Supplementary Information

The genome sequence of the Button Mushroom *Agaricus bisporus* reveals mechanisms governing adaptation to a humic-rich ecological niche

1. Supplementary Information Text

1.1. Genome Sequencing & Assembly

Agaricus bisporus var. *bisporus* H97 (homokaryon) was sequenced using Sanger sequencing on ABI 3730XL capillary machines. Three different sized libraries were used as templates for the plasmid subclone sequencing process and both ends were sequenced. 191,136 reads from the 2.9 kb sized library, 179,424 reads from the 6.4 kb sized library, and 36,192 reads from a 39.6 kb fosmids library were sequenced. A total of 406,752 reads were assembled using a modified version of Arachne (1) v.20071016 with parameters maxcliq1=100, correct1_passes=0 and BINGE_AND_PURGE=True. This produced 33 scaffold sequences, with L50 of 2.4 Mb, 18 scaffolds larger than 100 kb, and total scaffold size of 30.4 Mb. Two scaffold breaks where made based on integration of a genetic map. Each scaffold was screened against bacterial proteins, organelle sequences and GenBank using megablast against Genbank NR and blastp against a set of known microbial proteins. No scaffolds were identified as contamination. We classified additional scaffolds as unanchored rDNA (18), mitochondrion (1), and small repetitive (3). Additional scaffolds were removed if they consisted of greater than 95% 24mers that occurred 4 other times in the scaffolds larger than 50kb or if the scaffold contained only unanchored rDNA sequences. We also excluded 2 scaffolds smaller than 1 kb. The final assembly contains 29 scaffolds that cover 30.0 Mb of the genome with a contig L50 of 262.5 kb and a scaffold L50 of 2.3 Mb (Table S1)

Agaricus bisporus var. burnettii JB137-S8 (homokaryon) genome was sequenced using a combination of 454 (Roche) and Illumina sequencing platforms. All general aspects of library construction and sequencing can be found at the JGI website (http://jgi.doe.gov). The Illumina reads were quality trimmed to 57 bp based on an average quality per base position plot, randomly subsetted to 50X coverage depth, assembled as unpaired data with the Velvet assembler version 0.7.34 (2) with a hash length of 31, and the resulting contigs greater than or equal to 800bp in length were shredded into 1000bp chunks, if possible, with 800bp overlap. Reverse complemented shreds were also created at contig ends. 454 data and shredded Velvet assembled Illumina consensus, screened for contamination, were assembled with the Newbler assembler version 2.3-PreRelease-10/20/2009 to produce a draft with a final assembled coverage of 46X. The assembly was improved using one round of automated gap closure using gapResolution (DOE Joint Genome Institute http://jgi.doe.gov), which resulted in 131 closed gaps. The consensus of the closed gaps was stitched into the assembly and it was further improved using synteny to *Agaricus bisporus* H97_2 using Nucmer and mummerplot (3). 15 joins were identified and manually made with 600bp of gap in between each manual join. This resulted in a final improved assembly with 2016 scaffolds with an N/L50 of 8/1.23Mb, up from 2030 scaffolds with an N/L50 of 14/0.56Mb (Table S1).

The genome sequences have been deposited in the GenBank database [accession nos. AEOK00000000 (*A. bisporus* var. *bisporus* H97) and AEOL00000000 (*A. bisporus* var. *burnettii*)].

1.2. Anchoring the assembly on the genetic map

The linkage mapping data used for the scaffolds chromosome assignment was described in Foulongne-Oriol *et al.* (4, 5). This map was constructed from a haploid progeny derived from an intervarietal *A. bisporus* var. *bisporus* (U1-7) x *A. bisporus* var. *burnettii* (Jb3-83) hybrid. The *A. bisporus* linkage map was assumed to cover most of the genome of *A. bisporus* (4). Linkage groups (LGs) were assigned to chromosomes according to the

literature (6; 7). The markers for which sequences are publicly available were located on scaffolds with BLASTN (8). The genetic map and the genome assemblies were manually aligned. Maps are scaled to represent the relative physical length of each chromosome.

Ninety-two sequenced markers made possible the scaffold assignment (Fig. S1). The number of anchored markers ranged from two (LG XIII) to 11 (LG II, LG III) per chromosome. In total *ca.* 30 MB of the genome sequence (99 %) could be assigned to chromosomes. Most of the scaffolds could be genetically oriented, excepted scaffold 17, scaffold 18 and scaffold 19 for which only one marker could be aligned. The majority of the mapped markers were collinear with the sequence assembly. Inconsistencies were observed on LG I (scaffold 1). Scaffold 4 was split in two with one part assigned to LG II, and the other one to LG IV. Scaffold 29 included the rDNA repeats. On the linkage map, one rDNA tandem repeat locus was located on the distal part of LG IX.

The ratio of physical lengths to genetic distances averaged 33 kbp/cM. According to 77 segments bounded by two successive anchor markers, the physical length was correlated with the genetic distance ($r = 0.28$, $p = 0.0033$).

1.3. Transposable Elements & Simple Sequence Repeats

RepeatScout (9) was used to identify *de novo* repetitive DNA in the *A. bisporus* H97 genome. The default parameters (with l= 15) were used. RepeatScout generated a library of 755 consensus sequences. This library was then filtered as follows: 1) all the sequences less than 100 bp in size were discarded; 2) repeats having less than five copies in the genome were removed (as they may correspond to protein-coding gene families) and 3) repeats having significant hits to known proteins in UNIPROT (The UNIPROT consortium, 2008) other than proteins known as belonging to TEs were removed. The 216 consensus sequences remaining were annotated manually by a TBLASTX search (8) against RepBase (10); 32 consensus sequences showed homologies with Class 1 retrotransposons and 13 with Class 2 transposons (Fig. S3). The remaining 171 consensus sequences were not categorized. To identify full length LTR retrotransposons, a second *de novo* search was performed with LTR_STRUC (11). The program yielded 36 full-length candidates LTR retrotransposon sequences, which were checked for their homology using the BLASTN algorithm (8) against the sequences coming from the RepBase database. Among the 36 putative full length LTRs, 15 were attributed to *Gypsy*/*Ty3*-like elements and 13 to *Copia*/*Ty1*-like. Eight other elements did not exhibit a significant homology with known TE families. These sequences have been excluded for further analyses. The insertion age of full length LTRs was determined from the evolutionary distance between 5'- and 3'-solo LTR derived from a ClustalW alignment of the two solo LTR sequences using the Kimura correction in ClustalW (12). For the conversion of the nucleotide sequence distance to putative genome insertion age, a substitution rate of 1.3×10^{-8} mutations per site per year was used (13). The number of TE occurrences and the percent of genome coverage were assessed by masking the genome assembly using RepeatMasker (14) with the 216 consensus sequences coming from the RepeatScout pipeline and the 28 coming from the LTR_STRUC pipeline. RepeatMasker masked 11.17 % of the genome assembly; 3.76 % of the genome was masked by repeated elements belonging to unknown/uncategorized families and 3.5% by Class 1 retrotransposons *Gypsy* (Fig. S3).

