorrhizal *S. vomeracea* protocorms using the CTAB method. Hundred mg aliquots of well-developed protocorms were mechanically ground in liquid nitrogen and RNA was extracted in CTAB buffer (2% CTAB, 2% PVP, 100 mM Tris-HCl pH 8.0; 25 mM EDTA pH 8; 2 M NaCl) at 65°C. 2% PVPP was added to the buffer 1 h before the RNA extraction and 2% β-mercaptoethanol was added to the buffer just before use. The homogenate was incubated at 65°C for 5 min and extracted twice with chloroform:isoamylalcohol (24:1 v/v), each extraction followed by 10 min centrifugation at 5,000 rpm and room temperature. RNA was precipitated overnight in 10 M LiCl at 4°C. After centrifugation at 10,000 rpm at 4°C for 20 min, the pellet was dissolved in SSTE buffer (1 M NaCl, 0.5% SDS, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8) and extracted with an equal volume of phenol (pH 4.5-5; Roti-Phenol, Roth A980): chloroform:isoamyl alcohol (25:24:1, v/v/v), followed by chloroform:isoamyl alcohol (24:1, v/v). Each extraction step was followed by 10 min centrifugation at 10,000 rpm and 4°C. Two volumes of 100% ethanol were then added

and RNA was precipitated for 2 hours at -20°C. After centrifugation for 20 min at 10000 rpm and 4°C, the pellet was washed with 80% ethanol, centrifuged 10 min at 10000 rpm and 4°C, air dried on ice, re-suspended in DEPC-treated water and quantified using Nanodrop 1000 (Thermo Scientific, USA) and Qubit 2.0 (Life Technologies, Italy). RNA integrity was checked using a Bioanalyzer 2100 (Agilent Technologies, Italy).

Piloderma croceum-Quercus robur ectomycorrhiza. *P. croceum* (DSMZ 4824, ATCC MYA-4870) was cultivated at 23°C on MMN (modified Melin-Norkrans medium) agar¹⁰⁷ with 10 g l⁻¹ glucose, in the dark, or under 16:8 h day-to-night-cycle (photon flux density 100 µmol m⁻² s⁻¹). For ECM synthesis, a modified culture system of the oak clone DF159 (*Quercus robur* L.) established by Herrmann et al.¹⁰⁸ was used and harvest was performed as described¹⁰⁹. After eight weeks on agar or in ECM synthesis cultures, samples were submerged into liquid nitrogen and stored at -80°C. Total RNA was extracted from the samples using the MasterPure Plant RNA Purification Kit (Epicentre, Hessisch Oldendorf, Germany) according to the manufacturer's instructions. In total, 50 mg of fungal mycelium or ECM root tips were used for each extraction. RNA quantity and quality were estimated using a NanoDrop spectrophotometer (Thermo Scientific, Passau, Germany), gel electrophoresis, and a Nano Chip in a Bioanalyzer 2100 (Agilent, Böblingen, Germany).

Sebacina vermifera-Arabidopsis thaliana endophytic interaction. *Arabidopsis thaliana* seeds (Ecotype Columbia-0) were incubated for 5 min in 70% ethanol, surface sterilized for 5 min with 6% sodium hypochlorite and washed 6 times for 5 min in sterile water. After stratification for 3 days at 4°C in the dark on 1/10 PNM medium, *Arabidopsis* seedlings were grown for 14 days under sterile conditions in a phytochamber (Vötsch, Balingen-Frommern, Germany) at long day conditions (day: 16 h, 23°C, 350 µmol m⁻² s⁻¹; night: 8 h, 18°C). *S. vermifera* strain was grown on MYP (7 g malt extract, 1 g peptone, 0.5 g yeast extract and 12 g agar) agar plates or liquid MYP with 120 rpm shaking at 25°C. Seven-day-old *S. vermifera* culture was filtered through miracloth filter and the mycelium was washed with 0.9% NaCl. Mycelium was crushed for 10 seconds in fresh MYP using a sterile blender (Microtron MB 550, Kinematica AG). 20 ml of crushed mycelium was inoculated in 130 ml MYP and regenerated for 3 days at 25°C with 130 rpm shaking. For inoculation of *S. vermifera* with *Arabidopsis thaliana*, fourteen-day-old germlings were inoculated with 5 g crashed fungal mycelium in 5 ml 0.9% NaCl solution for 2 h or 0.9% NaCl mock treated. *S. vermifera* inoculated plants, of approximately the same size, were transferred to square petri dishes containing 1/10 PNM and the roots were treated with either 1 ml of crushed fungal biomass (1 g ml⁻¹ in 0.9% NaCl solution) per 20 seedlings or mock treated. The first four cm of the roots below the seed were excised and immediately frozen in liquid nitrogen for RNA extraction at 3, 7 and 14 days post inoculation (dpi). For each time point, roots from 80 to 100 plants were harvested and the experiments were performed in three independent biological repetitions. Total RNA from 200 mg of ground material was extracted using TRIzol (Invitrogen, Karlsruhe, Germany) following the manufacturer's instructions. Prior RNA extraction, plant materials harvested at 3, 7 and 14 dpi from the same

independent biological experiment were pooled together. As a control, total RNA from three independent biological replicates of seven-day-old *S. vermifera* grown in MYP medium was used. RNA samples were additionally precipitated with ethanol. In brief, 1/10 volume of 3 M NaOAc and 3 volumes of ethanol were added to RNA solution. After incubation at -20°C overnight, the RNA pellet was centrifuged at 13000 rpm for 30 min and washed once with 70% ethanol (diluted in DEPC ddH₂O) and spin down for 10 min. The RNA pellet was then air-dried and resuspended in RNase-free water with a final concentration of 1 µg/µl. Purity and quantity of RNA samples were measured using the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and Agilent RNA 6000 Nano Kit and Agilent 2100 Bioanalyzer following the manufacturer's protocol (Agilent, Santa Clara, USA).

Suillus luteus-*Pinus sylvestris* ectomycorrhiza. *Suillus luteus* was grown for 7 days at 23°C on sterilized cellophane sheets sitting in Fries medium plates (4 plugs/plate) and used further for inoculation. Free-living mycelium was harvested after 7 days of growth in the same Fries medium by peeling off the mycelium and snap frozen it in an Eppendorf tube in liquid nitrogen. Samples were then stored at -80°C. *Pinus sylvestris* seeds were kindly provided by Office National des Forêts, France. Four hundred of them were rinsed in a sieve with sterile MQ water, soaked in 34ml of a 12% TWEEN20 solution for 20 min, rinsed with distilled water, soaked in 10% H₂O₂ for 15 min, rinsed with distilled water, soaked in 70% ethanol for 5 min, rinsed with distilled water, and placed in a tray containing sieved, acid-washed and moistened perlite. Seedlings were grown for four weeks in a growth chamber (day:night regime of 18 h light at 22°C and 6 h dark at 15°C) and watered every week with Ingestad's nutrient solution slightly modified to induce P limitation in the pine seedlings. The macronutrient weight proportions were 100 N:9 P:54 K:6 Ca:6 Mg:9 S^{110,111}. *Suillus luteus* was inoculated on *Pinus sylvestris* seedlings using the sandwich technique described in Van Tichelen & Colpaert¹¹². Mycorrhizal root tips were harvested on 40 days-old seedlings. Two sets of about 50 ectomycorrhizal root tips were harvested from 10 seedlings (5 seedlings each). Each tip was harvested with tweezers and immediately frozen in an Eppendorf tube floating in liquid nitrogen after separation from the root. RNA extraction was performed immediately after harvesting.

For ECMs, total RNAs were extracted from the frozen material using the protocol of Chang et al.¹¹³ that is well adapted to pine samples containing high levels of high-polyphenols (http://ipmb.sinica.edu.tw/affy/document/Pine_Tree_Method.pdf). For mycelium, 100 mg of frozen sample was extracted using the Plant RNA mini kit. RNA quality was checked using Bioanalyzer.

Paxillus involutus-*Betula pendula* ectomycorrhiza. Mycelium patches and mycorrhizas were prepared as described in Wright et al.¹¹⁴. In brief, ECM association between birch and *P. involutus* was synthesized using a cellophane-agar Petri-dish system. Seeds of birch were surface-sterilized and then transferred aseptically to water agar plates until germination occurred. Nine-day-old seedlings were aseptically transferred to the edge of 9-day-old colonies of *P. involutus* growing on sheets of autoclaved cellophane placed

over MMN agar containing 5.55 mg of glucose. After 4 weeks, mycorrhizal seedlings were transferred to pots containing unamended sphagnum peat to which four plugs of *P. involutus* mycelium were added. These pots were enclosed within a propagator, to maintain high humidity, and placed in a growth chamber at a day/night temperature of 18/15°C with a 16-h photoperiod and 80% relative humidity for 4 weeks. After approximately 3 weeks, once the extramatrical mycelium had colonized approximately two-thirds of the microcosm, two nutrient patches were placed into each chamber in advance (1 cm) of the mycelial front. Each nutrient patch consisted of a shallow plastic dish filled with autoclaved fine quartz sand (Sigma-Aldrich Sweden AB, Stockholm, Sweden) to which 1.5 ml of 1 mM (NH₄)₂SO₄ was applied initially. Subsequently, 0.5 ml of 1 mM (NH₄)₂SO₄ was applied weekly for 4 weeks to allow extensive hyphal development within the patch and the formation of mycelial rhizomorphs which link the mycelium within the nutrient patch to the mycorrhizal root tips. In order to obtain a sufficient sample for total RNA extraction, material of each tissue type was pooled from three microcosms to produce one biological replicate. All material was immediately frozen in liquid N₂ and stored at –80°C until use. Total RNA was immediately isolated from each sample using the RNeasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions, except that PEG 6000 (20 mg ml^{–1}) (Merck KGaA, Darmstadt, Germany) and β-mercaptoethanol (10 μl l^{–1}) were added to the RLC buffer. Total RNA preparations were inspected using a BioAnalyzer and the Total RNA Nano Series II kit (Agilent).

Hebeloma cylindrosporum–*Pinus pinaster* ectomycorrhiza. The wild-type diploid strain TV98 IV3 of *H. cylindrosporum* was used in this work. Pure culture grown mycelia were six-day old thalli, obtained on the MNM supplemented with 0.5 g l^{–1} fructose¹¹⁵. Agar medium was covered with a cellophane sheath to prevent mycelial growth inside the agar medium and to make it easier mycelium recovery. Synthetic ectomycorrhizas were obtained as previously described¹¹⁶ and collected three weeks after inoculation. *H. cylindrosporum* total RNA from free living mycelia or ectomycorrhizas was extracted using the RNeasy Plant Mini Kit (Qiagen). PEG 8000 (20 mg ml^{–1}) and β-mercaptoethanol (10 μl l^{–1}) were added to the RLC buffer. Samples were subsequently purified using the Total RNA Purification – Nucleobond ARN/ADN 80 kit (Macherey-Nagel) by following the manufacturer's instructions.

Oidiodendron maius–*Vaccinium myrtillus* ericoid endomycorrhiza. *O. maius* strain Zn was grown on a modified MMN medium containing 0.5 g L^{–1} KH₂PO₄, 0.1 g L^{–1} BSA (bovine serum albumin, SIGMA), 0.066 g L^{–1} CaCl₂·2 H₂O, 0.025 g L^{–1} NaCl, 0.15 g L^{–1} MgSO₄·7 H₂O, 0.1 g L^{–1} thiamine-HCl, 0.001 g L^{–1} FeCl₃·6 H₂O, and 10 g L^{–1} agar. BSA and thiamine-HCl were added to the medium as filter sterilized solutions just before plating. The final pH was 4.7. Prior to fungal inoculation, sterile cellophane membranes were placed aseptically on the agar surface to provide a convenient means of removing the mycelium from the plate. Plates (15 cm in diameter, 2.5 cm in height) were inoculated with 5 mm fungal plugs. Fungal colonies were removed after 45 days by peeling the biomass from the cellophane membrane. Several plates were prepared in order to get enough fungal material to perform three separate RNA extractions for the

RNA sequencing. Endomycorrhizae were synthesized aseptically in plastic Petri plates (15 cm in diameter, 2.5 cm in height) containing the same medium reported above for fungal cultures. Control, non-inoculated *Vaccinium myrtillus* plants were grown on the same medium but 0,075 g L⁻¹ (NH₄)₂HPO₄ was added instead of the BSA. Axenic *V. myrtillus* seedlings were obtained from surface-sterilized (70% ethanol, v/v, plus 0.2% Tween 20 for 3 min; rinsed twice with sterile water and 0.25% sodium hypochlorite for 15 min, with three additional sterile water rinses) bilberry seeds (obtained by Les Semences du Puy, Le Puy-En-Velay, France) germinated on water and 1% agar Petri plates in the dark for 2 weeks before transfer to the growth chamber for 1 month. A conidia suspension in sterile deionised water was prepared from 1 month-old fungal cultures grown on Czapek-glucose solid medium (NaNO₃ 3 g L⁻¹, K₂HPO₄*3H₂O 1.31 g L⁻¹, MgSO₄*7H₂O 0.5 g L⁻¹, FeSO₄*7H₂O 0.01 g L⁻¹, KCl 0.5 g L⁻¹, glucose 20 g L⁻¹, agar 10 g L⁻¹). All reagents were purchased from Sigma. The medium was adjusted to pH 6 with the addition of 1M HCl. The conidia suspension was distributed in the bottom half of the MMN petri plates. Ten germinated *V. myrtillus* seedlings were then transferred aseptically in the MMN plates at the upper limit of the surface covered by the conidia suspension. Finally the plates were sealed and placed in a growth chamber (16-h photoperiod, light at 170 μmol m⁻² s⁻¹, temperatures at 23°C during the day and 21°C overnight). Several plates were prepared in order to get enough root material to perform three separate RNA extractions for the RNA sequencing. The root systems were observed after a 1.5-month incubation period and the percentage of mycorrhization was evaluated. Prior the observation, each *V. myrtillus* hair-root system from the synthesis plates was stained in 0.1 % (wt/vol) cotton blue (methyl blue) overnight and destained with 80% lactic acid. Whole roots were then mounted in the destaining solution, observed using a Nikon Eclipse E400 optical microscope, and photographed.

RNA was extracted from *O. maius* mycelia and from *O. maius*-inoculated *V. myrtillus* roots 45 days after cultures start. Mycelium (~100 mg) was collected, snap frozen in liquid nitrogen, mechanically ground and RNA was extracted in Tris-HCl extraction buffer (100 mM Tris-HCl pH 8, 100 mM NaCl, 20 mM Na-EDTA, 0.1% PVP, 1% Na-laurylsarcosine sodium salt dissolved in DEPC-treated deionised water), mixed 1:1 with phenol (pH 4.5-5; Roti-Phenol, Roth A980). A phenol:chloroform:isoamyl alcohol (25:24:1) extraction step and a chloroform extraction step followed. Each extraction step was followed by 5 min centrifugation at 14,000 rpm at 4°C. An isopropyl alcohol precipitation step at -80°C for 30 min was then performed followed by 30 min centrifugation at 14,000 rpm at 4°C. The pellet was then re-suspended in a 1:1 DEPC treated water:6 M LiCl solution and precipitated overnight at 4°C. After 30 min centrifugation at 14,000 rpm at 4°C the RNA was rinsed with 70% ethanol, centrifuged for 5 min at 14,000 rpm at 4°C, air dried on ice, re-suspended in DEPC-treated water and quantified using the nanodrop and the bioanalyzer.

Total RNA was extracted from *V. myrtillus* mycorrhizal roots using the CTAB method. *V. myrtillus* mycorrhizal roots (~100 mg) were collected, snap frozen in liquid nitrogen, mechanically ground and RNA was extracted in warm (65°C) CTAB buffer (2% CTAB, 2% PVP, 100 mM Tris-HCl pH 8.0; 25 mM EDTA pH 8; 2 M NaCl); 2% PVPP was added to the buffer 1 h before performing the extraction and 2% β-mercaptoethanol was added to the buffer just before use. The extraction suspension was incubated at 65°C for 5 min. Two chloroform/isoamylalcohol (24:1) extraction steps followed. Each extraction step was followed by 10 min centrifugation at 5,000 rpm at

room temperature. An overnight precipitation step in 10 M LiCl at 4°C followed. After a centrifugation at 10,000 rpm at 4°C for 20 min, pellet was dissolved in SSE buffer (1 M NaCl, 0.5% SDS, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8). Extraction steps with an equal volume of 1:1 phenol (pH 4.5-5; Roti-Phenol, Roth A980): chloroform/isoamyl alcohol (24:1) and of chloroform/isoamyl alcohol (24:1) followed. Each extraction step was followed by 10 min centrifugation at 10,000 rpm at 4°C. Two volumes of 100% ethanol were then added and a precipitation step of 2 hours at -20°C followed. After a centrifugation of 20 min at 10,000 rpm at 4°C, the pellet was washed with 80% ethanol, centrifuged 10 min at 10,000 rpm at 4°C, air dried on ice, re-suspended in DEPC-treated water and quantified using the nanodrop and the bioanalyzer.

Serpula lacrymans-*Pinus sylvestris* interaction. *S. lacrymans* was grown on MMN agar medium or on shavings of *Pinus sylvestris* sapwood as described in Eastwood *et al.*⁵⁵. Interaction with *Picea sylvestris* seedlings was carried out as described in Eastwood *et al.*⁵⁵.

RNA Sequencing

Preparation of libraries from total RNA and 2 x 100bp Illumina HiSeq sequencing (RNA-Seq) was performed at the IGA Technology Services facilities (Udine, Italy) for *O. maius*, *T. calospora*, *S. vermifera*, *P. croceum* and *P. involutus*) or at the JGI for *H. cylindrosporum* and *A. muscaria*. Raw reads were trimmed and aligned to the respective reference transcripts available at the JGI MycoCosm database (<http://genome.jgi-psf.org/programs/fungi/index.jsf>) using CLC Genomics Workbench v6. For mapping, the minimum length fraction was 0.9, the minimum similarity fraction 0.8 and the maximum number of hits for a read was set to 10. The unique and total mapped reads number for each transcript were determined, and then normalized to RPKM (Reads Per Kilobase of exon model per Million mapped reads). Intact pairs were counted as one, broken pairs were ignored. A summary of the aligned reads is given in Supplementary Table 11. The complete data sets were submitted to GEO (Superseries GSE63947)

Data analysis. To identify differentially regulated transcripts in mycorrhizal tissues compared to free-living mycelium the Baggerly *et al.*'s Test¹¹⁷ implemented in CLC Genomic workbench was applied to the data. Baggerley *et al.*'s test¹¹⁷ compares the proportions of counts in a group of samples against those of another group of samples. The samples are given different weights depending on their sizes (total counts). The weights are obtained by assuming a Beta distribution on the proportions in a group, and estimating these, along with the proportion of a binomial distribution, by the method of moments. The result is a weighted t-type test statistic. In addition Benjamini & Hochberg multiple-hypothesis testing corrections with False Discovery Rate (FDR) were used. Transcripts with a more than 5-fold change and a FDR corrected *p*-value <0.05 were kept for further analysis.

Additional information on RNA-Seq analysis can be found in the CLC Genomic Workbench online manual at:
http://www.clcsupport.com/clcgenomicsworkbench/current/index.php?manual=RNA_Seq_analysis.html

For *L. bicolor* and *S. lacrymans*, previously published^{55,91} microarray data were re-analyzed (GSE9784, GSE27839, GSE63929). Both microarrays had been designed using version 1 of the respective genomes. Therefore, the 60mer-probe sequences from each array were blasted against version 2 of the respective genome to identify the corresponding new gene models and only transcripts with perfect matches were kept in the data set. As for RNA-Seq data, transcripts with a more than 5-fold change in transcript concentration compared to control mycelium and a FDR corrected *p*-value <0.05 were considered as significantly differentially regulated.

Symbiosis-regulated transcripts analysis (Fig. 2).

To assess whether symbiosis-regulated transcripts were either conserved among fungal species or were lineage-specific (i.e., orphan genes with no similarity to known sequences in DNA databases), their protein sequences were queried against the protein repertoires of 55 fungal genomes using BLASTP with e-value $1e^{-5}$. Proteins were considered as orthologs of symbiosis-regulated transcripts pending they showed 70% coverage over the regulated sequence and at least 30% amino acid identity.

The genomes of the mycorrhizal *Meliniomyces bicolor* (Melbi), *Meliniomyces variabilis* (Melva), *Cenococcum geophilum* (Cenge), *Laccaria amethystina* (Lacam), *Cortinarius glaucopus* (Corgl), *Tricholoma matsutake* (Trima) and *Boletus edulis* (Boled) were added to the initial 49 genomes listed in Table S1. These genomes have been released after the main phylogenomic analyses (i.e., proteins clustering and organismal phylogeny). These genomes are available at the MycoCosm database (<http://genome.jgi-psf.org/programs/fungi/index.jsf>) and will be published elsewhere.

The sequence similarity matrices were clustered and the results displayed as heat maps using HeatPlus in R (<http://www.bioconductor.org/packages/release/bioc/html/Heatplus.html>). The hierarchical clustering was done by using a binary distance metric and ward clustering method (Supplementary Fig. 25 & 26).

The distribution of the symbiosis-regulated transcripts within each cluster were quantified according to their putative function as CAZymes, small (<300 AA) secreted proteins (SSPs), with KOG classification or without KOG (Supplementary Fig. 25 & 26). In addition to assess their enrichment in the set of symbiosis-regulated transcript set (if any), the percentage of secreted proteins and SSPs in symbiosis-regulated transcript clusters was compared to their relative percentage in the total gene repertoire.

1.8. MCL/CAFE

Multigene families were predicted from 742,804 predicted proteins found in the 49 genomes (Supplementary Table 1) using the MCL algorithm¹¹⁸ with an inflation parameter set to 2.0. As a result, 4,939 protein families phylogenetically relevant (containing at least ten species and 40 sequences) were identified. Multigene families were analyzed for evolutionary changes in protein family size using the CAFE program¹¹⁹. The program uses a random birth and death process to model gene gain and loss across a user specified tree structure. The distribution of family sizes generated under the random model provides a basis for assessing the significance of the observed family size differences among taxa (*p*-value 0.001). CAFE estimates for each branch in the tree whether a protein family has not changed, has expanded or contracted. The phylogenetic tree used is the one constructed according method in paragraph 1.4. None of the multigene families showed expansion or contraction in ALL mycorrhizal

symbionts. Also, as shown in Table S5, none of the multigene families are in expansion in the Boletales compare to the Agaricales.

1.9. CAZyme Annotation

Identification of carbohydrate active enzyme (CAZyme) and auxiliary redox enzyme (AA) families in the 49 genomes was achieved using the CAZy database (www.cazy.org) annotation pipeline⁷⁰. For each fungus, we have listed the number of representatives of each CAZy and AA family. Lignocellulose-degrading and secondary-metabolism genes found in the 49 fungal genomes representing white-rot, brown-rot, litter decaying, ectomycorrhizal, orchid mycorrhizal, ericoid and endophytic lifestyles are listed in Supplementary Table 7. In Supplementary Fig. 3, we only selected CAZymes families targeting cellulose (GH6, GH7, GH12, GH44, GH45), hemicellulose (GH10, GH11, GH26, GH30, GH51, GH74) and pectin (GH43, GH28, GH53, GH78, GH88, GH105, PL1, PL3, PL4, CE8, CE12). We performed a double clustering based on Bray-Curtis distances (i) between organisms according to their family distribution and (ii) between families according on their distribution pattern in the different genomes. Distances were computed using GINKGO¹²⁰ and the distance trees were constructed with FastME (<http://www.atgc-montpellier.fr/fastme/>).

1.10. Secretome

Secreted proteins were identified using a custom pipeline including the SignalP v4¹²¹, WolfPSort¹²², TMHMM, TargetP¹²³, and PS-Scan algorithms¹²⁴. Proteins predicted to contain a signal-peptide, which not localized within endoplasmic reticulum or mitochondria and displaying less than one transmembrane domain were considered as part of the secretome.

2. References

1. Brun A., Chalot, M., Finlay, R.D. & Söderström B Structure and function of the ectomycorrhizal association between *Paxillus involutus* (Batsch) Fr. and *Betula pendula* Roth.). I. Dynamics of mycorrhiza formation. *New Phytol* 129: 487-493(1995)
2. Gnerre, S. *et al.* High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proceedings of the National Academy of Sciences*, Vol. 108, No. 4. pp. 1513-1518, doi:10.1073/pnas.1017351108 (2011)
3. Zerbino, D.R. & Birney, E. *Genome Research*, Vol. 18, No. 5. pp. 821-829, doi:10.1101/gr.074492.107. (2008)
4. Terpolilli, J. *et al.* Genome sequence of *Ensifer medicae* strain WSM1369; an effective microsymbiont of the annual legume *Medicago sphaerocarpos*. *Stand Genomic Sci.* Dec 17;9(2):420-30. (2013)
5. Kang, Y.J. *et al.* Genome sequence of mungbean and insights into evolution within *Vigna* species. *Nat Commun.* Nov 11;5:5443. doi: 10.1038/ncomms6443 (2014)
6. Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, Vol. 437, No. 7057.), pp. 376-380, doi:10.1038/nature03959 (2005)
7. LaButti, K., Foster, B., Han, C., Bretin, T., & Lapidus, A. Gap Resolution: A Software Package for Improving Newbler Genome Assemblies. (2009).
8. Martin, F. *et al.* Perigord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* 464, 1033-1038, doi:10.1038/nature08867 (2010).
9. Grigoriev, I.V., Martinez, D.A. & Salamov, A. A. Fungal genomic annotation. In Aurora, DK, RM Berka, and GB Singh (Vol Eds), *Applied Mycology and Biotechnology: Vol. 6. Bioinformatics* (pp 123–142) Amsterdam: Elsevier. (2006)
10. Kuo, A., Bushnell, B. & Grigoriev, I.V. Fungal Genomics: Sequencing and Annotation. In Martin, FM (Vol Ed), *Advances in Botanical Research: Vol. 70. Fungi* (pp 1-52) Oxford, United Kingdom: Elsevier. (2014)
11. Grigoriev, I.V. *et al.* The Genome Portal of the Department of Energy Joint Genome Institute, *Nucleic Acids Res.* 40(1):D26-32 (2012)
12. Jurka, J. Repbase update: a database and an electronic journal of repetitive elements. *Trends in Genetics* 16:418-20. (2000)
13. Price, A.L., Jones, N.C & Pevzner, P.A. De novo identification of repeat families in large genomes. *Bioinformatics* 21:i351–i358. (2005)
14. Bateman, A. The Pfam protein families database. *Nucleic Acids Research* 32:138D-141. (2004)
15. Birney, E. & Durbin R. Using GeneWise in the *Drosophila* annotation experiment. *Genome Research* 10:547–548. (2000)
16. Salamov, A.A. & VV Solovyev Ab initio Gene Finding in *Drosophila* Genomic DNA. *Genome Research* 10:516 -522. (2000)
17. Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y.O. & Borodovsky, M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Research* 18:1979 -1990. (2008)
18. Kent, W. J. BLAT--the BLAST-like alignment tool. *Genome Research* 12:656-64 (2002)
19. Zdobnov, E.M. & Apweiler, R. InterProScan – an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17:847 -848 (2001)
20. Bairoch, A. The ENZYME database in 2000. *Nucleic Acids Research* 28:304-305. (2000)

21. Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H. & Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* 27:29-34 (1999)
22. Bairoch, A. The ENZYME database in 2000. *Nucleic Acids Research* 28:304-305 (2000)
23. Ashburner, M., Ball, C.A., *et al.* Gene Ontology: tool for the unification of biology. *Nature Genetics* 25:25-29 (2000)
24. Koonin, E.V. *et al.* A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biology* 5:R7-R7. (2004)
25. Nielsen, H., Brunak, S. & von Heijne, G. Machine learning approaches for the prediction of signal peptides and other protein sorting signals. *Protein Engineering* 12:3-9 (1999)
26. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E.L.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology* 305:567-580 (2001)
27. Grigoriev, I.V. *et al.* MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 42(1):D699-704 (2014)
28. Löytynoja, A. & Goldman, N. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science* 320, 1632-1635 (2008).
29. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 15;30, 3059-3066 (2002).
30. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688-2690 (2006).
31. dos Reis, M. *et al.* Phylogenomic datasets provide both precision and accuracy in estimating the timescale of placental mammal phylogeny. *Proceedings of the Royal Society B* 7;279, 3491-3500 (2012).
32. Sul, S.-J. & Williams, T. A randomized algorithm for comparing sets of phylogenetic trees. *Proceedings of the Fifth Asia Pacific Bioinformatics Conference (APBC'07)*, 121-130 (2007).
33. Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* 56, 564-577 (2007).
34. Lanfear, R., Calcott, B., Ho, S. Y. & Guindon, S. Partitionfinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution* 29, 1695-1701 (2012).
35. Castoe, T. A., Doan, T. M. & Parkinson, C. L. Data partitions and complex models in Bayesian analysis: the phylogeny of Gymnophthalmid lizards. *Systematic Biology* 53, 448-469 (2004).
36. Xi, Z. *et al.* Phylogenomics and a posteriori data partitioning resolve the Cretaceous angiosperm radiation of Malpighiales. *Proceedings of the National Academy of Sciences USA* 109, 17519-17524 (2012).
37. Philippe, H. *et al.* Resolving difficult phylogenetic questions: why more sequences are not enough. *PLoS biology* 9, e1000602 (2011).
38. Sukumaran, J. & Holder, M. T. DendroPy: a Python library for phylogenetic computing. *Bioinformatics*, 1569-1571 (2010).
39. Lartillot, N. A Bayesian Mixture Model for Across-Site Heterogeneities in the Amino-Acid Replacement Process. *Molecular Biology and Evolution* 21, 1095-1109, doi:10.1093/molbev/msh112 (2004).
40. James, T. Y. *et al.* Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443, 818-822, doi:10.1038/nature05110 (2006).

41. Floudas, D. *et al.* The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes. *Science* 336, 1715-1719, doi:10.1126/science.1221748 (2012).
42. Hibbett, D. S. *et al.* A higher-level phylogenetic classification of the Fungi. *Mycological Research* 111, 509-547, doi:10.1016/j.mycres.2007.03.004 (2007).
43. Matheny, P. B., Gossmann, J. A., Zalar, P., Kumar, T. K. A. & Hibbett, D. S. Resolving the phylogenetic position of the Wallemiomycetes: an enigmatic major lineage of Basidiomycota. *Canadian Journal of Botany* 84, 1794-1805, doi:10.1139/b06-128 (2006).
44. Aime, M. C. *et al.* An overview of the higher level classification of Pucciniomycotina based on combined analyses of nuclear large and small subunit rDNA sequences. *Mycologia* 98, 896-905 (2006).
45. Matheny, P. B. *et al.* Contributions of *rpb2* and *tefl* to the phylogeny of mushrooms and allies (Basidiomycota, Fungi). *Molecular Phylogenetics and Evolution* 43, 430-451, doi:10.1016/j.ympev.2006.08.024 (2007).
46. Padamsee, M. *et al.* The genome of the xerotolerant mold *Wallemia sebi* reveals adaptations to osmotic stress and suggests cryptic sexual reproduction. *Fungal Genetics and Biology* 49, 217-226, doi:http://dx.doi.org/10.1016/j.fgb.2012.01.007 (2012).
47. Lewis, P. O., Holder, M. T. & Holsinger, K. E. Polytomies and Bayesian phylogenetic inference. *Systematic Biology* 54, 241-253 (2005).
48. Nagy, L. G. *et al.* The evolution of defense mechanisms correlate with the explosive diversification of autodigesting *Coprinellus* mushrooms (Agaricales, Fungi). *Systematic Biology* 61, 595-607 (2012).
49. Hibbett, D. S. A phylogenetic overview of the Agaricomycotina. *Mycologia* 98, 917-925 (2006).
50. Sanderson, M. J. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. *Molecular Biology and Evolution* 19, 101-109 (2002).
51. LePage, B. A., Currah, R. S., Stockey, R. A. & Rothwell, G. W. Fossil ectomycorrhizae from the Middle Eocene. *American Journal of Botany* 84, 470-412 (1997).
52. Hibbett, D. S., Grimaldi, D. & Donoghue, M. J. Fossil mushrooms from Cretaceous and Miocene ambers and the evolution of homobasidiomycetes. *American Journal of Botany* 84, 981-991 (1997).
53. Taylor, J. W. & Berbee, M. L. Dating divergences in the Fungal Tree of Life: review and new analyses. *Mycologia* 98, 838-849 (2006).
54. Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7, 214, doi:doi:10.1186/1471-2148-7-214 (2007).
55. Eastwood, D. C. *et al.* The plant cell wall-decomposing machinery underlies the functional diversity of forest fungi. *Science* 333, 762-765, doi:10.1126/science.1205411 (2011).
56. Rothwell, G. W., Mapes, G., Stockey, R. A. & Hilton, J. The seed cone *Eathiestrobus* gen. nov.: fossil evidence for a Jurassic origin of Pinaceae. *Am J Bot* 99, 708-720, doi:10.3732/ajb.1100595 (2012).
57. Wang, X.-Q., Tank, D. C. & Sang, T. Phylogeny and divergence times in Pinaceae: evidence from three genomes. *Molecular Biology and Evolution* 17, 773-781 (2000).
58. Wang, H. *et al.* Rosid radiation and the rapid rise of angiosperm-dominated forests. *Proc Natl Acad Sci U S A* 106, 3853-3858, doi:10.1073/pnas.0813376106 (2009).

59. Henrissat, B. & Davies, G. J. Glycoside hydrolases and glycosyltransferases: families, modules, and implications for genomics. *Plant Physiology* 124, 1515-1519 (2000).
60. Ruiz-Dueñas, F. J. & Martínez, A. T. Microbial degradation of lignin: How a bulky recalcitrant polymer is efficiently recycled in nature and how we can take advantage of this. *Microbial Biotechnology* 2, 164-177 (2009).
61. Cullen, D. & Kersten, P. in *The Mycota III, Biochemistry and Molecular Biology*, 2nd Edition (eds R. Brambl & G.A. Marzluf) 249-273 (2004).
62. Kues, U. & Ruhl, M. Multiple multi-copper oxidase gene families in basidiomycetes – what for? . *Current Genomics* 12, 72-94 (2011).
63. Hofrichter, M., Ullrich, R., Pecyna, M. J., Liers, C. & Lundell, T. New and classic families of secreted fungal heme peroxidases. *Applied Microbiology and Biotechnology* 87, 871-897, doi:10.1007/s00253-010-2633-0 (2010).
64. Kersten, P. J. Glyoxal oxidase of *Phanerochaete chrysosporium*: Its characterization and activation by lignin peroxidase. *Proc. Natl. Acad. Sci. USA* 87, 2936-2940 (1990).
65. Kersten, P. J. & Kirk, T. K. Involvement of a new enzyme, glyoxal oxidase, in extracellular H₂O₂ production by *Phanerochaete chrysosporium*. *J. Bacteriol.* 169, 2195-2201 (1987).
66. Kersten90_glox.pdf
67. Baldrian, P. & Valášková, V. Degradation of cellulose by basidiomycetous fungi. *FEMS Microbiology Reviews* 32, 501-521, doi:10.1111/j.1574-6976.2008.00106.x (2008).
68. Leggio, L. L. & Larsen, S. The 1.62 Å structure of *Thermoascus aurantiacus* endoglucanase: completing the structural picture of subfamilies in glycoside hydrolase family 5. *FEBS Letters* 523, 103-108 (2002).
69. Mertz, B., Kuczenski, R. S., Larsen, R. T., Hill, A. D. & Reilly, P. J. Phylogenetic analysis of family 6 glycoside hydrolases. *Biopolymers* 79, 197-206 (2005).
70. Goedegebuur, F. *et al.* Cloning and relational analysis of 15 novel fungal endoglucanases from family 12 glycosyl hydrolase. *Current Genetics* 41, 89-98 (2002).
71. Aspeborg, H., Coutinho, P. M., Wang, Y., Brumer, H. & Henrissat, B. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). . *BMC Evolutionary Biology* 12, 186-201 (2012).
72. Harris, P. V. *et al.* Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry* 49, 3305-3316, doi:10.1021/bi100009p (2010).
73. Langston, J. A. *et al.* Oxidoreductive Cellulose Depolymerization by the Enzymes Cellobiose Dehydrogenase and Glycoside Hydrolase 61. *Applied and Environmental Microbiology* 77, 7007-7015, doi:10.1128/aem.05815-11 (2011).
74. Benoit, I. *et al.* Degradation of different pectins by fungi: correlations and contrasts between the pectinolytic enzyme sets identified in genomes and the growth on pectins of different origin. *BMC Genomics* 13, 321-331 (2012).
75. Cantarel, B. L. *et al.* The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* 37, D233-D238, doi:10.1093/nar/gkn663 (2009).
76. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* 30, 772-780 (2013).
77. MacClade 4: Analysis of Phylogeny and Character Evolution (Sinauer Associates, Sunderland, 2005).
78. Durand, D., V., H. B. & Vernet, B. A hybrid micro-macroevolutionary approach to gene tree reconstruction. *Journal of Computational Biology* 13, 320-335 (2006).

79. Schmidt, O. & Liese, W. Variability of wood degrading enzymes of *Schizophyllum commune*. *Holzforschung* 34, 67-72, doi:10.1515/hfsg.1980.34.2.67 (1980).
80. Hoegger, P. J., Kilaru, S., James, T. Y., Thacker, J. R. & Kues, U. Phylogenetic comparison and classification of laccase and related multicopper oxidase protein sequences. *FEBS Journal* 273, 2308-2326 (2006).
81. De Silva, D. M., Askwith, C. C., Eide, D. & Kaplan, J. The Fet3 gene product required for high affinity iron transport in yeast is a cell surface ferroxidase. *Journal Biological Chemistry* 270, 1098-1101 (1995).
82. Vanden Wymelenberg, A. *et al.* Structure, organization, and transcriptional regulation of a family of copper radical oxidase genes in the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. *Appl. Environ. Microbiol.* 72, 4871-4877 (2006).
83. Whittaker, M. M. *et al.* Glyoxal oxidase from *Phanerochaete chrysosporium* is a new radical-copper oxidase. *Journal Biological Chemistry* 271, 681-687 (1996).
84. Leuthner, B. *et al.* A H₂O₂-producing glyoxal oxidase is required for filamentous growth and pathogenicity in *Ustilago maydis*. *Molecular Genetics and Genomics* 272, 639-650 (2005).
85. Hamada, N. *et al.* Cloning and characterization of a new exo-cellulase gene, *cel3*, in *Irpex lacteus*. *FEMS Microbiology Letters* 172, 231-237 (1999).
86. Bey, M. *et al.* Cello-oligosaccharide oxidation reveals differences between two lytic polysaccharide monooxygenases (family GH61) from *Podospira anserina*. *Applied and Environmental Microbiology* 79, 488-496 (2013).
87. Beeson, W. T., Phillips, C. M., Cate, J. H. D. & Marletta, M. A. Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. *Journal of the American Chemical Society* 134, 890-892 (2012).
88. Koseki, T., Furuse, S., Iwano, K., Sakai, H. & Matsuzawa, H. An *Aspergillus awamori* acetylsterase : purification of the enzyme, and cloning and sequencing of the gene. *Biochemical Journal* 326, 485-490 (1997).
89. Crepin, V. F., Faulds, C. B. & Connerton, I. F. A non-modular type B feruloyl esterase from *Neurospora crassa* exhibits concentration-dependent substrate inhibition. *Biochemical Journal* 370, 417-427 (2003).
90. Li, X.-L., Skory, C. D., Cotta, M. A., Puchart, V. & Biely, P. Novel family of carbohydrate esterases, based on identification of the *Hypocrea jecorina* acetyl esterase gene. *Applied and Environmental Microbiology* 74, 7482-7489 (2008).
91. Martin, F. *et al.* The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452, 88-92, doi:http://www.nature.com/nature/journal/v452/n7183/supinfo/nature06556_S1.html (2008).
92. Bodeker, I. T., Nygren, C. M., Taylor, A. F., Olson, A. & Lindahl, B. D. ClassII peroxidase-encoding genes are present in a phylogenetically wide range of ectomycorrhizal fungi. *The ISME journal* 3, 1387-1395, doi:10.1038/ismej.2009.77 (2009).
93. Liers, C., Bobeth, C., Pecyna, M., Ullrich, R. & Hofrichter, M. DyP-like peroxidases of the jelly fungus *Auricularia auricula-judae* oxidize nonphenolic lignin model compounds and high-redox potential dyes. *Applied Microbiology and Biotechnology* 85, 1869-1879, doi:10.1007/s00253-009-2173-7 (2009).
94. Salvachúa, D., Prieto, A., Martínez, A. T. & Martínez, M. J. Characterization of a novel dye-decolorizing peroxidase (DyP)-type enzyme from *Irpex lacteus* and its application in enzymatic hydrolysis of wheat straw. *Applied and Environmental Microbiology* 79, 4316-4324 (2013).
95. Doddapaneni, H., Subramanian, V., Fu, B. & Cullen, D. A comparative genomic

- analysis of the oxidative enzymes potentially involved in lignin degradation by *Agaricus bisporus*. *Fungal Genetics and Biology* 55, 22-31 (2013).
96. Morin, E. *et al.* Genome sequence of the button mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche. *Proceedings of the National Academy of Sciences* 109, 17501-17506, doi:10.1073/pnas.1206847109 (2012).
 97. Binder, M. & Hibbett, D. S. Molecular systematics and biological diversification of Boletales. *Mycologia* 98, 971-981 (2006).
 98. Rineau, F. *et al.* The ectomycorrhizal fungus *Paxillus involutus* converts organic matter in plant litter using a trimmed brown-rot mechanism involving Fenton chemistry. *Environmental Microbiology* 14, 1477-1487 (2012).
 99. Nakasone, K.K.. Cultural studies and identification of wood- inhabiting Corticiaceae and selected Hymenomycetes from North America. *Mycologia Memoirs* 15, 1–412 (1990)
 100. Hampp, R., Schaeffer, C., Wallenda, T., Stölten, C., Johann, P. & Einig, W. Changes in carbon partitioning or allocation due to ectomycorrhiza formation: biochemical evidence. *Canadian Journal of Botany* 73(1): 548-556 (1995)
 101. Marx, D.H. The influence of ectotrophic ectomycorrhizal fungi on the resistance of pine roots to pathogenic infections. I. Antagonism of mycorrhizal fungi to pathogenic fungi and soil bacteria. *Phytopathology* 59: 153-163 (1969)
 102. Hampp R. *et al.* Axenic mycorrhization of wild type and transgenic hybrid aspen expressing T-DNA indoleacetic acid biosynthetic genes. *Trees* 11: 59-64 (1996)
 103. Tuominen, H., Sitbon, F., Jacobsson, C., Sandberg, G., Olsson, O. & Sundberg B.. Altered growth and wood characteristics in transgenic hybrid aspen expressing *Agrobacterium tumefaciens* T-DNA indoleacetic acid-biosynthetic genes. *Plant Physiology* 109: 1179-1189 (1995)
 104. Nehls, U., Wiese, J., Gутtenberger, M. & Hampp, R. Carbon allocation in ectomycorrhizas: identification and expression analysis of an *Amanita muscaria* monosaccharide transporter. *Molecular Plant Microbe Interactions* 11(3): 167-176 (1998)
 105. Girlanda, M. *et al.* Photosynthetic Mediterranean meadow orchids feature partial mycoheterotrophy and specific mycorrhizal associations. *American Journal of Botany* 98 1148-1163. (2011).
 106. Perotto S. *et al.* Gene expression in mycorrhizal orchid protocorms suggests a friendly plant-fungus relationship. *Planta*. 239 (6): 1337-49. (2014).
 107. Molina, R., Palmer, J. G., & Schenck, N. C. Methods and principles of mycorrhizal research. *Methods and principles of mycorrhizal research*. (1982).
 108. Herrmann, S., Munch, J.C. & Buscot, F. A gnotobiotic culture system with oak microcuttings to study specific effects of mycobionts on plant morphology before, and in the early phase of, ectomycorrhiza formation by *Paxillus involutus* and *Piloderma croceum*. *New Phytologist* 138: 203-212 (1998).
 109. Tarkka, M.T. *et al.* OakContigDF159.1, a reference library for studying differential gene expression in *Quercus robur* during controlled biotic interactions: use for quantitative transcriptomic profiling of oak roots in ectomycorrhizal symbiosis. *New Phytologist* 199: 529-540 (2013).
 110. Ingestad, T. & Kähr, M. Nutrition and growth of coniferous seedlings at varied relative nitrogen addition rate. *Physiologia Plantarum*, 65(2), 109-116 (1985)
 111. Colpaert, J. V., Van Tichelen, K. K., Van Assche, J. A., & Van Laere, A. Short term phosphorus uptake rates in mycorrhizal and non-mycorrhizal roots of intact *Pinus sylvestris* seedlings. *New Phytologist*, 143(3), 589-597 (1999)

112. Colpaert, J. V., Vandenkoornhuyse, P., Adriaensen, K., & Vangronsveld, J. Genetic variation and heavy metal tolerance in the ectomycorrhizal basidiomycete *Suillus luteus*. *New Phytologist*, 147(2), 367-379 (2000)
113. Chang, S., Puryear, J., & Cairney, J. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter*, 11(2), 113-116. (1993)
114. Wright, D. P., Johansson, T., Le Quéré, A., Söderström, B. & Tunlid, A. Spatial patterns of gene expression in the extramatrical mycelium and mycorrhizal root tips formed by the ectomycorrhizal fungus *Paxillus involutus* in association with birch (*Betula pendula*) seedlings in soil microcosms. *New Phytologist*, 167: 579–596. doi: 10.1111/j.1469-8137.2005.01441.x (2005)
115. Gay, G. Effect of the ectomycorrhizal fungus *Hebeloma hiemale* on adventitious root formation in derooted *Pinus halepensis* shoot hypocotyls. *Can. J. Bot.* 68:1265-1270 (1990)
116. Dore, J., Marmeisse, R., Combier, J.P. & Gay, G. A fungal conserved gene from the basidiomycete *Hebeloma cylindrosporum* is essential for efficient ectomycorrhiza formation. *Mol. Plant Microbe Interac.* DOI : <http://dx.doi.org/10.1094/MPMI-03-14-0087-R> (2014)
117. Baggerly, K., Deng, L., Morris, J., & Aldaz, C. Differential expression in SAGE: accounting for normal between-library variation. *Bioinformatics*, 19(12):1477-1483. (2003)
118. Enright, A.J., Van Dongen, S & Ouzounis, C.A. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Research* 30(7):1575-1584 (2002).
119. De Bie, T., Cristianini, N., Demuth, J.P., & Hahn, M.W. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics*. 22:1269-1271 (2006).
120. De Caceres, M., Oliva, F., Font, X. & Vives, S. Ginkgo, a program for non-standard multivariate fuzzy analysis. *Advances in Fuzzy Sets and Systems* 2:41-56 (2007)
121. Petersen, T.N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785-786 (2011)
122. Horton, P. et al. WoLF PSORT: Protein Localization Predictor. *Nucleic Acids Res.* 35:W585-7 (2007)
123. Emanuelsson, O. *et al.* Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* 2:953–971 (2007)
124. Nielsen H., Engelbrecht, J., Brunak, S. & von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.*

3. Supplementary Tables 1-11

Supplementary Table 1. Genomes included in this study.