MISA (http://pgrc.ipk-gatersleben.de/misa/) with default parameters was used to identify mono- to hexanucleotide simple sequence repeat (SSR) motifs. A total of 3,134 SSRs have been identified in the *A. bisporus* H97 genome corresponding to 2,030 mono-, 354 di-, 701 tri-, 22 tetra-, 10 penta- and 17

hexanucleotide motifs. The relative abundance of SSRs was calculated as the number of SSRs per Mb. For all 3,134 SSRs, the relative abundance was 104 SSRs/Mb.

1.4. Genome Annotation

Both genomes were annotated using the JGI annotation pipeline, which takes multiple inputs (scaffolds, ESTs, and known genes) and runs several analytical tools for gene prediction and annotation, and deposits the results in the JGI Genome Portal (http://jgi.doe.gov/Agaricus) for further analysis and manual curation. Genomic assembly scaffolds were masked using RepeatMasker (14) and the RepBase library of 234 fungal repeats (10). tRNAs were predicted using tRNAscan-SE (15). Using the repeat-masked assembly, several gene prediction programs falling into three general categories were used: 1) *ab initio* - FGENESH (16); GeneMark (17), 2) *homology-based* - FGENESH+; Genewise (18) seeded by BLASTx alignments against GenBank's database of non-redundant proteins (NR: http://www.ncbi.nlm.nih.gov/BLAST/), and 3) *EST-based* - EST_map (http://www.softberry.com/) seeded by *A. bisporus* H97 EST contigs. Genewise models were extended whenever possible using scaffolds data to find start and stop codons. EST BLAT alignments (19) were used to extend, verify, and complete the predicted gene models. The resulting set of models was then filtered for the best models, based on EST and homology support, to produce a non-redundant representative set. This representative set was subject to further analysis and manual curation. Measures of model quality include proportions of models complete (with start and stop codons) (>87% of models), consistent with ESTs (>61% of models covered over ≥75% of exon length), supported by similarity with proteins from the NCBI NR database (>83% of models). Quality metrics for gene models are summarized in Table S2.

All predicted gene models were functionally annotated using SignalP (20), TMHMM (21), InterProScan (22), BLASTp (8) against GenBank's database of non-redundant proteins, and hardware-accelerated double-affine Smith-Waterman alignments (deCypherSW; http://www.timelogic.com/decypher_sw.html) against SwissProt (http://uniprot.org), KEGG (23), and KOG (24). KEGG hits were used to assign EC numbers (http://enzyme.expasy.org/), and Interpro and SwissProt hits were used to map GO terms (http://www.geneontology.org/).

The different species chosen for multigene families prediction cover Agaricomycotina, Tremellomycetes, Pucciniomycotina and Ustilagomycotina phyla, with representatives from Agaricales, Boletales, Russulales, Polyrales, Tremellales, Pucciniales, Sporidiobolales, Ustilaginales and Malasseziales. Protein sets from *Agaricus bisporus* H97 (10,438 models), *Coprinus cinereus* okayama 7 (13,394 models), *Laccaria bicolor* (19,036 models), *Schizophylum commune* (13,181 models), *Pleurotus ostreatus* PC15 (11,579 models), *Serpula lacrymans* (12,917 models), *Heterobasidion annosum* (12,193 models), *Postia placenta* (12,415 models), *Phanerochaete chrysosporium* (10,048 models), *Cryptococcus neoformans grubii* h99 (6,967 models), *Tremella mesenterica* (8,313 models), *Melampsora larici populina* (16,841 models), *Puccinia graminis* (18,241 models), *Sporobolomyces roseus* (5,536 models), *Ustilago maydis* (6,522 models) and *Malassezia globosa* (4,274 models) were retrieved from the Joint Genome Institute Portal and the Broad Institute Portal.

Multigene families were predicted from 181,895 predicted proteins found in *A. bisporus* and the other representatives using MCL algorithm (25) with inflation parameter set to 3.0. As a result, 5,058 protein families phylogenetically relevant (containing at least two species and 10 sequences) were identified. The *A. bisporus*

bigger gene families are coding for proteins containing a Cytochrome p450 domain, Protein kinase domain, GMC oxidoreductase or Fungal hydrophobin domain (Table S5).

Multigene families were analyzed for evolutionary changes in protein family size using the CAFE program (26) (Fig. S5). 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. A phylogenetic tree was constructed with 203 core genes representatives (FUNYBASE) of the following fungi: *A. bisporus* and other basidiomycetes.

In *A. bisporus*, 496 families were expanded, 3,811 showed no change and 751 families had undergone contraction by comparison to a putative common ancestor Basidiomycota having 5,058 gene families. Tables S6 and S7 present contracted and expanded gene families with lower p-value in *A. bisporus* genome, respectively. The PFAM domains for these families were associated by homology searches using hmmscan from HMMER3 package (27) and the PFAM database (28).

Manual curation of the automated annotations was performed trough by using the web-based interactive editing tools of the JGI Genome Portal to assess predicted gene structures, assign gene functions, and report supporting evidence.

Secreted proteins were identified using a custom pipeline including the TargetP (http://www.cbs.dtu.dk/services/TargetP/) and SignalP (http://www.cbs.dtu.dk/services/SignalP/) algorithms (20, 29). The secreted peptidases were identified in the genome using the MEROPS database (30; http://merops.sanger.ac.uk).

The gene models coding for membrane transporters were identified and curated by using the Transport Classification Database (TCD) (31; http://www.tcdb.org/). This approach combines BLAST searches against a curated membrane transport protein database (TCD), as well as HMM searches and COG-based searches against membrane transporter protein families. Membrane transporters were assigned to protein families based on sequence similarities.