ABBREVIATED TAXON NAME	TAXON	CLASSIFICATION
Agabi_varb	<i>Agaricus bisporus</i> var <i>bisporus</i>	Agaricales (Basidiomycota)
Amamu1	<i>Amanita muscaria</i>	Agaricales (Basidiomycota)
Amath1	<i>Amanita thiersii</i>	Agaricales (Basidiomycota)
Aspid1	<i>Aspergillus nidulans</i>	Eurotiales (Ascomycota)
Aurde2	<i>Auricularia delicata</i>	Auriculariales (Basidiomycota)
Batde5	<i>Batrachochytrium dendrobatidis</i>	Rhizophydiales (Chytridiomycetes)
Botbo1	<i>Botryobasidium botryosum</i>	Cantharellales (Basidiomycota)
Conpu1	<i>Coniophora puteana</i>	Boletales (Basidiomycota)
Copci1	<i>Coprinopsis cinerea</i>	Agaricales (Basidiomycota)
Crypa2	<i>Cryphonectria parasitica</i>	Diaporthales (Ascomycota)
Dacsp1	<i>Dacryopinax</i> sp.	Dacrymycetales (Basidiomycota)
Fomme1	<i>Fomitiporia mediterranea</i>	Hymenochaetales (Basidiomycota)
Fompi3	<i>Fomitopsis pinicola</i>	Polyporales (Basidiomycota)
Galma1	<i>Galerina marginata</i>	Agaricales (Basidiomycota)
Glotr1_1	<i>Gloeophyllum trabeum</i>	Gloeophyllales (Basidiomycota)
Gymlu1	<i>Gymnopus luxurians</i>	Agaricales (Basidiomycota)
Hebcy2	<i>Hebeloma cylindrosporum</i>	Agaricales (Basidiomycota)
Hetan2	<i>Heterobasidium annosum</i>	Russulales (Basidiomycota)
Hydpi2	<i>Hydnomerulius pinastris</i>	Boletales (Basidiomycota)
Hypsu1	<i>Hypholoma sublateritium</i>	Agaricales (Basidiomycota)
Jaaar1	<i>Jaapia argillacea</i>	Jaapiales (Basidiomycota)
Lacbi2	<i>Laccaria bicolor</i>	Agaricales (Basidiomycota)
Mellp1	<i>Melampsora laricis-populina</i>	Uredinales (Basidiomycota)
Oidma1	<i>Oidiodendron majus</i>	Incertae sedis (Ascomycota)
Paxin1	<i>Paxillus involutus</i>	Boletales (Basidiomycota)
Paxru1	<i>Paxillus rubicundulus</i>	Boletales (Basidiomycota)
Phchr1	<i>Phanerochaete chrysosporium</i>	Polyporales (Basidiomycota)
Phybl2	<i>Phycomyces blakesleeanae</i>	Mucorales (Mucoromycotina)
Picst3	<i>Pichia stipitis</i>	Saccharomycetales (Ascomycota)
Pilcr1	<i>Piloderma croceum</i>	Atheliales (Basidiomycota)
Pirin1	<i>Piriformospora indica</i>	Sebacinales (Basidiomycota)
Pismi1	<i>Pisolithus microcarpus</i>	Boletales (Basidiomycota)
Pisti1	<i>Pisolithus tinctorius</i>	Boletales (Basidiomycota)
PleosPC15	<i>Pleurotus ostreatus</i>	Agaricales (Basidiomycota)
Plicr1	<i>Plicaturopsis crispa</i>	Amylocorticiales (Basidiomycota)
Punst1	<i>Punctularia strigosozonata</i>	Corticiales (Basidiomycota)
Schco2	<i>Schizophyllum commune</i>	Agaricales (Basidiomycota)
Sclci1	<i>Scleroderma citrinum</i>	Boletales (Basidiomycota)
Sebve1	<i>Sebacina vermifera</i>	Sebacinales (Basidiomycota)

SerlaS7_9	<i>Serpula lacrymans</i>	Boletales (Basidiomycota)
Sphst1	<i>Sphaerobolus stellatus</i>	Phallales (Basidiomycota)
Stano2	<i>Stagonospora nodorum</i>	Pleosporales (Ascomycota)
Suilu1	<i>Suillus luteus</i>	Boletales (Basidiomycota)
Trave1	<i>Trametes versicolor</i>	Polyporales (Basidiomycota)
Treme1	<i>Tremella mesenterica</i>	Tremellales (Basidiomycota)
Trire2	<i>Trichoderma reesei</i>	Hypocreales (Ascomycota)
Tubme1	<i>Tuber melanosporum</i>	Tuberales (Ascomycota)
Tulca1	<i>Tulasnella calospora</i>	Cantharellales (Basidiomycota)
Ustma1	<i>Ustilago maydis</i>	Ustilaginales (Basidiomycota)

Supplementary Table 2. Summary of genome sequencing and assembly

Genome	Library Type	Sequencing	Run type	Assembler*
<i>Amanita muscaria</i>	270bp, Illumina Std PE Unamplified (2)	Illumina	2x150	AllPathsLG (R42328) 75x+75x
<i>Amanita muscaria</i>	4kb LFPE	Illumina	2x150	
<i>Gymnopus luxurians</i>	300bp, Illumina Std PE Unamplified	Illumina	2x100	Velvet (0.7.55) Newbler (2.5) -ml41
<i>Gymnopus luxurians</i>	454 Titanium PE - 4kb	454		
<i>Gymnopus luxurians</i>	454-Rapid	454		
<i>Hebeloma cylindrosporum v2</i>	300bp, Illumina fragment	Illumina	2 x 76	Velvet (0.7.55) Newbler (2.5) -ml41 AllPathsLG (R42137)
<i>Hebeloma cylindrosporum v2</i>	454 Titanium PE - 3.5kb	454		
<i>Hebeloma cylindrosporum v2</i>	454-Rapid (2)	454		
<i>Hebeloma cylindrosporum v2</i>	40kb	Sanger		
<i>Hebeloma cylindrosporum v2</i>	3kb	Pacbio		
<i>Hydnomerulius pinastri</i>	300bp, Illumina Std PE Unamplified (2)	Illumina	2x100	AllPathsLG R38445 38x+38x
<i>Hydnomerulius pinastri</i>	9kb CLIP PE	Illumina	2x100	
<i>Hypholoma sublateritium</i>	300bp, Illumina Std PE Unamplified (2)	Illumina	2x100	AllPathsLG (R37753) 187.5x+62.5x
<i>Hypholoma sublateritium</i>	4kb CLIP PE	Illumina	2x100	
<i>Laccaria amethystina LAAM-08-1</i>	270bp fragment	Illumina	2x150	AllpathsLG (R41043) 75x+75x
<i>Laccaria amethystina LAAM-08-1</i>	4kb LFPE	Illumina	2x100	
<i>Oidiodendron maius Zn</i>	300bp, Illumina fragment	Illumina	2x100	Velvet (0.7.55) Newbler (2.5) -ml41
<i>Oidiodendron maius Zn</i>	454 Titanium PE - 4kb	454		
<i>Oidiodendron maius Zn</i>	454-Rapid	454		
<i>Paxillus involutus ATCC 200175</i>	454 Titanium PE - 4kb (2)	454		Velvet (0.7.55) Newbler (2.5) -ml31
<i>Paxillus involutus ATCC 200175</i>	454-Rapid	454		
<i>Paxillus involutus ATCC 200175</i>	40kb	Sanger		

<i>Paxillus rubicundulus</i>	270bp fragment (2)	Illumina	2x150	AllPathsLG (R44008) 50x+50x
<i>Paxillus rubicundulus</i>	4kb LFPE	Illumina	2x100	
<i>Piloderma croceum</i> F 1598	250bp frag	Illumina	2x150	AllPathsLG (R38445) 38x+112x
<i>Piloderma croceum</i> F 1598	4kb (2) & 9kb CLIP PE	Illumina	2x100	
<i>Pisolithus microcarpus</i>	270bp frag	Illumina	2x150	AllPathsLG (R39750) 50x+50x+5x
<i>Pisolithus microcarpus</i>	4kb CLIP (2)	Illumina	2x100	
<i>Pisolithus microcarpus</i>	454 Titanium PE - 4kb	454		
<i>Pisolithus microcarpus</i>	40kb (2)	Sanger		
<i>Pisolithus tinctorius</i>	10kb CLIP PE	Illumina	2x100	AllPathsLG (R41043) 50x+50x
<i>Pisolithus tinctorius</i>	270bp Illumina Std PE Unamplified	Illumina	2x150	
<i>Pisolithus tinctorius</i>	4kb & 8Kb CLIP PE	Illumina	2x100	
<i>Plicaturopsis crispa</i>	454 PE - 4kb & 8kb	454		Newbler (2.5) -ml31
<i>Plicaturopsis crispa</i>	454 Rapid	454		
<i>Plicaturopsis crispa</i>	40Kb	Sanger		
<i>Scleroderma citrinum</i>	270bp Fragment	Illumina	2x150	AllPathsLG (R40776) 38x+112x
<i>Scleroderma citrinum</i>	4kb CLIP PE	Illumina	2x150	
<i>Sebacina vermifera</i> MAFF 30583	4kb & 8kb CLIP PE	Illumina	2x100	AllPathsLG (R41043) 38x+112x
<i>Sebacina vermifera</i> MAFF 30583	Illumina Std PE Unamplified (3)	Illumina	2x100	
<i>Sphaerobolus stellatus</i>	Illumina Stranded Std PE	Illumina	2x100	AllPathsLG (R39750) 100x+100x
<i>Sphaerobolus stellatus</i>	300bp, Illumina Std PE Unamplified	Illumina	2x150	
<i>Sphaerobolus stellatus</i>	4kb & 9kb CLIP PE	Illumina	2x100	
<i>Suillus luteus</i>	270bp, Illumina Std PE Unamplified	Illumina	2x150	Velvet (0.7.55) AllPathsLG (R42137) 100x+50x
<i>Tulasnella calospora</i> AL13/4D	270bp	Illumina	2x150	AllPathsLG (R39750) 100x+100x
<i>Tulasnella calospora</i> AL13/4D	4kb CLIP	Illumina	2x100	

* For AllPathsLG, coverage of combined fragment and combined LMP libraries are shown in the last column; HAPLOIDIFY=True PLOIDY=2" was used for diploid organisms, or PLOIDY=1 for haploid. For Newbler, the following parameters were used -mi99 and -wi with variable -ml as shown in the last column. Velvet was used prior to Newbler to assemble Illumina short reads.

Supplementary Table 3. Summary statistics of genome assemblies

Species	assembly size (Mbp)	# contigs	contig N50 / L50 (# / kbp)	# scaffolds	scaffold N50 / L50 (# / kbp)	min - max scaffold (kpb)	total gap length (%)	total repeat length (%)	GC content (%)
<i>Amanita muscaria</i> Koide BX008	40.7	3814	266 / 30	1101	54 / 168	1 - 1492	12.0	6.1	47.6
<i>Gymnopus luxurians</i> FD-317 M1	66.3	1848	114 / 129	383	30 / 635	2 - 3435	1.6	7.7	45.1
<i>Hebeloma cylindrosporum</i> h7	38.2	526	34 / 276	176	12 / 1069	1 - 3308	5.3	0.8	48.4
<i>Hydnomerulius pinastri</i> MD-312	38.3	2315	182 / 51	603	16 / 689	1 - 2445	7.9	2.3	50.8
<i>Hypholoma sublateritium</i> FD-334 SS-4	48.0	1329	78 / 146	704	44 / 299	1 - 1727	3.1	1.1	51.0
<i>Laccaria amethystina</i> LaAM-08-1	52.2	4756	379 / 28	1229	111 / 121	1 - 842	8.6	6.5	46.9
<i>Oidiodendron maius</i> Zn	46.4	387	28 / 489	100	8 / 1952	2 - 6606	0.4	6.9	47.1
<i>Paxillus involutus</i> ATCC 200175	58.3	6222	815 / 14	2681	29 / 381	2 - 2811	15.7	8.4	50.2
<i>Paxillus rubicundulus</i> Ve08.2h10	53.0	7170	858 / 15	6945	847 / 15	1 - 236	0.1	25.7	48.0
<i>Piloderma croceum</i> F 1598	59.3	4469	407 / 32	715	33 / 529	1 - 1657	11.8	5.9	46.3
<i>Pisolithus microcarpus</i> 441	53.0	5476	592 / 19	1064	89 / 152	1 - 947	10.5	14.8	49.0
<i>Pisolithus tinctorius</i> Marx 270	71.0	5915	612 / 22	610	36 / 491	1 - 2370	13.3	29.8	48.9
<i>Plicaturopsis crispa</i> FD-325 SS-3	34.5	955	53 / 166	316	7 / 1834	2 - 4120	3.3	1.5	53.7
<i>Scleroderma citrinum</i> Foug A	56.1	3919	324 / 39	938	63 / 243	1 - 1201	6.1	22.3	48.6
<i>Sebacina vermifera</i> MAFF 305830	38.1	2457	317 / 32	546	29 / 319	1 - 1897	6.9	3.9	48.9
<i>Sphaerobolus stellatus</i> SS14	176.4	13156	1486 / 19	1460	85 / 585	1 - 3052	30.7	10.9	44.4
<i>Suillus luteus</i> UH-Slu-Lm8-n1	37.0	2113	182 / 51	1944	165 / 58	1 - 567	0.2	2.4	48.0
<i>Tulasnella calospora</i> AL13/4D	62.4	6848	730 / 19	1335	113 / 160	1 - 927	16.6	4.9	51.4

Supplementary Table 4. Summary statistics of annotated genomes.

Species	# genes	average # exons / gene	average protein length (aa)	# genes w/ Pfam domain (%)	# distinct Pfam domains	# genes w/ signal peptide (%)	# genes in multi- gene family	# multi- gene families	average # genes / family	# genes in dupl. blocks
<i>Amanita muscaria</i> Koide BX008	18,153	4.54	328	6244 (34)	2293	2430 (13)	11526	2213	5.21	189
<i>Gymnopus luxurians</i> FD-317 M1	22,057	5.26	364	8581 (39)	2411	4348 (20)	16573	3022	5.48	303
<i>Hebeloma cylindrosporum</i> h7	15,382	5.24	386	6032 (39)	2370	2358 (15)	8965	2020	4.44	116
<i>Hydnomerulius pinastri</i> MD-312	13,270	5.84	406	6462 (49)	2343	2325 (18)	8989	1952	4.61	273
<i>Hypholoma sublateritium</i> FD-334 SS-4	17,911	5.29	387	6729 (38)	2363	3505 (20)	11671	2502	4.66	216
<i>Laccaria amethystina</i> LaAM-08-1	21,066	4.49	319	6286 (30)	2324	2911 (14)	11791	5436	2.17	90
<i>Oidiodendron maius</i> Zn	16,703	2.97	429	9598 (57)	2490	3196 (19)	11061	2143	5.16	50
<i>Paxillus involutus</i> ATCC 200175	17,968	4.99	356	7271 (40)	2276	2289 (13)	13854	2355	5.88	232
<i>Paxillus rubicundulus</i> Ve08.2h10	22,065	3.81	264	5676 (26)	2291	2396 (11)	13371	2991	4.47	42
<i>Piloderma croceum</i> F 1598	21,583	4.75	332	7135 (33)	2352	2813 (13)	13497	2667	5.06	247
<i>Pisolithus microcarpus</i> 441	21,064	4.04	282	6222 (30)	2250	2091 (10)	14127	2471	5.72	664
<i>Pisolithus tinctorius</i> Marx 270	22,701	4.11	290	6401 (28)	2271	2508 (11)	15735	2574	6.11	909
<i>Plicaturopsis crispa</i> FD-325 SS-3	13,626	5.84	410	6406 (47)	2368	2682 (20)	8477	1868	4.54	52
<i>Scleroderma citrinum</i> Foug A	21,012	4.33	299	6512 (31)	2291	2392 (11)	13662	2191	6.24	221
<i>Sebacina vermifera</i> MAFF 305830	15,312	4.94	410	6642 (43)	2284	2442 (16)	10216	2039	5.01	565
<i>Sphaerobolus stellatus</i> SS14	35,274	4.67	309	10896(31)	2362	4891 (14)	30848	5063	6.09	2613
<i>Suillus luteus</i> UH-Slu-Lm8-n1	18,316	4.79	327	6586 (36)	2325	2374 (13)	11275	2133	5.29	86
<i>Tulasnella calospora</i> AL13/4D	19,659	4.65	341	8180 (42)	2277	2583 (13)	13950	2659	5.25	773

Supplementary Table 5: Gene families in expansion in Boletales compare to the Agaricales (a) and in Agaricales compared to Boletales (b)

cluster nb	Boletales										Agaricales										PFAM
	Pismi	Pisti	Sclci	Paxin	Paxru	Hydpi	Suilu	Conpu	Serla	Hebcy	Galma	Hypsu	Lacbi	Copci	Agabi	Amamu	Amath	Gymlu	Schco	Pleos	
91	33	20	30	110	45	61	41	16	10	0	2	0	3	5	3	9	0	8	0	0	50S ribosome-binding GTPase
127	30	41	32	22	113	10	4	3	10	1	2	3	3	2	0	1	0	1	2	1	NO PFAM
207	6	30	8	27	94	5	10	2	2	0	0	2	0	0	0	0	0	0	0	1	NO PFAM
313	14	20	15	22	13	5	3	6	4	2	4	2	2	2	2	2	2	3	4	2	membrane-bound O-acyltransferase family
407	21	41	13	9	6	6	9	6	10	0	6	4	0	0	0	0	0	0	0	4	NO PFAM
412	6	5	5	5	79	6	5	2	4	2	2	0	6	2	0	1	0	2	1	0	DDE superfamily endonuclease
458	5	9	15	11	3	10	11	5	4	6	1	2	1	0	0	0	0	2	3	0	Ribosomal protein L10e family
609	5	5	7	9	5	7	5	5	4	2	3	2	6	1	2	1	2	2	2	3	Phosphatase
779	9	7	3	7	2	8	5	6	7	0	0	0	0	1	1	0	0	0	2	0	NO PFAM
805	5	6	2	14	4	8	3	3	2	1	1	1	1	1	1	1	1	2	1	1	NO PFAM
1115	4	2	5	3	4	12	3	5	9	0	1	0	0	0	1	0	0	2	1	1	NO PFAM
1287	12	10	11	3	6	3	7	4	6	0	0	0	0	0	1	0	0	0	0	0	Kelch motif

a

cluster nb	Agaricales												Boletales								PFAM
	Hebcy	Galma	Hypsu	Lacbi	Copci	Agabi	Amamu	Amath	Gymlu	Schco	Pleos	Pismi	Pisti	Sclci	Paxin	Paxru	Hydpi	Suilu	Conpu	Serla	
15	86	254	137	54	26	120	118	1	76	2	3	0	2	0	0	1	3	19	70	12	Tyrosine kinase
730	4	14	6	6	4	6	4	4	6	7	6	0	0	2	1	1	1	1	2	1	NO PFAM
843	4	13	3	4	3	3	3	3	3	3	3	3	1	1	1	1	2	2	2	2	Tetraspanin

b

Supplementary Table 6. Node ages in genome-based phylogeny inferred using r8s. For node numbers, see Supplemental Figs. 12-21.

Node	MYR (Mean)	Taxa defined by the node
1	683	Fungi
3	635	Dikarya
4	497	Ascomycota
5	CAL	Pezizomycotina
6	257	
7	229	
8	184	
9	118	
10	500	Basidiomycota
11	459	Ustilaginomycotina-Pucciniomycotina
12	436	Agaricomycotina
13	347	Agaricomycetes-Dacrymycetes
14	294	Agaricomycetes
15	255	Cantharellales
16	273	
17	89	Sebacinales
18	251	
19	231	
20	193	
21	157	
22	126	Polyporales
23	107	
24	147	

25	136	Gloeophyllales-Jaapiales-Corticiales
26	104	
27	140	Agaricomycetidae-Russulales
28	125	Agaricomycetidae
29	108	Agaricales
30	101	
31	CAL	
32	92	
33	53	Amanitaceae
34	84	
37	68	Hymenogastraceae
36	77	
38	37	
38	43	
39	115	Boletales-Atheliales-Amylocorticiales
40	100	Atheliales-Amylocorticiales
41	84	Boletales
42	74	
43	CAL	"core Boletales"
44	44	
45	30	
46	8	
47	26	Sclerodermatineae
48	11	

Supplementary Table 7. Lignocellulose-degrading and secondary-metabolism genes found in 49 fungal genomes representing white-rot, brown-rot, litter-decaying, ectomycorrhizal, orchid mycorrhizal, ericoid and endophytic lifestyles. Subphyla: Agar = Agaricomycotina; Ustil = Ustilaginomycotina; Puccin = Pucciniomycotina; Peziz = Pezizomycotina; Sacc = Saccharomycotina; Mucor = Mucoromycotina; Chytrid = Chytridiomycota. Species: Agabi, *Agaricus bisporus* var *bisporus*; Amamu, *Amanita muscaria* var. *guessowii*; Amath, *Amanita thiersii*; Aspnid, *Aspergillus nidulans*; Aurde, *Auricularia delicata*; Batde, *Batrachochytrium dendrobatidis*; Botbo, *Botryobasidium botryosum*; Conpu, *Coniophora puteana*; Copci, *Coprinopsis cinerea*; Crypa, *Cryphonectria parasitica*; Dacsp, *Dacryopinax* sp.; Fomme, *Fomitiporia mediterranea*; Fompi, *Fomitopsis pinicola*; Galma, *Galerina marginata*; Glotr, *Gloeophyllum trabeum*; Gymlu, *Gymnopus luxurians*; Hebcy, *Hebeloma cylindrosporum*; Hetan, *Heterobasidion annosum*; Hydpi, *Hydnum ericoides*; Hypsu, *Hypholoma sublateritium*; Jaaar, *Jaapia argillacea*; Lacbi, *Laccaria bicolor*; Mellp1, *Melampsora larici-populina*; Oidma, *Oidiodendron maius*; Paxin, *Paxillus involutus*; Paxru, *Paxillus rubicundulus*; Phchr, *Phanerochaete chrysosporium*; Phybl, *Phycomyces blakesleeana*; Pilcr, *Piloderma croceum*; Pirin, *Piriformospora indica*; Pismi, *Pisolithus microcarpus*; Pisti, *Pisolithus tinctorius*; PleosPC15, *Pleurotus ostreatus*; Plicr, *Plicaturopsis crispa*; Punst, *Punctularia strigosozonata*; Schco, *Schizophyllum commune*; Scldi, *Scleroderma citrinum*; Sebve, *Sebacina vermifera*; Serla, *Serpula lacrymans*; Stano, *Stagonospora nodorum*; Suilu, *Suillus luteus*; Trave, *Trametes versicolor*; Treme, *Tremella mesenterica*; Trire, *Trichoderma resei*; Tubme, *Tuber melanosporum*; Tulca, *Tulasnella calospora*; Ustma, *Ustilago maydis*. Ecology: Orch = orchid symbiont; ECM = ectomycorrhizal; Eric = ericoid mycorrhiza; Root = root endophyte; BR = brown rot; WR = white rot; S/L/O = soil, litter or other saprotroph; Mycpar = mycoparasite; PlPath = plant pathogen; AnPath = animal pathogen. Relative abundance of the different enzymes within each family is represented by a colour scale, from the minimum (blue) to maximum number of copies (red) per species. Notice a striking white/brown-rot and ECM fungi grouping with respect to the lignin-attacking PODs and the CAZymes that target crystalline cellulose (CBM1, GH6, and GH7), and a continuum with other lignin-targeting enzymes.

Species	Ecol	POD	Lac	Lac/Fet3	Fet3	DyP	HTP	CRO 3-5	CRO 1,2,6	GLX	GH43	GH10	GH5_7_30	GH28	GH3	CE1	CE16	GH12	GH5_5	GH5_22	GH6	GH7	LPMO	GH5_9	GH5_15	GH5_12	GH5_49_50	GH5_A	GH5_B	CBM1
Tulca	Orch	0	0	0	1	2	2	0	4	0	13	13	14	7	12	3	15	0	18	3	7	27	33	8	0	3	2	0	0	110
Sebve	Orch	0	5	0	1	2	3	5	5	0	3	10	5	6	8	8	12	1	7	3	5	5	34	10	0	2	2	0	0	50
Pirin	Root	0	2	0	1	2	2	3	5	0	6	18	5	5	8	9	7	1	4	2	2	1	25	6	0	2	2	0	0	64
Pilcr	ECM	1	11	6	3	4	11	0	5	0	2	0	2	8	11	2	8	2	1	1	0	1	1	4	3	2	2	0	0	0
Lacbi	ECM	1	11	0	4	2	5	3	8	0	0	0	2	6	2	0	4	3	1	0	0	0	11	10	3	3	2	0	0	1
Suil	ECM	0	16	1	1	2	5	1	3	0	1	0	3	3	10	0	4	2	2	1	0	0	2	5	3	2	1	1	0	0
Paxin	ECM	0	13	1	1	0	4	1	3	0	0	1	5	3	4	0	6	0	1	3	0	0	5	3	2	2	3	1	0	0
Hebcy	ECM	3	2	0	1	2	7	1	6	0	2	1	2	2	5	0	5	3	1	1	0	1	3	7	2	2	2	0	0	2
Paxru	ECM	0	14	1	1	0	4	1	3	0	0	1	3	2	4	0	8	0	2	3	0	0	3	3	2	2	2	0	0	0
Pismi	ECM	0	8	0	1	0	4	1	3	0	3	0	3	1	5	0	4	0	0	0	0	0	4	3	3	2	1	0	0	0
Amamu	ECM	0	18	0	1	2	4	1	3	0	3	0	1	0	4	0	6	0	0	1	0	0	2	6	1	3	1	0	0	0
Pisti	ECM	0	12	0	1	1	4	1	3	0	3	0	1	1	8	0	2	0	0	0	0	0	4	3	1	2	1	0	0	0
Scic	ECM	0	10	0	1	1	4	5	3	0	0	0	2	3	3	0	5	1	0	0	0	0	3	3	2	2	1	0	0	0
Dacsp	BR	0	0	3	2	0	6	0	3	0	5	3	2	6	9	0	3	1	3	2	0	0	0	5	1	2	1	5	2	1
Glotr	BR	0	4	0	1	0	6	0	2	0	6	3	3	10	10	1	6	2	2	2	0	0	4	6	1	2	2	0	1	1
Fompi	BR	1	5	1	1	0	4	1	3	0	7	2	2	12	12	0	7	2	3	3	2	0	0	4	5	1	2	0	2	0
Serla	BR	0	4	1	1	0	3	1	2	0	2	1	5	7	10	0	4	2	3	2	1	0	5	4	1	2	1	0	2	8
Conpu	BR	0	6	1	1	0	2	0	6	0	6	3	3	13	13	0	7	4	5	2	2	2	10	4	2	2	1	0	1	3
Hydpi	BR	0	9	1	1	1	4	1	4	0	3	3	4	9	13	1	12	3	5	2	1	3	15	3	1	2	2	2	2	16
Agabi	S/L/O	2	12	0	1	0	24	1	4	3	4	2	2	5	8	1	11	2	3	2	1	1	11	7	1	2	2	0	0	13
Copci	S/L/O	1	17	0	0	4	13	1	5	0	4	6	5	3	7	4	2	1	1	2	5	6	35	11	0	2	2	0	0	44
Amth	S/L/O	0	15	0	1	1	4	1	3	2	6	4	2	5	9	1	4	3	3	2	1	1	16	7	2	2	2	0	1	10
Jaaar	WR	1	1	1	1	1	8	0	4	0	3	5	12	7	10	2	8	8	11	2	3	5	15	5	1	2	2	0	1	24
Botbo	WR	0	0	0	1	3	7	0	5	0	1	11	8	2	7	2	7	1	3	2	3	7	32	5	3	2	0	0	1	29
Gymli	S/L/O	5	16	1	2	13	19	1	4	5	11	5	6	19	16	5	12	3	1	2	1	7	13	8	3	2	1	0	3	32
Schco	WR	0	2	1	0	0	3	0	2	0	19	5	1	3	12	11	11	1	2	1	1	2	22	7	1	2	2	0	0	5
Pilcr	WR	7	5	0	1	0	3	0	5	0	3	2	3	10	10	4	11	2	2	1	2	1	9	4	2	2	2	0	0	9
Galma	WR	23	8	0	1	5	24	2	10	4	6	9	10	18	12	2	10	4	4	11	3	8	19	8	1	2	1	0	0	52
Hypsu	WR	14	11	0	2	2	13	1	6	3	3	7	3	7	9	3	7	1	6	2	1	4	14	11	1	2	2	0	0	28
PleosPC15	WR	9	11	0	1	4	4	3	7	4	8	3	5	6	12	2	9	2	4	2	3	16	29	6	1	2	1	0	0	31
Hetan	WR	7	14	2	1	1	5	1	4	0	4	2	3	8	12	1	6	4	3	2	1	1	10	4	0	2	1	1	0	17
Punst	WR	11	12	0	1	5	8	1	5	3	7	5	3	13	14	2	12	2	3	1	1	5	14	5	1	2	2	0	1	26
Trave	WR	26	7	1	2	2	3	1	3	5	3	6	2	11	13	3	8	5	3	2	1	4	18	9	1	2	2	0	1	23
Phchr	WR	16	0	4	1	0	3	3	3	1	4	6	3	4	11	4	1	2	2	2	1	8	15	6	1	2	2	0	1	28
Fomme	WR	17	11	0	1	3	4	1	3	0	7	4	3	16	8	0	6	3	3	2	2	2	13	5	1	2	2	0	2	10
Aurde	WR	18	5	1	1	11	16	1	5	2	28	4	3	10	13	4	12	1	4	2	2	8	20	15	6	3	0	2	3	42
Spbst	WR	62	29	4	1	31	151	1	3	2	17	18	5	17	23	7	24	5	9	2	4	11	37	6	1	3	2	1	3	0
Treme	Mycpar	0	0	2	1	0	0	0	2	0	0	0	0	0	3	0	0	0	0	1	0	0	0	5	0	3	2	0	0	0
Ustma	PIPath	0	0	2	1	0	3	0	0	0	3	2	0	1	3	1	2	0	0	0	0	0	0	6	0	2	0	0	0	0
Mellp1	PIPath	0	0	14	1	0	16	0	0	0	8	6	0	3	3	0	1	10	4	0	0	8	4	9	2	3	1	9	0	0
Tubme	ECM	0	3	1	1	0	0	0	0	0	1	1	0	2	5	0	1	1	1	0	0	0	2	2	1	2	0	0	0	3
Oidma	Eric	0	15	5	3	1	4	3	0	0	16	2	1	21	31	5	6	7	7	2	2	6	8	1	1	1	1	1	0	35
Trire	S/L/O	0	1	2	2	1	4	0	0	0	2	1	1	4	13	2	2	2	2	0	1	2	3	1	1	1	0	0	1	14
Aspnid	AnPath	0	3	0	0	0	5	1	0	0	18	3	6	10	20	2	3	1	2	0	2	3	9	2	1	1	0	0	0	6
Crypa	PIPath	2	10	3	1	0	2	0	0	0	17	4	2	21	16	3	7	5	5	0	2	5	12	1	0	2	0	0	0	12
Stano	PIPath	2	1	1	1	0	10	1	0	0	13	7	1	4	15	10	2	3	4	1	4	5	26	4	1	1	1	0	2	10
Picst	S/L/O	0	0	0	2	0	0	0	0	0	0	1	0	0	5	0	0	0	0	0	0	0	0	2	0	1	1	0	0	0
Phybl	S/L/O	0	0	0	2	0	0	0	0	0	0	0	0	9	3	0	1	0	0	0	0	0	0	2	0	2	1	0	0	1
Batde	AnPath	0	0	0	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	2	0	0

Supplementary Table 8. Protein sequences excluded due to low quality or potential bacterial origin. A question mark in the HGT column indicates genes potentially transmitted horizontally from bacteria.

Protein ID	family	HGT			
Amamu1_13216	CE1	?	Sphst1_786025	DyP	
Amamu1_13215	CE1	?	Sphst1_247319	GH3	
Amamu1_166350	CE1	?	Sphst1_262543	GH6	
Amamu1_166373	CE1	?	Sphst1_174727	LPMO	
Amath1_903	GH7		Sphst1_264402	LPMO	
Aspnid_9599	CE1		Sphst1_23465	LPMO	
Aspnid1_8388	GH3		Sphst1_29363	LPMO	
Aurde1_63162	GH3		Sphst1_31786	LPMO	
Aurde1_66718	GH28		Sphst1_34171	LPMO	
Aurde1_175635	GH5		Sphst1_38256	LPMO	
Aurde1_153595	GH5	?	Sphst1_35388	LPMO	
Aurde1_150400	GH5	?	Sphst1_100577	LPMO	
Conpu1_159524	GH5		Sphst1_271155	LPMO	
Copci1_10186	DyP		Sphst1_267507	LPMO	
Crypa1_107997	DyP		Sphst1_267896	LPMO	
Fomme1_159422	GH28		Sphst1_29203	LPMO	
Fompi1_162677	GH28		Sphst1_249714	LPMO	
Galma1_219084	GH28		Sphst1_267903	LPMO	
Galma1_149097	GH28		Sphst1_274941	HTP	
Galma1_140770	LPMO		Sphst_266044	MCO	
Hebcy2_447072	MCO		Sphst1_274845	POD	
Hebcy2_77249	GH5		Sphst_258413	POD	
Hydpi2_96533	CE1		Sphst_258450	POD	
Hydpi2_53107	GH7		Stano2_173_SNOG 00220	CRO	
Jaaar1_180063	HTP		Stano2_11781_SNO G_05919	LPMO	
Lacbi2_320856	DyP		Stano2_5590_SNO G_12127	LPMO	
Lacbi2_596890	GH28		Stano2_6291	MCO	
Lacbi2_450012	LPMO		Stano2_8728_SNO G_15642	GH5	
Lacbi2_308609	LPMO		Suulu1_792934	GH3	
Lacbi2_304867	GH5		Trave1_112830	HTP	
Lacbi2_298855	GH5		Treme1_73633	GH5	?
Lacbi2_700975	CRO		Tubme1_3144	GH3	
Oidma1_26049	GH3		Tubme1_5647	CRO	
Oidma_40995	GH7		Tulca1_27383	GH28	
Paxin1_101898	GH5		Tulca1_34361	GH28	
Pilcr_816775	MCO		Tulca1_18595	GH3	
Pilocr16425	HTP		Tulca1_75071	GH43	
Pirin1_81614	GH5		Tulca1_27305	LPMO	
Pirin1_76361	GH5	?	Tulca1_246632	GH43	
PleosPC15_2_1047 909	GH3				
Sebve1_265109	GH5				
SerlaS7_9_2_10482 80	GH5				
Sphst1_263845	CE1				

Supplementary Table 9. Estimated gene copy numbers at internal nodes in organismal phylogeny estimated with Notung at edge weight threshold = 90%. For node numbers, see Supplemental Figs. 14-23.

clades	nodes	GH6	GH7	GH61	GH10	GH28	GH43	GH12	GH3	CE1	CE16	POD	MCO			HTP	DyP	CRO		GH5								
													LAC	Lac/ Fet3	Fet3			CRO1 26- GLX	CRO WSC	GH5_12	GH5_15	GH5_22	GH5_5	GH5_7_30	GH5_49_5_0	GH5_9	GH5_A	GH5_B
Ascomycota	1					2			3		1				1					2					1	1	2	
	2					2			3		1				1					2					1	1	2	
	3	5	4	4	4	16	26	2	22	6	6	1	2	3	1	5	1	1	3	1	1	3	2	2	2	3	2	
	4	3	4	4	3	14	18	2	22	6	5	1	1	2	2	4	1	1	3	1	1	3	2	1	2	3	2	
	5	3	4	4	3	14	19	2	23	6	5	1	2	3	3	4	1	1	3	2	1	3	2	1	2	3	2	
	6	4	4	11	4	20	23	3	29	7	6	1	3	4	3	8	1	1	2	2	1	4	4	1	2	3	2	
	7	4	4	16	4	11	18	2	26	6	5	1	2	2	3	9	0	1	1	2	1	4	4	1	2	0	2	
	8	2	5	10	3	22	20	8	29	6	7	1	8	6	3	5	1	1	2	2	1	5	3	1	1	1	1	
	9	2	5	7	3	19	14	7	20	4	7	1	8	5	3	4	1	0	2	2	0	5	3	0	1	0	1	
	10	4	4	4	3	16	23	2	16	4	7	1	2	5	2	5	1	5	1	2	2	1	3	2	3	6	4	2
Puc-Ust	11	0	2	1	3	2	5	2	4	1	3	0	0	3	2	4	0	4	0	2	1	0	1	0	2	6	3	2
Cantharellales	12	4	4	4	2	16	23	1	16	4	7	1	2	5	2	5	1	5	1	2	2	1	3	2	3	6	3	2
	13	4	4	4	3	19	24	2	17	4	7	1	2	6	2	5	1	5	1	2	2	2	4	5	4	6	5	2
	14	5	6	27	6	23	24	2	18	4	9	1	2	6	2	5	3	6	1	2	3	2	5	6	4	6	3	2
	15	3	5	26	5	5	5	1	7	1	5	0	0	0	2	3	3	4	0	2	2	1	4	2	2	5	0	2
Sebacinales	16	4	5	26	6	23	25	2	20	6	9	1	3	6	1	5	3	6	1	2	3	1	5	6	4	6	3	2
	17	2	2	23	8	5	5	1	8	4	8	0	2	0	1	3	2	5	1	2	0	2	4	4	2	6	0	0
Polyporales	18	4	7	25	7	24	28	2	23	6	11	2	3	6	2	10	7	7	1	2	3	2	5	6	3	6	3	4
	19	5	7	22	8	24	28	2	25	6	13	5	5	7	2	10	8	7	1	2	2	2	4	5	3	7	3	4
	20	5	5	21	8	26	27	2	24	6	12	5	5	6	1	7	7	7	1	2	2	2	4	5	3	7	3	4
	21	4	5	21	8	27	27	2	23	6	14	6	5	6	1	7	8	7	2	2	2	2	4	5	3	7	3	4
	22	1	3	15	4	9	8	2	12	2	13	5	4	2	1	3	1	4	1	2	1	2	3	2	3	6	0	1
	23	1	2	13	3	10	5	2	10	2	10	4	4	1	1	3	1	4	1	2	1	2	3	2	3	6	0	1
	24	4	3	21	8	26	26	2	22	6	13	6	5	6	1	7	8	7	2	2	2	2	4	7	3	6	3	4
Cort-Jaap-Gloeo	25	1	1	15	5	16	10	2	12	2	11	2	4	1	1	8	3	4	1	2	1	2	3	6	2	4	0	2
	26	1	1	11	3	9	6	2	10	1	9	1	3	1	1	7	1	2	0	2	1	2	2	6	2	5	0	2
Agaricales	27	4	3	18	5	22	21	2	19	5	11	7	5	5	1	5	8	7	2	2	2	2	4	7	3	6	3	4
	28	4	3	18	5	21	20	2	17	5	12	7	5	5	1	7	7	7	2	2	2	2	4	7	3	6	2	4
	29	4	3	17	5	13	19	2	14	5	11	7	6	2	1	6	6	6	2	2	1	2	4	6	3	6	0	2
	30	3	2	16	5	13	17	2	14	4	11	2	6	2	1	7	6	5	2	2	1	2	3	5	3	7	0	2
	31	2	1	15	3	10	16	2	12	2	10	2	5	2	1	5	4	4	1	2	1	2	3	3	3	7	0	1
	32	2	2	16	5	8	8	2	10	4	11	1	6	0	1	6	4	5	2	2	1	2	3	5	2	7	0	2
	33	1	1	15	3	5	6	1	8	1	5	0	5	0	1	3	1	4	1	2	1	2	3	2	2	6	0	1
	34	2	2	17	5	9	6	2	10	4	11	1	5	0	1	8	4	5	2	2	1	2	3	5	2	7	0	0
	35	2	2	17	5	9	5	2	9	4	10	1	6	0	1	8	4	6	2	2	1	2	3	5	2	7	0	0
	36	2	2	16	4	7	4	2	9	1	10	2	6	0	1	7	4	7	2	2	1	2	3	5	2	8	0	0
Amylocorticiales	37	2	4	16	7	9	5	2	10	1	8	5	6	0	1	8	3	10	1	2	1	2	4	5	2	8	0	0
	38	2	4	15	7	9	5	2	10	1	8	5	6	0	1	8	3	10	1	2	1	2	4	5	2	8	0	0
	39	1	1	11	2	13	6	2	11	3	11	6	3	4	1	7	2	4	1	2	2	2	2	5	2	4	2	3
	40	1	1	9	2	11	4	2	9	1	9	6	3	3	1	7	2	4	0	2	2	1	1	4	2	4	0	0
	41	1	1	10	3	10	3	2	13	2	7	0	4	1	1	3	1	3	1	2	1	1	3	4	3	4	2	3
	42	1	1	10	3	10	3	2	13	2	7	0	4	1	1	3	1	4	1	2	1	1	3	3	3	4	2	3
	43	1	1	10	3	8	2	2	10	2	7	0	4	1	1	3	1	4	1	2	1	1	3	3	2	4	2	2
	44	1	1	10	3	8	2	2	9	2	8	0	7	1	1	4	1	4	1	2	1	1	3	3	2	3	2	2
	45	1	1	11	3	8	2	2	9	2	8	0	7	1	1	4	1	4	1	2	1	1	3	3	2	3	2	2
	46	0	0	3	1	3	0	0	4	0	6	0	12	1	1	4	0	3	1	2	2	1	1	3	3	3	1	0
Boletales	47	0	0	3	0	2	1	1	4	0	3	0	6	0	1	3	1	3	1	2	1	0	0	1	1	3	0	0
	48	0	0	3	0	1	2	0	3	0	3	0	7	0	1	3	1	3	1	2	1	0	0	2	1	3	0	0

Supplementary Table 10. Subclassification of Class II peroxidase proteins (POD). **A.** Overview of POD diversity in newly sequenced

genomes.	MnP	MnP atypical	VP	VP atypical	LiP	GP	not assigned
Plicr	6					1	
Gymlu	5						
Galma		16		1	4?	1	1
Hypsu		14					
Hebcy		3					
Pilcr						1	

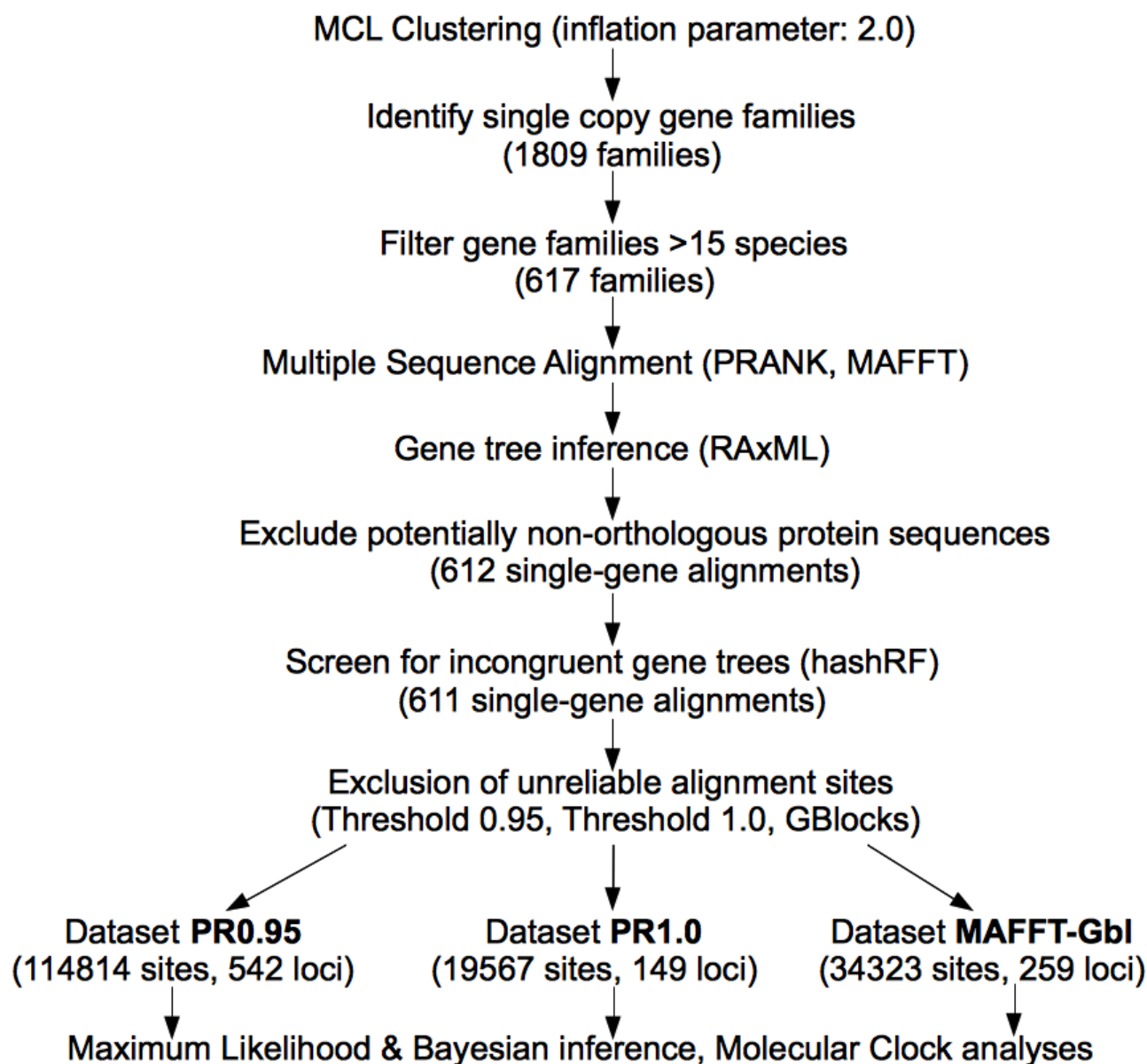
B. Presence of long-range electron transfer Try residue and manganese-binding residues and functional classifications of PODs.