1.5. EST Sequencing, Clustering, and Assembly

Total RNA of *A. bisporus* var. *bisporus* strain U1 (heterokaryon) was used to extract PolyA$^+$ RNA using oligo (DT) magnetic beads (Absolutely mRNA™ Purification kit, Stratagene). PolyA RNA was reversed transcribed using Superscript III (Invitrogen) using a $dT_{15}VN_2$ primer. Second strand cDNA was synthesized by nick translation with E. *coli* DNA ligase, E. *coli* DNA polymerase I, and RNase H and blunt end repaired using T4 polymerase (Invitrogen). The dscDNA was fragmented and 300-800 base pair fragments were gel purified using a 2% agarose gel. The purified fragments were then used to create the 454 single stranded cDNA library as described below (454 library preparation kit, Roche).

The fragment ends were polished using T4 ligase and T4 polynucleotide kinase (Roche). Adaptors containing primer sequences and a biotinylated tag were ligated to the fragment ends (Roche). The fragments with properly ligated adapters were immobilized onto magnetic streptavidin-coated beads (Roche). Nicks or gaps between the adapters and the dscDNA fragments were repaired using the fill-in polymerase (Roche). The non biotinylated

strands of the immobilized dscDNA fragments were melted off to generate the single stranded cDNA library for 454 sequencing.

The ESTs were evaluated for the presence of polyA tails (which if present were removed) then evaluated for length, removing ESTs with fewer than 50 bases remaining. Additionally, ESTs consisting of more than 50% low complexity sequence were removed from the final set of ESTs. For clustering, ESTs were evaluated with malign (JGI tool), a kmer based alignment tool, which clusters ESTs based on sequence overlap (kmer = 16, seed length requirement = 32 alignment ID >= 98%). EST clusters were then each assembled using CAP3 (32) to form consensus sequences.

Three separate RNA samples were used to generate libraries: vegetative, undifferentiated mycelium on casing and compost substrates, and fruiting bodies. They produced, respectively, 442,365, 291,949, and 405,357 ESTs. Together with 470 ESTs from external sources, they all were clustered and assembled into 79,271 consensus sequences and 42,693 singletons. ESTs have been submitted and released by NCBI Short Read Archive: Accession n° SRA037617.

1.6. Transcriptome analysis

Gene expressions were profiled in the commercial (heterokaryon) strain A15. Casing compost and fruit body samples were grown in standard conditions. The casing sample consisted of a mixture of mycelium aggregates and primordia. The culture samples refer to axenic culture and the media used was compost extract medium in agar plates (33). The 'Fruiting bodies' sample represents the mature mushroom fruit body (including the stipe, cap and *pilei pellis* (skin) tissues). The 'Compost' sample represents the mycelium growing in wheat straw compost. This mycelium is both fine (likely to be nutritional) and strands or cords (some secondary structure and thicker (clearly visible to the naked eye). The 'Casing' sample consists of fine and stranded/corded mycelium, and stages to fruit body primordia (including aggregates and hyphal knots, undifferentiated and differentiated primordia). The main components of compost are wheat straw, horse and chicken manure, gypsum and soy meal. It originates from a pre-treatment that includes bacterial pre-digestion and also the presence of other fungi during cultivation. The samples for RNA extraction were collected on separate occasions from separate mushroom houses. Four biological replicates of each developmental stage were analyzed.

RNA was prepared from fruiting body and culture samples using a standard Trizol protocol. RNA was extracted from compost and casing samples using a method based on the FastRNA Pro Soil-Direct kit (MP Biochemicals). RNA was quantified using a NanoDrop-1000 spectrophotometer and quality was monitored with the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA).

Custom arrays (Agilent ID 027120) were developed using 10,438 CDS (filtered model set) from the H97 v2 gene annotation; 5 x 60-mer oligos per CDS and the 8 x 60K randomised format were designed using the Agilent eArray software. Cyanine-3 (Cy3) labeled cRNA was prepared from 0.6 ug RNA using the Quick Amp Labelling kit (Agilent) according to the manufacturer's instructions, followed by RNAeasy column purification (QIAGEN, Valencia, CA). Dye incorporation and cRNA yield were checked with the NanoDrop ND-1000 Spectrophotometer. 600 ng of Cy3-labelled cRNA (specific activity >10.0 pmol Cy3/ug cRNA) was fragmented at 60°C for 30 minutes in a reaction volume of 25 µl containing 1x Agilent fragmentation buffer and 2x Agilent blocking agent following the manufacturers instructions. On completion of the fragmentation reaction, 25 µl of

2x Agilent hybridization buffer was added to the fragmentation mixture and hybridized to Agilent arrays (ID 027120) for 17 hours at 65°C in a rotating Agilent hybridization oven. After hybridization, microarrays were washed 1 minute at room temperature with GE Wash Buffer 1 (Agilent) and 1 minute with 37°C GE Wash buffer 2 (Agilent) then 10 seconds in acetonitrile and 30 seconds in Stabilization and drying solution (Agilent). Slides were scanned immediately after washing on the Agilent's High-Resolution C Scanner (G2505C US94100321) using one color scan setting for 8 x 60K array slides (Scan resolution 3um). The scanned images were analyzed with Feature Extraction Software (Agilent) using default parameters (protocol GE1_107_Sep09 and Grid: 027120_D_F_20100129) to obtain background subtracted and spatially detrended Processed Signal intensities. Features flagged in Feature Extraction as Feature Non-uniform outliers were excluded.

Raw array data were Robust multichip average normalized using the ARRAYSTAR software (DNASTAR, Inc. Madison, WI, USA). A Student t-test with false discovery rate (FDR) multiple testing corrections was applied to the data using the ARRAYSTAR software (DNASTAR). Transcripts with a significant p-value (<0.05) were considered as differentially expressed. The complete expression dataset is available as series GSE39569 at the Gene Expression Omnibus at NCBI (http://www.ncbi.nlm.nih.gov/geo/). In addition a supplemental table summarizing the array data is available for download on http://mycor.nancy.inra.fr/Agaricus/.

Comparison of ratios of compost/culture transcript profiles was used to identify the most highly upregulated transcripts found in mycelium grown on compost during vegetative growth. To assess the effect of a humic environment on gene expression the ratio of humic/non-humic expression was calculated as: expression in compost/(average expression in culture, casing and fruitbodies). The comparison of compost/fruitbody transcript profiles highlights developmental stage differences during mushroom formation.