	LRET W	Mn BD E	Mn BD E	Mn BD D	classification
Plicr1 52464	-	+	+	+	MnP
Plicr1 544796	-	+	+	+	MnP
Plicr1 172282	-	+	+	+	MnP
Plicr1 461658	-	+	+	+	MnP
Gymlu1 623402	-	+	+	+	MnP
Plicr1 148817	-	+	+	+	MnP
Plicr1 43914	-	+	+	+	MnP
Galma1 75730	-	+	-	+	MnP atypical
Galma1 75958	-	+	-	+	MnP atypical
Galma1 103111	-	+	-	+	MnP atypical
Galma1 106691	-	+	-	+	MnP atypical
Galma1 106718	-	+	-	+	MnP atypical
Galma1 143334	-	+	-	+	MnP atypical
Galma1 145790	-	+	-	+	MnP atypical
Galma1 134156	-	+	-	+	MnP atypical
Galma1 162345	-	+	-	+	MnP atypical
Galma1 138251	-	+	-	+	MnP atypical
Hypsu1 86525	-	+	-	+	MnP atypical
Hypsu1 167926	-	+	-	+	MnP atypical
Hypsu1 184750	-	+	-	+	MnP atypical
Hypsu1 132620	-	+	-	+	MnP atypical
Galma1 271169	-	+	-	+	MnP atypical
Hebcy2 20759	-	+	-	+	MnP atypical
Galma1 155556	-	+	-	?	uncertain
Hebcy2 27571	-	+	-	+	MnP atypical
Hebcy2 143582	-	+	-	+	MnP atypical
Galma1 220286	+	+	-	+	VP atypical
Galma1 104378	+	+	-	-	uncertain
Galma1 134763	+	+	-	-	uncertain
Galma1 147216	+	+	-	-	uncertain
Galma1 265964	+	+	-	-	uncertain
Galma1 101039	-	-D	-	+	MnP atypical
Galma1 142466	-	-D	-	+	MnP atypical
Hypsu1 43989	-	-D	-	+	MnP atypical
Hypsu1 47473	-	-D	-	+	MnP atypical
Hypsu1 207400	-	-D	-	+	MnP atypical
Hypsu1 220069	-	-D	-	+	MnP atypical
Hypsu1 296328	-	-D	-	+	MnP atypical
Hypsu1 47481	-	-D	-	+	MnP atypical
Hypsu1 47480	-	-D	-	+	MnP atypical
Hypsu1 47465	-	-D	-	+	MnP atypical
Hypsu1 48449	-	-D	-	+	MnP atypical
Hypsu1 207174	-	-D	-	+	MnP atypical
Galma1 281593	-	-D	-	+	MnP atypical
Galma1 263611	-	-D	-	+	MnP atypical
Galma1 1328628	-	-D	-	+	MnP atypical
Gymlu1 148876	-	+	+	+	MnP long/extra long
Gymlu1 182393	-	+	+	+	MnP long/extra long
Gymlu1 1021783	-	+	+	+	MnP long/extra long
Gymlu1 180143	-	+	+	+	MnP long/extra long
Pilcr1 223440	-	-	-	-	generic
Plicr1 177549	-	-	-	-	generic
Galma1 77799	-	-	-	-	generic

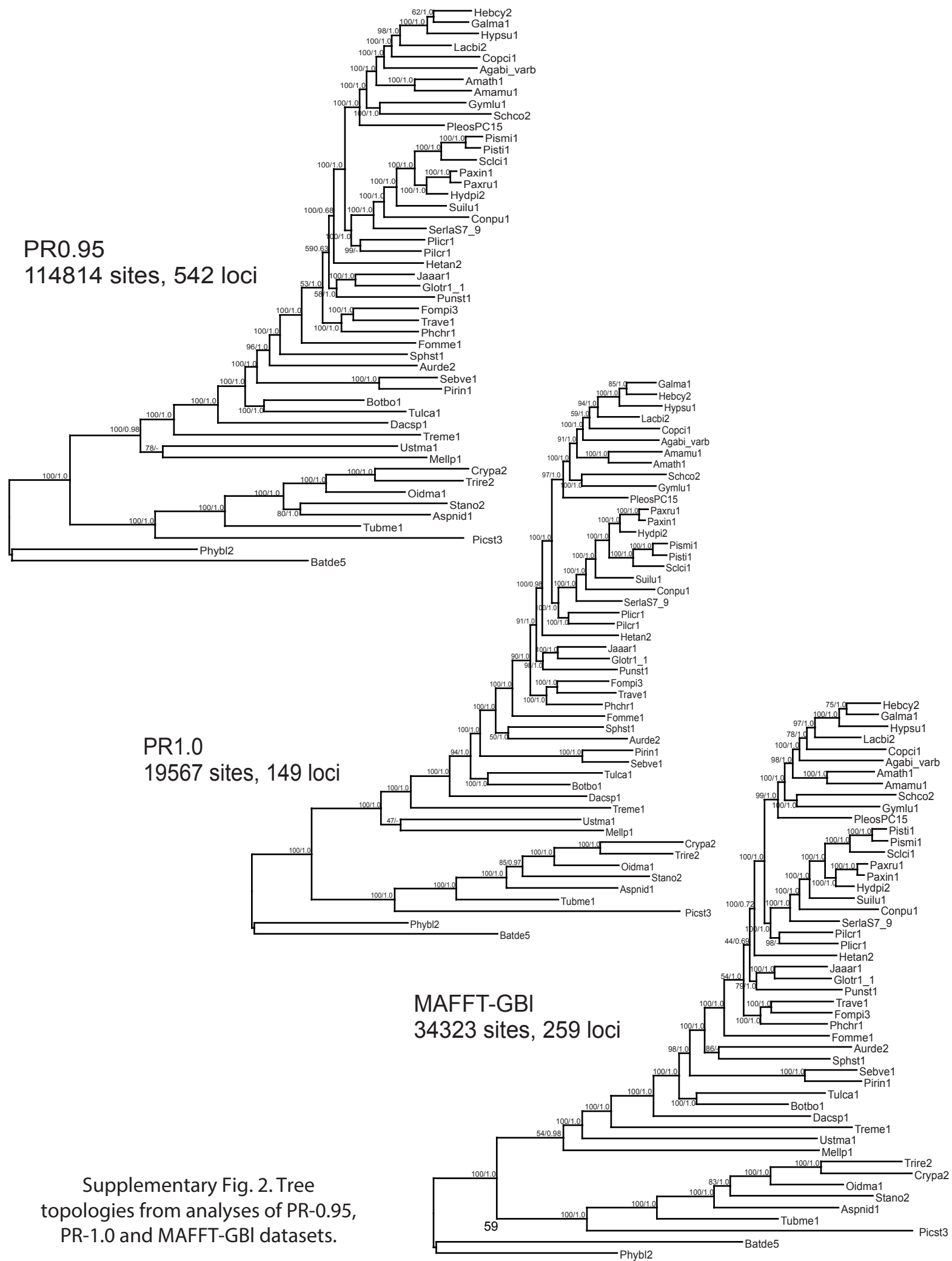
Supplementary Table 11: Summary of reads alignments to reference genomes

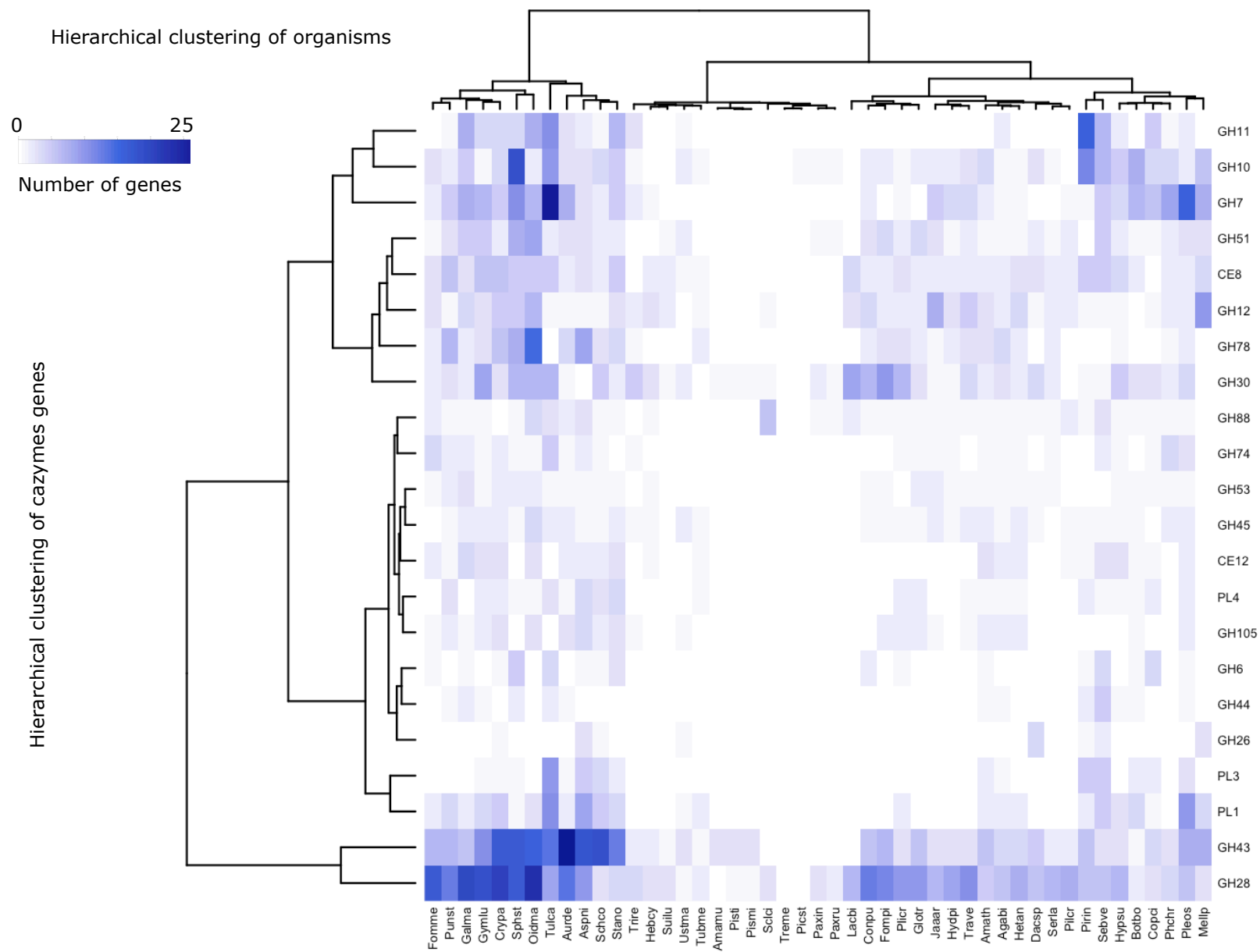
	trimmed reads	aligned	not aligned	% aligned
<i>Hebcy</i>				
Hebcy-FLM1	36570502	25941232	10629270	70.9
Hebcy-FLM2	35884216	24245219	11638997	67.6
Hebcy-FLM3	33590534	22461347	11129187	66.9
Hebcy-MYC1	104061378	37765489	66295889	36.3
Hebcy-MYC2	90962146	29343075	61619071	32.3
Hebcy-MYC3	99771880	33159064	66612816	33.2
<i>Pilcr</i>				
Picr-FLM1	117190140	20060065	97130075	82.9
Picr-FLM2	30900110	5220624	25679486	83.1
Picr-FLM3	47383442	8542865	38840577	82.0
Pilcr-MYC1	43519072	36570077	6948995	16.0
Pilcr-MYC2	41281936	35366718	5915218	14.3
Pilcr-MYC3	42000778	35773633	6227145	14.8
<i>Oidma</i>				
Oidma-FLM1	52297206	11668722	40628484	77.7
Oidma-FLM2	59135170	13425192	45709978	77.3
Oidma-FLM3	40110112	8728011	31382101	78.2
Oidma-MYC1	57687306	56343363	1343943	2.3
Oidma-MYC2	42932494	41905372	1027122	2.4
Oidma-MYC3	72438460	71062609	1375851	1.9
<i>Tulca</i>				
Tulca-FLM1	32577878	16863378	15714500	48.2
Tulca-FLM2	58791552	17267873	41523679	70.6
Tulca-FLM3	47093050	11452822	35640228	75.7
Tulca-MYC1	49032928	44511901	4521027	9.2
Tulca-MYC2	53076342	48303122	4773220	9.0
Tulca-MYC3	55877332	50685727	5191605	9.3
<i>Suilu</i>				
Suilu-FLM1	59038940	8817266	50221674	85.1
Suilu-FLM2	56159382	8098246	48061136	85.6
Suilu-MYC1	63685682	52251365	11434317	18.0
Suilu-MYC2	59977216	46892014	13085202	21.8
<i>Paxin</i>				
Paxin-PATCH1	44642676	11657331	32985345	73.9
Paxin-PATCH2	56284686	21878462	34406224	61.1
Paxin-MYC1	60909940	32826978	28082962	46.1
Paxin-MYC2	61537308	36491699	25045609	40.7
<i>Sebve</i>				
Sebve-FLM1	78325404	17862721	60462683	77.2
Sebve-FLM2	84171348	18807318	65364030	77.7
Sebve-FLM3	29640496	6462064	23178432	78.2
Sebve-MYC1	43403650	42617003	786647	1.8
Sebve-MYC2	56060040	55234927	825113	1.5
Sebve-MYC3	59492588	58781740	710848	1.2

4. Supplementary Figures 1 to 28

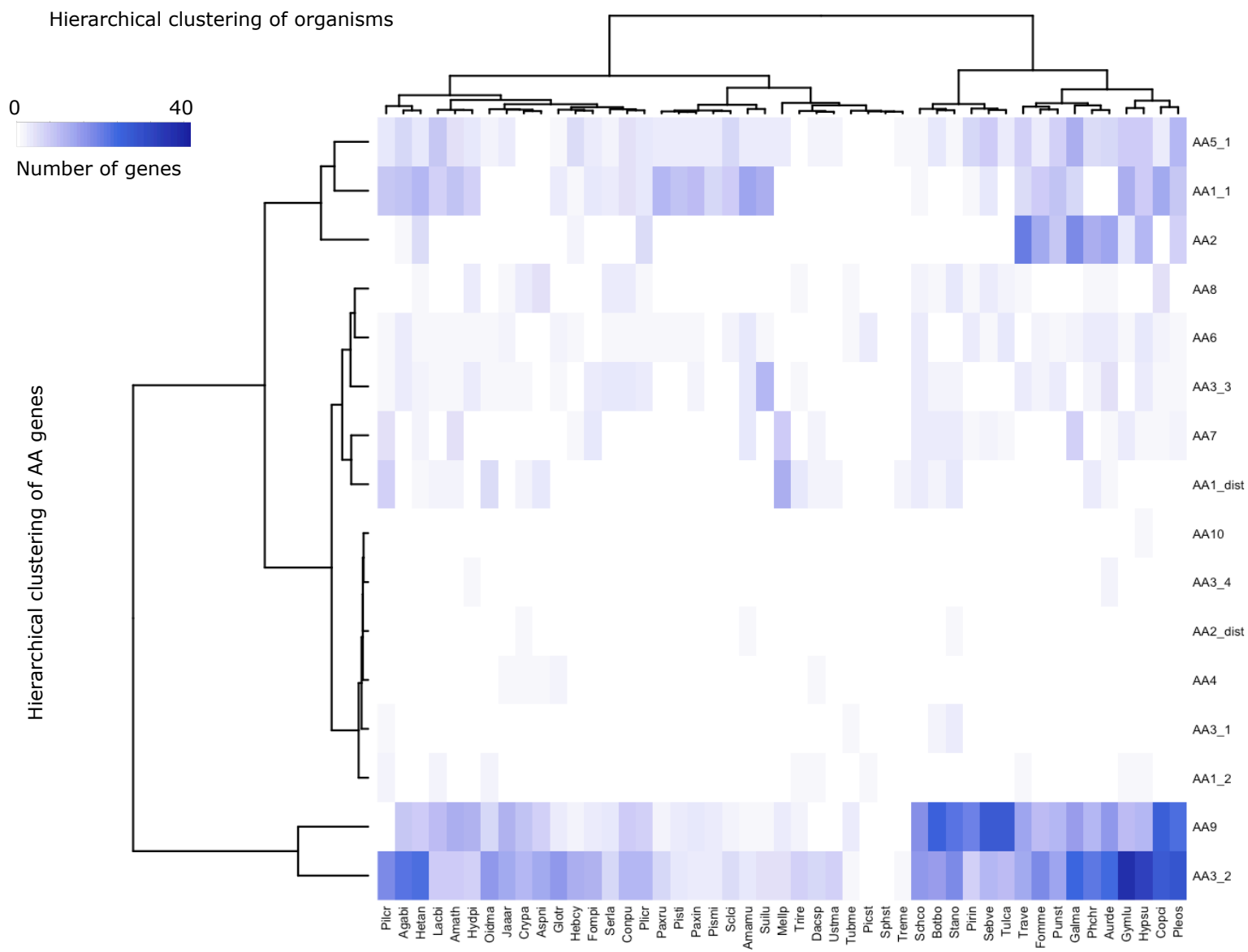


Supplementary Fig. 1. Phylogenomic datasets used in organismal phylogenetic of data collection, filtering steps and the number of loci retained in each step. For analyses and the pipeline explanation, please refer to the text.



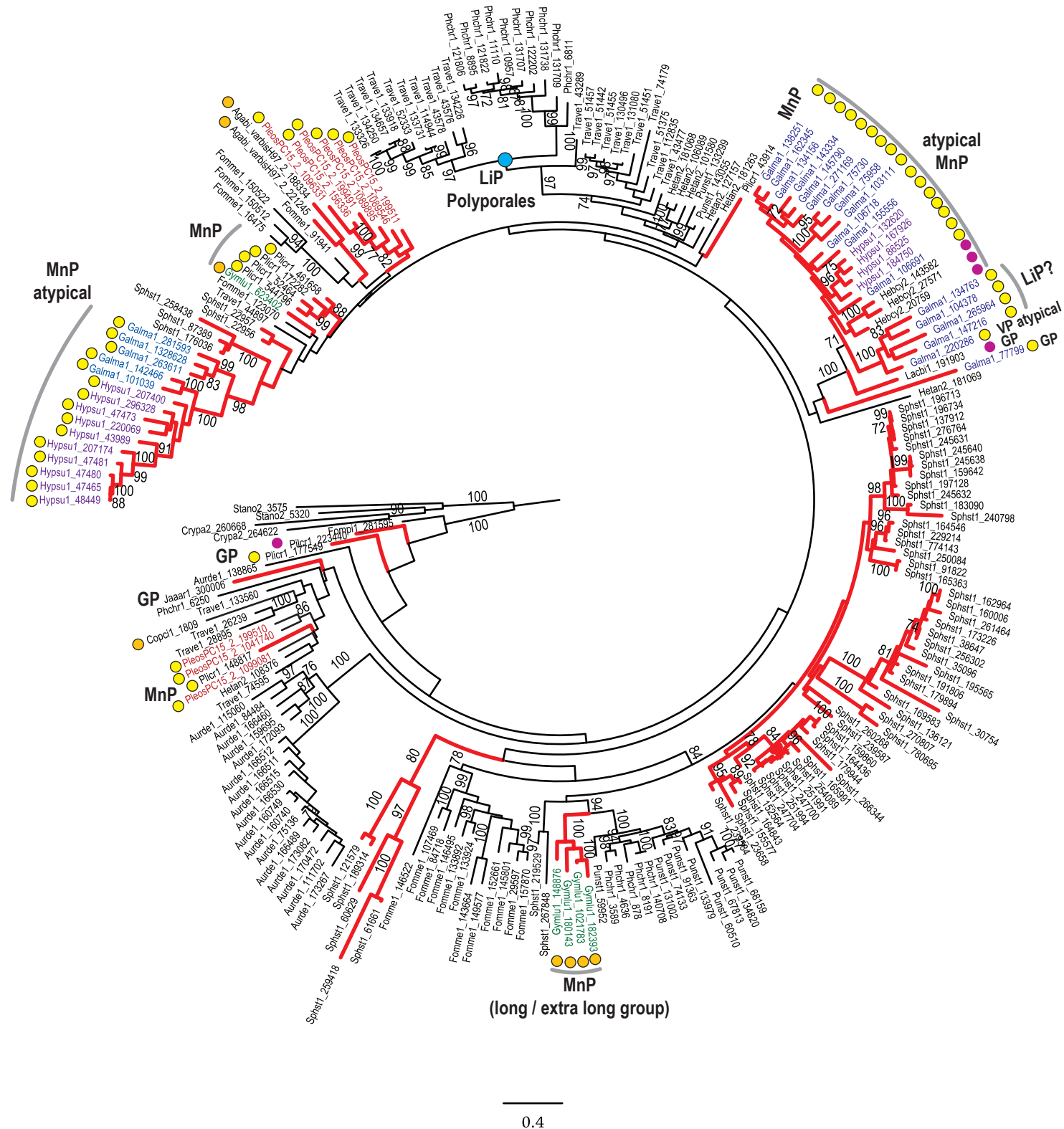


a

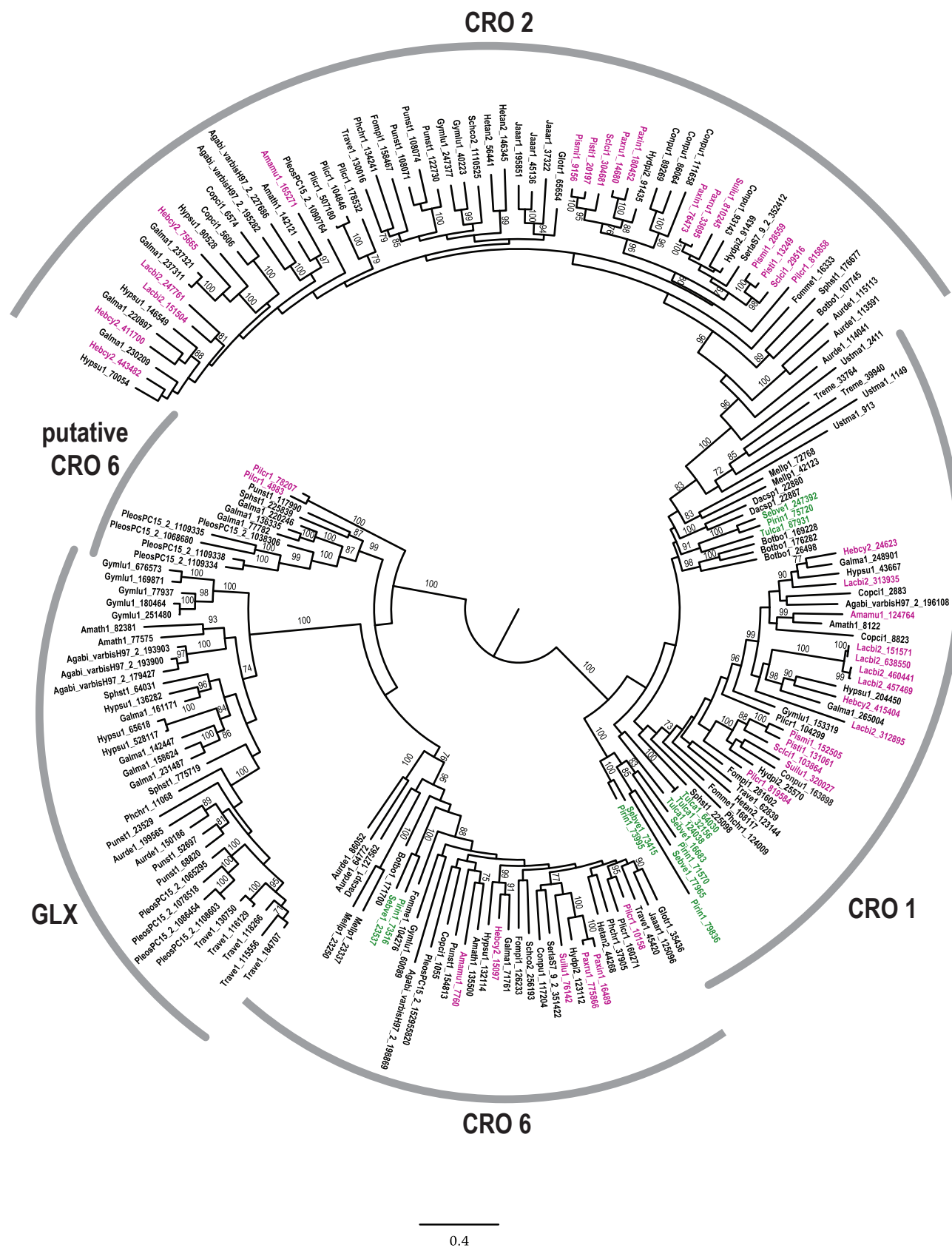


b

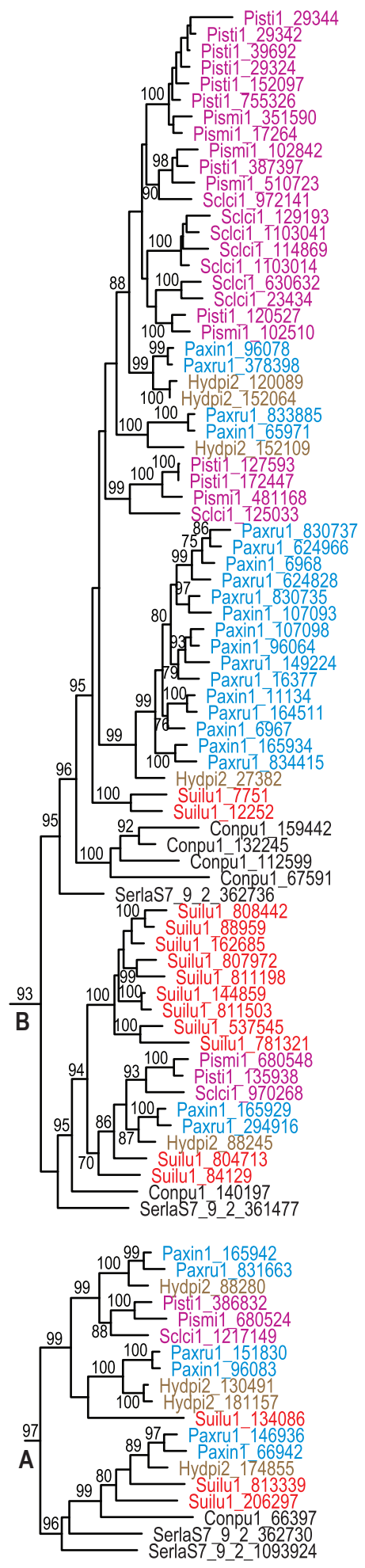
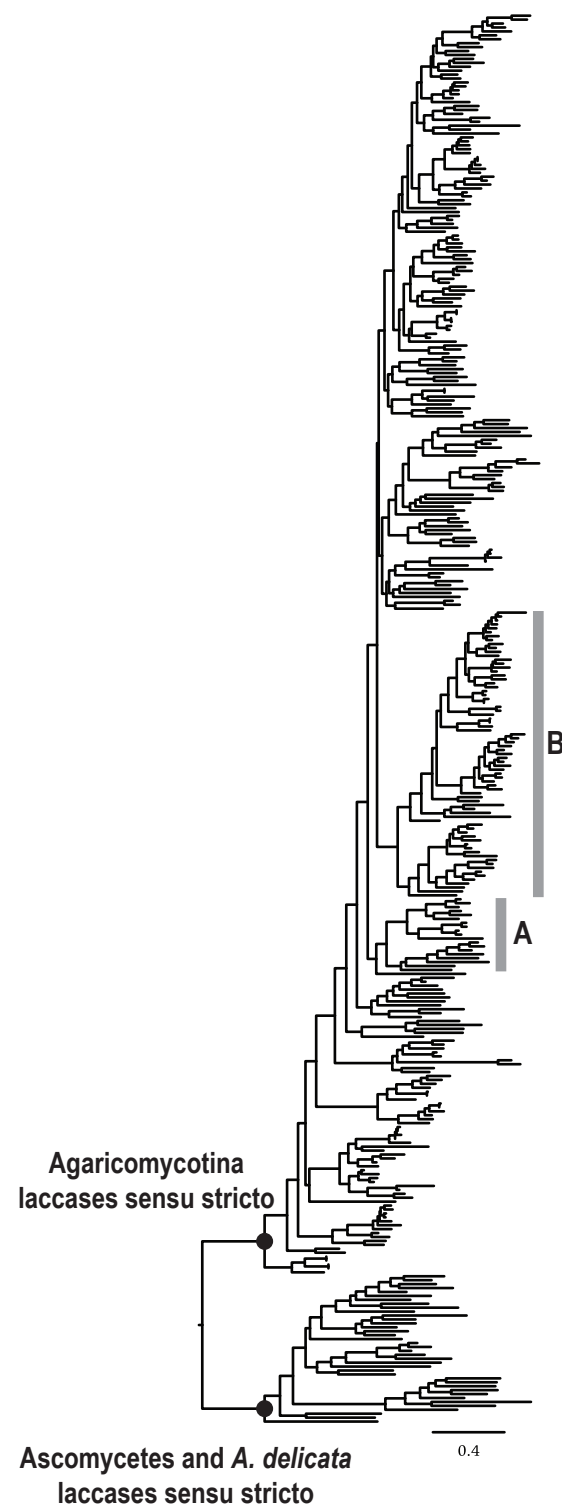
Supplementary Fig. 3 Double-hierarchical clustering of selected CAZymes (a) and auxiliary redox enzyme (AA) families (b). Gene families and organisms are hierarchically clustered. Selected CAZymes are families targeting Cellulose (GH6, 7, 12, 44, 45), targeting hemicellulose (GH 10, 11, 26, 30, 51, 74, 43) and families targeting pectin (GH43, 28, 53, 78, 88, 105, PL1, 3, 4, CE8, 12).



Supplementary Fig. 4. Phylogeny of POD protein sequences. Bootstrap values above 70 are shown. Clades with newly generated sequences are highlighted in red. The Agaricales and Amylocorticiales sequences are indicated with colored circles depending on the nutritional strategy of the species (purple: ECM, yellow: white rot, orange: soil saprotrophs). In Agaricales, most POD genes are found in white rot species compared to ECM and soil saprotrophs. Subclassification for newly generated sequences (excluding *S. stellatus*); MnP: manganese peroxidase, LiP: lignin peroxidase, VP: versatile peroxidase, GP: generic peroxidase. The light blue circle indicates the clade of Polyporales LiP sequences.

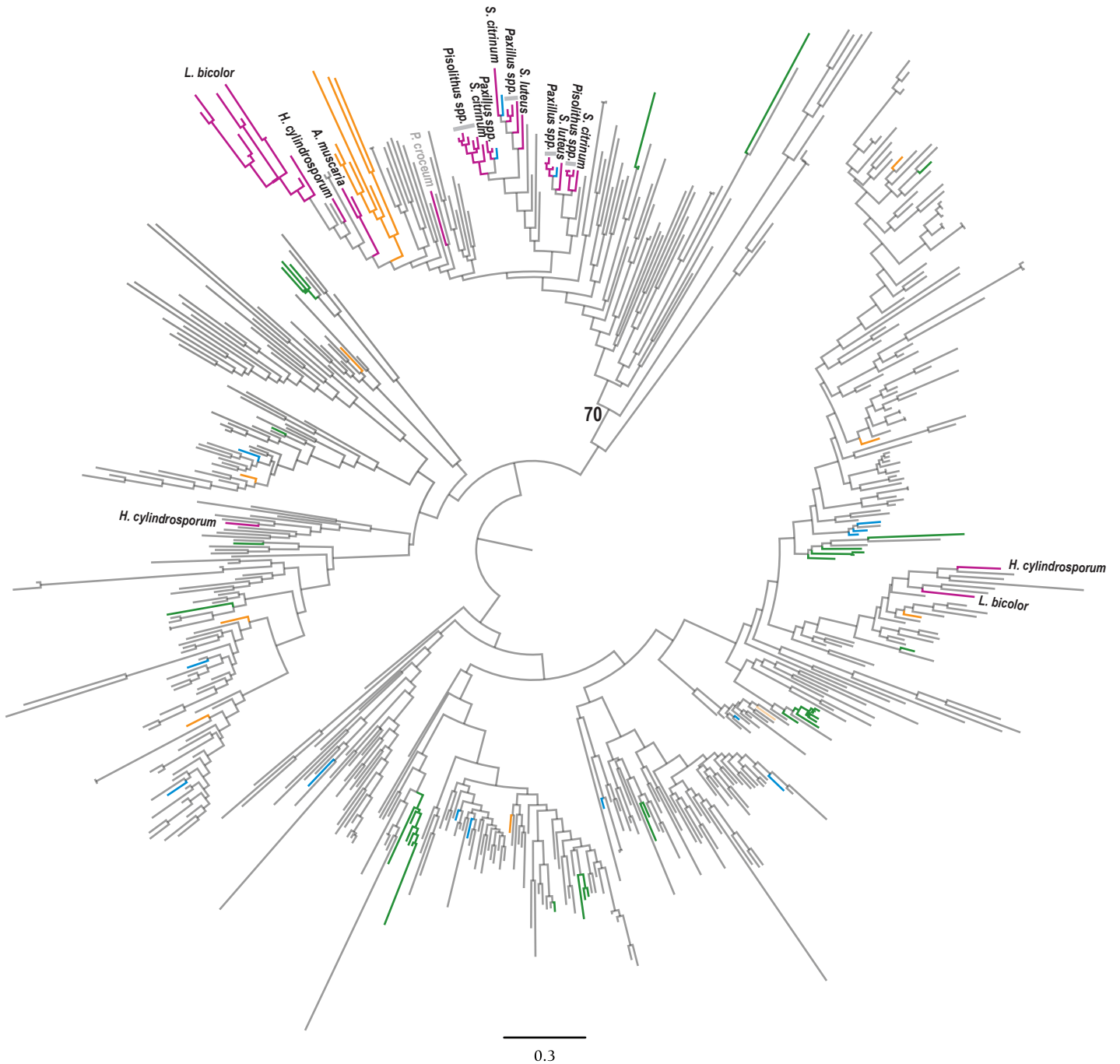


Supplementary Fig. 5. Phylogeny of copper radical oxidases 1, 2, 6, GLX protein sequences. Bootstrap values above 70 are shown. Sequences from ECM Agaricomycetidae genomes are indicated in purple. Sequences from the symbiotic *S. vermifera*, *T. calospora* and *P. indica* are indicated in green. CRO 1, 2 and 6 include sequences from species with various nutritional strategies, whereas GLX include sequences only from white rot or soil saprotrophs.

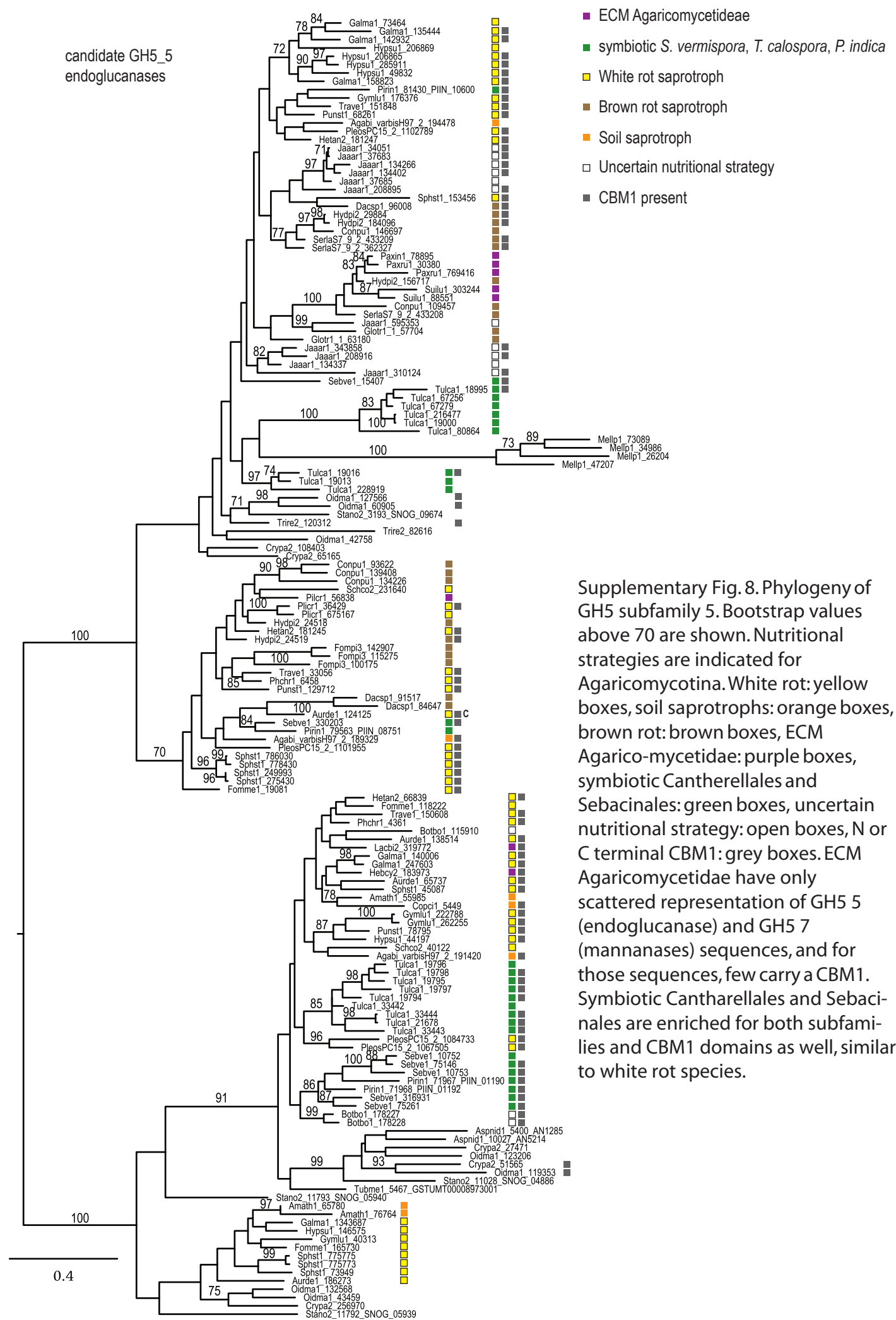


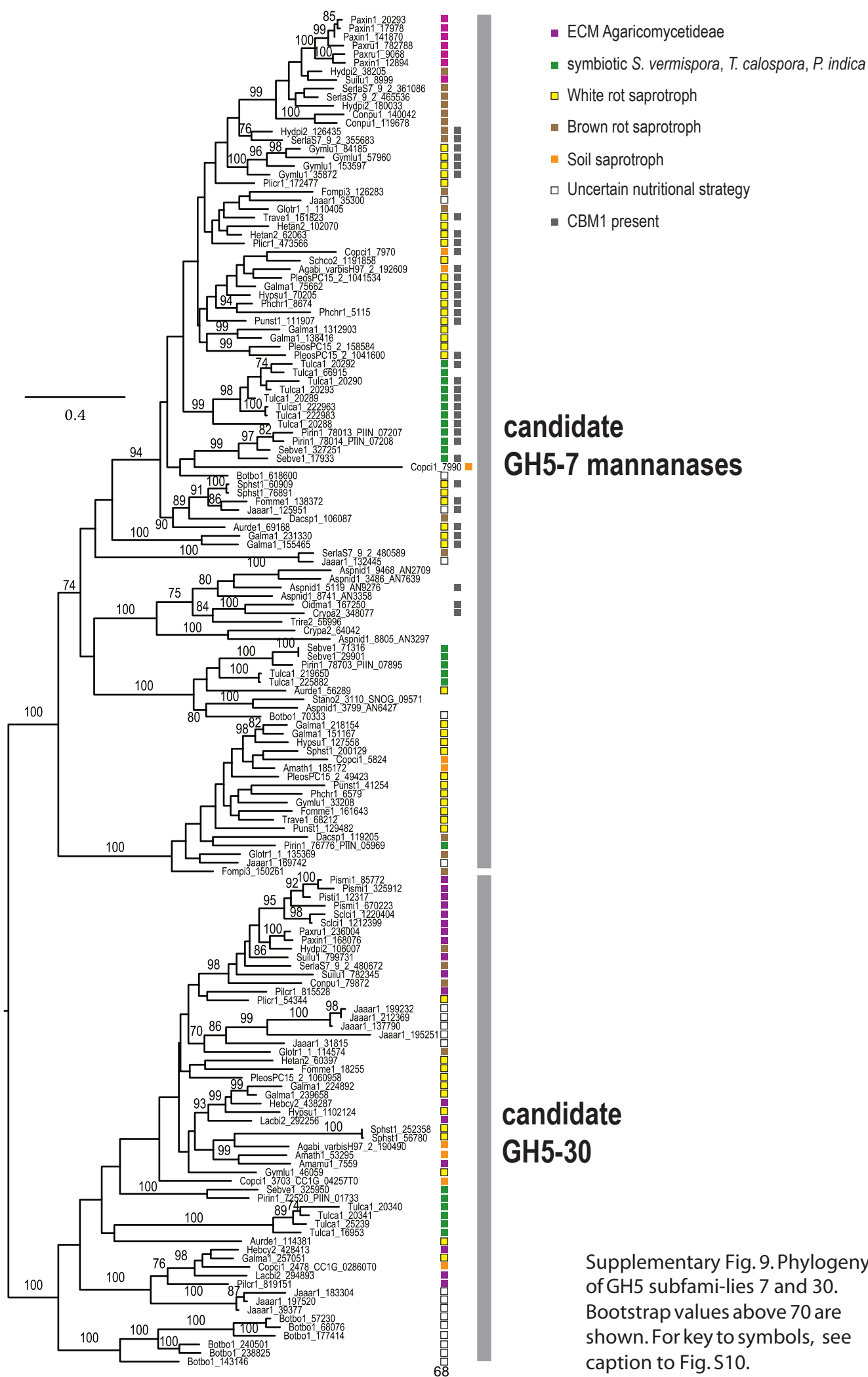
Supplementary Fig. 6. Phylogeny of basidiomycete and ascomycete laccases. Phylogenetic analysis of basidiomycete and ascomycete laccases sensu stricto (left) and detailed representation of the two clades (A and B) of Boletales laccases sensu stricto (right). Bootstrap values above 70 are shown. Sequences from the Sclerodermatineae, *S. luteus* and *Paxillus* species are in purple, red and blue respectively. Sequences from the brown rot *H. pinastri* are colored in brown, while the basal brown rot *S. lacrymans* and *C. puteana* are colored black.