1.7. Growth and utilization of carbohydrates

Growth and utilization of carbohydrates. Growth of *A. bisporus* was performed using minimal medium with the carbon sources as indicated in the text and at www.fung-growth.org. Minimal medium consisted of 20.5 mM MOPS, 2 mM KH_2PO_4, 1 mM $MgSO_4$, 0.5 mM $CaCl_2$, 0.134 mM EDTA, 25 µM FeO_4, 5 µM $ZnSO_4$, 5 µM $MnSO_4$, 4.8 µM H_3BO_3, 2.4 µM KI, 52 nM Na_2MoO_4, 4 nM $CuSO_4$, 4 nM $CoCl_2$, 0.5 µM thiamine.HCL, 0.1 µM D(+) biotine and 20 mM NH_4CL and was set at pH 6.8. Carbon sources were added to a final concentration of 25 mM for monosaccharides and 1% for polysaccharides.

A. bisporus var. bisporus H97, H39 and their dikaryon U1 and *A. bisporus var. burnettii* JB137-s8 were grown on minimal medium containing 25 mM of the carbon sources as indicated in Fig S7. The commercial heterokaryon U1 was similar to H39 on most monosaccharides. U1 and H39 showed more dispersed growth on D-glucose, D-mannose, L-rhamnose, and better growth on D-xylose, L-arabinose and D-galacturonic acid that H97. U1 grew best on D-galactose, while less dispersed growth was observed for H97 and no growth for H39. *A. bisporus var. burnettii* JB137-s8 differed significantly from the other strains in that it grew best on D-glucose and D-mannose, with slower growth on D-galactose, D-xylose and L-arabinose, and only residual growth on L-rhamnose. It was not able to grow on D-galacturonic acid (Fig S7).

Effect of humic acids on growth and utilization of carbohydrates. Growth of *A. bisporus* was performed using minimal medium (see above) with the carbon sources and with adding humic extracts (Fig S7). Humic extracts were presented as a pine (cypress) soil extract and compost (phase II) extract and they were added to a final

concentration of 0.1%. *A. bisporus var. bisporus* H97, H39 and their dikaryon U1, *A. bisporus var. burnettii* JB137-s8 and the monokaryotic strain of *C. cinereus* were grown on minimal medium containing 25 mM of monosaccharides alone or with 0.1 % of humic extracts. The U1 heterokaryon grew similarly on most monosaccharides and polysaccharides with or without humic extracts, suggesting that the humic extracts do not stimulate vegetative mycelial growth under these conditions. On some carbon sources a small reduction in growth was observed in the presence of the cyprus litter extract (Fig S7). In contrast, growth of the monokaryons H97 and H39 was stimulated by the humic extracts, but the extent of this depended on the carbon source that was available. *A. bisporus* H97 grew better on glucose with addition of humic extracts. Growth of H39 was more dispersed on xylose, arabinose and glucose with pine soil and compost extracts than media without the extracts (Fig. S7). Both H97 and H39 grew better on guar gum, pectin, soluble starch and inulin when the extracts were added. This suggests a synergy between the presence of humic extracts and degradation of the hemicellulose analogue guar gum to stimulate colonization of the substrate. Overall the positive effect of the compost extract was stronger than the Cyprus litter extract, which is likely due to the adaptation of the commercial *A. bisporus* strains to this substrate.

1.8. Prediction of CAZymes and lignolytic oxidoreductases

To determine whether *A. bisporus* sugar-cleaving capabilities resemble those of other fungi, we have undertaken a comparison of the glycoside hydrolases (GH) and polysaccharide lyases (PL) repertoires of 25 completely sequenced fungi (Fig. S8). For each fungus we have listed the number of representatives of each CAZy family (families defined in the CAZy database; www.cazy.org; 34) and then 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 GINKGO (35) and the distance trees were constructed with FastME (36).

Then, we searched *A. bisporus var bisporus* and *A. bisporus var burnettii* proteomes for 27 gene families encoding oxidoreductases and carbohydrateactive enzymes (CAZymes) that are known to be implicated in wood decay (37)(Table S10). To estimate the significance of the observed family size differences among taxa, we performed an analysis using the CAFE program (26) with a p-value <0.001. To assess if *Agaricus bisporus* is sharing the same distribution for these 27 gene families with brown rot or white rot fungi, we carry out a Correspondance Analysis, with R package FactoMineR. (Fig. S10)

1.9. Genome comparisons between *A. bisporus* var. *bisporus* and *A. bisporus* var. *burnettii*

These two varieties differ in their ecology, modes of reproduction, and morphological characteristics. Understanding their genome structural differences could potentially help us to reveal the genetic basis for such differences and help to guide future hybridization and breeding programs. Segmental duplications were found for 54 genes and 31 genes in 15 and 9 blocks respectively for the var. *bisporus* and var. *burnettii* strains, with maximum 6 and 5 genes per block in the two strains. However, the amounts of tandem repeats were larger than segmental duplications. In the *A. bisporus* var. *bisporus* strain H97, 1083 genes were included in 363 tandems with 15 genes maximum for an individual repeat. A similar number was found in the var. *burnettii* strain with 1032 genes in 352 tandems with 9 genes maximum for an individual repeat. The tandem repeats include

cytochromes P450, aldo/keto-reductases, hydrophobins, coesterases, methyltransferases, GMC_oxred, Cu-oxidases, Ricin_B_lectin (6-12 genes each). Several domains were found duplicated in only one of the two strains (eg, var. *burnettii*-specific NADH-ubiquinone/plastoquinone oxidoreductases and LAGLIDADG endonucleases). Over 390,000 SNPs were detected between the two strains with 40% in protein coding regions.

We compared the genome structures between strains H97 and JB137 based on their respectively assembled 29 and 2016 scaffolds. Because of the large difference in the number of scaffolds between the two assembled genomes (Fig. S2) and their incomplete correspondence with chromosomes and linkage groups, inferences about inter-chromosomal translocations between the genomes are not possible at present. Instead, our analysis is focused on comparing the unambiguously assembled scaffolds to identify the inversions and translocations between homologous scaffolds of the two strains. In this analysis, we first used each individual scaffold of strain H97 as a query to identify homologous sequences in strain JB137. The corresponding scaffold(s) in strain JB137 that matched each queried scaffold of strain H97 was recorded. Each matching scaffold pair between these two strains was then compared using the "Dot Plot" tab of the "Synteny" function at the JGI website to identify inversions and translocations.