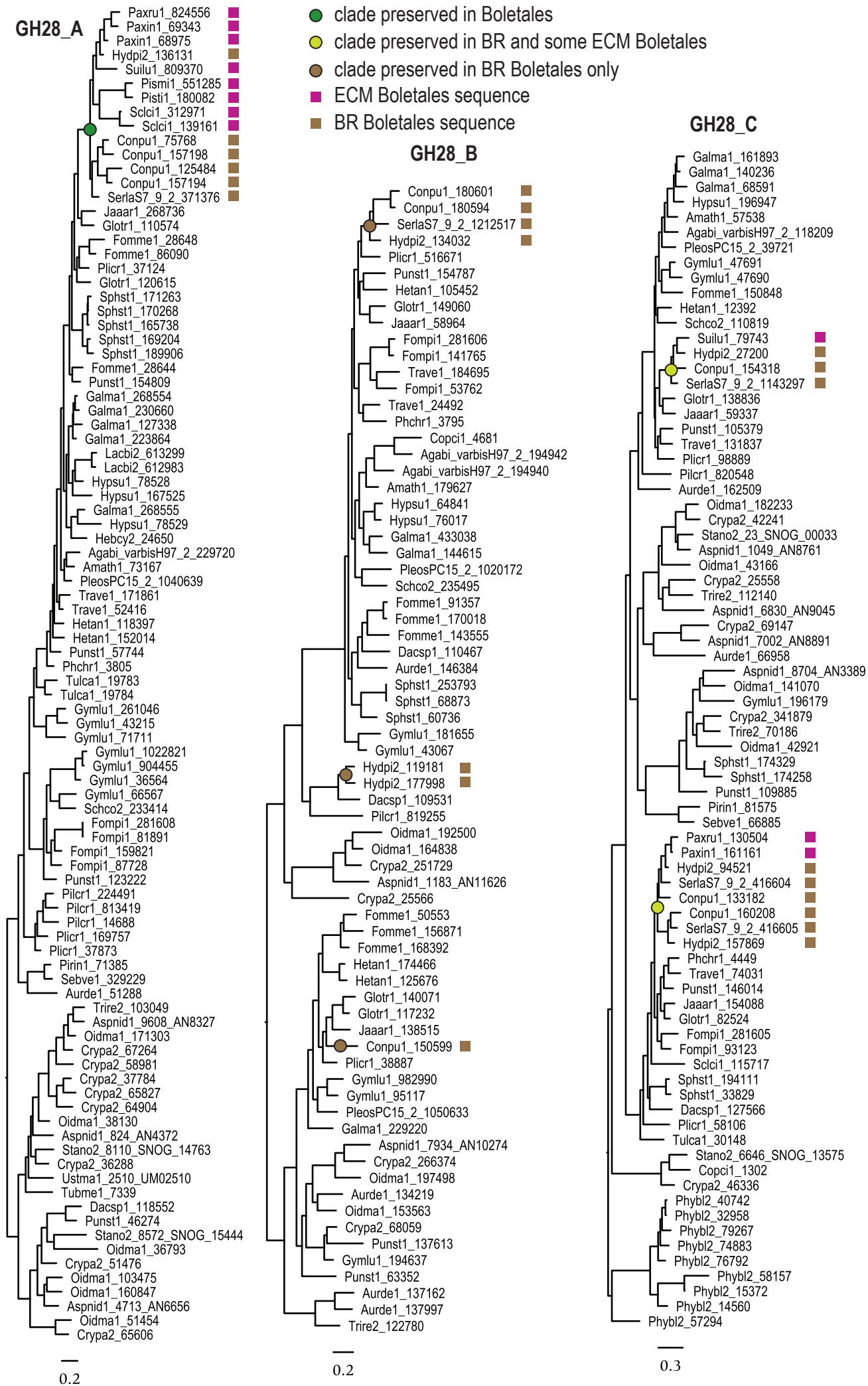
- *H. pinastri* GH61 genes (brown rot)
- ectomycorrhizal Agaricomycetidaeae GH61 genes
- *Amanita thiersii* GH61 genes (saprotroph)
- *S. vermifera* GH61 genes (orchid mycorrhizal)



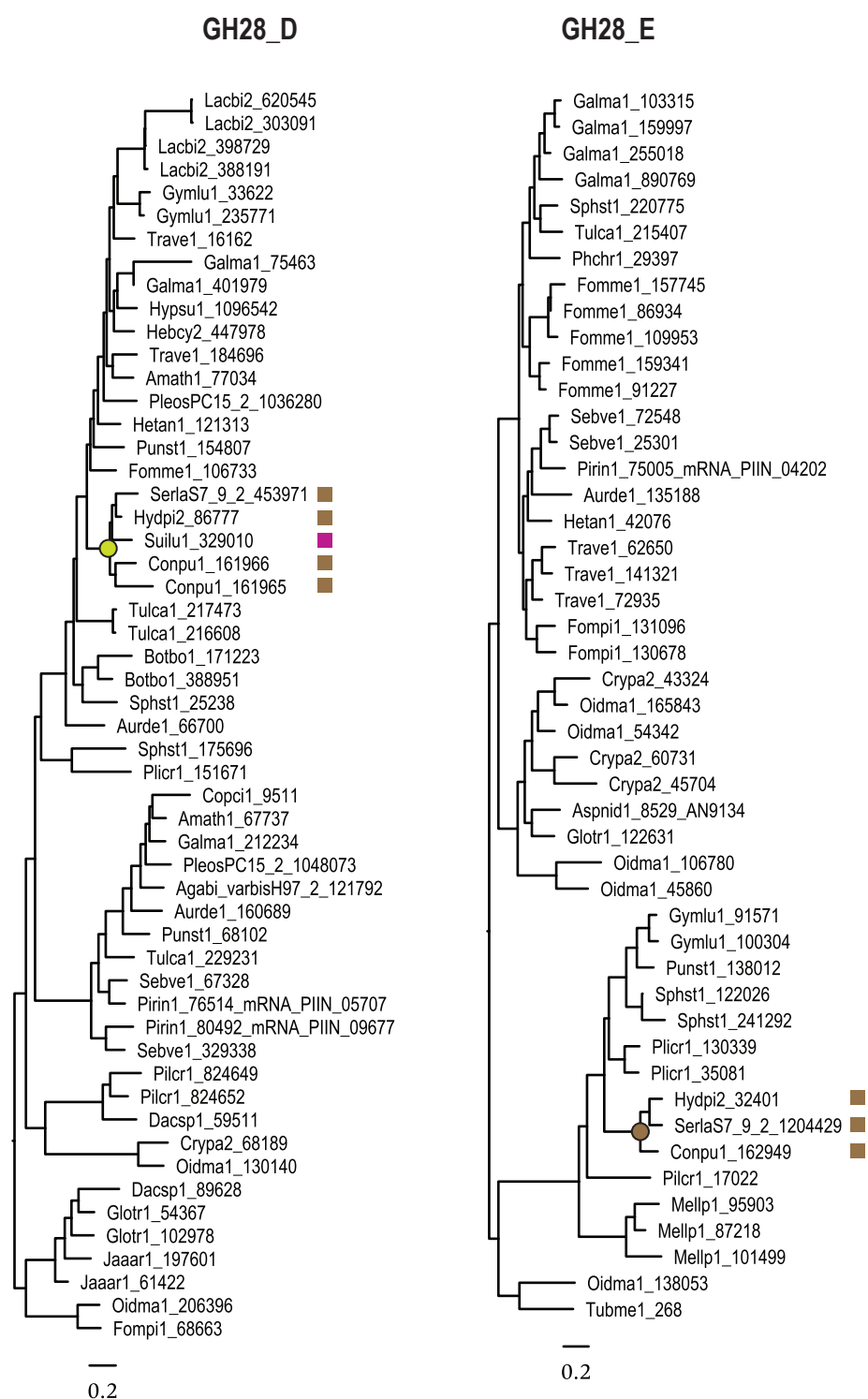
Supplementary Fig. 7. Phylogeny of LPMO protein sequences. Branches with sequences from ECM species in Agaricomycetidae, the orchid mycorrhizal *S. vermifera*, the soil saprotroph *A. thiersii* and the brown rot *H. pinastri* are highlighted in purple, green, yellow and blue respectively. ECM species in Agaricomycetidae maintain sequences in one clade of the LPMO phylogeny. *Hebeloma cylindrosporum* and *L. bicolor* are the only exceptions, with three sequences out of this clade. The rest of the highlighted species have numerous scattered sequences across the LPMO phylogeny suggesting that they have maintained many of the ancestral copies of LPMO.







Supplementary Fig. 10a. For caption see Fig. S10b.



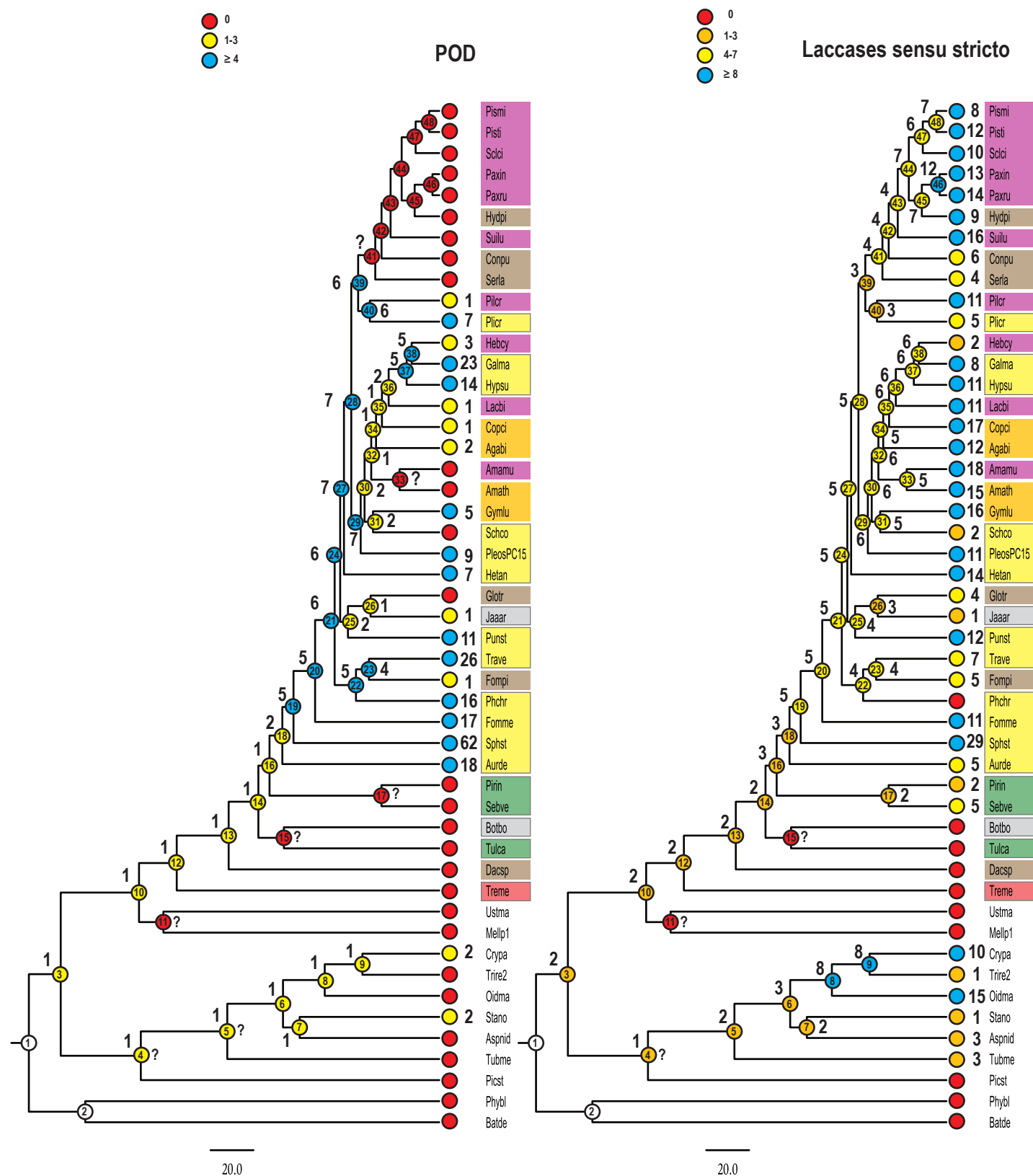
Supplementary Fig. 10. Phylogeny of GH28 subfamilies. Supplementary Fig. 10a (previous page): GH28 subfamilies A, B and C.

Supplementary Fig.10b (this page): GH28 subfamilies D and E. ECM and brown rot Boletales sequences are indicated with purple and brown squares respectively. Brown, light green and dark green circles indicate branches with sequences from brown rot Boletales only, from brown rot and only some of the ECM Boletales, and from all Boletales of the study, respectively. There are 8 branches including Boletales sequences for GH28 pectinases, but only one has sequences from all ECM species, while three have sequences from some but not all ECM species. The last four branches have sequences retained only in brown rot Boletales. *Hydnomerulius pinastri* has retained sequences in 7 of these clades.

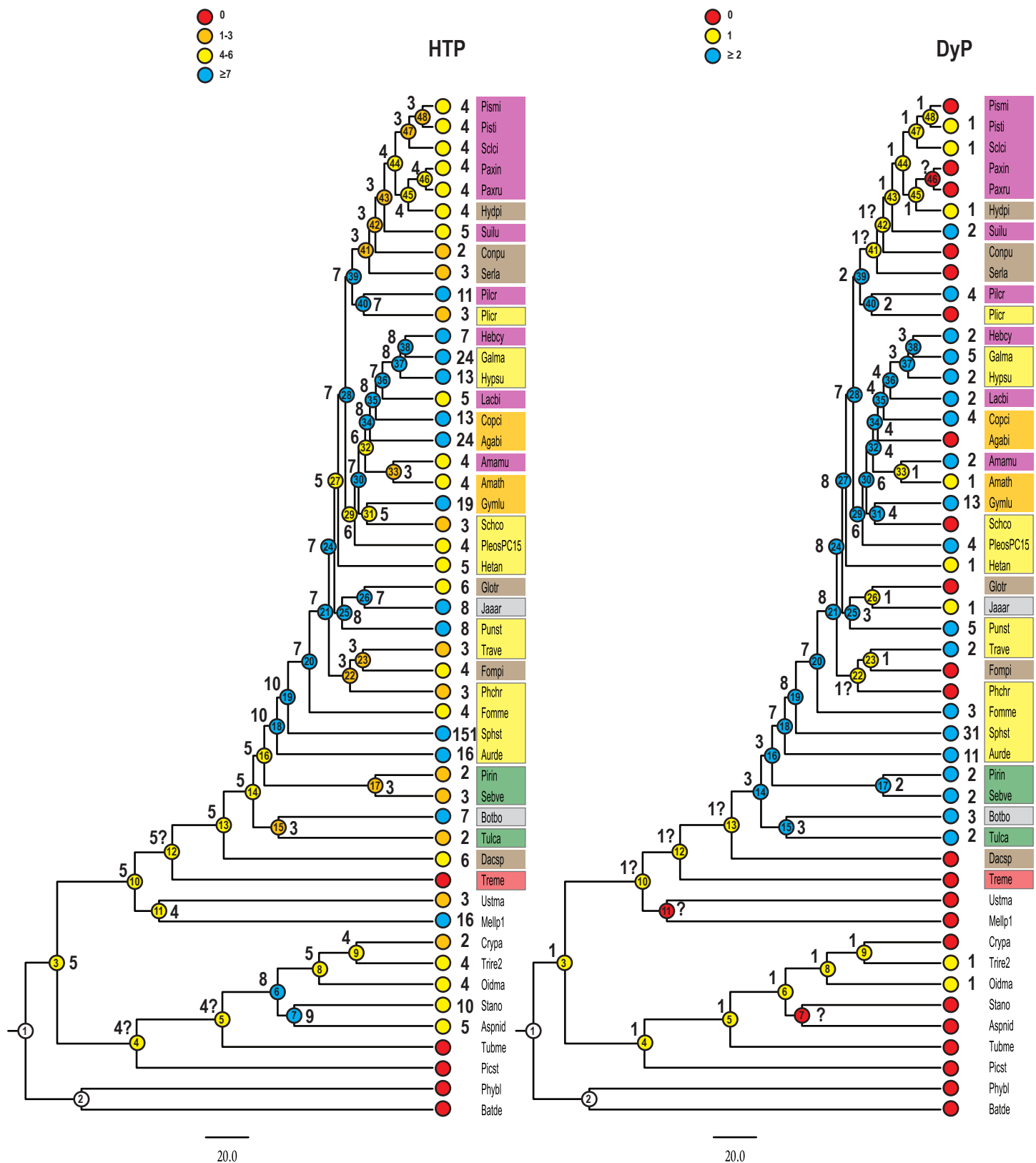
- clade preserved in Boletales
- clade preserved in BR and some ECM Boletales
- clade preserved in BR Boletales only
- ECM Boletales sequence
- BR Boletales sequence



Supplementary Fig. 11. Phylogeny of GH3 B protein sequences. As in GH28, there are 4 clades including Boletales sequences, but only one has sequences from all ECM Boletales and the rest have been retained in brown rot Boletales only. *Hydnomerulius pinastri* has retained copies in three of them

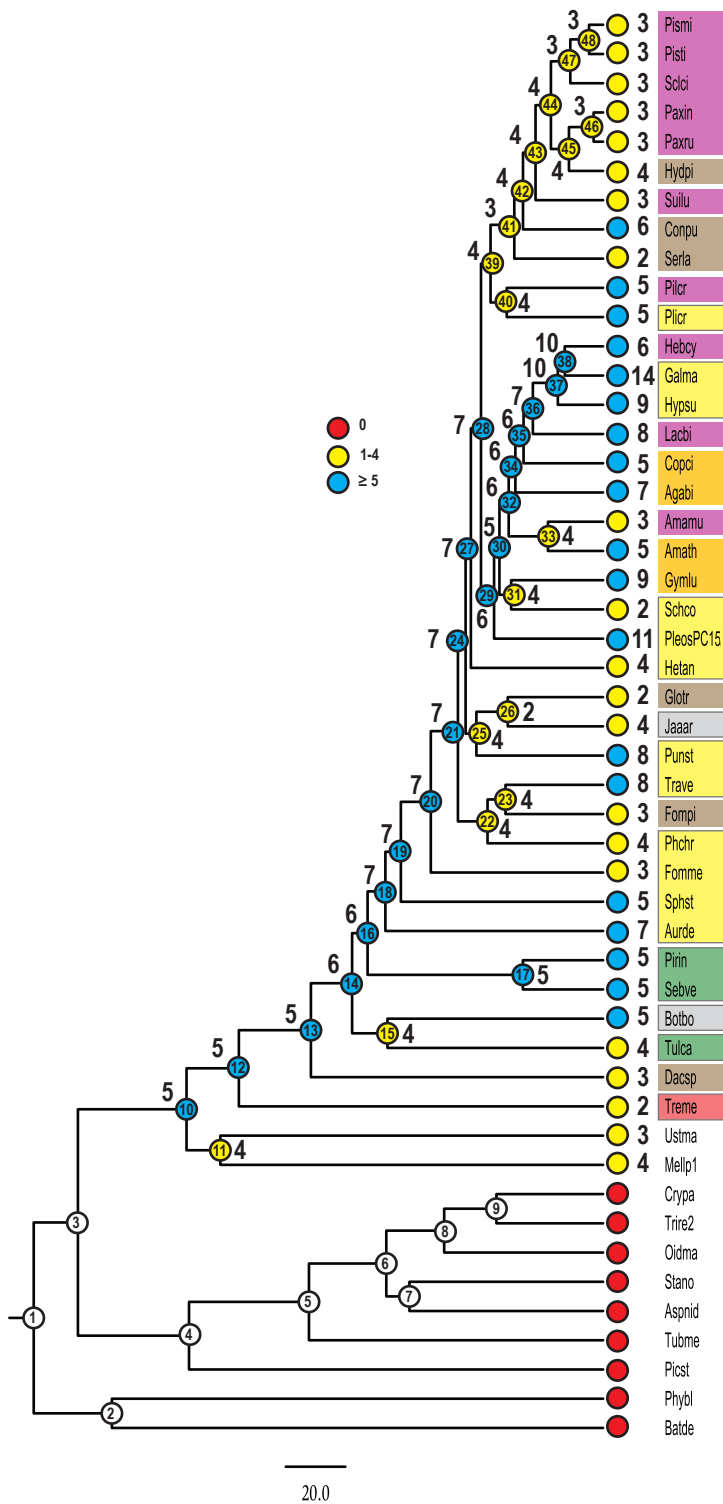


Supplementary Fig.12. Evolution of POD gene copy numbers. Organismal phylogeny from analysis of the PR0.95 dataset, with gene copy numbers at internal nodes estimated with Notung. Nodes or tips with gene copy numbers at or above the average for Agaricomycotina are in blue, those with below average copy numbers are yellow or orange, and nodes or tips with no copies of the gene family are red. Numbers within circles are node numbers. Numbers at the tips and adjacent to nodes represent the number of genes observed or reconstructed (respectively). Numbers with question marks indicate uncertain estimates through reconciliation under the strictest parsimonious scenario of gene losses.

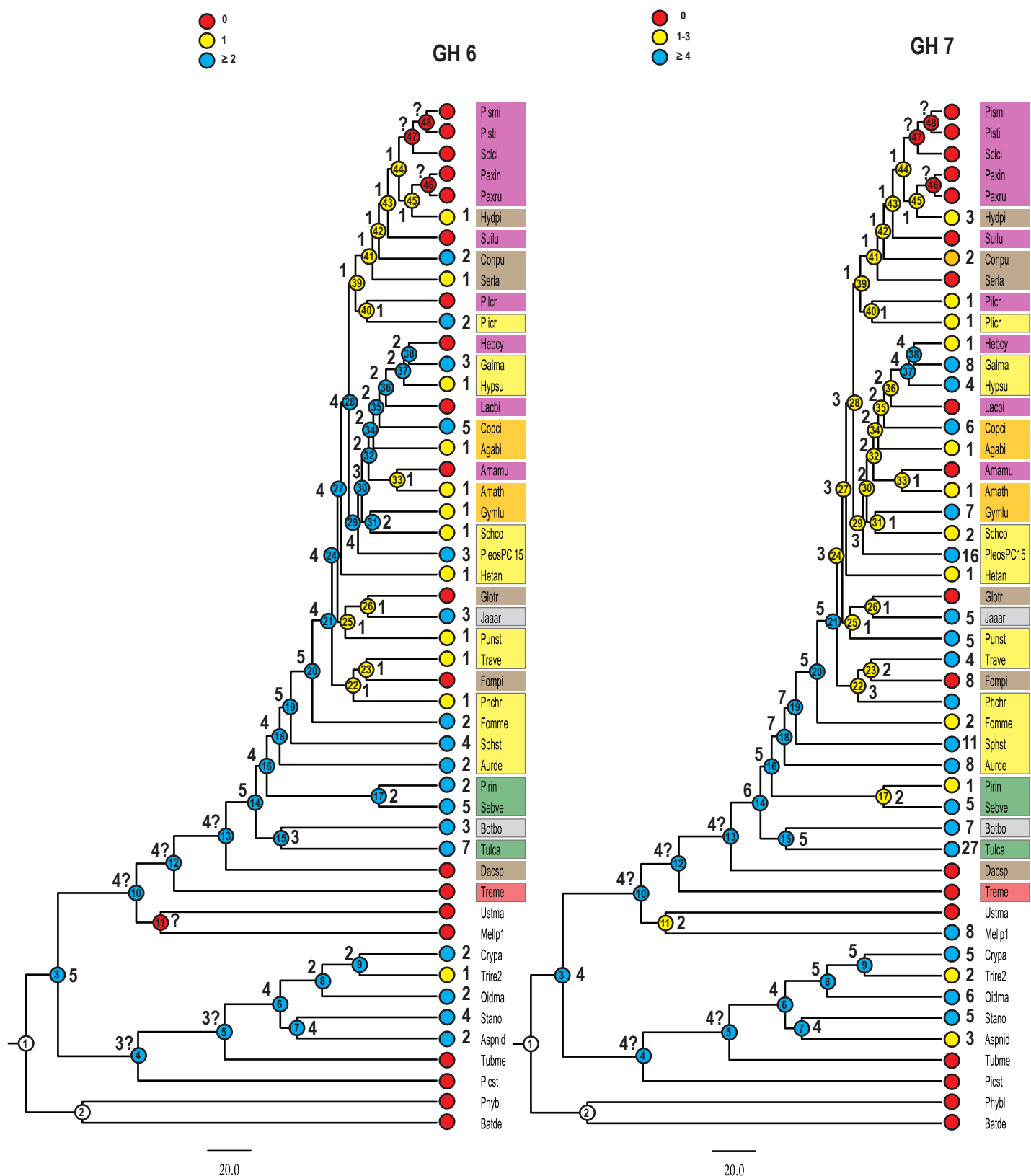


Supplementary Fig. 13. Evolution of HTP and DyP gene copy numbers. For explanation of symbols, see caption to Supplementary Fig. 12.

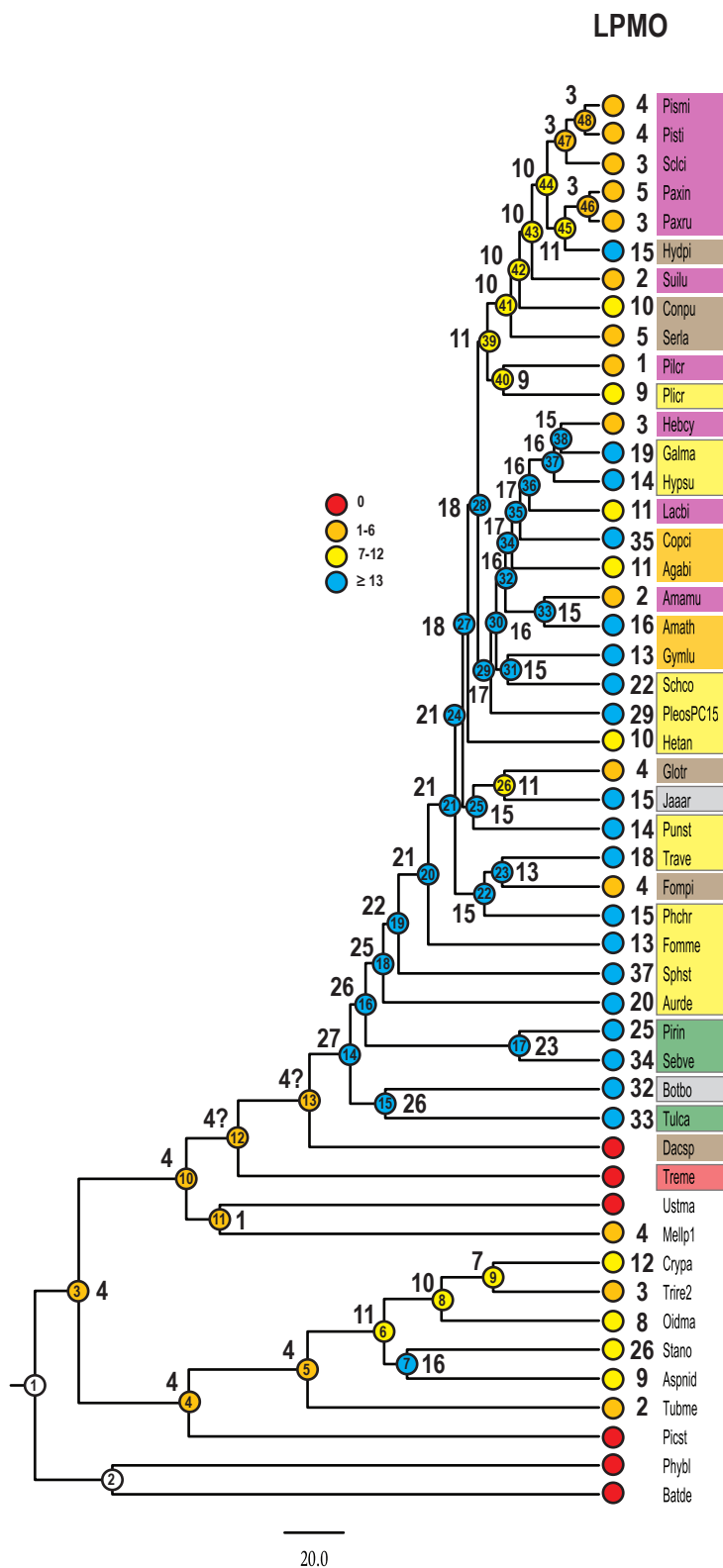
CRO 1-2-6-GLX



Supplementary Fig. 14. Evolution of CRO 1, 2, 6 and GLX gene copy numbers. For explanation of symbols, see caption to Supplemental Fig. 12.



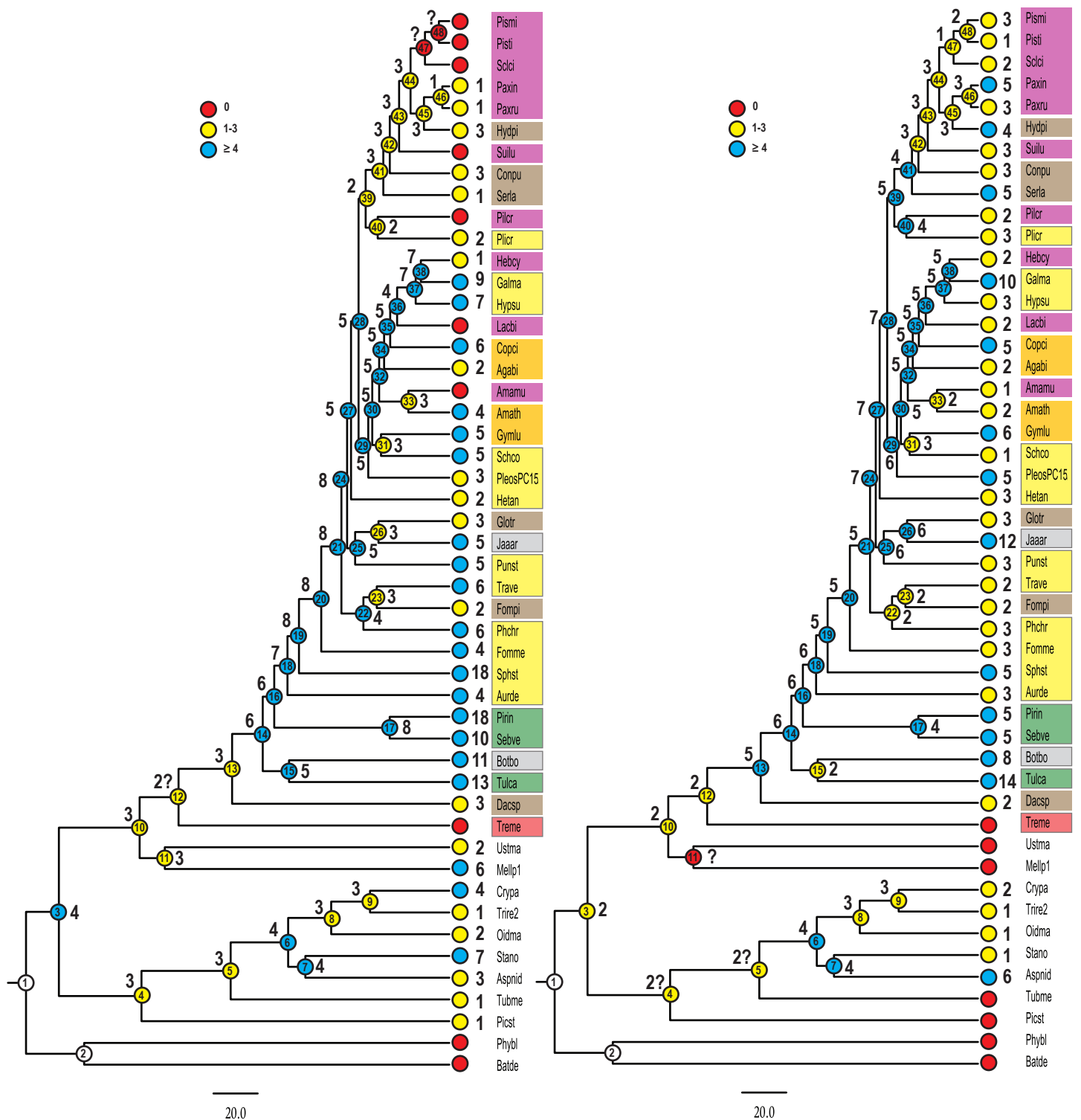
Supplementary Fig. 15. Evolution of GH6 and GH7 gene copy numbers. For explanation of symbols, see caption to Supplemental Fig. 12.



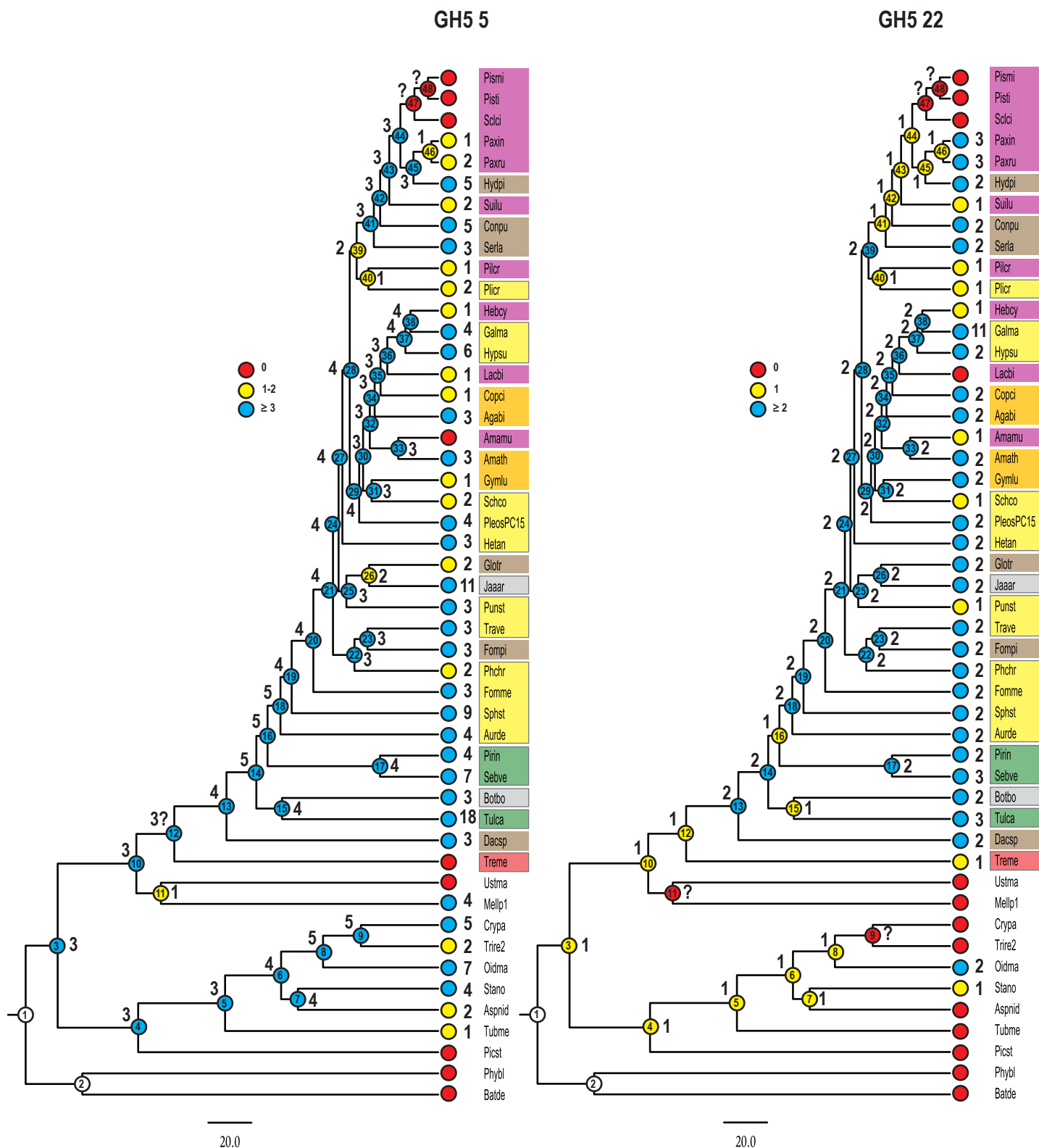
Supplementary Fig. 16. Evolution of LPMO gene copy numbers. For explanation of symbols, see caption to Supplementary Fig. 12.

GH 10

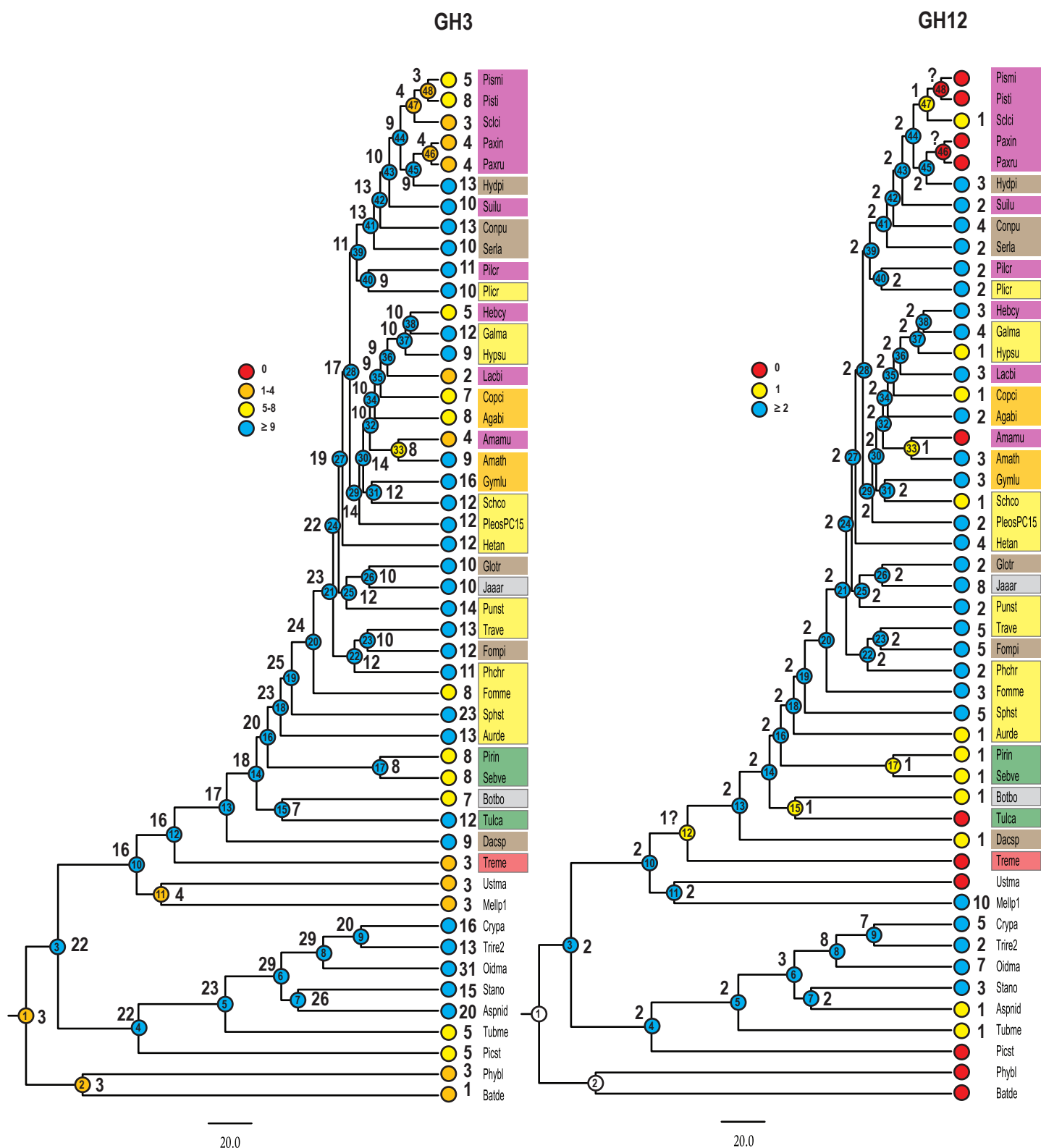
GH5 7 & 30



Supplementary Fig. 17. Evolution of GH10, GH5_7 and GH5_30 gene copy numbers. For explanation of symbols, see caption to Supplementary Fig. 12.

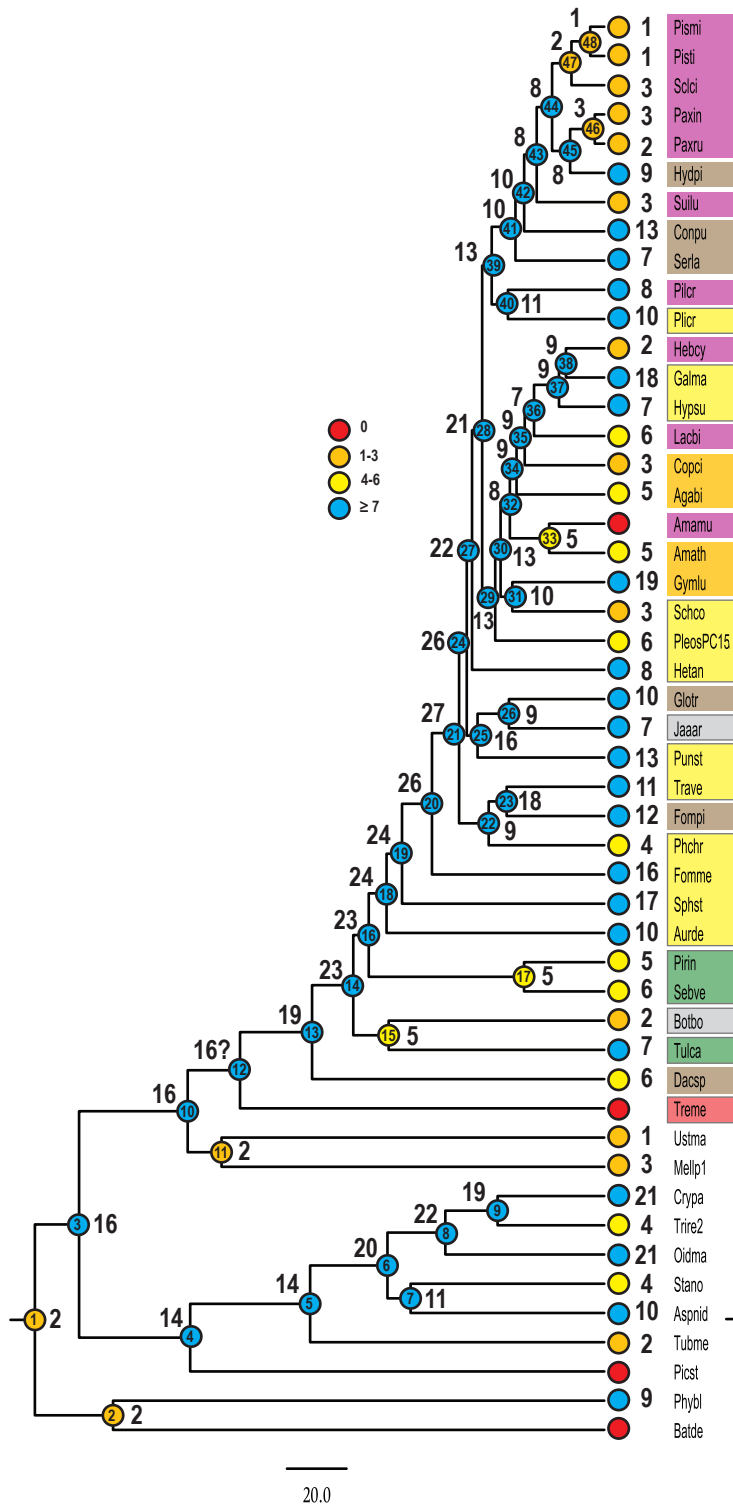


Supplementary Fig. 18. Evolution of GH5 and GH22 gene copy numbers. For explanation of symbols, see caption to Supplementary Fig. 12.

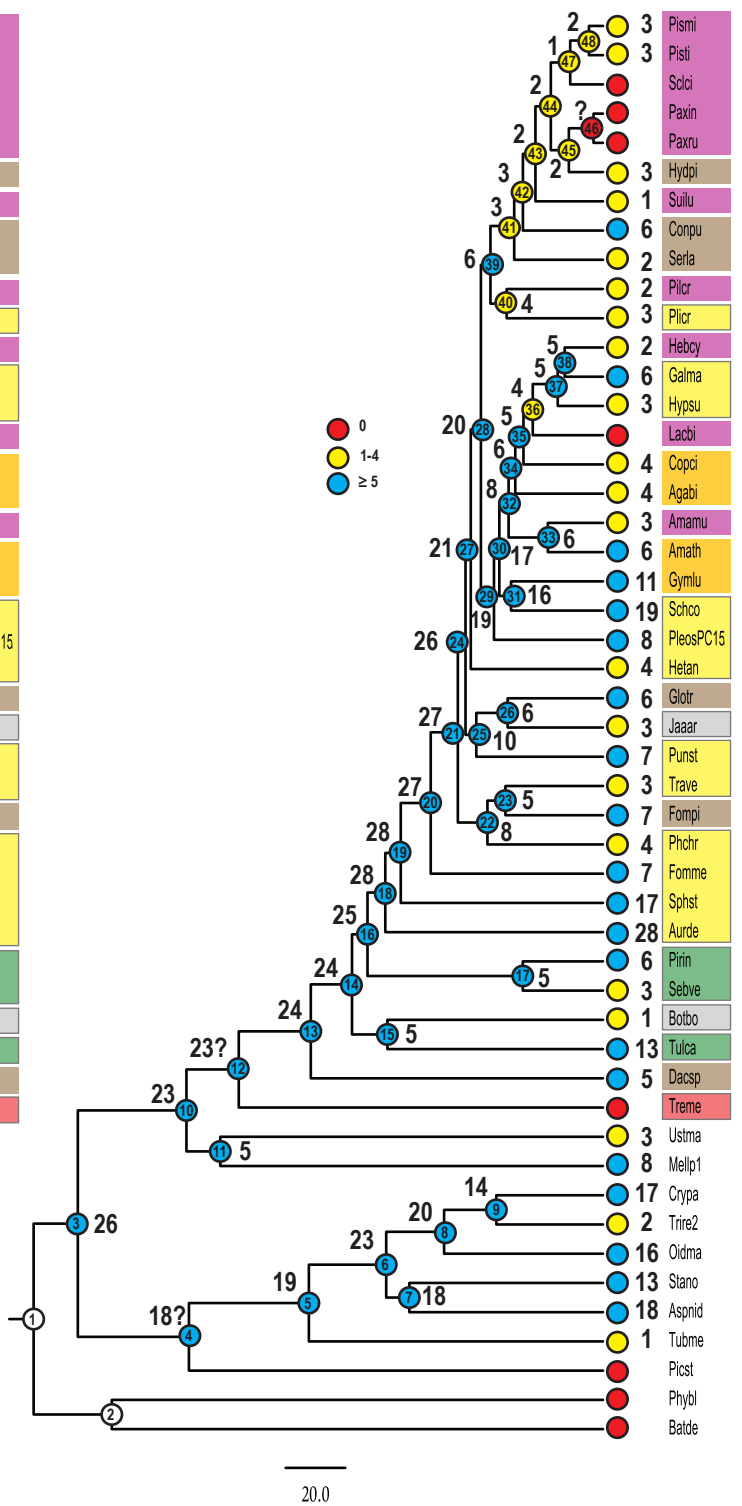


Supplementary Fig. 19. Evolution of GH3 and GH12 gene copy numbers. For explanation of symbols, see caption to Supplementary Fig. 12.

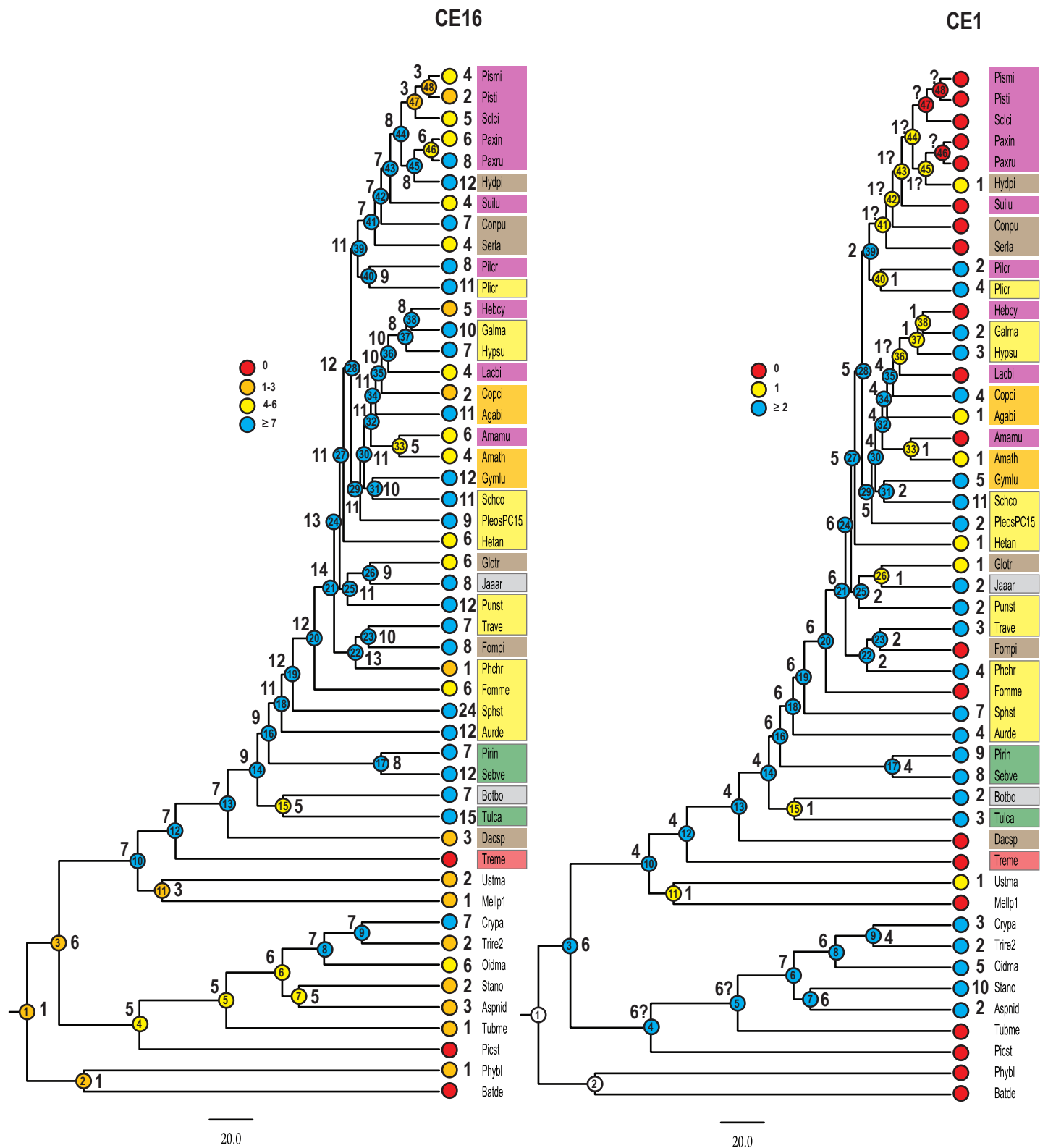
GH 28



GH 43

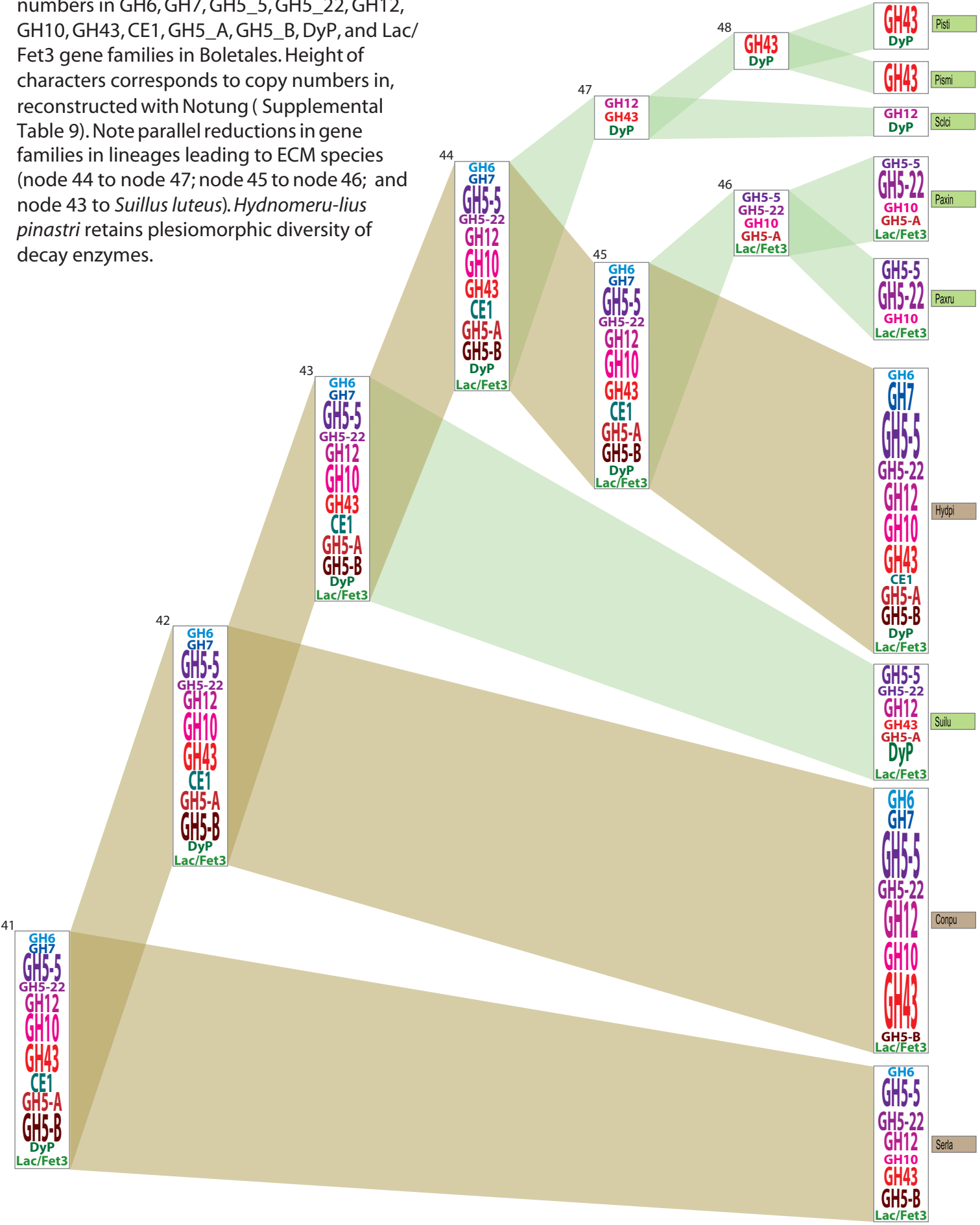


Supplementary Fig. 20. Evolution of GH28 and GH43 gene copy numbers. For explanation of symbols, see caption to Supplementary Fig. 12.

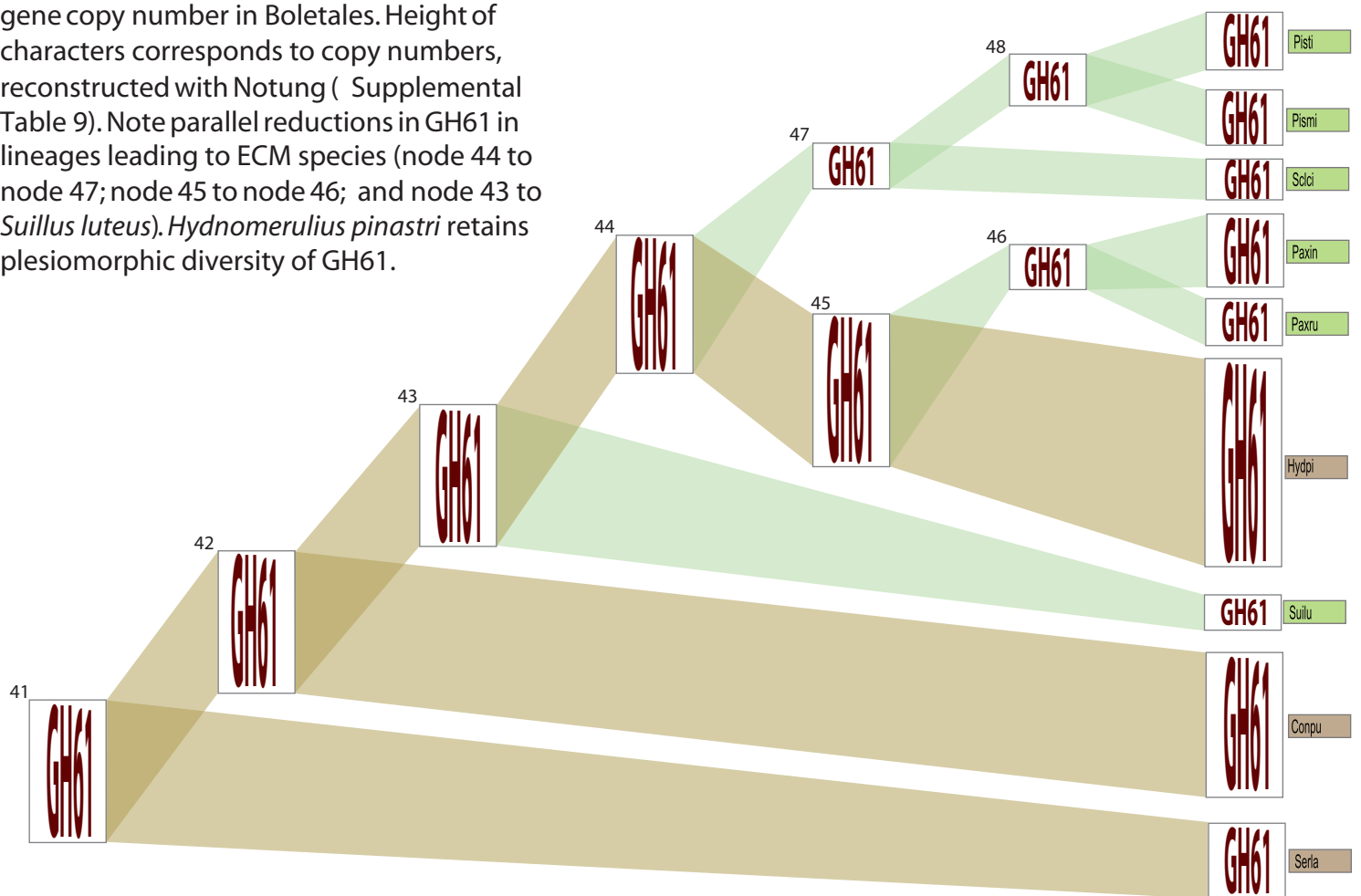


Supplementary Fig. 21. Evolution of CE16 and CE1 gene copy numbers. For explanation of symbols, see caption to Supplementary Fig. 12.

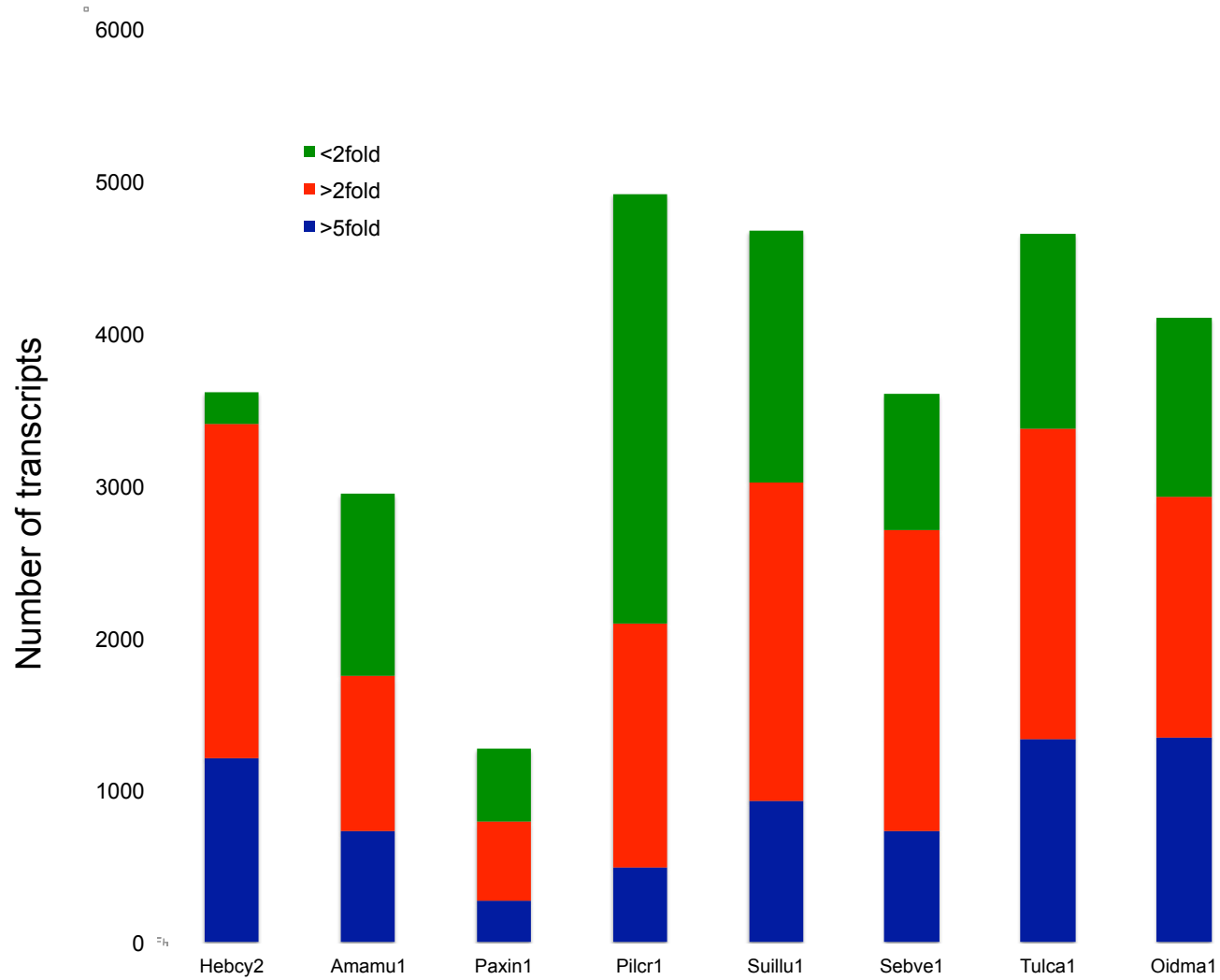
Supplementary Fig. 22. Changes in copy numbers in GH6, GH7, GH5_5, GH5_22, GH12, GH10, GH43, CE1, GH5_A, GH5_B, DyP, and Lac/Fet3 gene families in Boletales. Height of characters corresponds to copy numbers in, reconstructed with Notung (Supplemental Table 9). Note parallel reductions in gene families in lineages leading to ECM species (node 44 to node 47; node 45 to node 46; and node 43 to *Suillus luteus*). *Hydnomeru-lius pinastri* retains plesiomorphic diversity of decay enzymes.

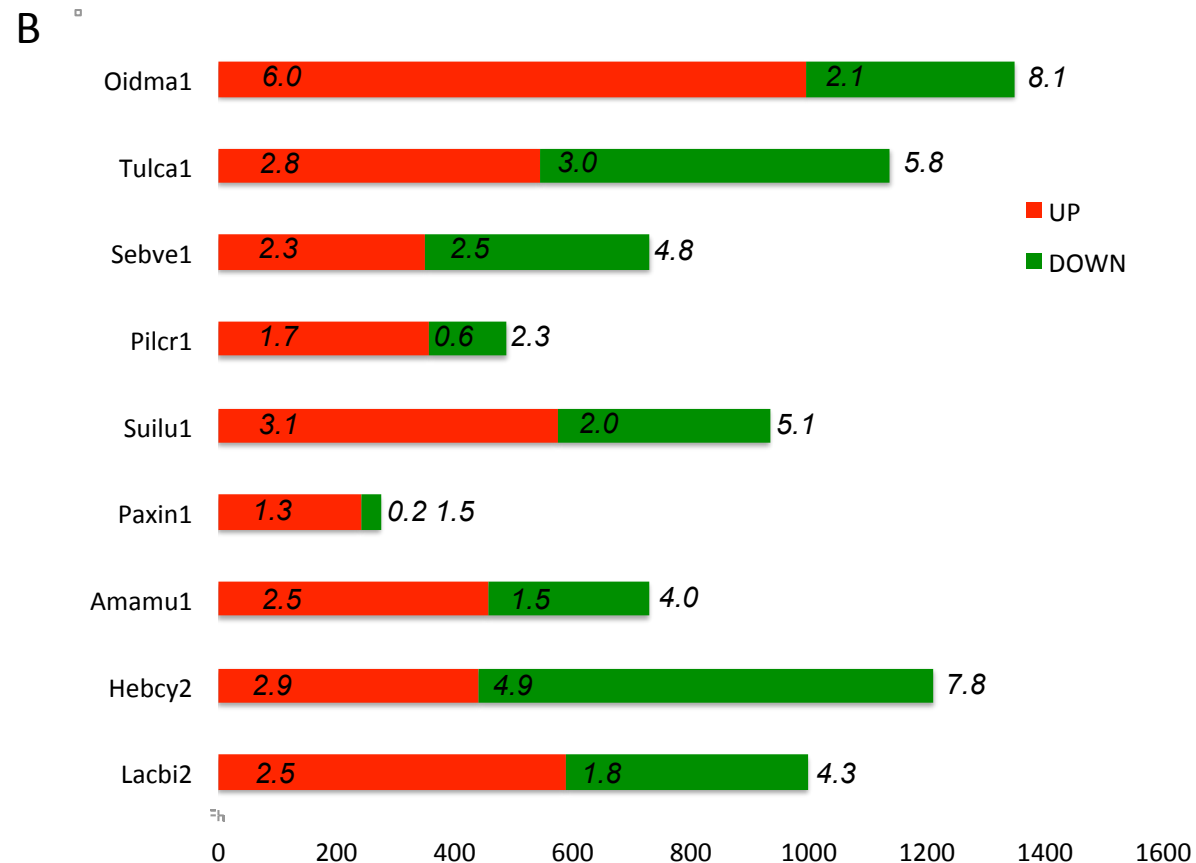


Supplementary Fig. 23. Changes in GH61 gene copy number in Boletales. Height of characters corresponds to copy numbers, reconstructed with Notung (Supplemental Table 9). Note parallel reductions in GH61 in lineages leading to ECM species (node 44 to node 47; node 45 to node 46; and node 43 to *Suillus luteus*). *Hydnomerulius pinastri* retains plesiomorphic diversity of GH61.



A



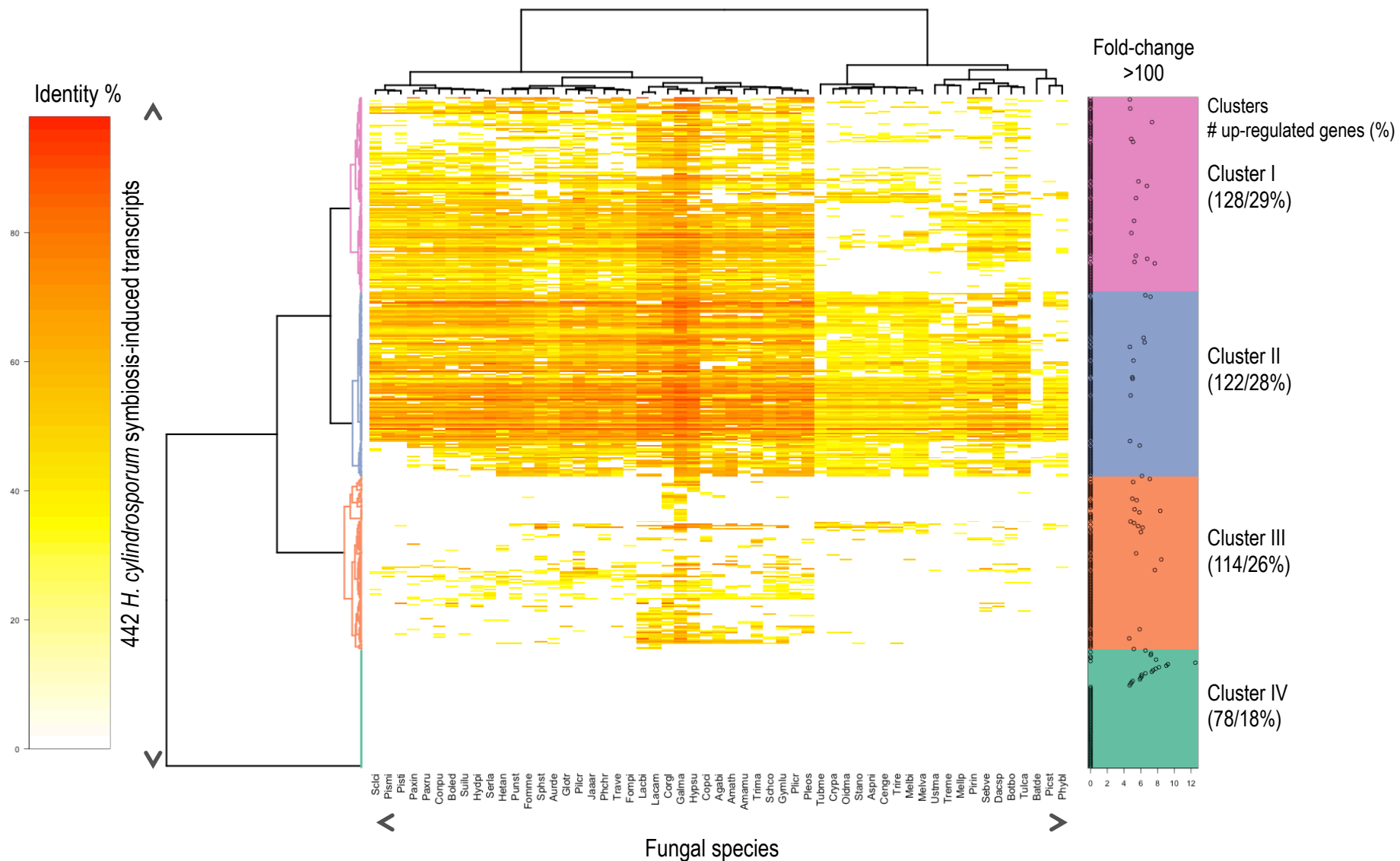


Supplementary Fig. 24. Number of transcripts regulated during mycorrhiza formation.

A Number of transcripts significantly (FDR corrected p-value<0.05) regulated, >2-fold and >5fold compared to free-living mycelium.

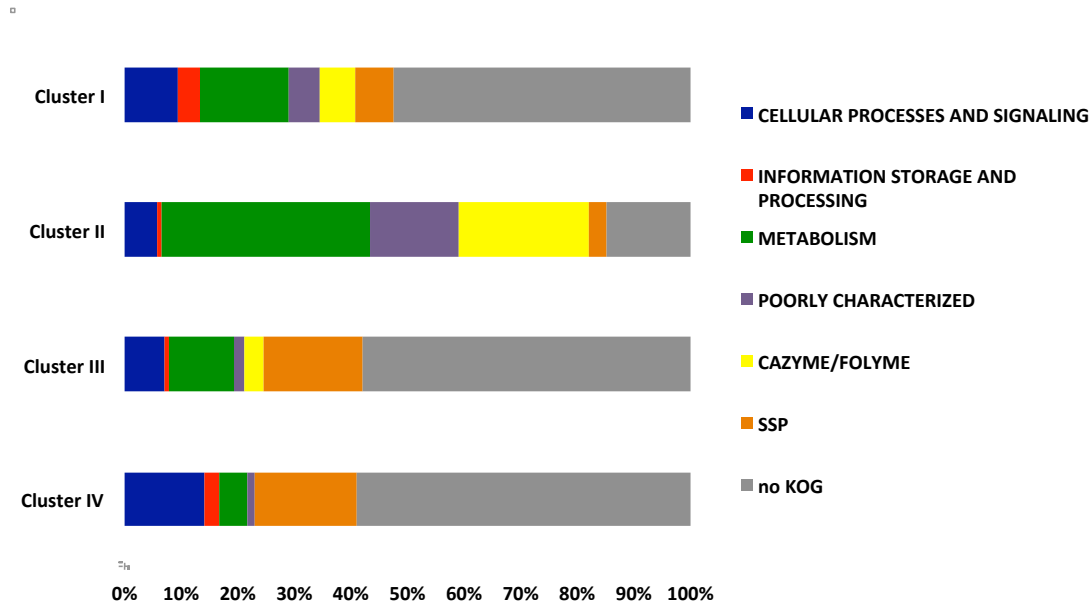
B The number of transcripts with higher transcript concentrations in mycorrhizal tissue then in free-living mycelium (UP; >5fold, FDR corrected p-value<0.05), and with lower transcript concentrations in mycorrhiza then in mycelium (DOWN) are shown. In addition, in the bars the percentage of up- and down-regulated transcripts as well as the percentage of regulated transcripts is given.

A-1 *Hebeloma cylindrosporum*



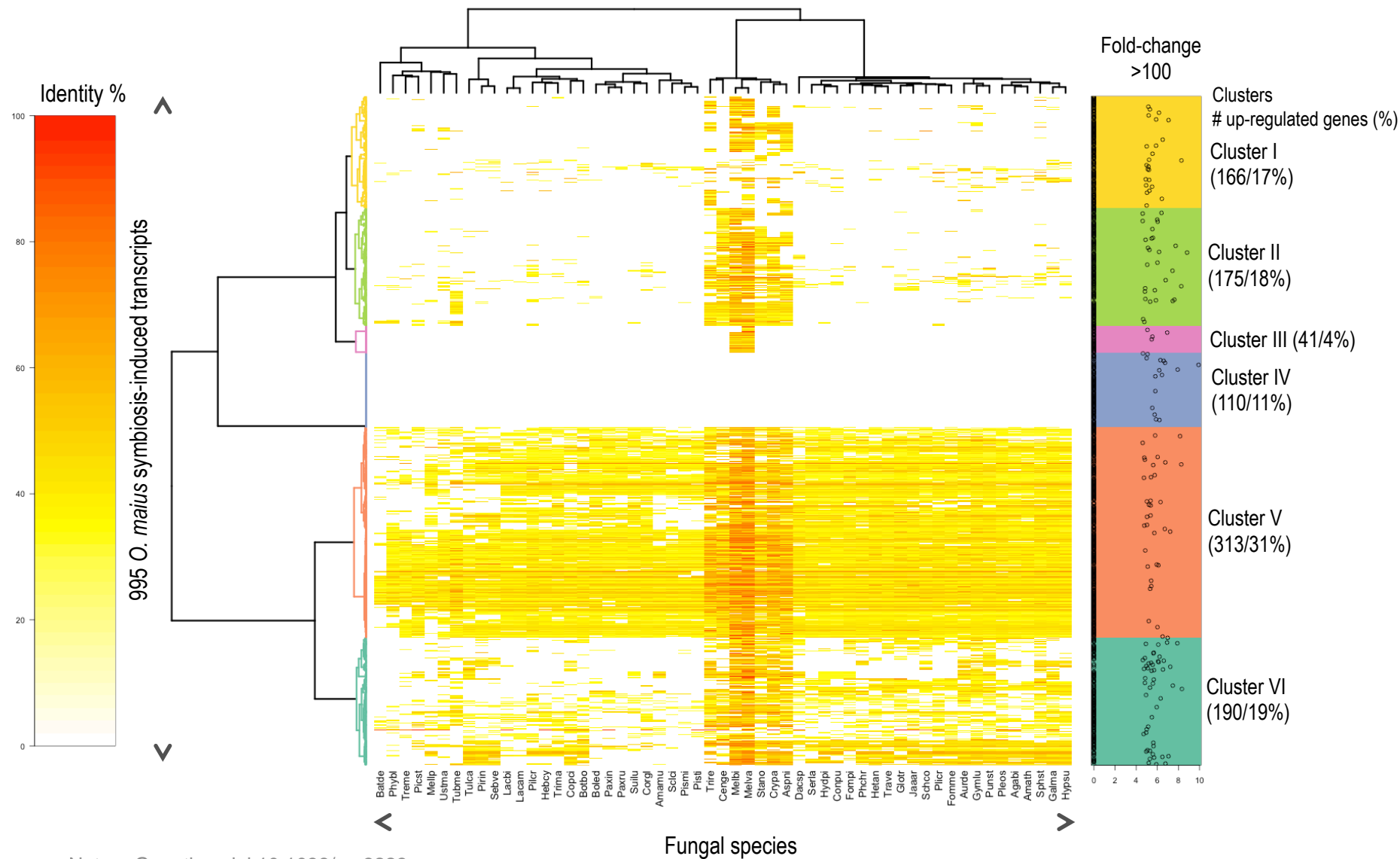
Nature Genetics: doi:10.1038/ng.3223

A-2

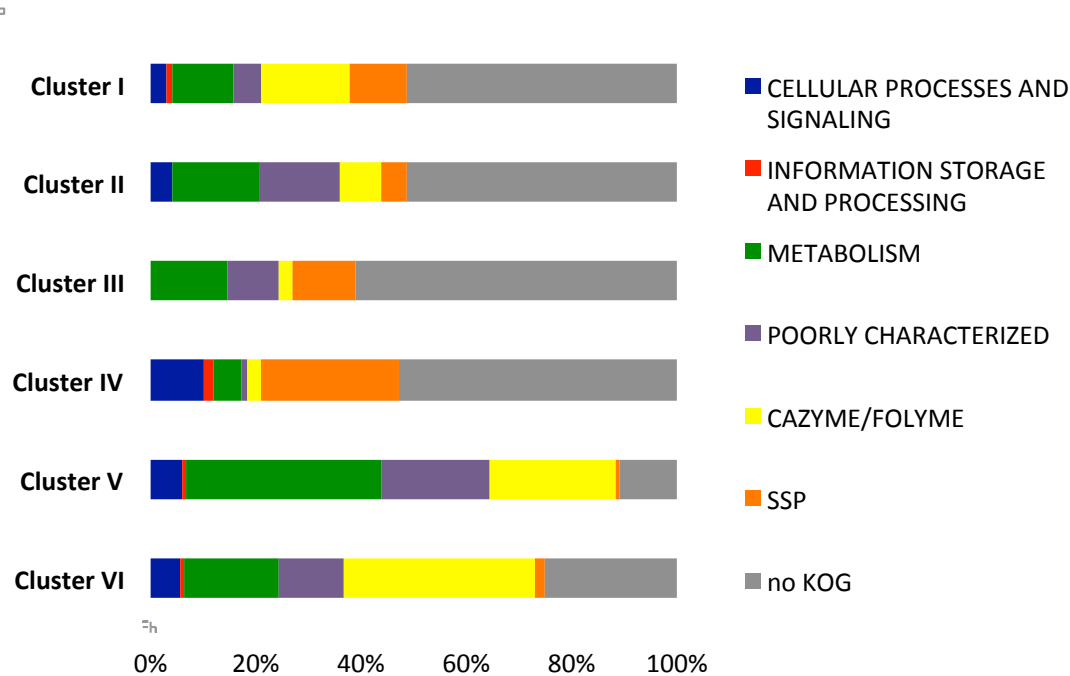


	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	17	yes	< 2.2e-16	7	yes	0.000648
Cluster II	20	yes	2.90E-11	3	no	1.92E-01
Cluster III	25	yes	1.50E-14	18	yes	2.52E-14
Cluster IV	21	yes	1.01E-07	18	yes	1.52E-10
Total up	21	yes	< 2.2e-16	11	yes	< 2.2e-16
Genome	4			2		

B-1 *Oidiodendron maius*

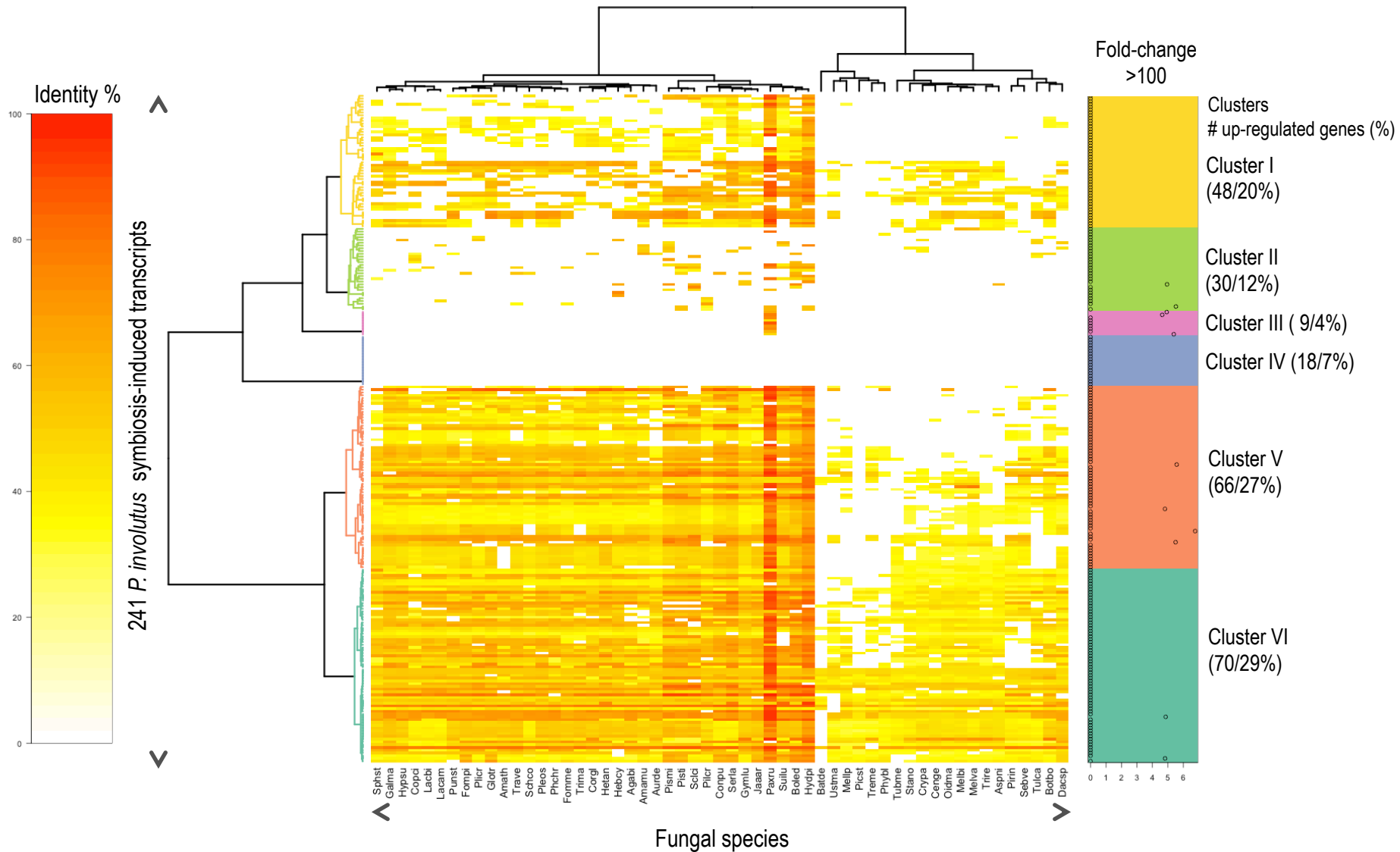


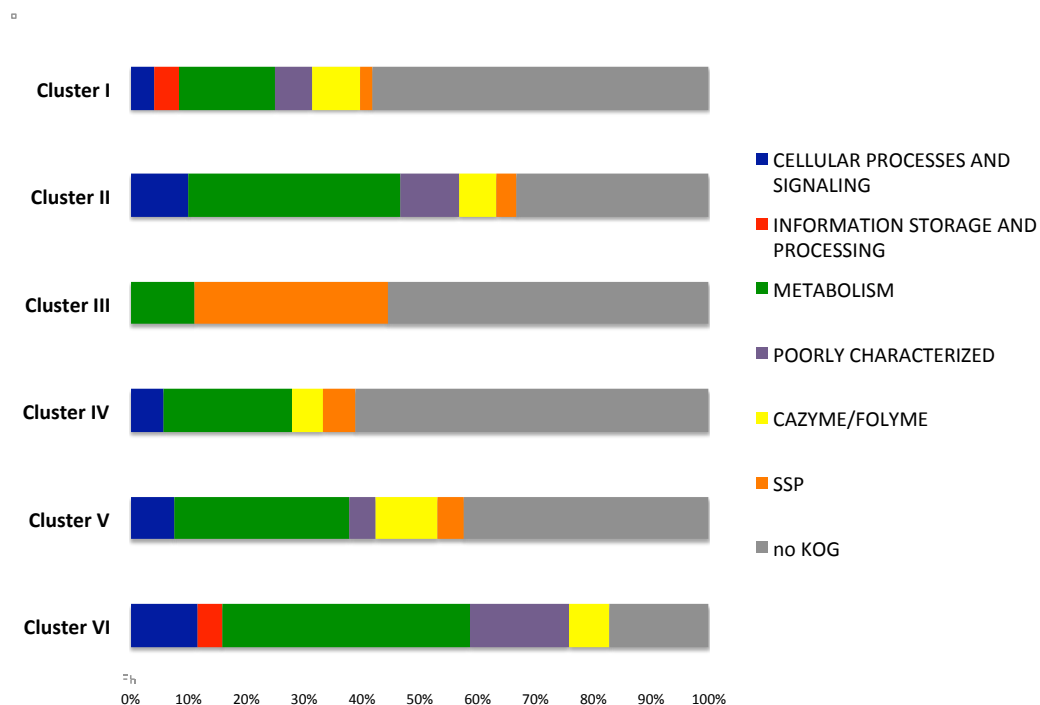
B-2



	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	27	yes	< 2.2e-16	11	yes	5.55E-08
Cluster II	18	yes	< 2.2e-16	5	no	0.05012
Cluster III	15	yes	4.11E-07	12	yes	0.002401
Cluster IV	33	yes	< 2.2e-16	26	yes	< 2.2e-16
Cluster V	26	yes	< 2.2e-16	1	no	0.9946
Cluster VI	42	yes	< 2.2e-16	2	no	0.8173
Total up	28	yes	< 2.2e-16	7	yes	1.69E-14
Genome	7			2		