A total of 74 one-to-one matching scaffold block pairs ranging from ~4.5kb to over 2.9 Mbp were identified between the two genomes. Within the 74 pairs of blocks, 12 were completely syntenic with no inversion or translocation, 30 contained both inversions and translocations, 5 contained only inversions, and 27 contained only translocations. The numbers of inversions and translocations between the two sequenced genomes of *A. bisporus* were significantly greater than those found between closely related strains in other species (38). Chromosomal size polymorphisms due to chromosomal rearrangements have also been observed among strains within the same variety, *A. bisporus* var. *bisporus* (39). Such rearrangements likely contribute to the relatively low rate of recombination and low viability of haploid meiotic spores from crosses involving parental strains from either within the same variety (6) or between the two different varieties (4).

2. References

1. Jaffe DB, Butler J, Gnerre1 S, Mauceli E, Lindblad-Toh K et al. (2003) Whole-genome sequence assembly for mammalian genomes: Arachne 2. *Genome Res* 13: 91-96.
2. Zerbino DR and Birney E (2008) Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Res* 18: 821-829.
3. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M et al. (2004) Versatile and open software for comparing large genomes. *Genome Biol* 5: R12.
4. Foulongne-Oriol M, Spataro C, Cathalot V, Monllor S, Savoie JM (2010) An expanded genetic linkage map of an intervarietal *Agaricus bisporus* var. *bisporus* x *A. bisporus* var. *burnettii* hybrid based on AFLP, SSR and CAPS markers sheds light on the recombination behaviour of the species. *Fungal Genet Biol* 47: 226-236.
5. Foulongne-Oriol M, Dufourcq R, Spataro C, Devesse C, Broly A et al. (2011) Comparative linkage mapping in the white button mushroom Agaricus bisporus provides foundation for breeding management. *Curr Genet* 57: 39-50.
6. Kerrigan RW, Royer JC, Baller LM, Kohli Y, Horgen PA et al. (1993) Meiotic Behavior and Linkage Relationships in the Secondarily Homothallic Fungus *Agaricus-Bisporus*. *Genetics* 133: 225-236
7. Sonnenberg AS, de Groot PW, Schaap PJ, Baars JJ, Visser J et al. (1996) Isolation of expressed sequence tags of *Agaricus bisporus* and their assignment to chromosomes. *Appl Environ Microbiol* 62: 4542-4547.
8. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215***:* 403–410.
9. Price AL, Jones NC, Pevzner PA (2005) *De novo* identification of repeat families in large genomes. To appear in Proceedings of the 13 Annual International conference on Intelligent Systems for Molecular Biology (ISMB-05). Detroit, Michigan
10. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O et al. (2005) Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. *Genome Res*. 110: 462–467.
11. Mc Carthy E and Mc Donald JF (2003) LTR_STRUC: a novel search and identification program for LTR retrotransposons**.** *Bioinformatics* 19: 362-367
12. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignments through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680.
13. Ma J and Bennetzen JL (2004) Rapid recent growth and divergence of rice nuclear genomes. *Proc. Natl. Acad. Sci. USA* 101: 12404-12410.
14. Smit AFA, Hubley R, Green P (1996-2010) *RepeatMasker Open-3.0*. http://www.repeatmasker.org
15. Lowe TM and Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*
16. Salamov AA and Solovyev VV (2000) Ab initio gene finding in Drosophila genomic DNA. *Genome Res* 10: 516-522.
17. Isono K, McIninch JD, Borodovsky M (1994) Characteristic features of the nucleotide sequences of yeast mitochondrial ribosomal protein genes as analyzed by computer program GeneMark. *DNA Res*. 1: 263-269.
18. Birney E and Durbin R. 2000. Using GeneWise in the Drosophila annotation experiment. *Genome Res* 10, 547-548.
19. Kent WJ (2002) BLAT--the BLAST-like alignment tool. *Genome Res*
20. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* 10:1-6.
21. Melen K, Krogh A, von Heijne G (2003) Reliability measures for membrane protein topology prediction algorithms. *J. Mol. Biol* 327:735-744.
22. Zdobnov EM and Apweiler R (2001) InterProScan--an integration platform for the signature-recognition methods in InterPro. *Bioinformatics* 17: 847-848.
23. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M et al. (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36: D480–D484
24. Koonin EV, Fedorova ND, Jackson JD, Jacobs AR, Krylov DM et al. (2004) A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol* 5: R7
25. Enright AJ, Van Dongen S, Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30: 1575-1584.
26. De Bie, T., Cristianini N., Demuth J.P., and Hahn M.W. CAFE: a computational tool for the study of gene family evolution. Bioinformatics. 22:1269-1271 (2006).
27. Eddy S.R. A New Generation of Homology Search Tools Based on Probabilistic Inference. Genome Inform. 23:205-211 (2009).
28. Finn R.D., Mistry J., Tate J., Coggill P., Heger A., Pollington J.E., Gavin O.L., Gunesekaran P., Ceric G., Forslund K., Holm L., Sonnhammer E.L., Eddy S.R., Bateman A. The Pfam protein families database. Nucleic Acids Research Database Issue 38:D211-222 (2010).
29. Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300: 1005-1016.
30. Rawlings ND, Barrett AJ, Bateman A (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* 40 : D343-D350.
31. Saier MH, Yen MR, Noto K, Tamang DG, Elkan C (2009) The Transporter Classification Database: recent advances, *Nucleic Acids Res* 37: D274-D278.
32. Huang, X. & Madan, A. CAP3: A DNA Sequence Assembling Program. *Genome Research* 9, 868-877 (1999).
33. Calvo-Bado L, Noble R, Challen M, Dobrovin-Pennington A, Elliott T (2000) Sexuality and Genetic Identity in the Agaricus Section Arvenses. Applied and Environmental Microbiology Vol 66, 728-734
34. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res* 37, D233–D238.
35. De Caceres, M., Oliva, F., Font, X. & Vives, S. (2007) Ginkgo, a program for non-standard multivariate fuzzy analysis. *Advances in Fuzzy Sets and Systems* 2 (1): 41-56.
36. Desper R., Gascuel O. (2002) Fast and accurate phylogeny reconstruction algorithms based on the minimum-evolution principle. *Journal of Computational Biology*. 9(5):687-705.
37. Floudas D, Binder M, Riley R, Barry K, Blanchette RA et al. (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336 (6089):1715-9
38. Sun S, Xu J (2009) Chromosomal rearrangements between serotype A and D strains in *Cryptococcus neoformans*. PLoS ONE 4: e5524.
39. Royer, D. J., W. E. A. Hintz, R. W. Kerrigan, and P. A. Horgen. 1992. Electrophoretic karyotype analysis of the button mushroom, Agaricus bisporus. Genome. 35: 694-698.

40. Murata, Y, Fujii, M, Zolan, ME, Kamada, T (1998). Molecular analysis of pcc1, a gene that leads to A-regulated sexual morphogenesis in Coprinus cinereus. Genetics, 149: 1753–1761.
41. Muraguchi H, Fujita T, Kishibe Y, Konno K, Ueda N, Nakahori K, Yanagi SO, Kamada T (2008) The exp1 gene essential for pileus expansion and autolysis of the inky cap mushroom Coprinopsis cinerea (Coprinus cinereus) encodes an HMG protein. Fung Genet Biol 45:890–896

Supplementary Tables

Table S1. Final assembly statistics for *Agaricus bisporus* genomes

Sequenced strain	***A. bisporus* var. *bisporus***	***A. bisporus* var. *burnettii***
Assembler (Sequencing strategy)	Arachne (Sanger)	Newbler (454)+Velvet (Illumina)
Main genome scaffold total	29	2016
Main genome contig total	254	2527
Main genome scaffold sequence total	30.2 Mb	32.6 Mb
Main genome contig sequence total	30.0 Mb	31.2 Mb
Main genome scaffold N/L50	6/2.3 Mb	8/1.23 Mb
Main genome contig N/L50	35/262.5 Kb	39/0.163 Kb
Number of scaffolds > 50 Kb	19	52
% main genome in scaffolds > 50 Kb	99.7%	87.9%

Table S2. Characteristics of predicted gene models.

Sequenced strain	***A. bisporus* var. *bisporus***	***A. bisporus* var. *burnettii***
Avg. gene length, bp	1764	1669
Avg. protein length, aa	426	410
Avg. exon frequency	6.05 exons/gene	5.72 exons/gene
Avg. exon length, bp	234	236
Avg. intron length, bp	71	69

Table S3. Predicted gene models and supporting lines of evidence

Sequenced strain	***A. bisporus* var. *bisporus***	***A. bisporus* var. *burnettii***
# gene models	10,438	11,289
% complete (with start and stop codons)	90%	87%
% genes with homology support	83%	79%
% genes with PFAM domains	52%	62%
Models with exons covered by ESTs ≥75%	69%	61%

Table S4. Number of orthologs (% amino acid identity) between *Agaricus bisporus* strains and other Agaricomycetes

	A. bisporus* var. *bisporus	***A. bisporus* var. *burnettii***
Agaricus bisporus (non-self)	8649 (96.4%)	8649 (96.4%)
Coprinopsis cinerea	6059 (57%)	6074 (57%)
Laccaria bicolor	5927 (60%)	5942 (60%)
Phanerochaete chrysosporium	5139 (58%)	5134 (58%)
Postia placenta	5090 (58%)	5062 (58%)

Table S5. The top 10 most abundant (MCL) protein families (excluding TE-related families) in *A. bisporus* var. *bisporus* genome as compared to other fungi.

Family ID	PFAM description	AGABI	COPCI	LACBI	SCHCO	PLEOS	SERLA	HETAN	POSPL	PHACH	CRYNE	TREME	MELLA	SPORO	USTMA	MALGO	Total
2	Cytochrome P450	**64**	69	26	49	86	72	87	119	61	0	0	4	0	4	1	642
1	Protein kinase domain	**39**	37	46	44	47	38	39	43	42	30	33	62	34	29	29	592
11	GMC oxidoreductase	**32**	35	11	19	42	9	31	26	32	3	2	6	5	9	5	267
3	WD domain, G-beta repeat	**31**	36	76	35	41	43	45	38	31	22	20	22	21	19	21	501
43	Fungal hydrophobin	**28**	31	16	11	21	17	16	1	12	0	0	0	0	2	0	155
4	Major Facilitator Superfamily	**25**	33	13	40	37	19	33	26	28	39	16	22	18	16	2	367
5	DEAD/DEAH box helicase	**25**	24	27	28	25	25	25	24	23	25	24	25	26	26	24	376
8	Ras family	**23**	19	27	20	19	16	18	21	20	15	18	27	18	18	11	290
39	GTPase of unknown function	**23**	29	47	0	29	0	0	15	10	0	0	0	0	0	0	153
15	Cytochrome P450	**22**	14	26	27	27	26	30	51	17	2	3	5	1	5	2	258

Abbreviations: AGABI, *A. bisporus*; COPCI, *C. cinerea*; LACBI, *Laccaria bicolor;* SCHCO, *S. commune*; PLEOS, *P. ostreatus*; SERLA, *S. lacrymans;* HETAN, *H. annosum*; POSPL, *P. postia*; PHACH, *P.chrysosporium*; CRYNE, *C. neoformans*; TREME, *T. mesenterica*; MELLA, *M. larici-populina*; SPORO, *S. roseus*; USTMA, *U. maydis* and MALGO, *M. globosa*.

Table S6. The top 15 protein families in contraction in *A. bisporus* var. *bisporus* genome as compared to other fungi. The table lists 15 MCL families with the lower size in the *A. bisporus* lineage (CAFE analysis, $P<0.001$). Annotations are based on searches of *A. bisporus* protein sequences against the PFAM database.