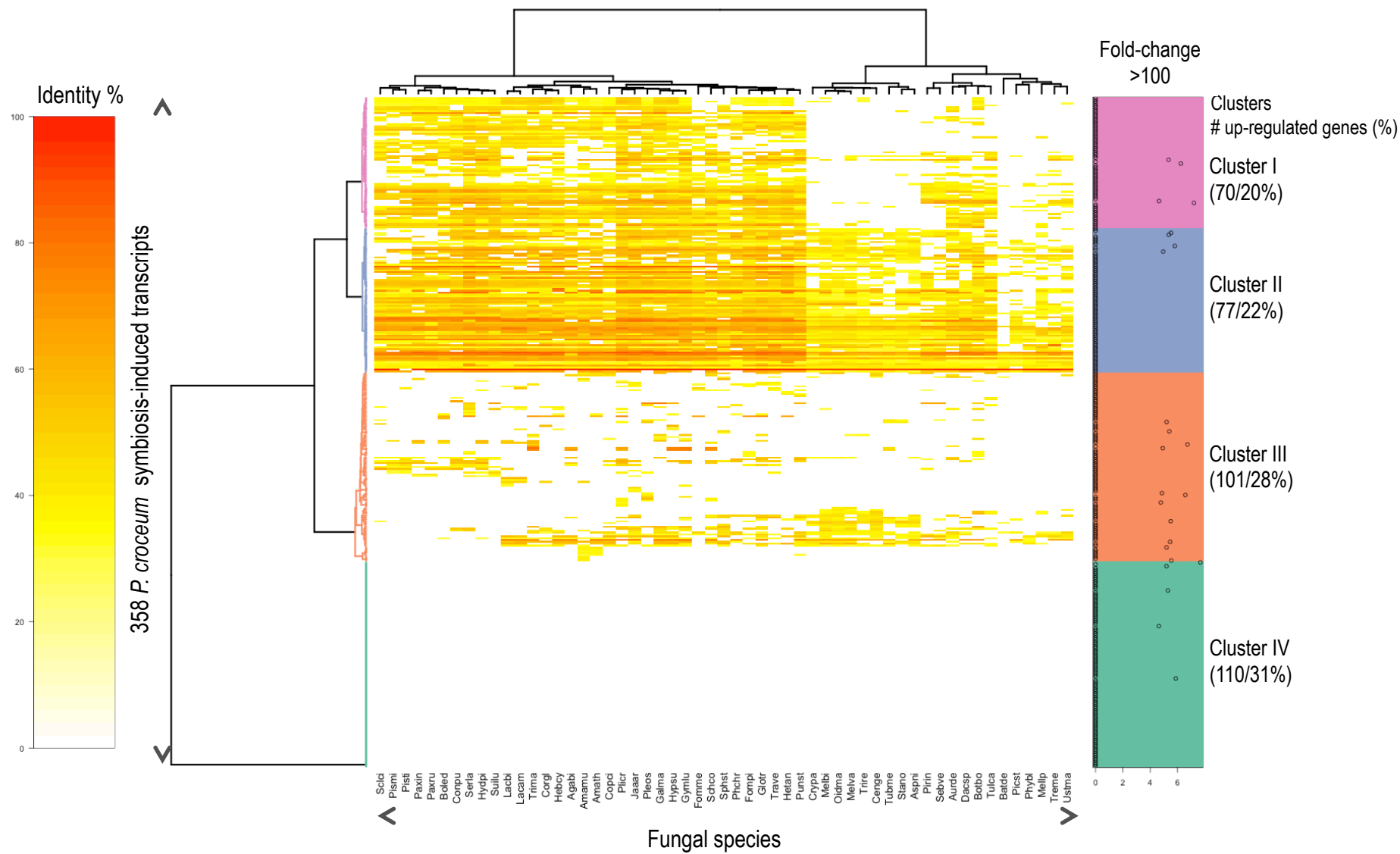
C-1 *Paxillus involutus*



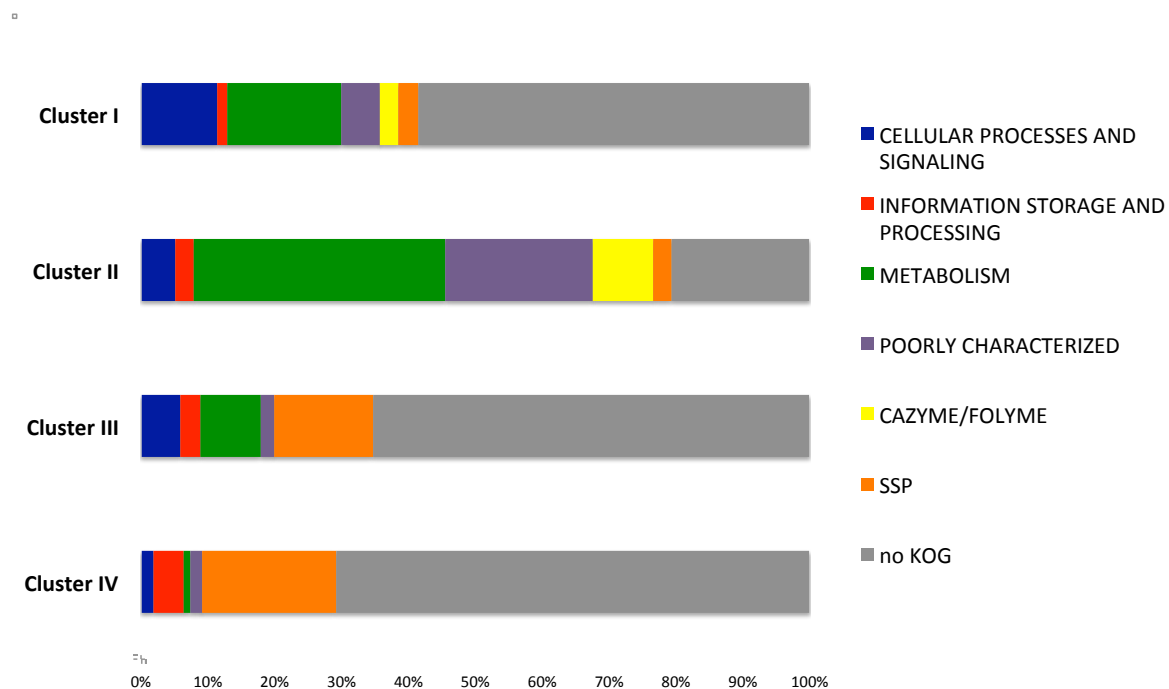


	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	10	yes	2.25E-07	2	no	0.4528
Cluster II	7	yes	0.003045	3	no	0.3138
Cluster III	33	yes	1.65E-06	33	yes	0.0001519
Cluster IV	11	yes	0.001094	6	no	0.2022
Cluster V	15	yes	3.67E-15	5	no	0.04933
Cluster VI	11	yes	2.28E-11	0	no	1
Total up	12	yes	< 2.2e-16	4	yes	0.003334
Genome	3			1		

D-1 *Piloderma croceum*

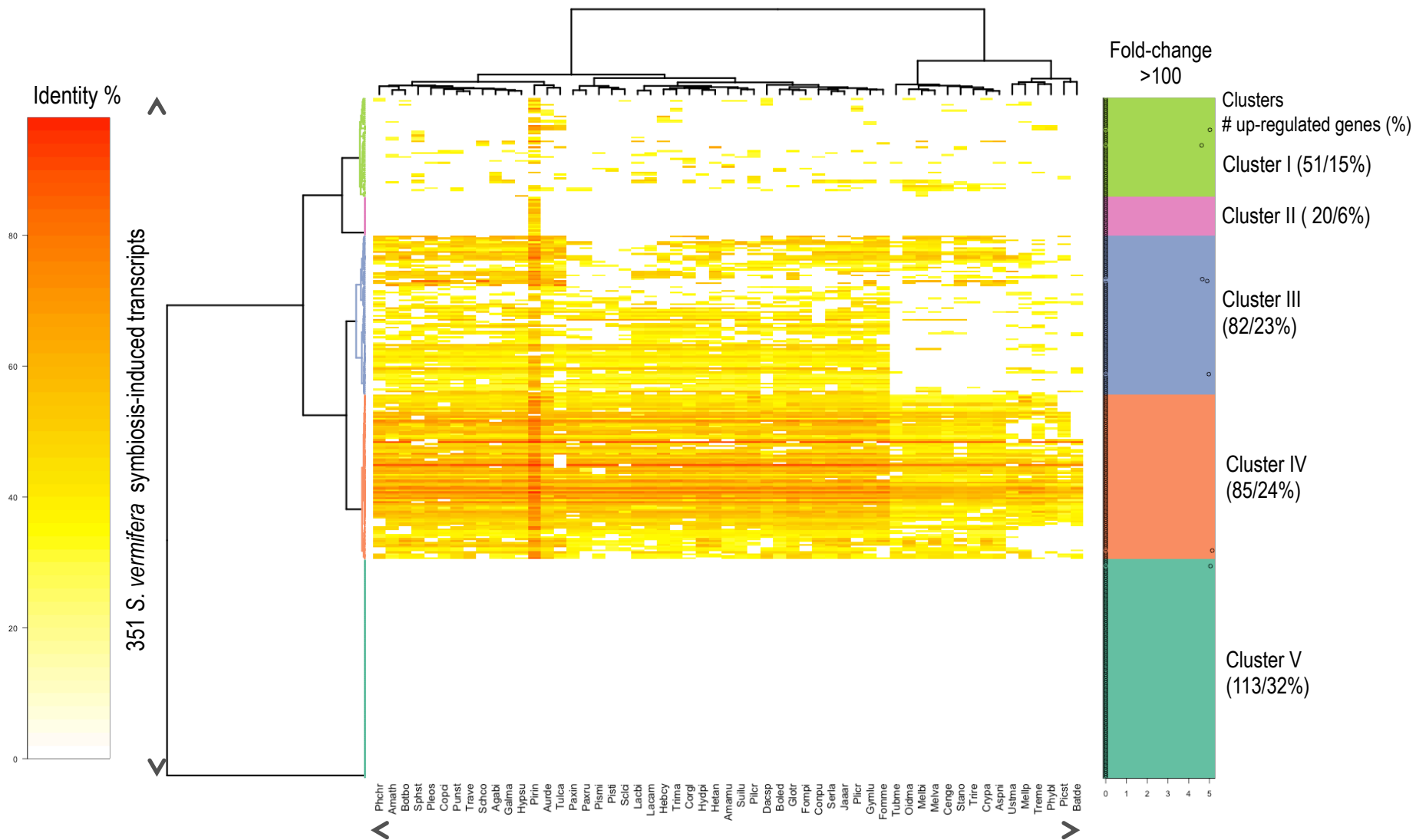


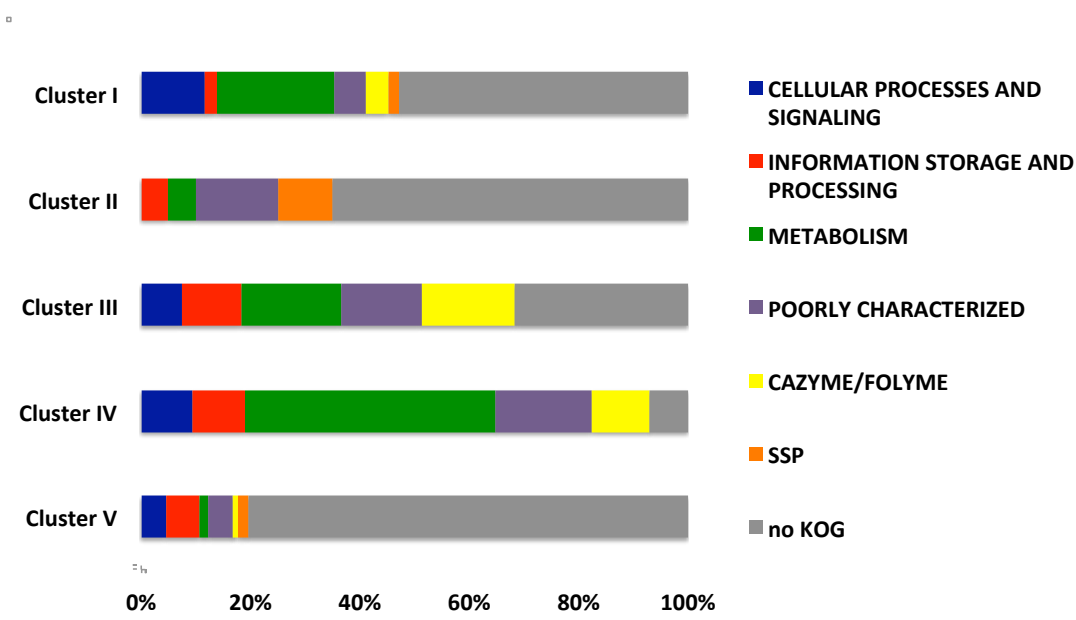
D-2



	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	9	no	0.03675	3	no	0.3675
Cluster II	10	yes	0.005686	3	no	0.4129
Cluster III	18	yes	1.40E-08	15	yes	4.73E-10
Cluster IV	20	yes	3.32E-11	20	yes	< 2.2e-16
Total up	15	yes	< 2.2e-16	11	yes	< 2.2e-16
Genome	4			2		

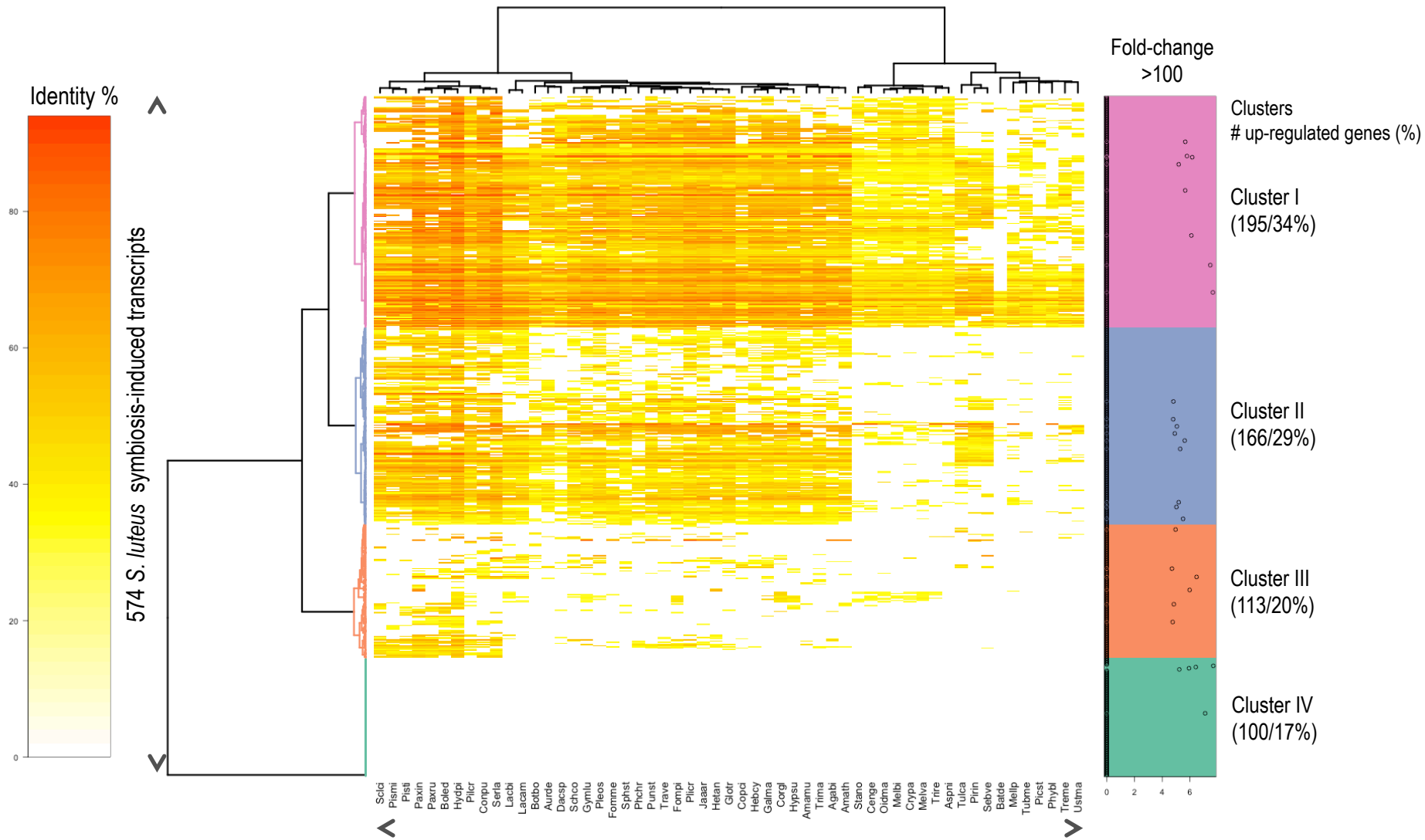
E-1 *Sebacina vermifera*

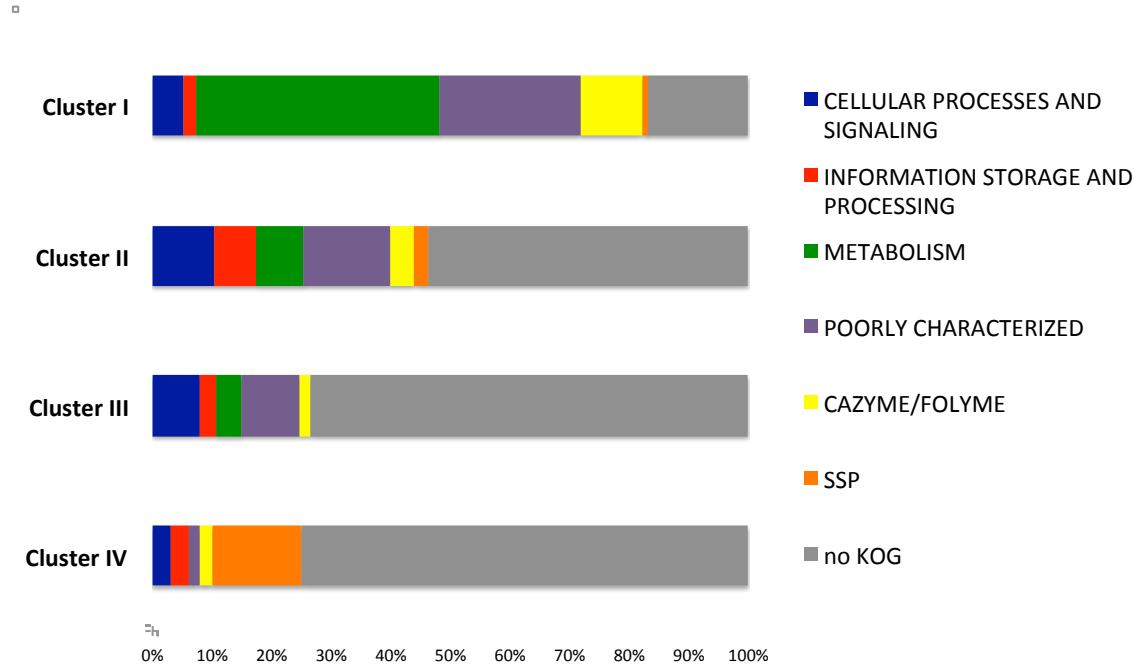




	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	12	no	0.03348	2	no	0.6105
Cluster II	10	no	0.2466	10	no	0.05099
Cluster III	21	yes	2.49E-07	0	no	1
Cluster IV	9	no	0.0494	0	no	1
Cluster V	4	no	0.6309	2	no	0.6153
Total up	11	yes	2.04E-06	1	no	0.7729
Genome	5			2		

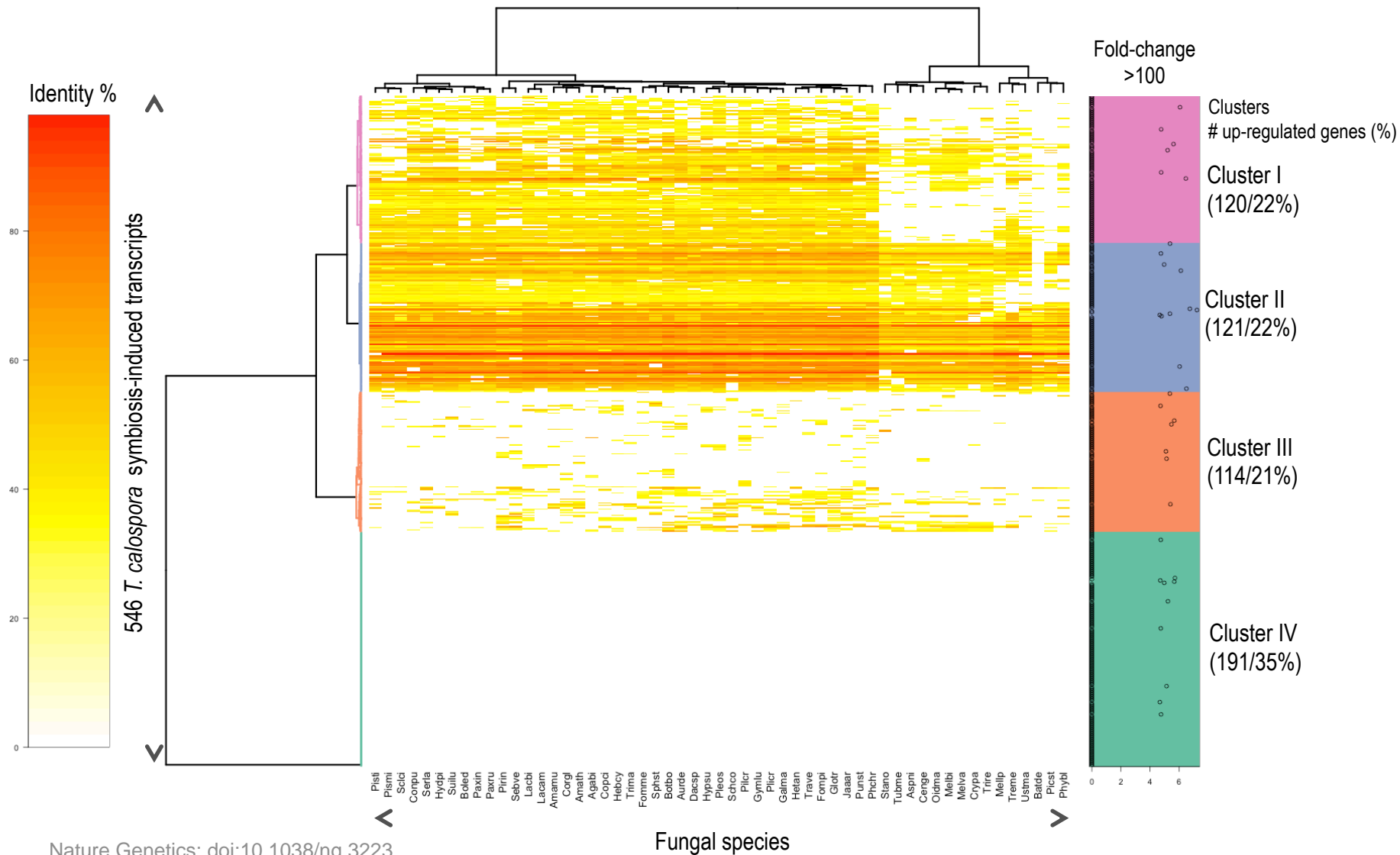
F-1 *Suillus luteus*

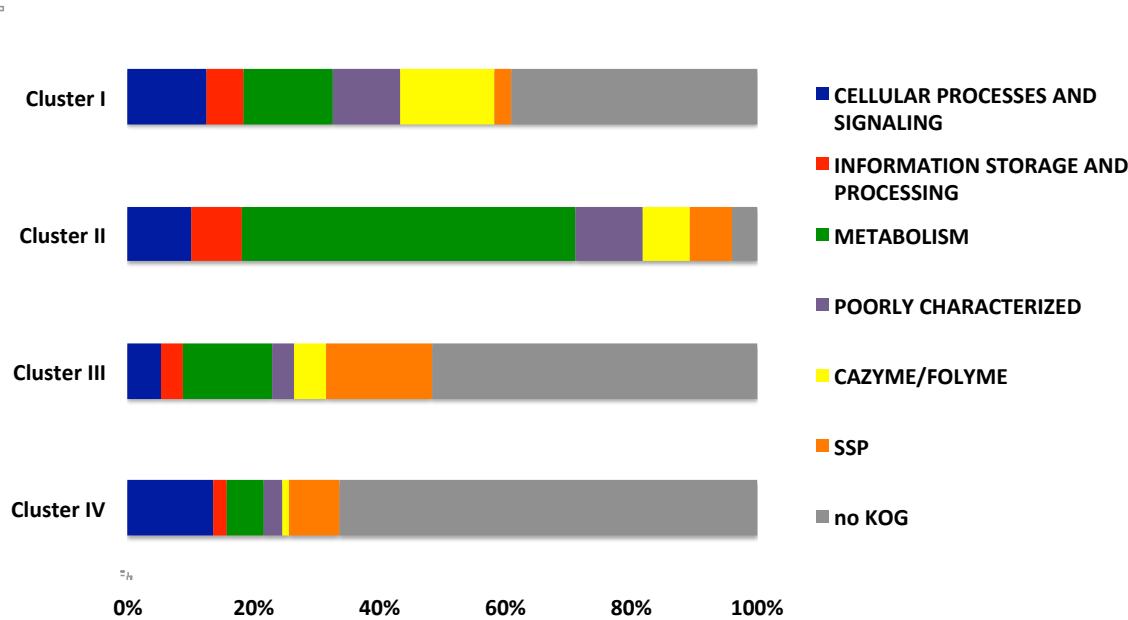




	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	8	yes	0.0003986	1	no	0.7232
Cluster II	8	yes	0.000807	2	no	0.1704
Cluster III	1	no	0.9598	0	no	1
Cluster IV	15	yes	1.16E-07	15	yes	3.11E-12
Total up	8	yes	1.50E-09	4	yes	2.06E-05
Genome	3			1		

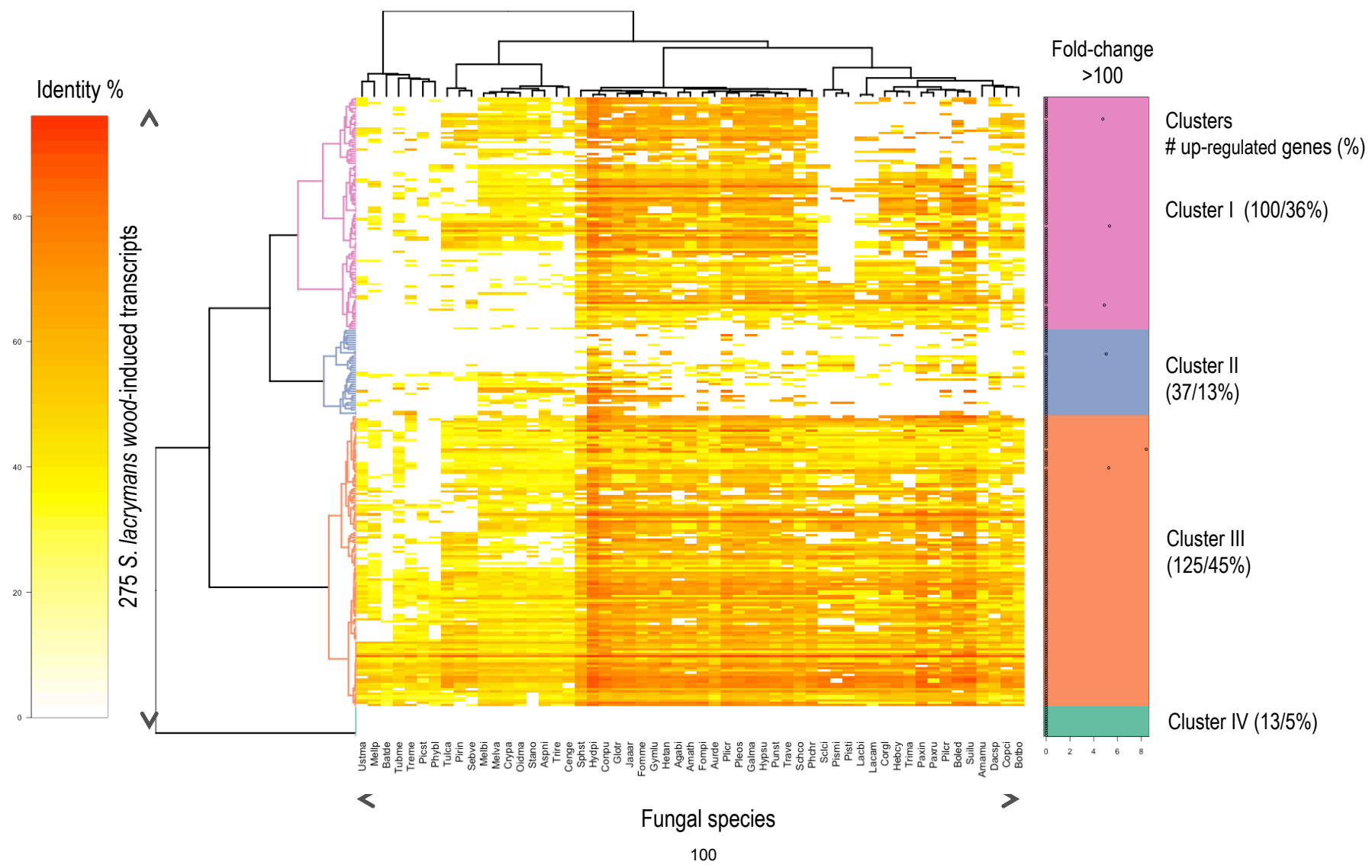
G-1 *Tulasnella calospora*

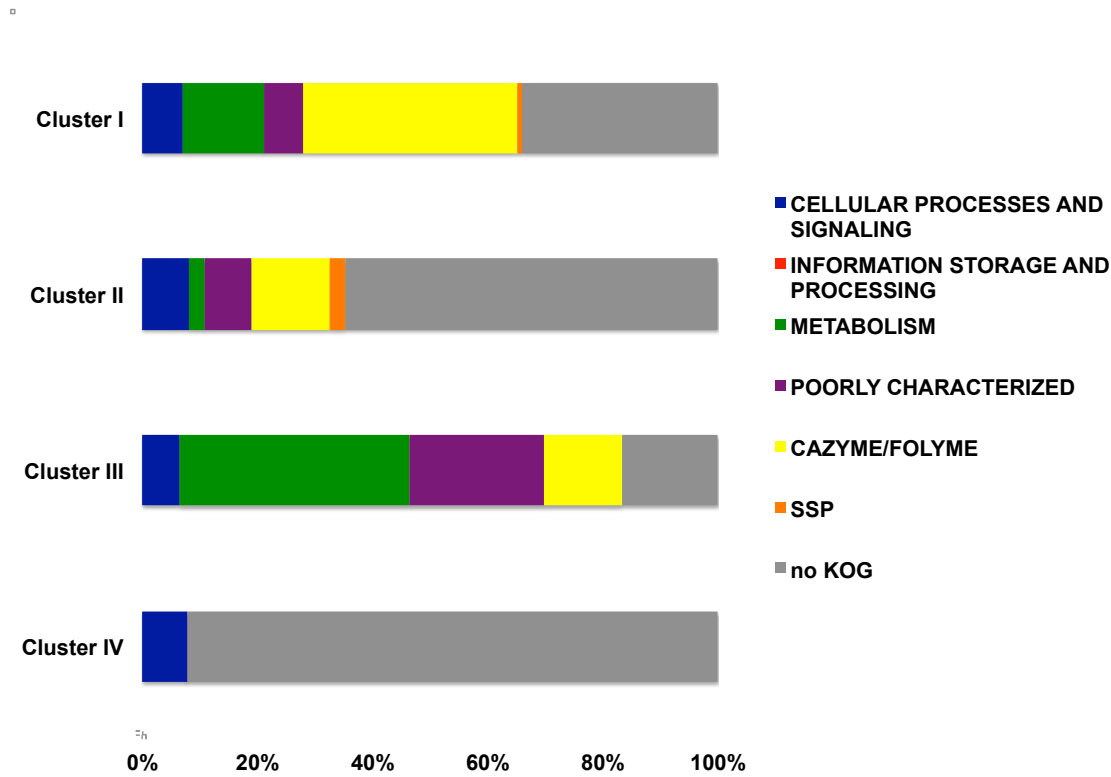




	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	13	yes	1.79E-05	3	no	0.3888
Cluster II	7	no	0.1021	1	no	0.8985
Cluster III	24	yes	2.77E-14	17	yes	3.59E-13
Cluster IV	9	yes	0.001397	8	yes	3.20E-06
Total up	12	yes	< 2.2e-16	7	yes	3.23E-12
Genome	4			2		

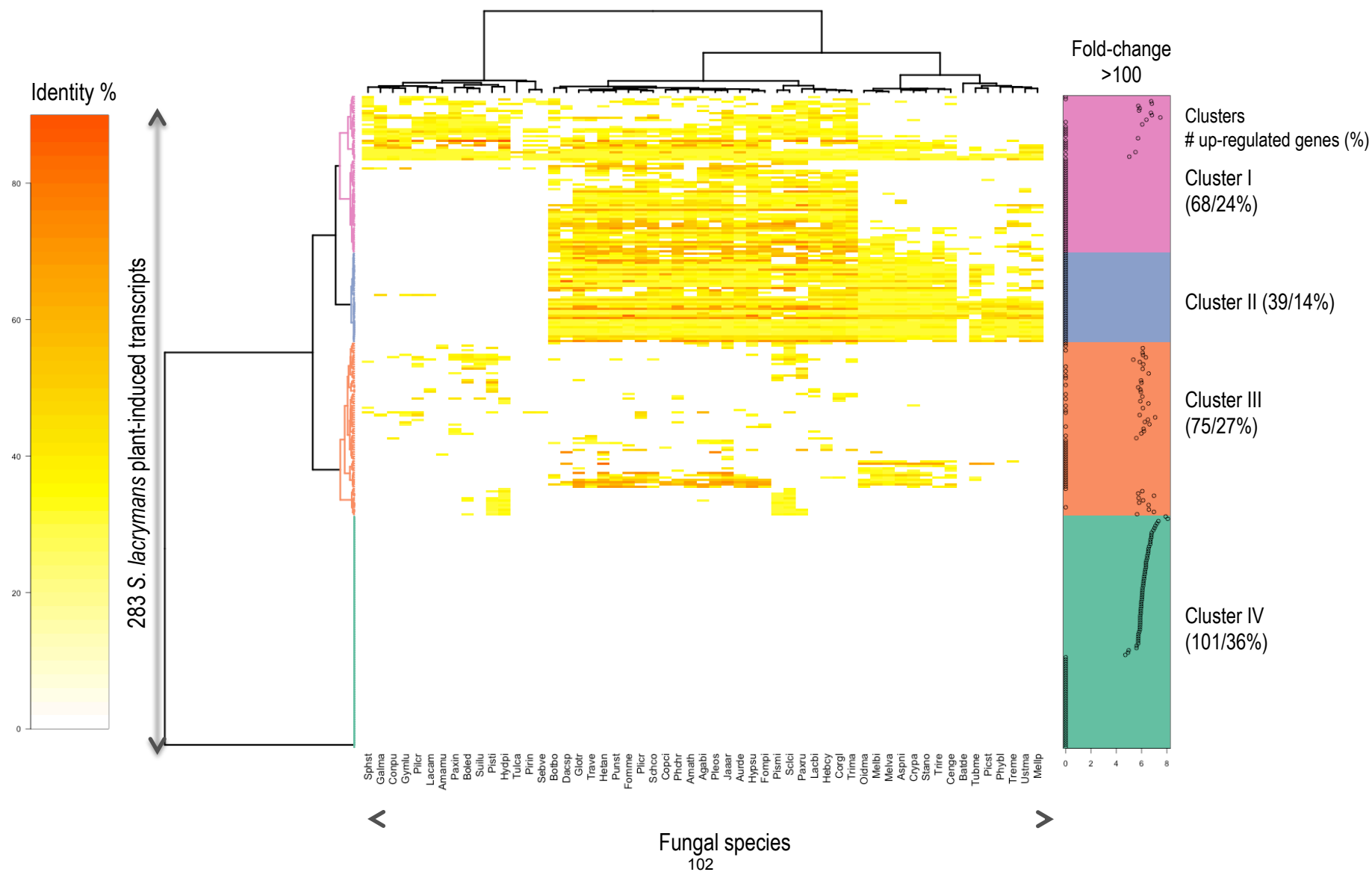
H-1 *Serpula lacrymans* mycelium grown on wood

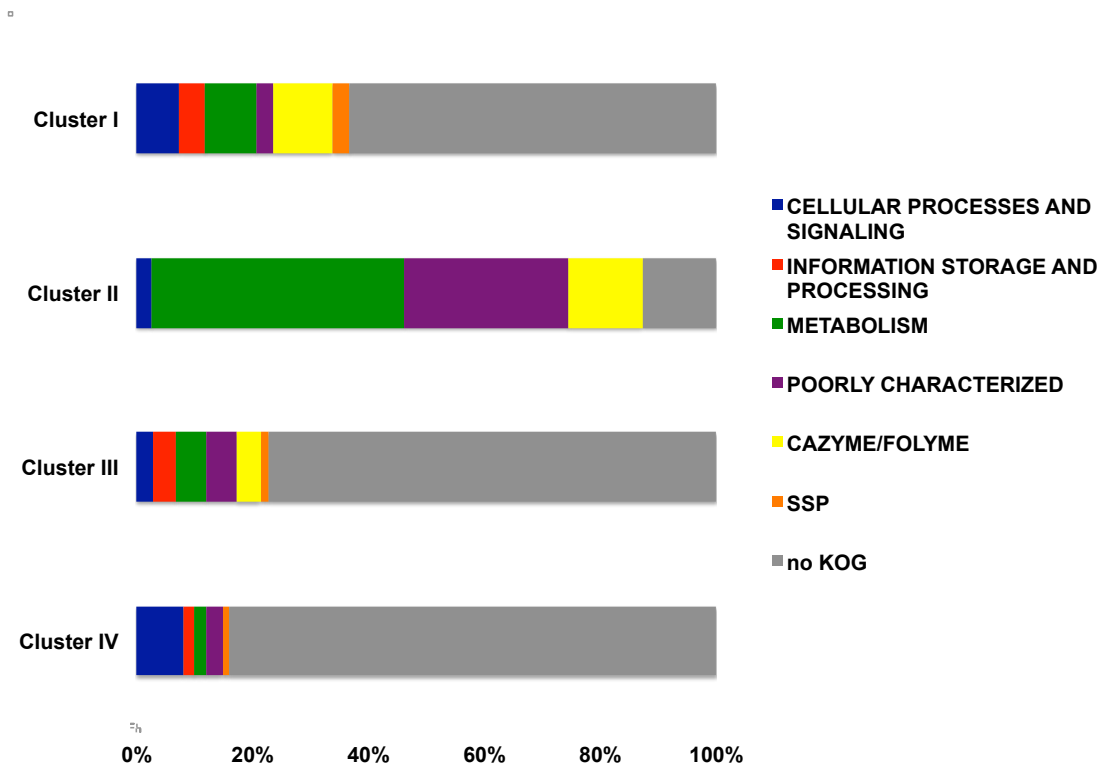




	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	39	yes	< 2.2e-16	1	no	0.7924
Cluster II	22	yes	2.47E-07	3	no	0.4402
Cluster III	22	yes	< 2.2e-16	0	no	1
Cluster IV	0	no	1	0	no	1
Total up	27	yes	< 2.2e-16	0.7	no	0.9301
Genome	2			2		

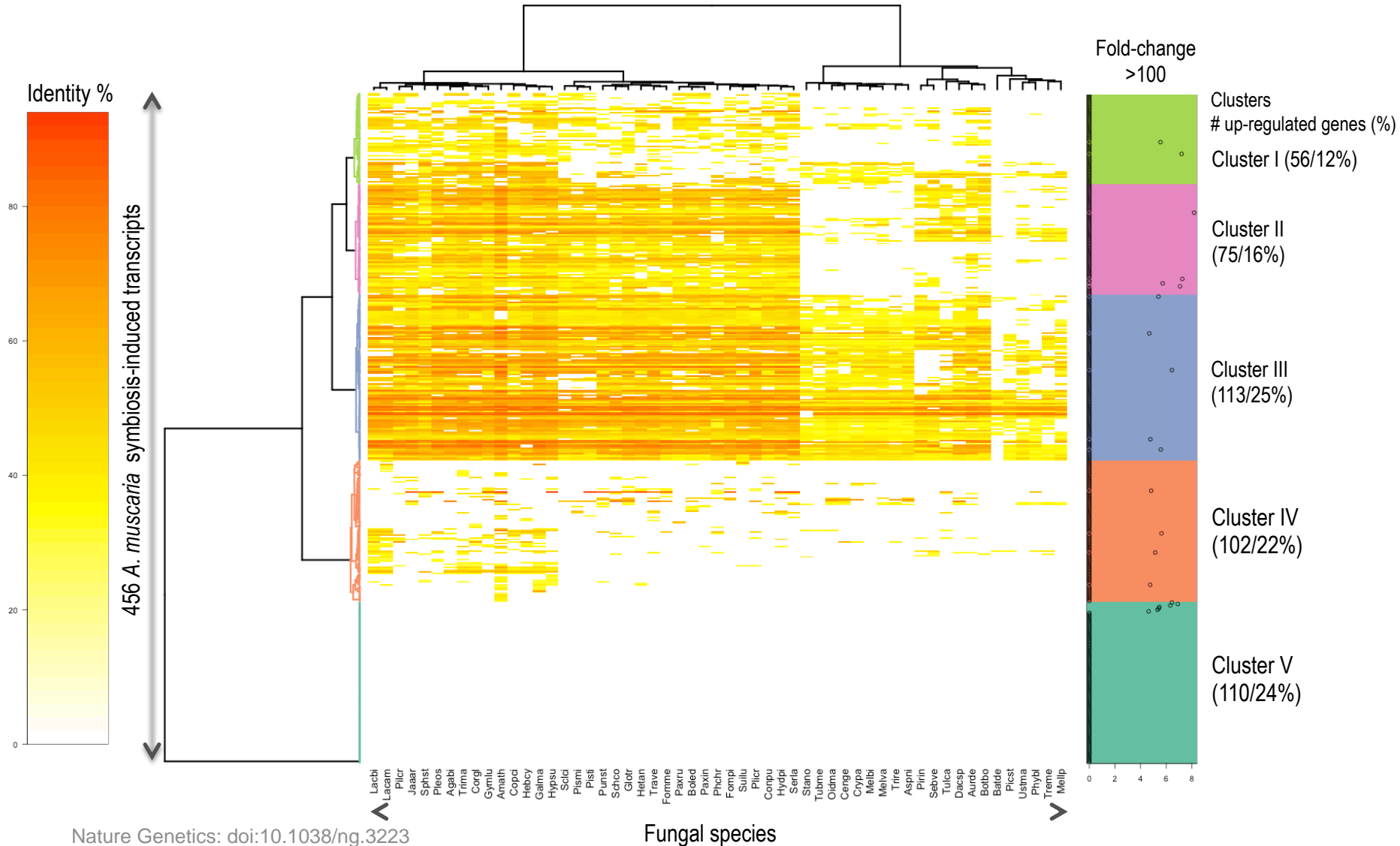
I-1 *Serpula lacrymans* pseudomycorrhiza

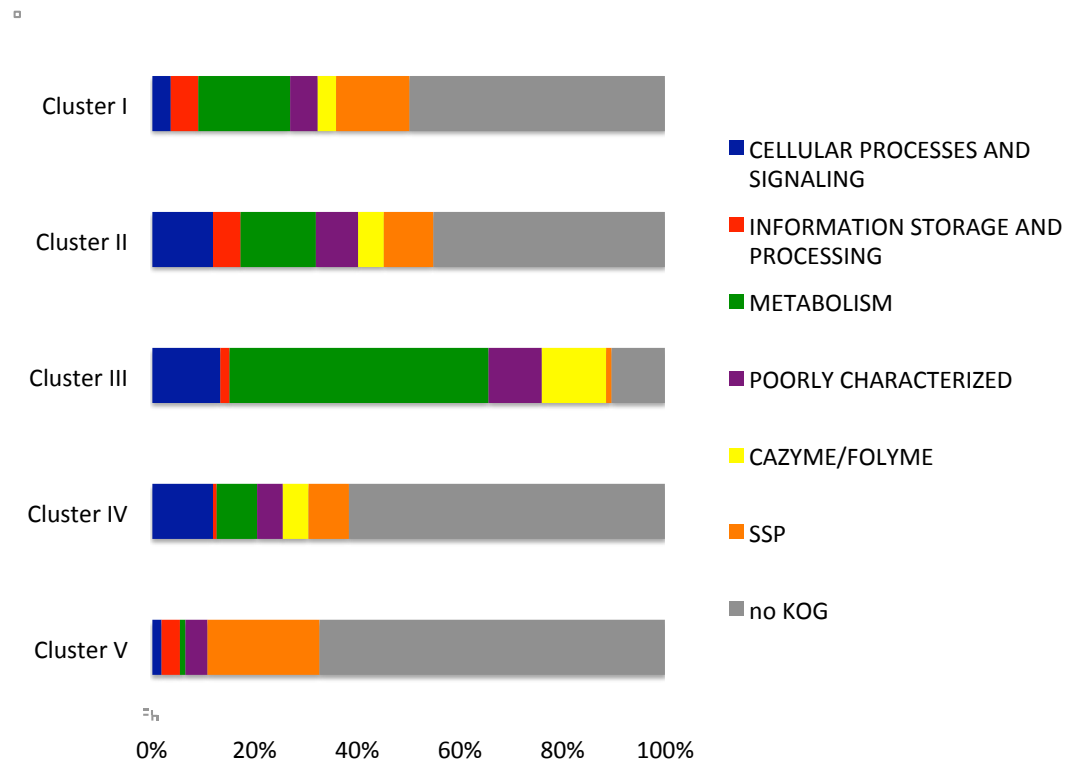




	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	15	yes	3.54E-07	3	no	0.2852
Cluster II	18	yes	5.38E-06	0	no	1
Cluster III	7	no	0.01141	1	no	0.6921
Cluster IV	2	no	0.5475	1	no	0.7957
Total up	8	yes	2.97E-10	1	no	0.6447
Genome	2			2		

J-1 *Amanita muscaria*





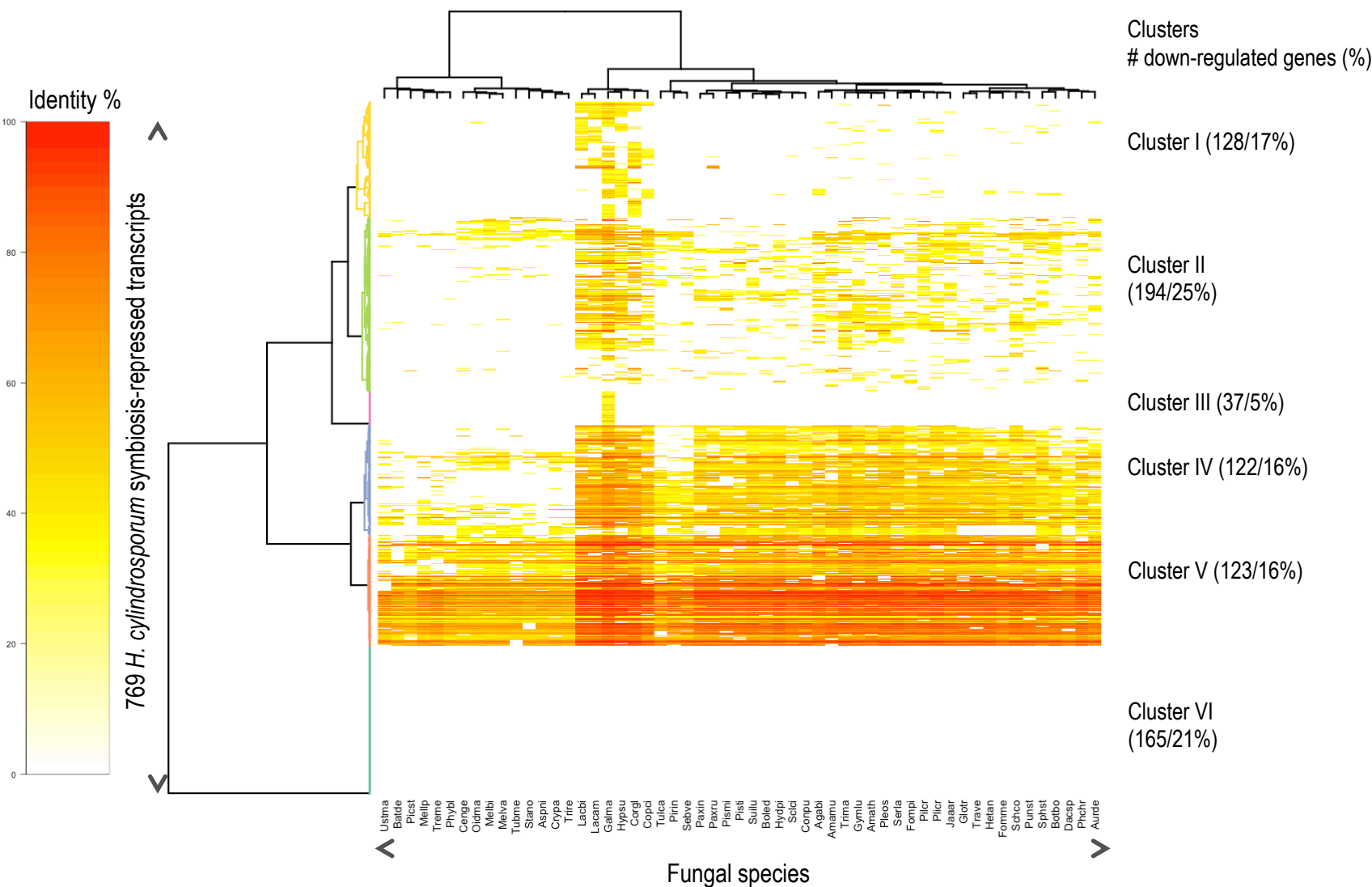
	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	14	yes	0.0008137	14	yes	9.65E-06
Cluster II	15	yes	7.04E-05	9	yes	0.0005337
Cluster III	11	yes	0.0007417	1	no	0.8849
Cluster IV	20	yes	4.90E-10	8	yes	0.0007029
Cluster V	23	yes	9.65E-14	22	yes	< 2.2e-16
Total up	17	yes	< 2.2e-16	11	yes	< 2.2e-16
Genome	4			2		

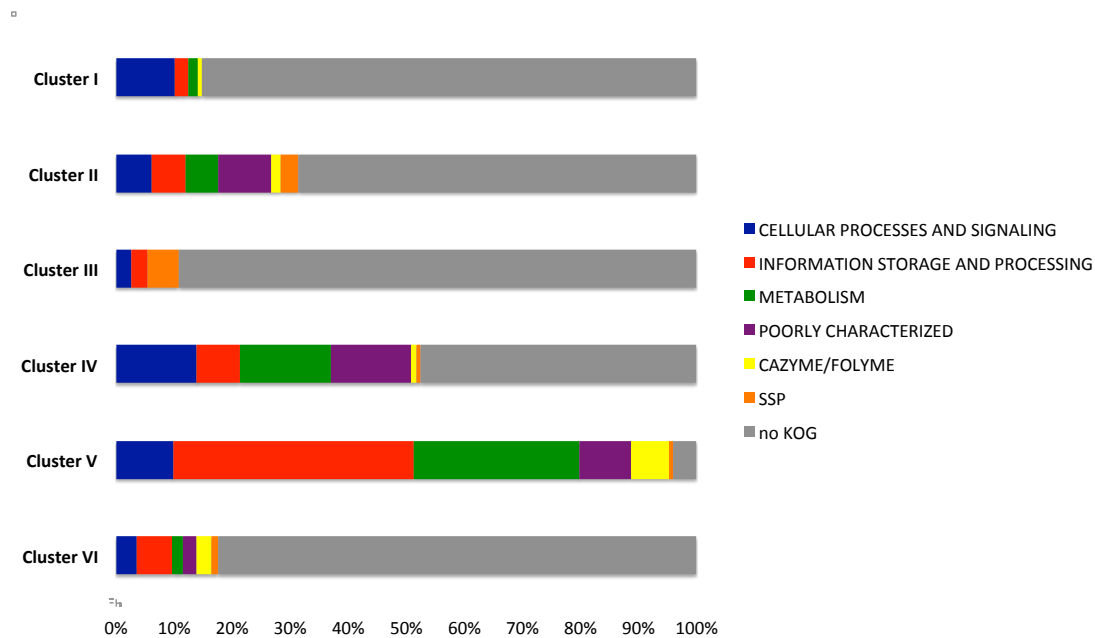
Fig. S25. A-J Functional analysis of symbiosis-induced transcripts.