Family ID	PFAM description	Nr description	AGABI	COPCI	LACBI	SCHCO	PLEOS	SERLA	HETAN	POSPL	PHACH	CRYNE	TREME	MELLA	SPORO	USTMA	MALGO	Total	p-value
54	NO PFAM	Hypothetical protein	**4**	13	53	19	45	1	5	1	2	0	0	0	0	0	0	143	0.000000
72	NO PFAM	Hypothetical protein	**0**	46	53	5	6	12	0	0	0	0	0	0	0	0	0	122	0.000000
40	NO PFAM	Hypothetical protein	**7**	23	27	40	5	6	9	12	30	0	0	0	0	0	0	159	0.000010
117	Protein kinase domain	other/AgaK1 protein kinase	**0**	5	15	10	4	23	2	29	6	0	0	0	0	0	0	94	0.000012
42	NO PFAM	Hypothetical protein	**5**	12	18	16	21	37	35	10	2	2	0	0	0	0	0	158	0.000042
213	F-box protein	Hypothetical protein	**1**	12	4	34	3	2	3	3	3	0	0	0	0	0	0	65	0.000294
166	NADH:flavin oxidoreductase / NADH oxidase family	NADPH dehydrogenase	**0**	2	6	17	9	5	9	15	2	3	1	0	1	0	0	70	0.000406
432	NO PFAM	Hypothetical protein	**0**	3	19	8	1	5	4	0	4	0	0	0	0	0	0	44	0.000406
597	NO PFAM	Hypothetical protein	**0**	2	11	11	1	11	0	1	0	0	0	0	0	0	0	37	0.000406
139	NO PFAM	Hypothetical protein	**2**	1	34	11	7	9	3	6	13	0	0	0	1	0	0	87	0.000627
56	tetratricopeptide repeat	Hypothetical protein	**5**	41	35	1	4	10	30	1	5	1	1	0	1	1	1	137	0.000898
122	F-box protein	Hypothetical protein	**1**	29	10	1	5	1	28	4	13	0	0	0	0	0	0	92	0.001161
61	Short-chain dehydrogenase	Retinol dehydrogenase	**3**	9	11	11	10	15	6	18	6	2	4	5	3	2	2	107	0.001173
103	NO PFAM	Predicted protein	**0**	1	67	0	36	0	0	0	0	0	0	0	0	0	0	104	0.004124
198	NO PFAM	Predicted protein	**0**	4	55	7	0	1	0	0	0	0	0	0	0	0	0	67	0.004124

Abbreviations: AGABI, *A. bisporus*; COPCI, *C. cinerea*; LACBI, *Laccaria bicolor;* SCHCO, *S. commune*; PLEOS, *P. ostreatus*; SERLA, *S. lacrymans;* HETAN, *H. annosum*; POSPL, *P. postia*; PHACH, *P.chrysosporium*; CRYNE, *C. neoformans*; TREME, *T. mesenterica*; MELLA, *M. larici-populina*; SPORO, *S. roseus*; USTMA, *U. maydis* and MALGO, *M. globosa*.

Table S7. The top 15 protein families in expansion in *A. bisporus* var. *bisporus* genome as compared to other fungi. The table lists MCL families that are in expansion in the *A. bisporus* lineage (CAFE analysis, $P<0.001$). Annotations are based on searches of *A. bisporus* protein sequences against the PFAM database.

Family ID	PFAM description	Nr description	AGABI	COPCI	LACBI	SCHCO	PLEOS	SERLA	HETAN	POSPL	PHACH	CRYNE	TREME	MELLA	SPORO	USTMA	MALGO	Total	p-value
26	NO PFAM	NWD2 protein	**169**	23	3	0	7	0	1	3	1	0	0	0	0	0	0	207	0.000000
32	NO PFAM	Hypothetical protein	**20**	3	0	0	3	149	3	0	4	0	0	0	0	0	0	182	0.000000
402	Patatin	Kinesin light chain	**18**	1	3	7	2	0	0	0	0	0	0	1	0	0	0	32	0.000000
546	NO PFAM	Predicted protein	**38**	0	0	0	0	1	0	0	0	0	0	0	0	0	0	39	0.000000
688	Peroxidase, family 2	Aromatic peroxygenase	**26**	5	1	0	0	0	0	0	0	0	0	0	0	0	0	32	0.000000
783	NO PFAM	Hypothetical protein	**29**	1	0	0	0	1	0	0	0	0	0	0	0	0	0	31	0.000000
964	NO PFAM	Hypothetical protein	**17**	4	1	1	0	0	0	4	0	0	0	0	0	0	0	27	0.000000
458	NO PFAM	Hypothetical protein	**10**	1	2	1	4	20	1	2	2	0	0	0	0	0	0	43	0.000010
668	NO PFAM	MutS protein homolog	**9**	5	1	0	1	5	0	9	4	0	0	0	0	0	0	34	0.000057
27	Aldehyde dehydrogenase family	Methylmalonate-semialdehyde dehydrogenase	**21**	9	11	10	12	11	10	14	14	11	7	8	9	14	4	165	0.000126
226	NO PFAM	Hypothetical protein	**11**	5	5	9	6	3	4	3	15	0	0	0	0	1	1	63	0.002189
43	Hydrophobin	Hydrophobin	**28**	31	16	11	21	17	16	1	12	0	0	0	0	2	0	155	0.010667
770	3'-5' exonuclease	Putative exonuclease	**10**	0	20	0	1	0	0	0	0	0	0	0	0	0	0	31	0.010830
118	Glutathione S-transferase	Putative Beta-etherase	**16**	13	3	9	7	22	8	9	5	0	0	0	0	0	0	92	0.014126
11	GMC oxidoreductase family	Pyranose dehydrogenase	**32**	35	11	19	42	9	31	26	32	3	2	6	5	9	5	267	0.018344

Abbreviations: AGABI, *A. bisporus*; COPCI, *C. cinerea*; LACBI, *Laccaria bicolor;* SCHCO, *S. commune*; PLEOS, *P. ostreatus*; SERLA, *S. lacrymans;* HETAN, *H. annosum*; POSPL, *P. postia*; PHACH, *P.chrysosporium*; CRYNE, *C. neoformans*; TREME, *T. mesenterica*; MELLA, *M. larici-populina*; SPORO, *S. roseus*; USTMA, *U. maydis* and MALGO, *M. globosa*.

Table S8. The 45 most highly up-regulated transcripts of *A. bisporus* var. *bisporus* growing on compost by comparison to defined agar culture medium. Up-regulation ratios are also showed for casing and fruit body. Non-significant ratios (p-value FDR, modified t test >0.05) are marked in grey.