Symbiosis-induced genes were blasted (BLASTP) against 55 fungal genomes to find homologs. Homologs are coloured from yellow to red depending on the percentage of similarity. The heatmap represents a double-hierarchical clustering of the symbiosis-induced transcripts orthologs in the 55 fungal genomes. Data were visualized and clustered using R (package HeatPlus). The hierarchical clustering was done by using a binary distance metric and ward clustering method. Transcripts up-regulated >100-fold are shown as circles on the right panel, next to it is the number of transcripts in each cluster (A-1-J-1). For each cluster the percentages of putative functional categories are given as bargrams. A table shows the percentage of transcripts coding for secreted proteins and small secreted proteins in each cluster. A Fisher exact test was applied to test if these categories were enriched compared to the number of these genes in the respective genome (A2-J2).

A UP-regulated transcripts in *Hebeloma cylindrosporum*-*Pinus pinaster* mycorrhiza compared to free-living mycelium; B UP-regulated transcripts in *Oidiodendron maius*-*Vaccinium myrtillus* mycorrhiza compared to free-living mycelium; C UP-regulated transcripts in *Paxillus involutus*-*Fagus sylvatica* mycorrhiza compared to free-living mycelium; D UP-regulated transcripts in *Piloderma croceum*-*Quercus robur* mycorrhiza compared to free-living mycelium; E UP-regulated transcripts in *Sebacina vermifera*-*Arabidopsis thaliana* mycorrhiza compared to free-living mycelium; F UP-regulated transcripts in *Suillus luteus*-*Pinus sylvestris* mycorrhiza compared to free-living mycelium; G UP-regulated transcripts in *Tulasnella calospora*-*Serapias vomeracea* mycorrhiza compared to free-living mycelium; H UP-regulated transcripts in *Serpula lacrymans* mycelium grown on shavings of *Pinus sylvestris* sapwood compared to control mycelium grown on glucose (MMN medium). I UP-regulated transcripts in *Serpula lacrymans* interacting with roots of *Picea sylvestris* (pseudomycorrhiza) compared to control mycelium grown on glucose (MMN medium); J UP-regulated transcripts in *Amanita muscaria*-*Populus tremula x tremuloides* mycorrhiza compared to free-living mycelium.

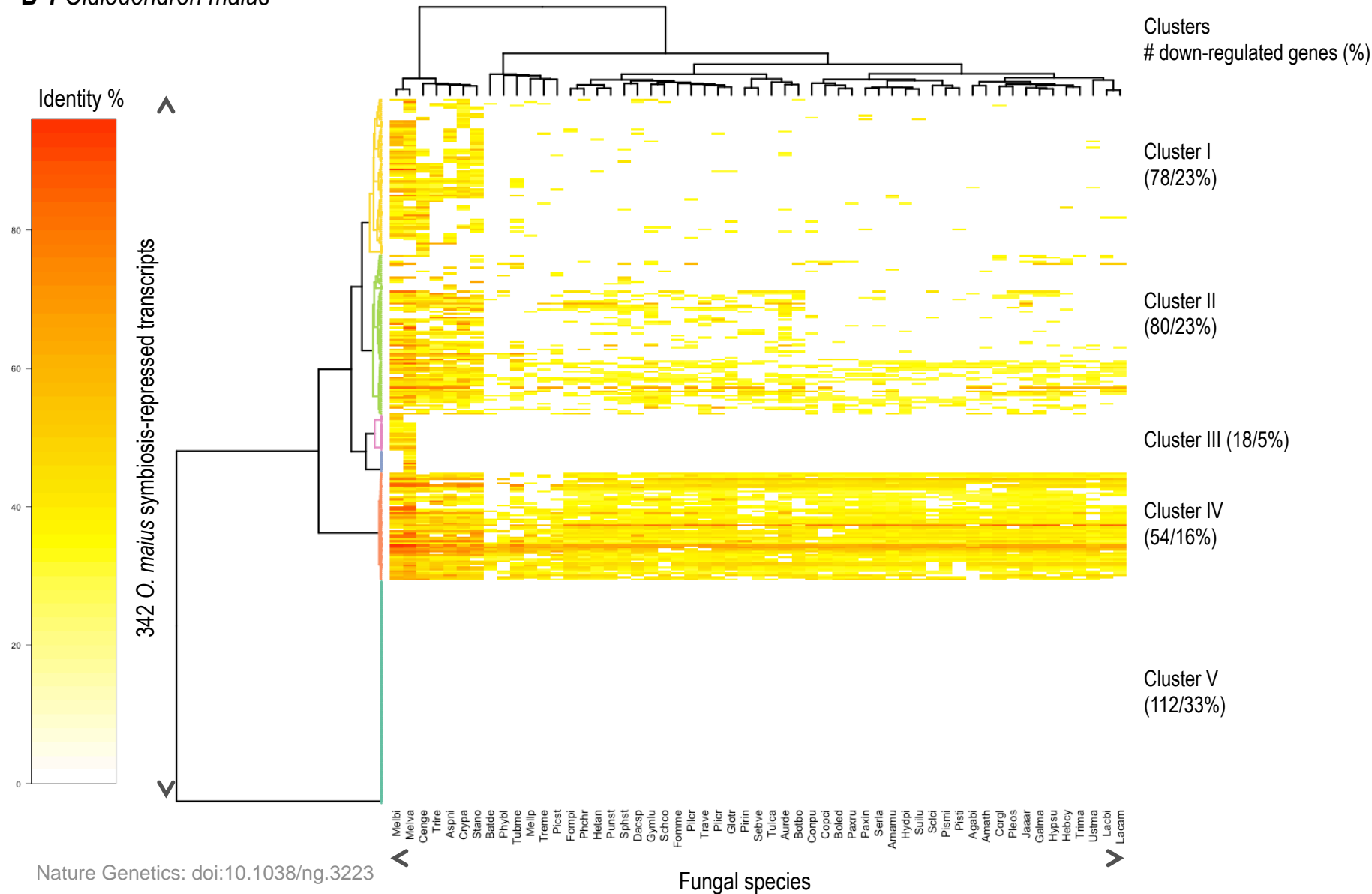
A-1 *Hebeloma cylindrosporum*

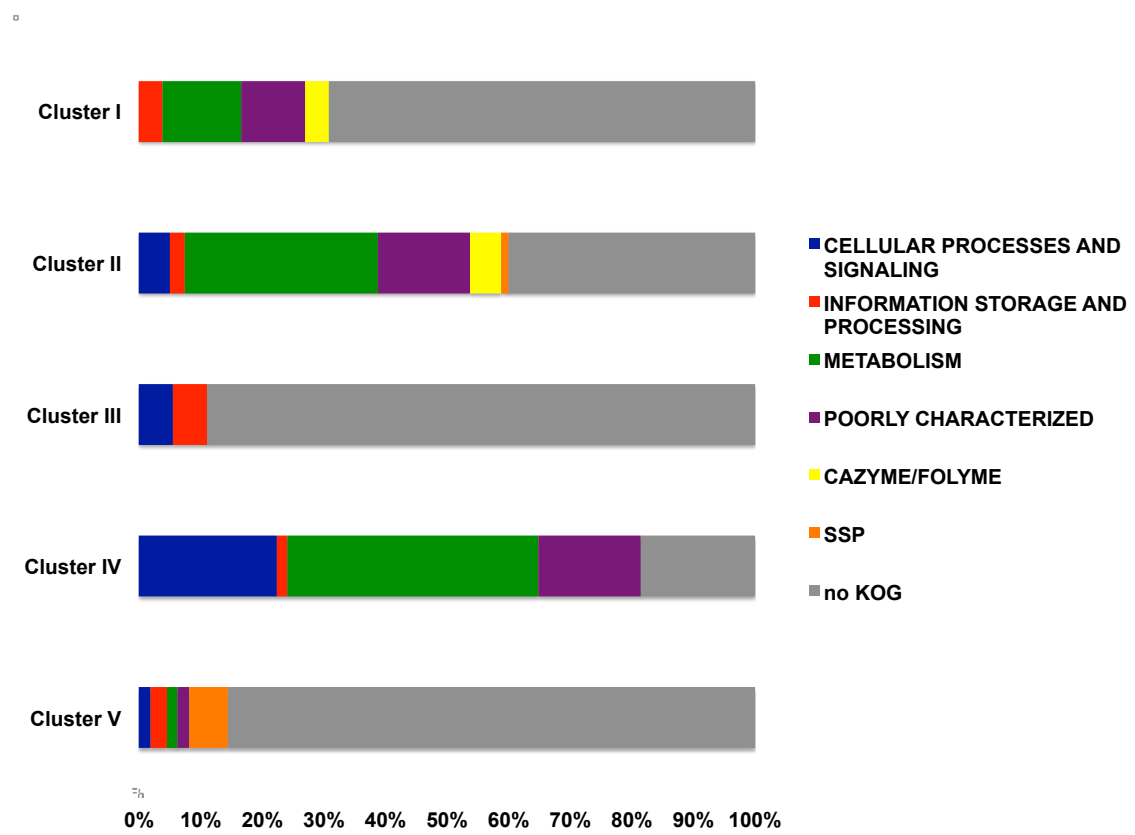




	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	2	no	0.9065	0	no	1
Cluster II	7	no	0.0626	3	no	0.1538
Cluster III	5	no	0.4601	5	no	0.1505
Cluster IV	8	no	0.0312	1	no	0.8996
Cluster V	4	no	0.5861	1	no	0.9015
ClusterVI	2	no	0.9705	1	no	0.8151
Total down	5	no	0.2564	2	no	0.4715006
Genome	4			2		

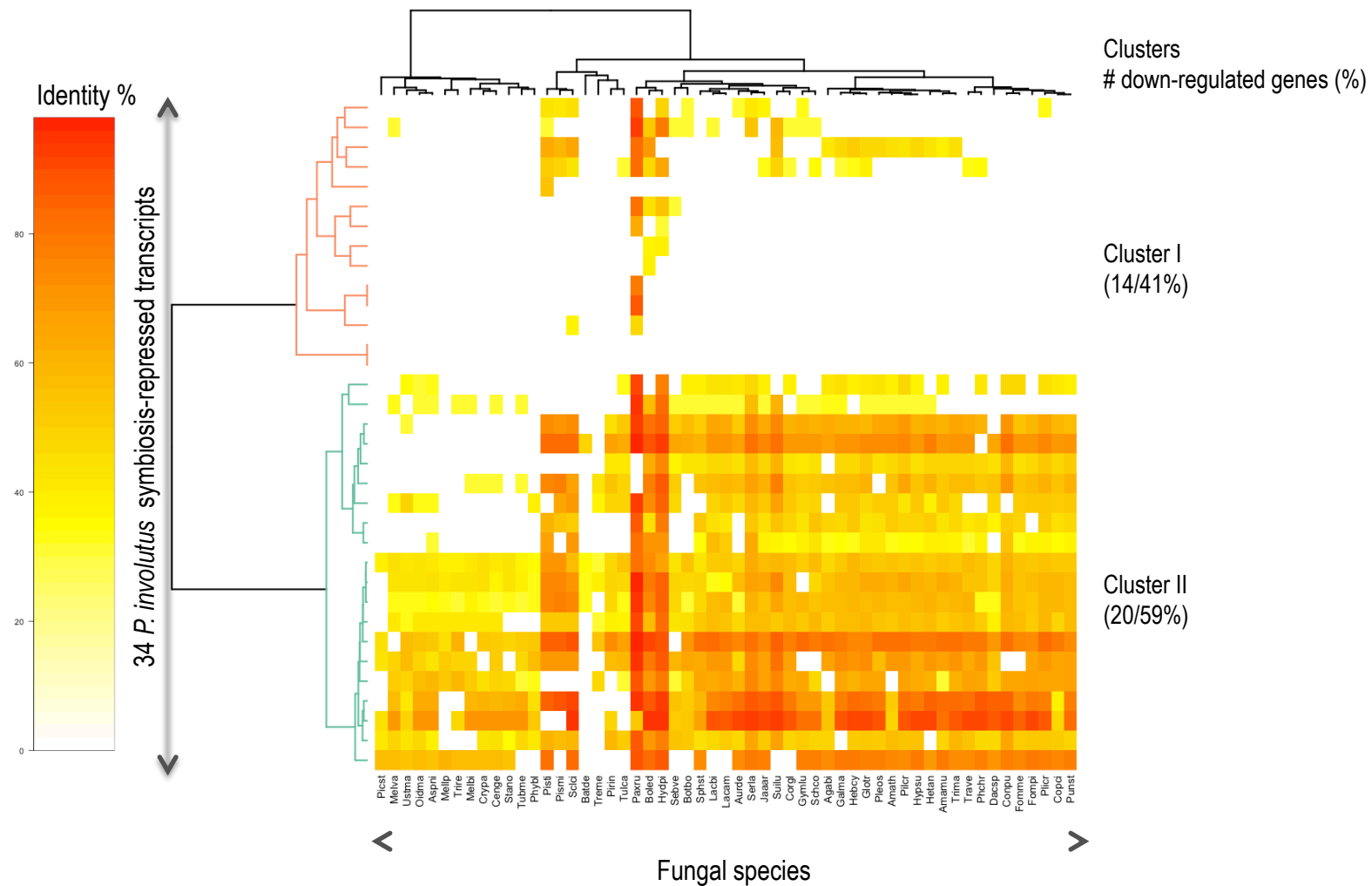
B-1 *Oidiodendron maius*

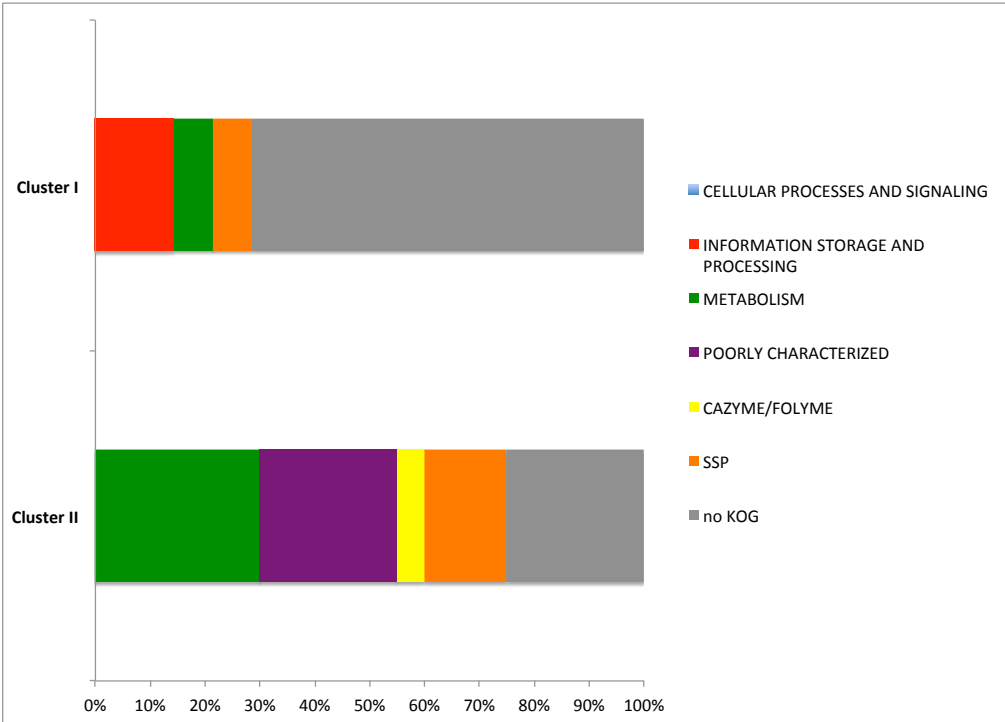




	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	3	no	0.9757	0	no	1
Cluster II	9	no	0.32	1	no	0.8459
Cluster III	0	no	1	0	no	1
Cluster IV	0	no	1	0	no	1
Cluster V	6	no	0.67	6	no	0.01503
Total down	5	no	0.9688	2	no	0.5333
Genome	7			2		

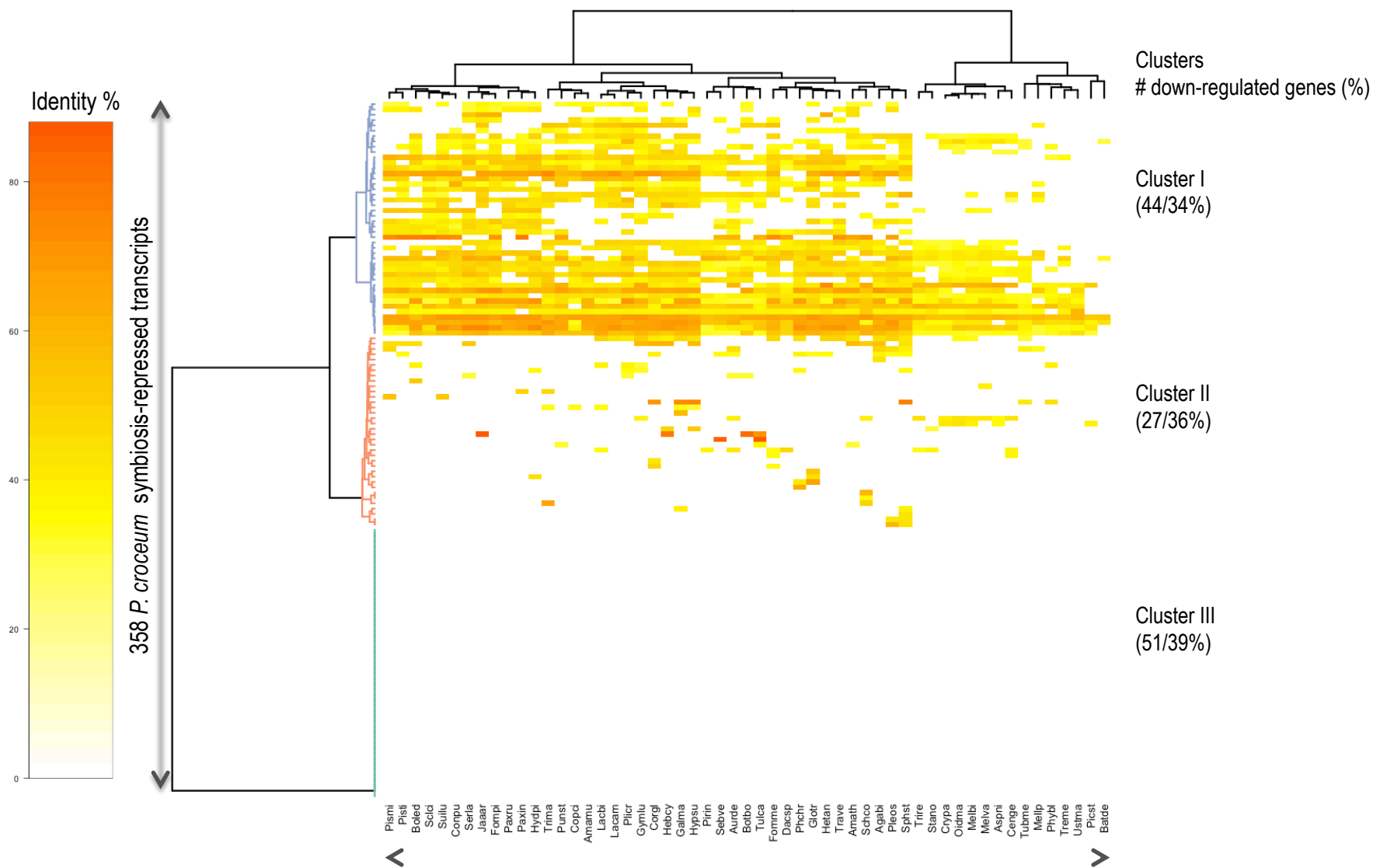
C-1 *Paxillus involutus*



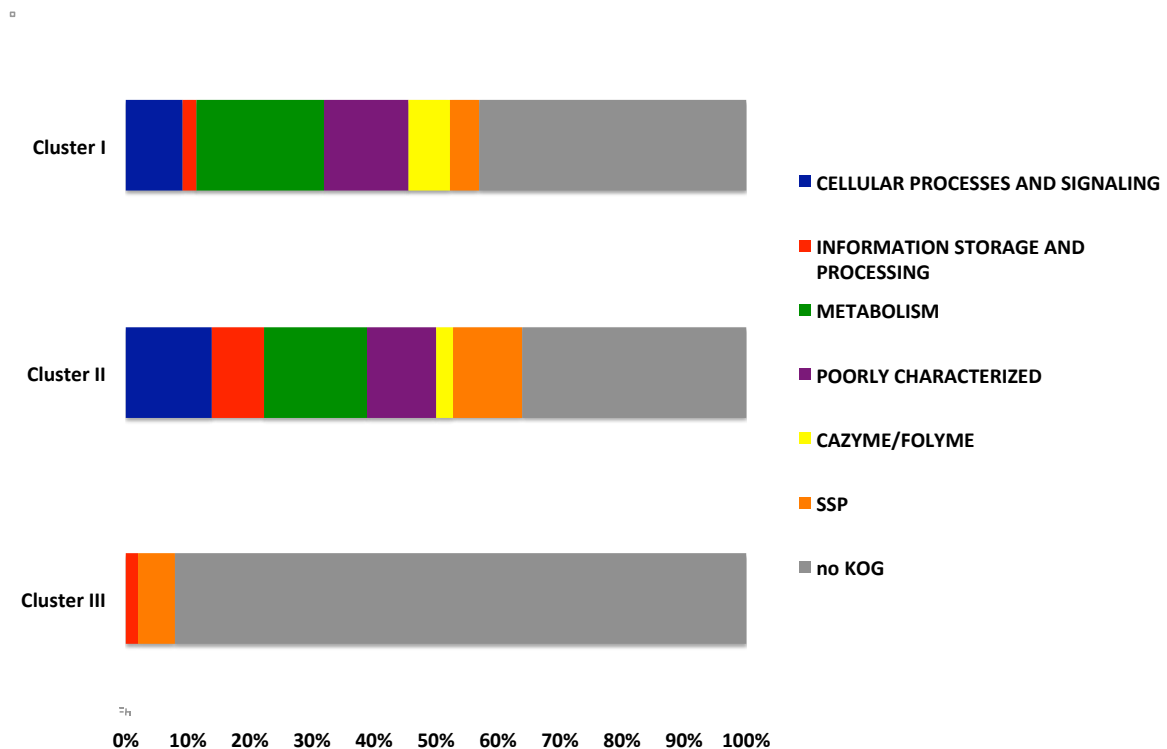


	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	7	no	0.32	7	no	0.1611
Cluster II	25	yes	0.0001601	15	yes	0.001863
Total down	18	yes	0.0002742	12	yes	0.0008131
Genome	3			1		

D-1 *Piloderma croceum*

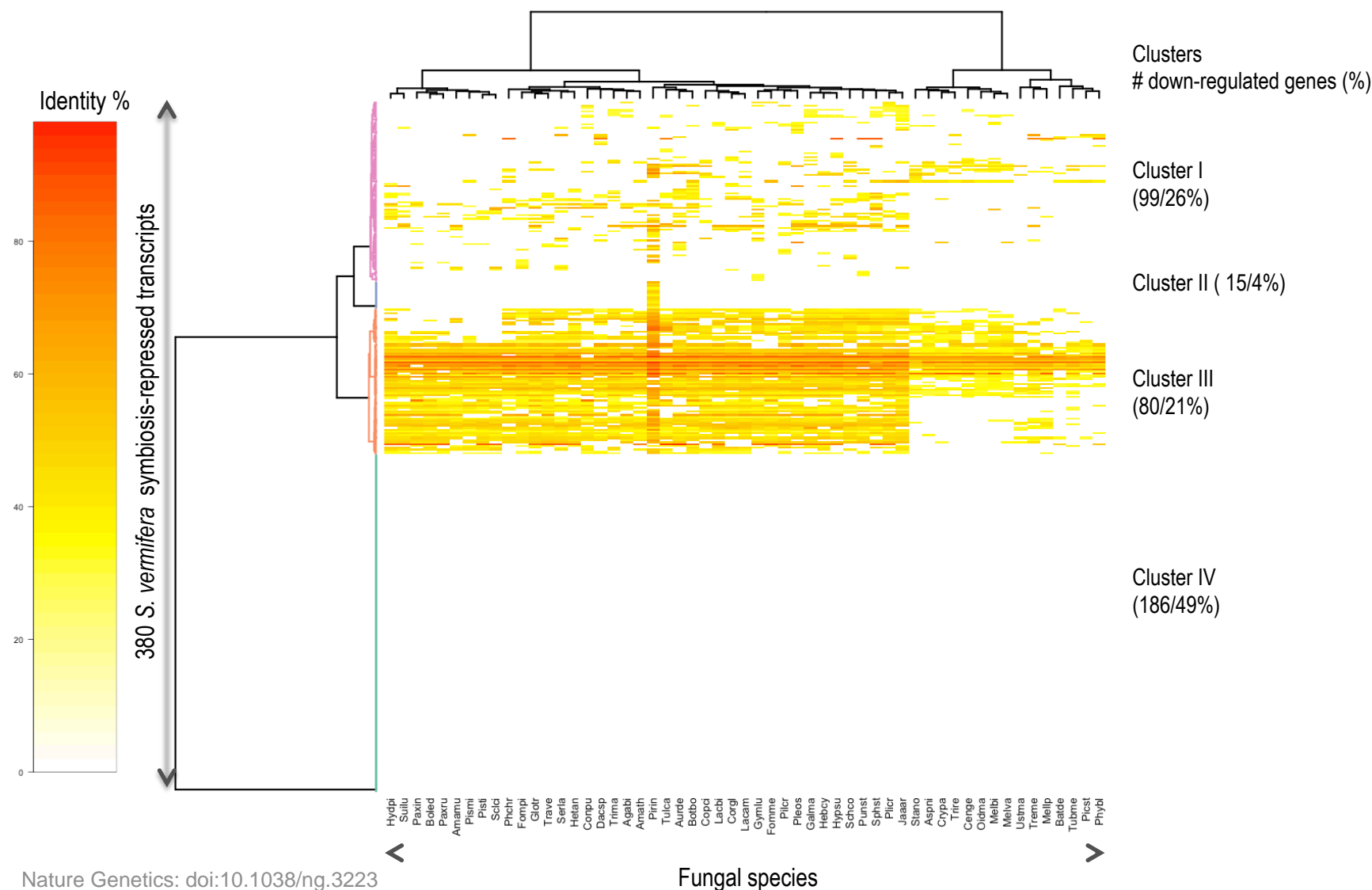


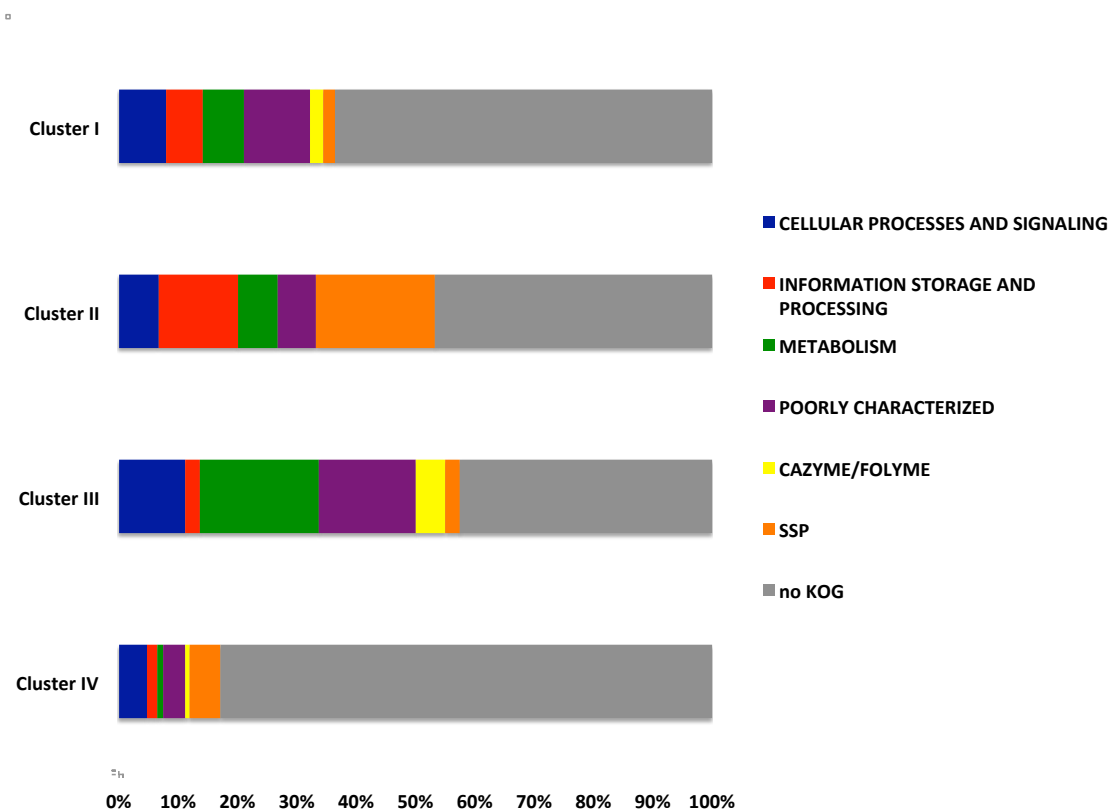
Nature Genetics: doi:10.1038/ng.3223



	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	9	no	0.06861	5	no	0.1924
Cluster II	17	yes	0.001486	11	yes	0.004098
Cluster III	6	no	0.06861	6	no	0.06655
Total down	10	yes	0.0007462	7	yes	0.0006986
Genome	4			2		

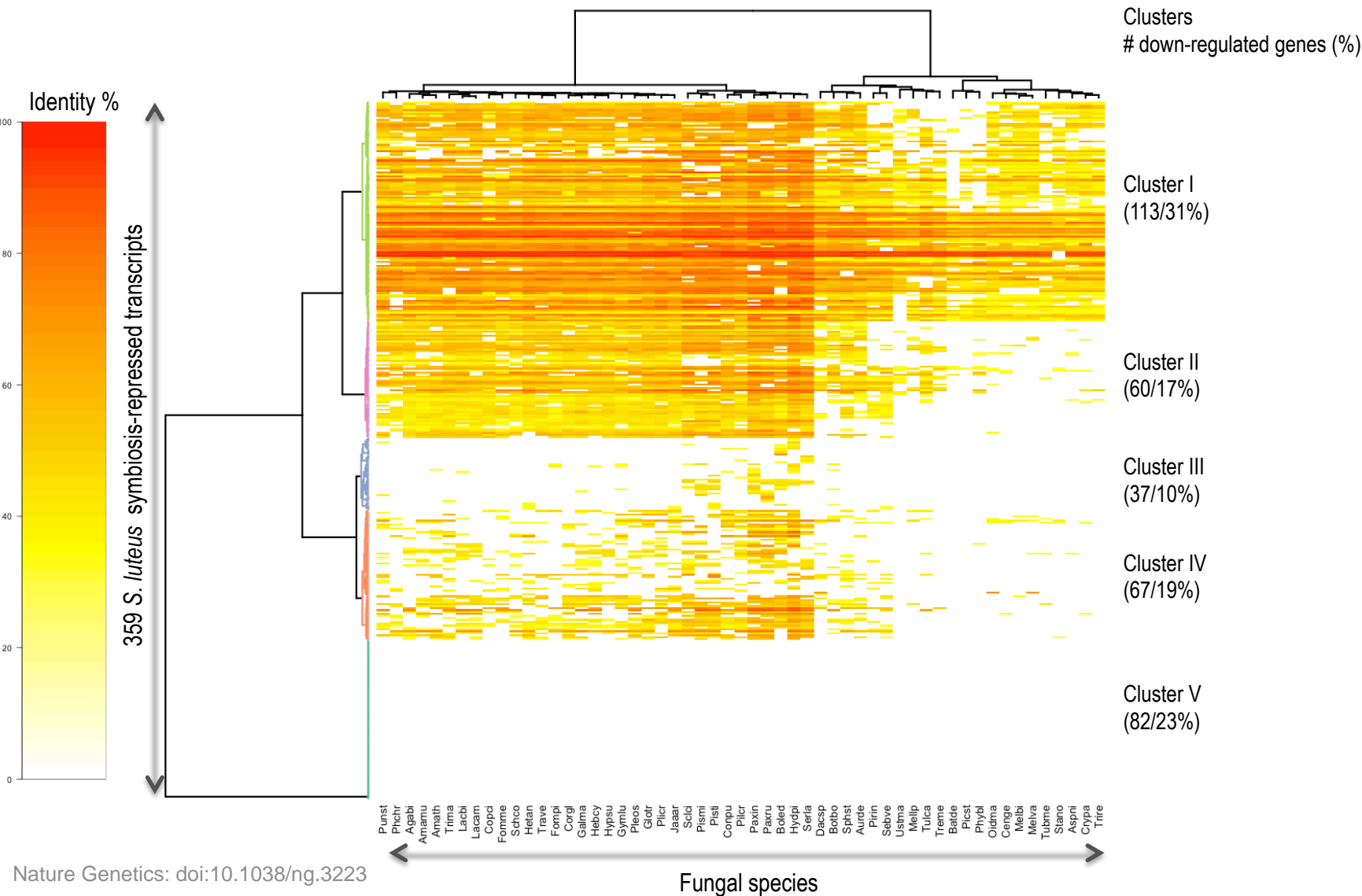
E-1 *Sebacina vermifera*



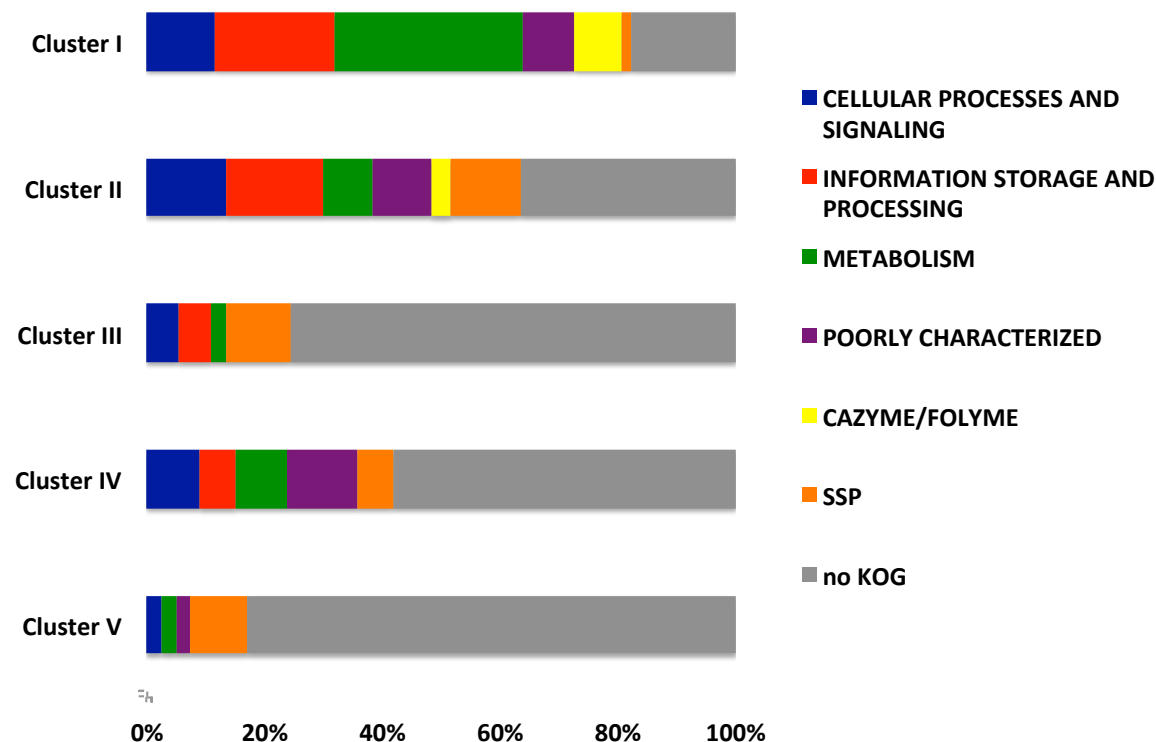


	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	7	no	0.1932	2	no	0.5432
Cluster II	20	no	0.03197	20	yes	0.002339
Cluster III	11	no	0.01367	3	no	0.4315
Cluster IV	6	no	0.2743	5	yes	0.002265
Total down	8	yes	0.004776	4	yes	0.0006494
Genome	5			2		

F-1 *Suillus luteus*

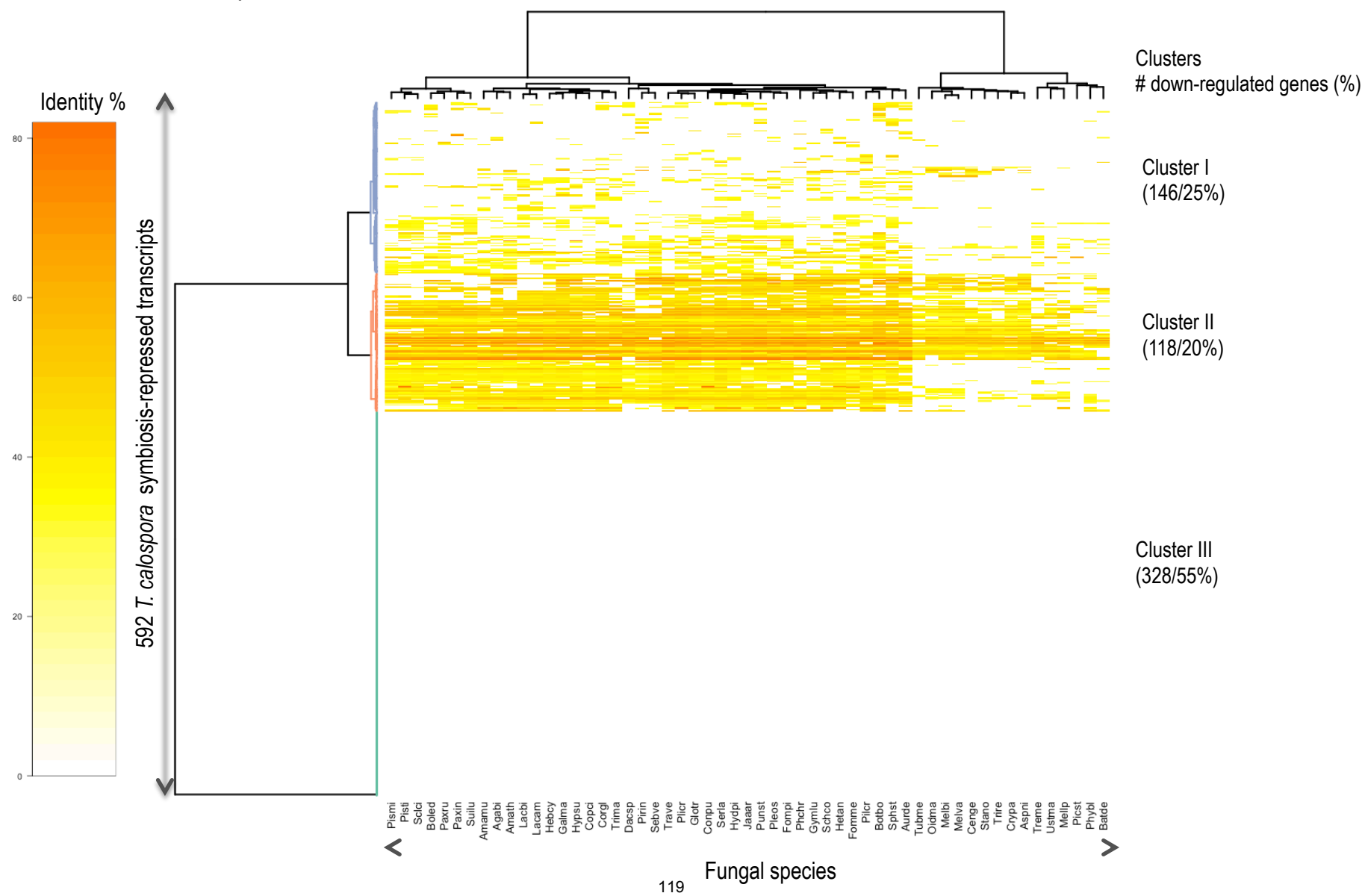


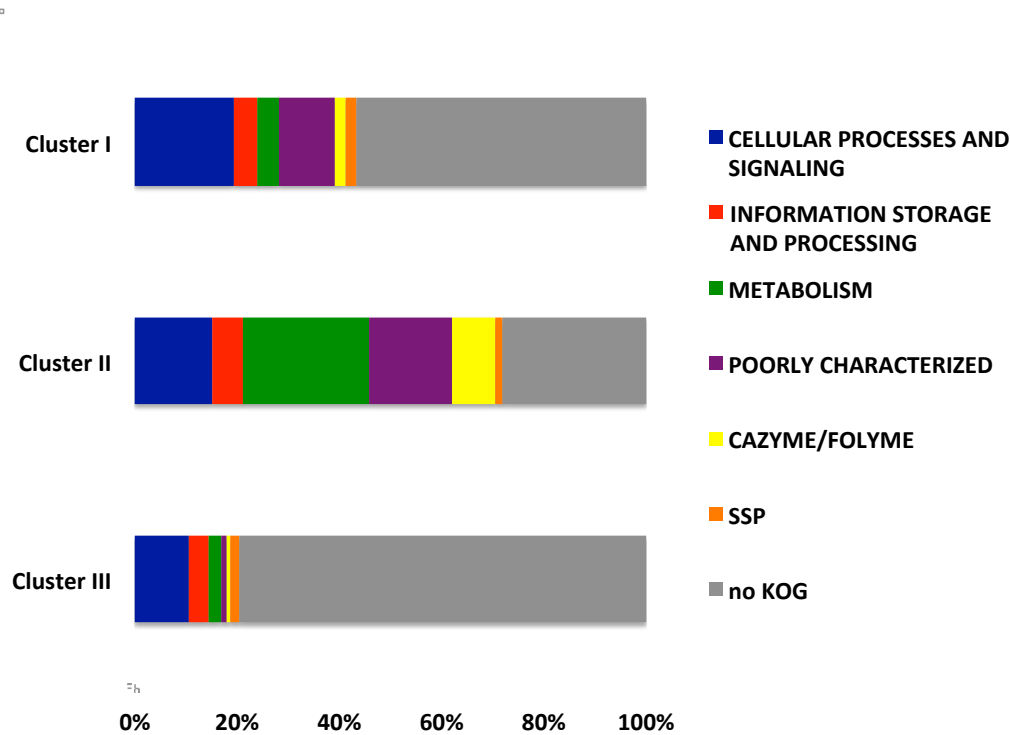
□



	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	17	yes	3.23E-10	2	no	0.433
Cluster II	18	yes	7.27E-07	12	yes	1.23E-05
Cluster III	11	no	0.01923	8	no	0.01216
Cluster IV	7	no	0.03924	6	no	0.01125
Cluster V	11	yes	0.0004655	10	yes	1.13E-05
Total down	13	yes	< 2.2e-16	7	yes	5.64E-11
Genome	3			1		

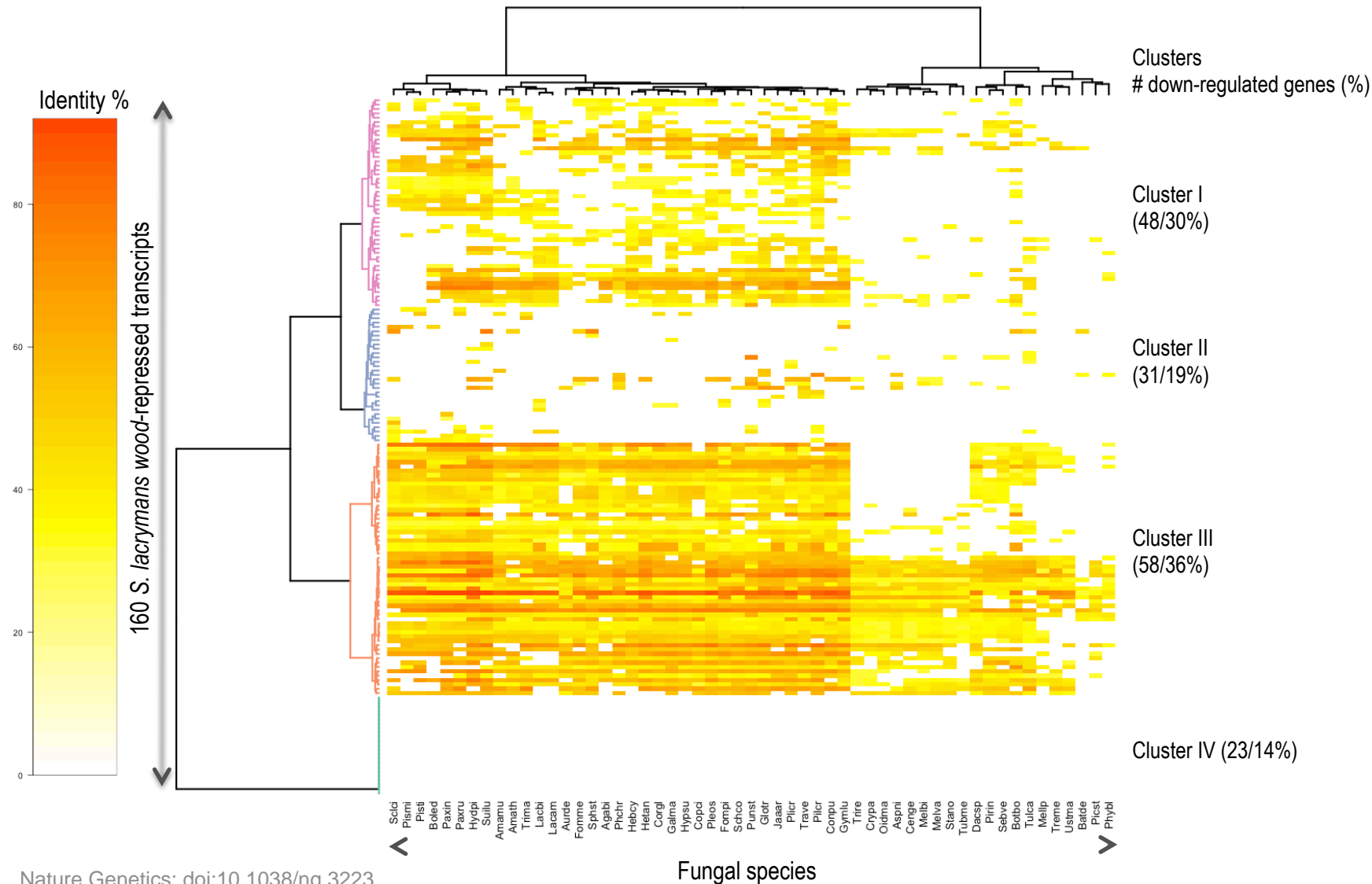
G-1 *Tulasnella calospora*

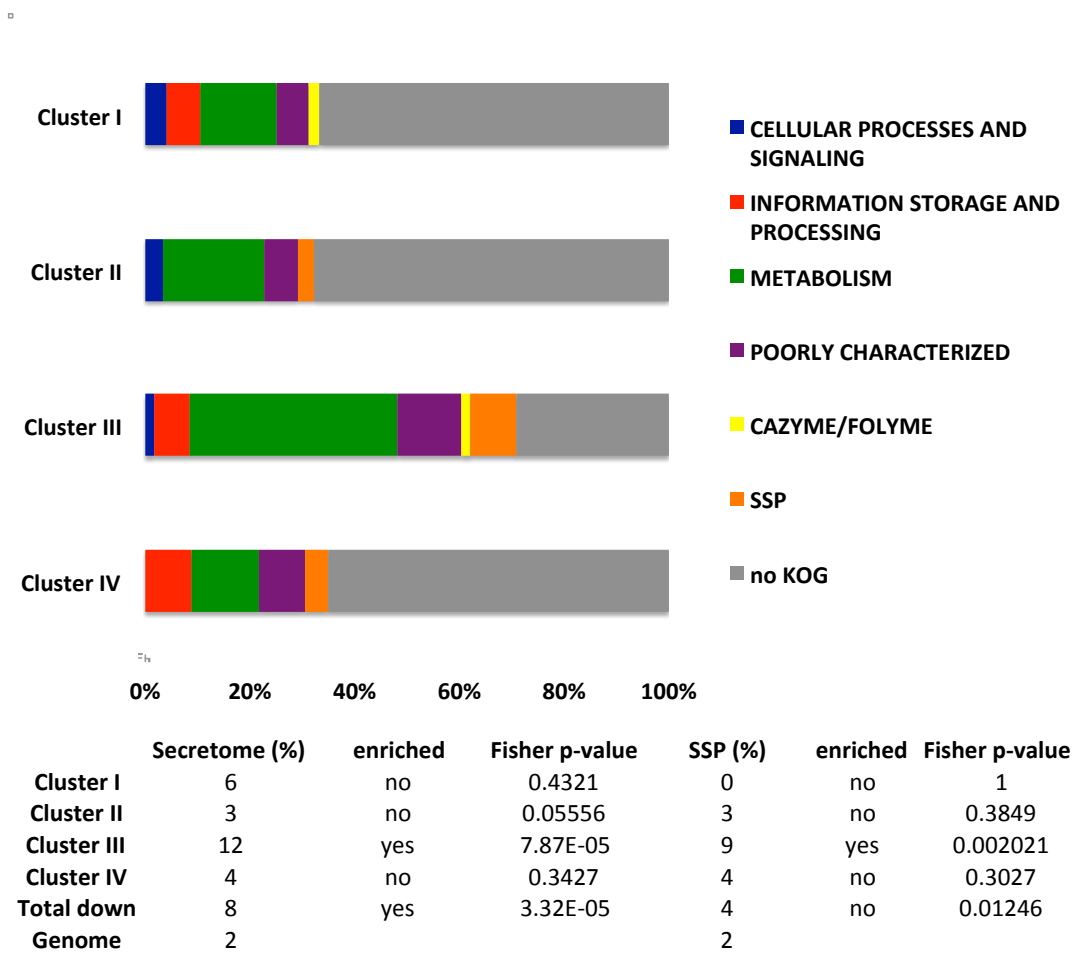




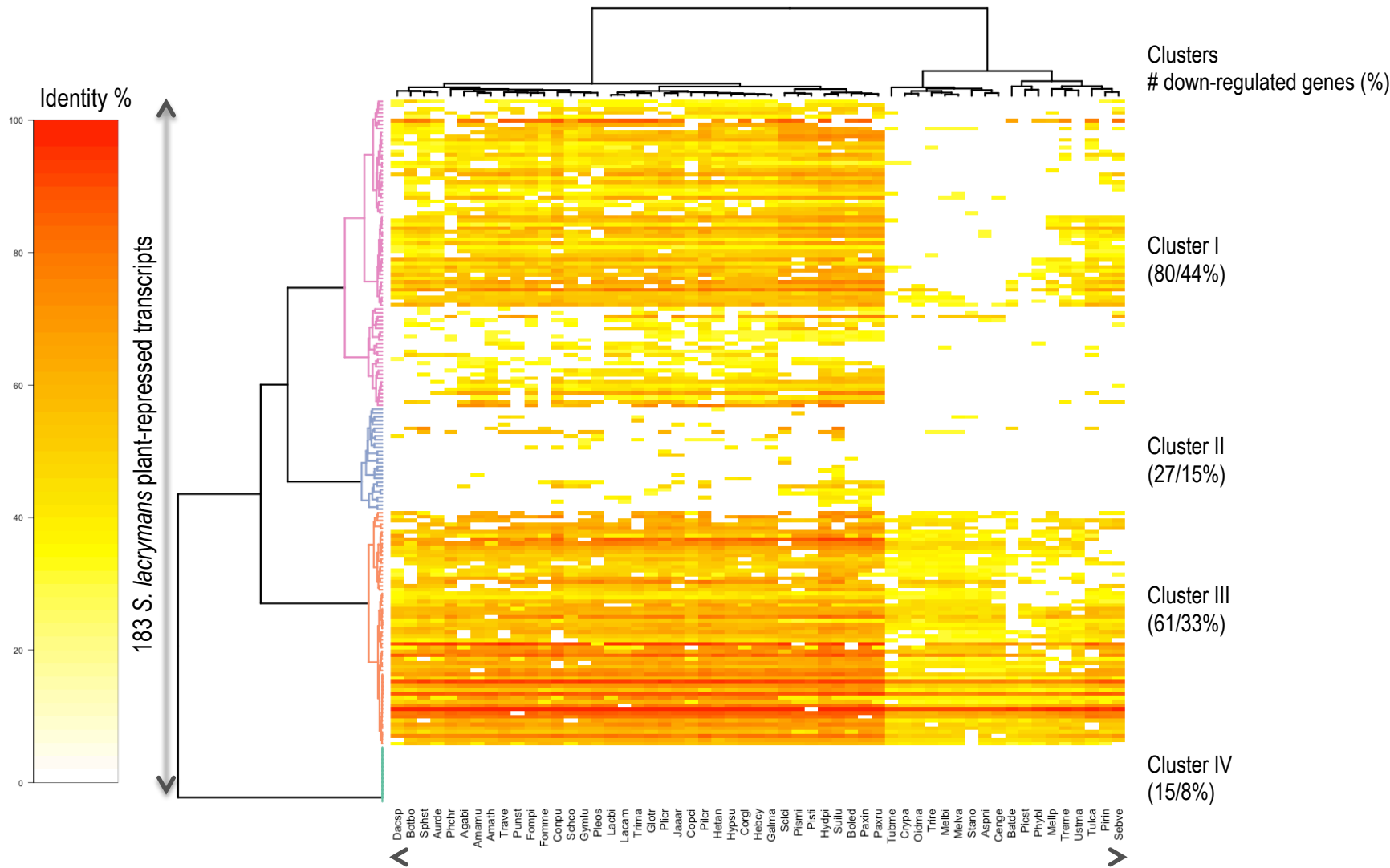
	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	5	no	0.3462	2	no	0.5149
Cluster II	8	no	0.01741	2	no	0.6498
Cluster III	3	no	0.8302	2	no	0.577
Total down	5	no	0.2312	2	no	0.5493
Genome	4			2		

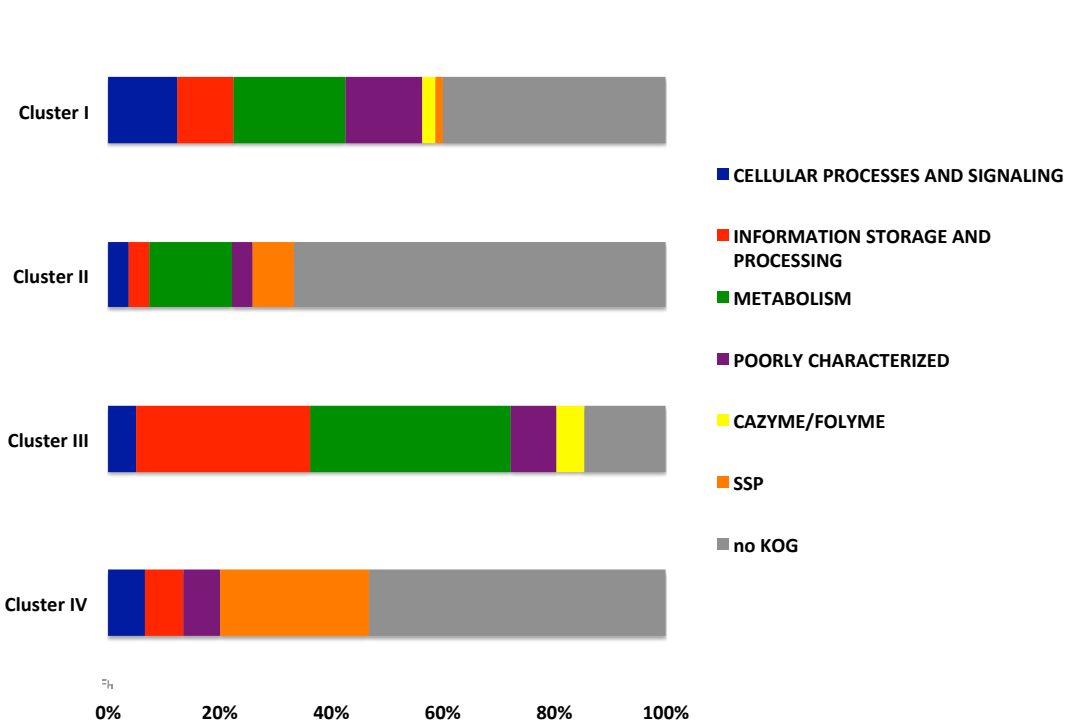
H-1 *Serpula lacrymans* mycelium grown on wood





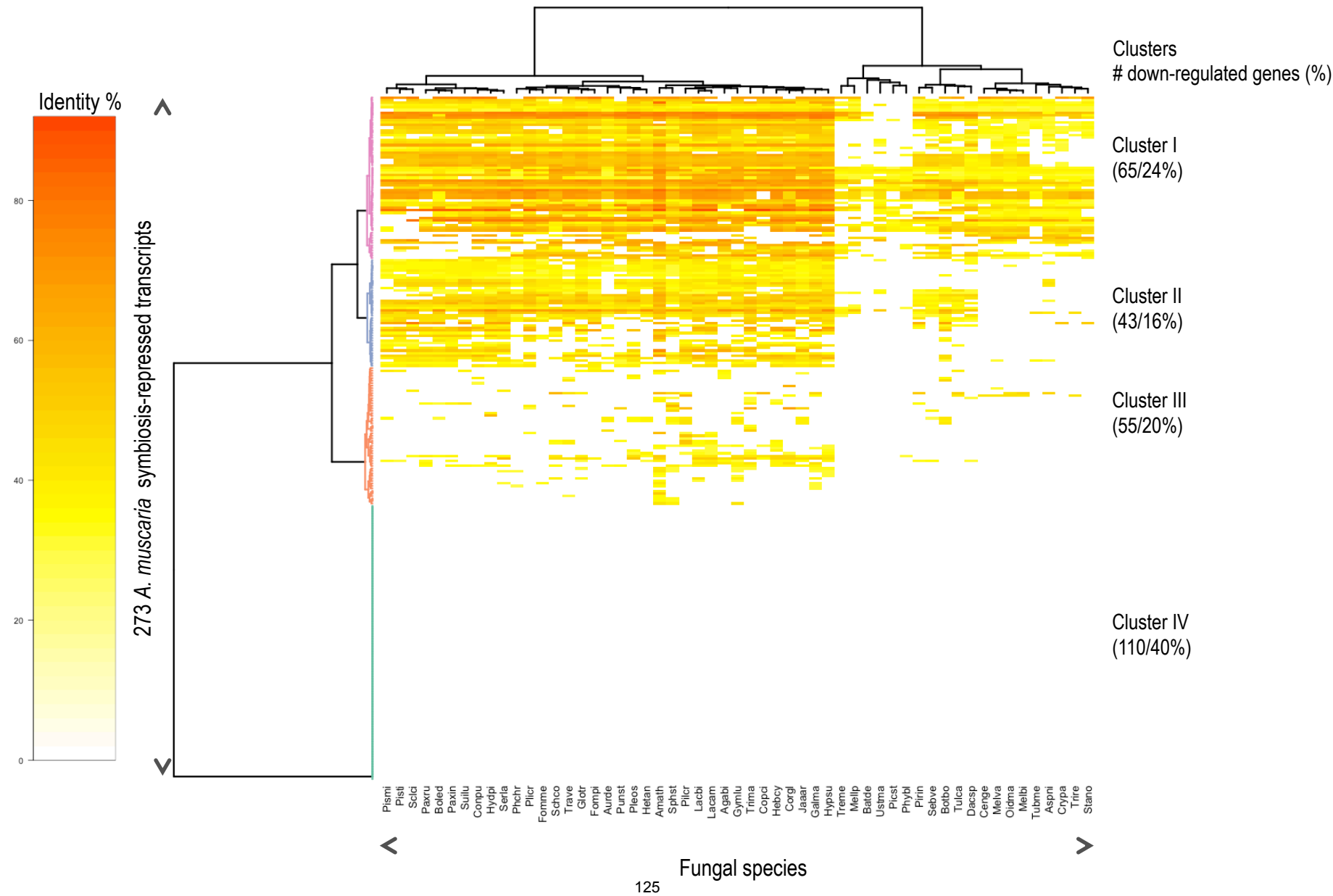
I-1 *Serpula lacrymans* pseudomycorrhiza

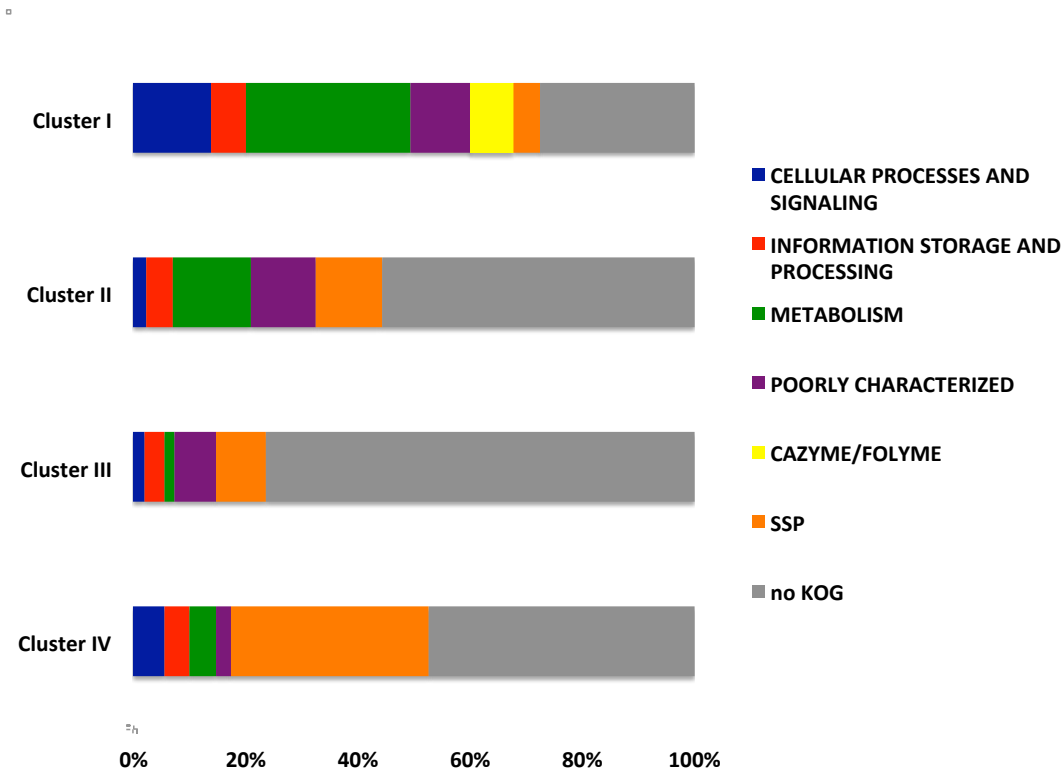




	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	5	no	0.05667	1	no	0.7154
Cluster II	7	no	0.0849	7	no	0.0654
Cluster III	2	no	0.6719	0	no	1
Cluster IV	27	yes	0.000121	27	yes	6.74E-05
Total down	6	yes	0.0004916	4	no	0.02419
Genome	2			2		

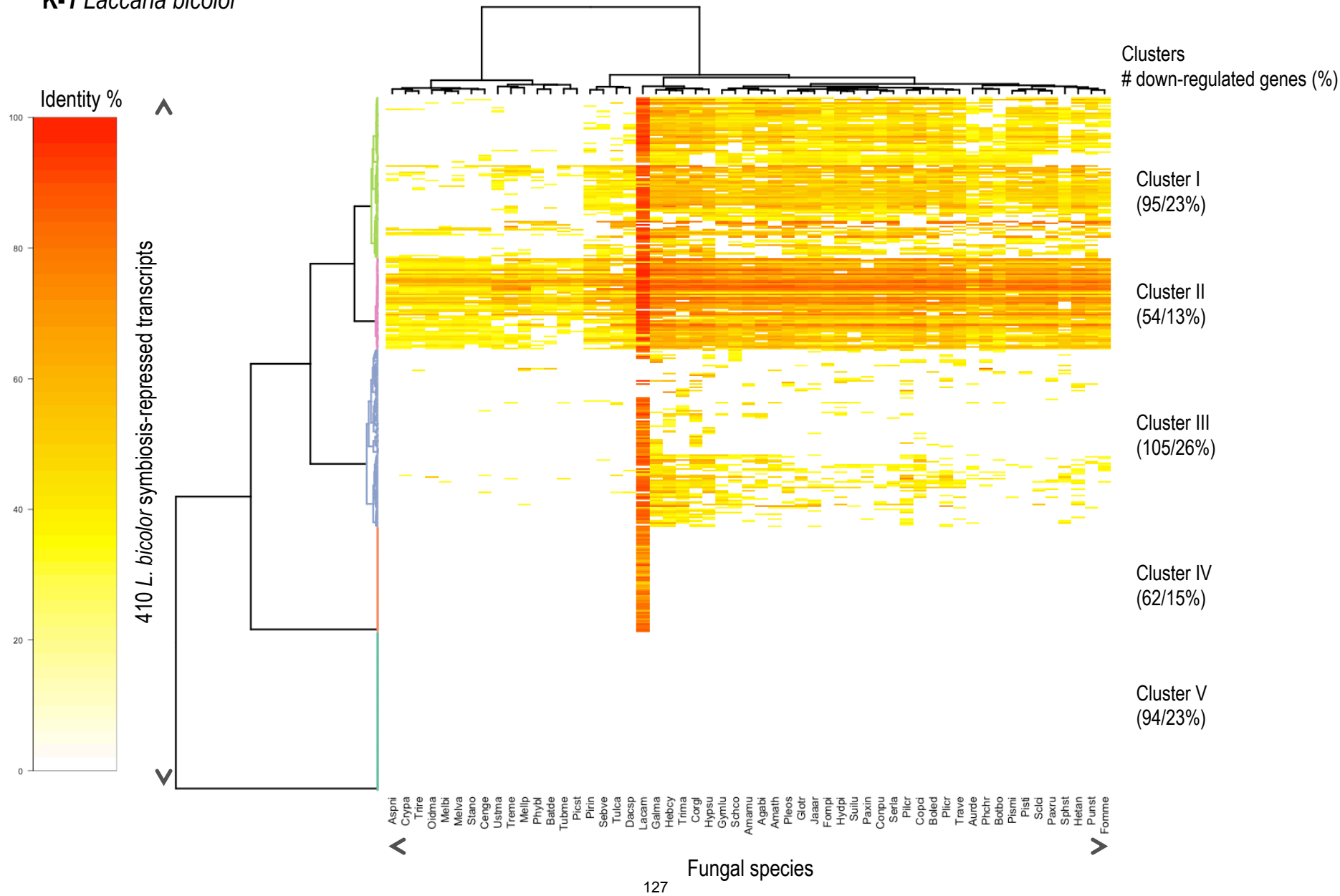
J-1 *Amanita muscaria*

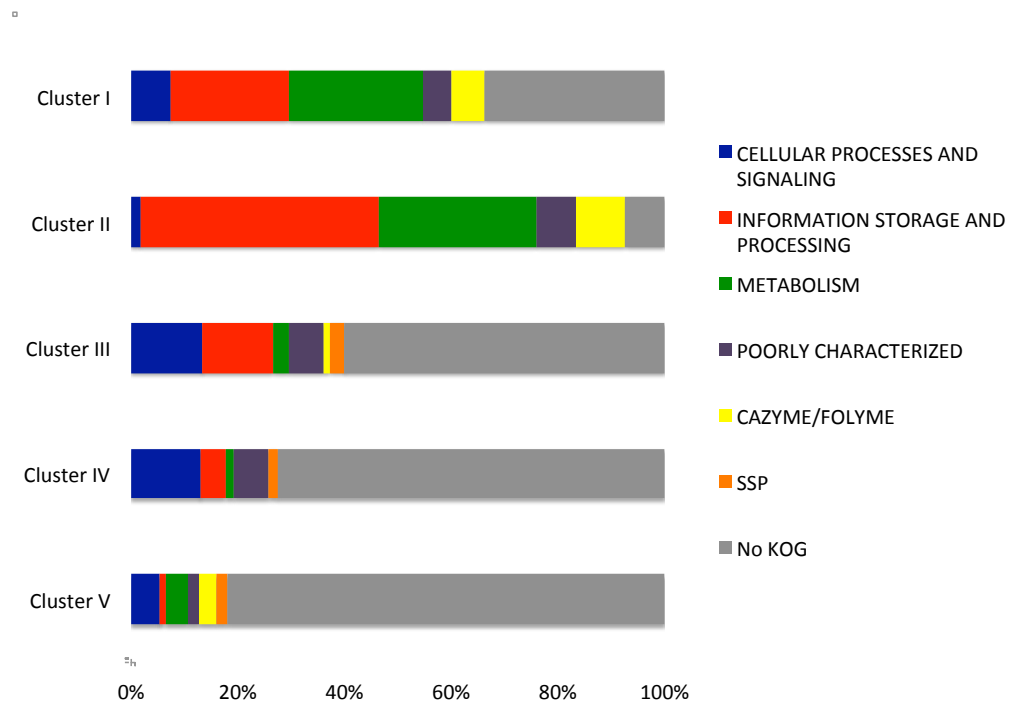




	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	17	yes	1.78E-05	5	no	0.1247
Cluster II	14	yes	0.004093	12	yes	0.001249
Cluster III	11	no	0.01358	9	yes	0.003755
Cluster IV	38	yes	< 2.2e-16	35	yes	< 2.2e-16
Total down	24	yes	< 2.2e-16	19	yes	< 2.2e-16
Genome	4			2		

K-1 *Laccaria bicolor*



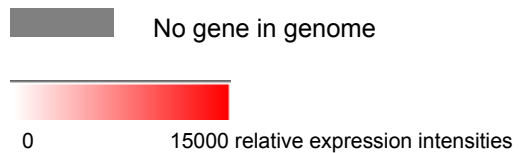
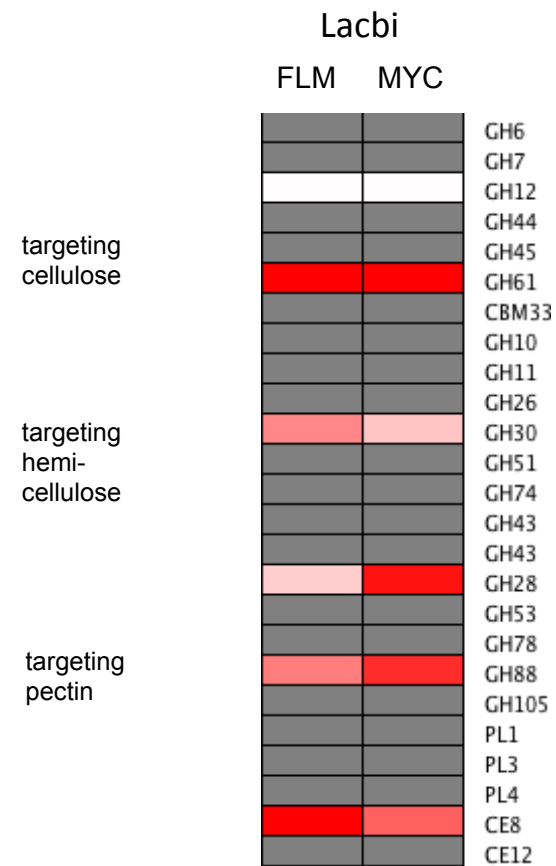


	Secretome (%)	enriched	Fisher p-value	SSP (%)	enriched	Fisher p-value
Cluster I	6	no	0.14	0	no	1
Cluster II	9	no	0.04889	0	no	1
Cluster III	4	no	0.5449	3	no	0.3362
Cluster IV	2	no	0.9026	2	no	0.7045
Cluster V	3	no	0.6782	2	no	0.5487
Total down	5	no	0.1882	1	no	0.8084
Genome	4			2		

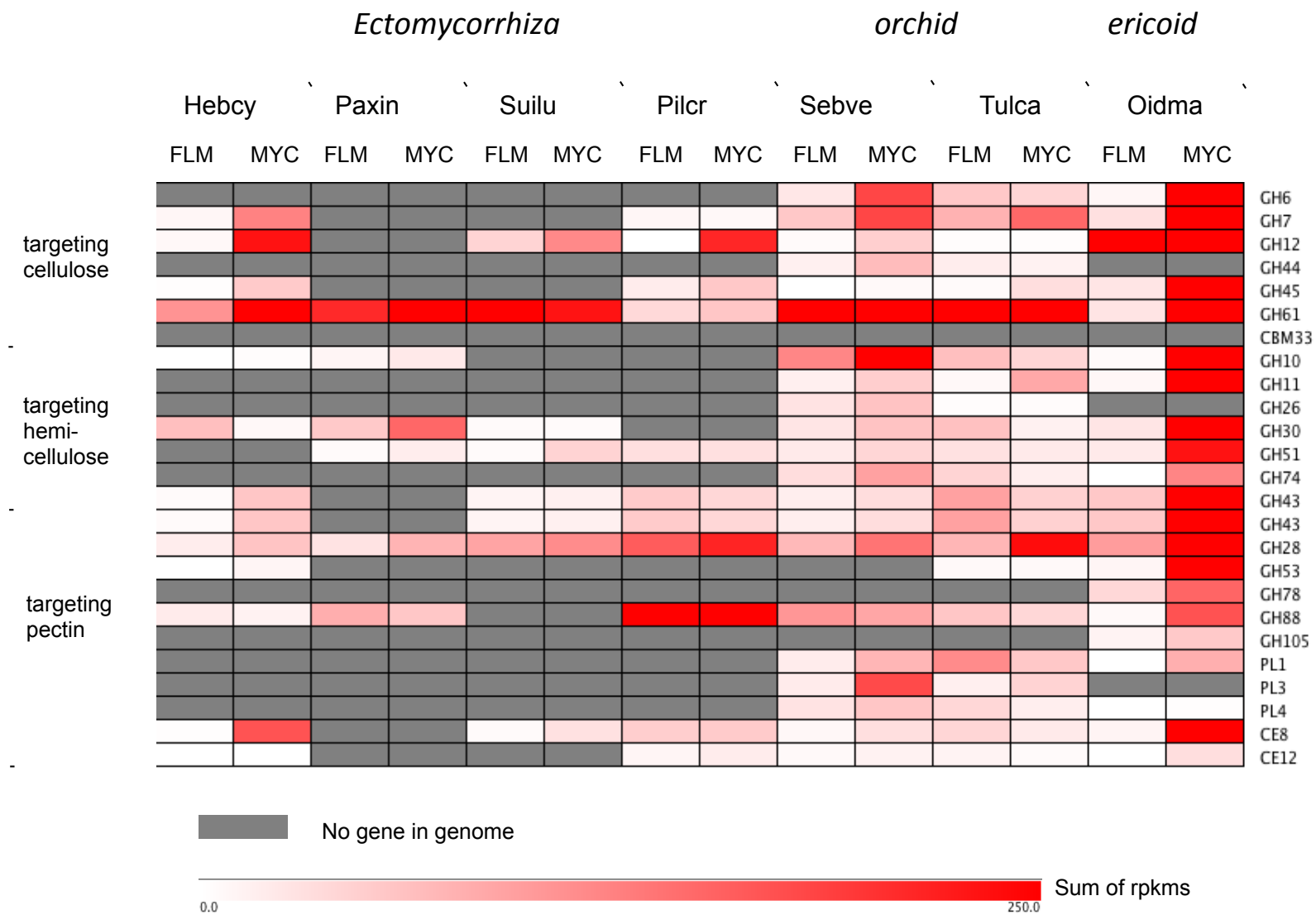
Supplementary Fig. 26. A-K Functional analysis of symbiosis-repressed transcripts.

Symbiosis-repressed genes were blasted (BLASTP) against 55 fungal genomes to find homologs. Homologs are coloured from yellow to red depending on the percentage of identity. . The heatmap represents a double-hierarchical clustering of the symbiosis-repressed transcripts orthologs in the 55 fungal genomes. Data were visualized and clustered using R (package HeatPlus). The hierarchical clustering was done by using a binary distance metric and ward clustering method. On the right panel, the number of transcripts in each cluster is shown (A-1-J-1). For each cluster the percentages of putative functional categories are given as bargrams. A table shows the percentage of transcripts coding for secreted proteins and small secreted proteins in each cluster. A Fisher exact test was applied to test if these categories were enriched compared to the number of these genes in the respective genome (A-2-K-2).

A DOWN-regulated transcripts in *Hebeloma cylindrosporum*-*Pinus pinaster* mycorrhiza compared to free-living mycelium; B DOWN-regulated transcripts in *Oidiodendron maius*-*Vaccinium myrtillus* mycorrhiza compared to free-living mycelium; C DOWN-regulated transcripts in *Paxillus involutus*-*Fagus sylvatica* mycorrhiza compared to free-living mycelium; D DOWN-regulated transcripts in *Piloderma croceum*-*Quercus robur* mycorrhiza compared to free-living mycelium; E DOWN-regulated transcripts in *Sebacina vermifera*-*Arabidopsis thaliana* mycorrhiza compared to free-living mycelium; F DOWN-regulated transcripts in *Suillus luteus*-*Pinus sylvestris* mycorrhiza compared to free-living mycelium; G DOWN-regulated transcripts in *Tulasnella calospora*-*Serapias vomeracea* mycorrhiza compared to free-living mycelium; H DOWN-regulated transcripts in *Serpula lacrymans* mycelium grown on shavings of *Pinus sylvestris* sapwood compared to control mycelium grown on glucose (MMN medium). I DOWN-regulated transcripts in *Serpula lacrymans* interacting with roots of *Picea sylvestris* (pseudomycorrhiza) compared to control mycelium grown on glucose (MMN medium); J DOWN-regulated transcripts in *Amanita muscaria*-*Populus tremula x tremuloides* mycorrhiza compared to free-living mycelium; K DOWN-regulated transcripts in *Laccaria bicolor*-*Populus trichocarpa* mycorrhiza compared to free-living mycelium

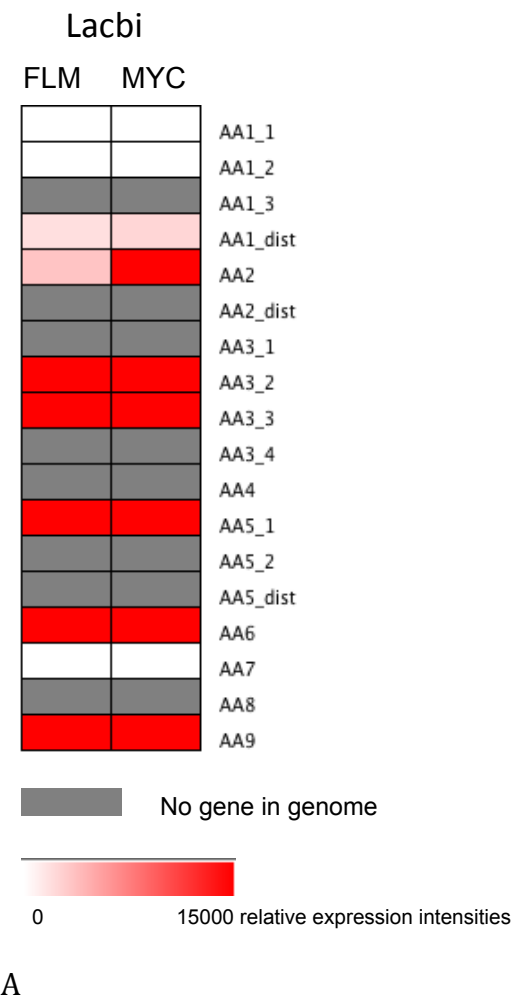


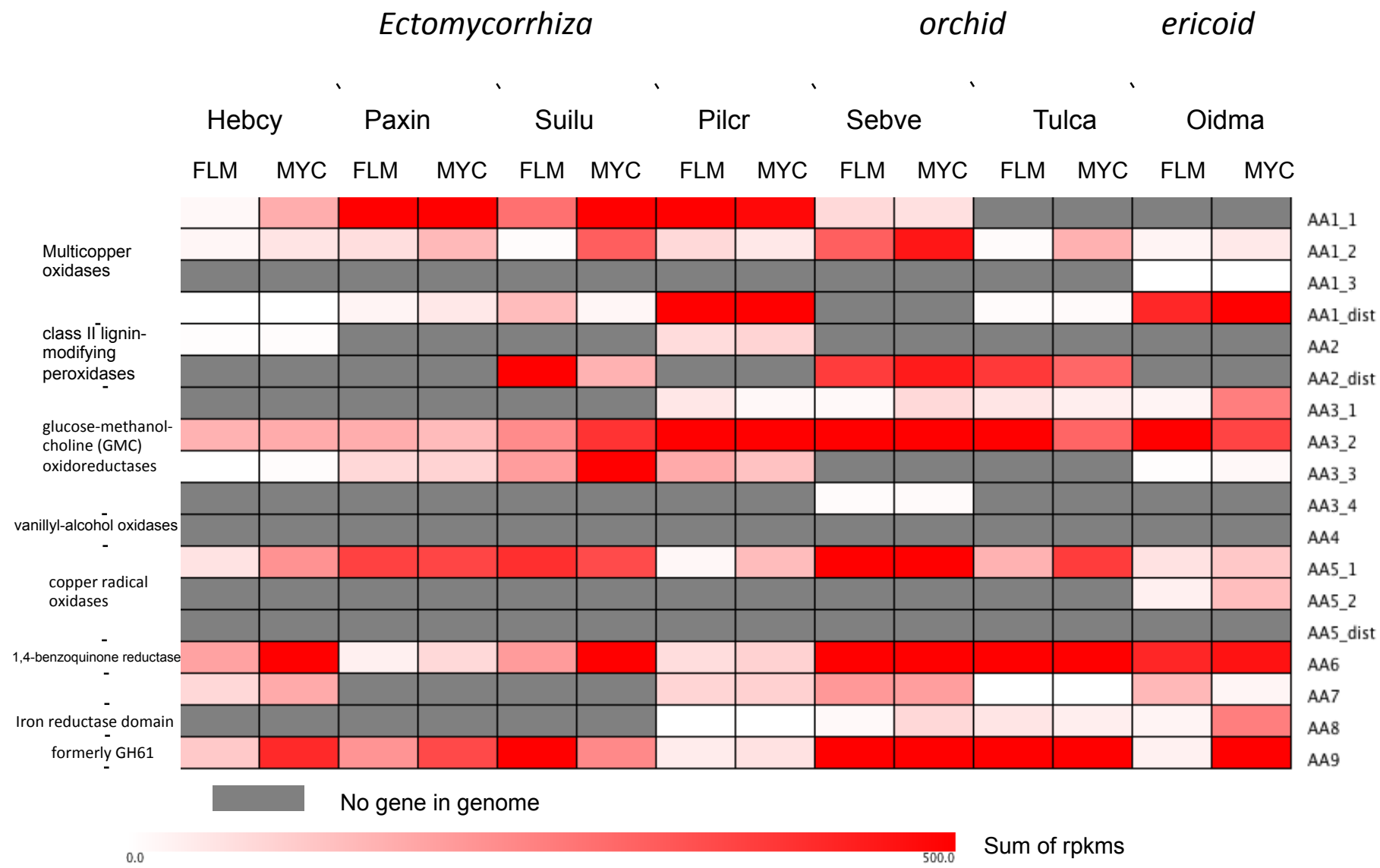
A



B

Supplementary F ig. 27A-B Expression of selected Cazymes in mycorrhizal tissue and free-living mycelium. A Lacbi; microarray data. B Hebcy, Paxin, Suilu, Pilcr, Sebve, Tulca, Oidma; RNA- Seq data. The sum of expression (sum of microarray expression values or sum of rpkms) is given for each Cazyme family and graduated in red. Grey indicates no gene for the respective Cazyme family in the respective genome.





Supplementary Fig. 28 A-B Expression of Auxiliary Activities (AA) families in mycorrhizal tissue and free-living mycelium.
A Lacbi; microarray data. B Hebcy, Paxin, Suilu, Pilcr, Sebve, Tulca, Oidma; RNA- Seq data. The sum of expression (sum of microarray expression values or sum of rpkms) is given for each AA family and graduated in red. Grey indicates no gene for the respective AA family in the respective genome.