Protein ID	ratio Compost/Culture	ratio Casing/Culture	ratio Fruit Body/Culture	Description
193563	1492.1	299.6	15.8	Chloroperoxidase
190202	1439.7	5.5	0.2	Cutinase
194280	1382.0	14.6	5.3	Glycoside hydrolase, family 12
227728	1335.5	0.9	1.4	Glycoside hydrolase, family 12
190200	1235.9	0.6	0.8	Cutinase
194662	864.3	5.5	1.2	Hypothetical protein
146117	820.9	7.2	2.2	Glycosyl hydrolase, family 88
190390	672.0	2.1	3.5	Glycoside hydrolase, family 6|1, 4-beta cellobiohydrolase
196181	664.7	0.7	1.8	Glycoside hydrolase, family 11
226235	641.2	2.6	1.8	Glycoside hydrolase, family 61
183005	616.9	0.7	0.3	Cutinase
122979	578.8	27.0	0.4	Hypothetical protein
195916	536.7	7.9	0.6	Glucanase
183646	523.6	0.8	1.2	Glucose-methanol-choline oxidoreductase
195861	489.2	1.6	0.4	Sugar transporter
194696	482.5	8.9	2.9	Glycoside hydrolase, family 61
194576	461.7	1.3	1.1	Alpha-L-arabinofuranosidase
188060	270.0	2.9	1.9	Glycoside hydrolase, family 20
192849	263.2	1.4	1.4	Glycoside hydrolase
211295	239.1	6.9	0.1	Glycoside hydrolase family 13
72717	235.5	2.3	1.5	Rhamnogalacturonase B
192976	223.2	7.8	4.3	Hypothetical protein
122757	204.8	1.1	0.8	Glycoside hydrolase, family 61
123457	188.3	14.2	0.4	Hypothetical protein
194618	178.9	1.5	2.8	Glucanase
133541	178.7	1.5	1.6	Glycoside hydrolase, family 10
212650	158.4	0.4	0.7	Peptidase S8, subtilisin
70106	142.8	8.1	10.2	Glycoside hydrolase, family 27
68978	140.5	1.8	1.1	Peptidase S10, serine carboxypeptidase
180108	137.1	2.6	4.8	Beta-lactamase
77253	136.2	4693.2	8588.5	Fungal hydrophobin
79647	134.0	2.1	1.2	Glycoside hydrolase, family 11
194521	127.9	2.2	1.2	Glycoside hydrolase, family 7
180408	122.1	2.3	1.4	Protein DUF1680
209748	122.1	2.3	2.0	Cellulose-binding region
181624	121.4	71.9	31.8	Hypothetical protein
139877	118.7	1.8	2.3	Pectate lyase
149407	114.1	1.2	0.3	Hypothetical protein
194321	113.5	1.5	4.1	Cutinase
175634	110.7	1.5	0.8	Hypothetical protein
207786	109.3	0.9	1.8	Chloroperoxidase
139148	108.1	1.7	0.8	Multicopper oxidase|Cupredoxin
194744	106.0	1.1	6.3	Glycoside hydrolase, family 1
139552	105.1	43.1	2.5	Alcohol dehydrogenase
194648	104.6	1.9	0.1	Peptidase S8, subtilisin

Table S9. Up-regulation of transcripts coding for carbohydrate-active enzymes (CAZymes) potentially involved in degradation of polysaccharides in *A. bisporus* var. *bisporus* mycelium grown on agar medium (Culture), casing layer (Casing), compost or in fruiting body. Genes showing an up-regulation >30-fold on compost are presented. Non-significant ratios (p-value FDR, modified t test >0.05) are marked in grey.

CAZy module(s)	Protein_ID	Culture	Casing	Compost	Fruit Body	ratio Casing/Culture	*P value*	ratio Compost/Culture	*P value*	ratio Fruit body/Culture	*P value*
CE5	190202	44	241	63346	9	5.48	*0.2420*	1447.25	*0.0002*	0.20	*0.0119*
GH12	194280	14	204	19348	74	14.57	*0.0019*	1370.85	*0.0013*	5.21	*0.0199*
GH12	227728	35	30	46743	50	0.86	*0.5800*	1343.04	*0.0001*	1.44	*0.2170*
CE5	190200	8	5	9887	6	0.63	*0.1550*	1277.22	*0.0025*	0.81	*0.5500*
GH105	146117	18	129	14776	39	7.17	*0.0026*	836.16	*0.0001*	2.22	*0.0892*
CBM1-GH6	190390	51	109	34273	176	2.14	*0.0133*	672.22	*0.0002*	3.46	*0.0006*
GH11	196181	84	58	55831	151	0.69	*0.2860*	668.23	*0.0010*	1.80	*0.0177*
GH61-CBM1	226235	5	13	3206	9	2.60	*0.0482*	631.88	*0.0257*	1.82	*0.0267*
CE5	183005	126	93	77730	35	0.74	*0.6760*	617.99	*0.0006*	0.28	*0.1300*
GH16	195916	175	1386	93918	113	7.92	*0.1220*	536.28	*0.0003*	0.64	*0.2040*
GH61	194696	28	248	13510	81	8.86	*0.0009*	482.33	*0.0270*	2.88	*0.0064*
GH51	194576	61	77	28165	70	1.26	*0.6020*	464.34	*0.0002*	1.15	*0.5870*
GH20	188060	8	23	2160	15	2.88	*0.0062*	276.83	*0.0005*	1.98	*0.0125*
GH5	192849	131	188	34480	179	1.44	*0.0289*	262.61	*0.0002*	1.37	*0.0286*
GH13	211295	91	630	21755	11	6.92	*0.1210*	238.07	*0.0002*	0.12	*0.0023*
PL4	72717	169	388	39793	251	2.30	*0.0181*	235.71	*0.0002*	1.49	*0.0734*
GH61	122757	45	51	9215	35	1.13	*0.6320*	202.84	*0.0273*	0.76	*0.3690*
GH16	194618	246	358	44006	698	1.46	*0.1010*	178.85	*0.0003*	2.84	*0.0030*
CBM1-GH10	133541	490	751	87557	782	1.53	*0.0245*	178.74	*0.0004*	1.60	*0.0442*
GH27	70106	419	3396	59817	4281	8.11	*0.0000*	142.61	*0.0001*	10.21	*0.0766*
GH11-CBM1-UNK	79647	140	290	18762	170	2.07	*0.0295*	133.98	*0.0570*	1.21	*0.0742*
GH7-CBM1	194521	1681	3694	214989	1981	2.20	*0.0026*	127.90	*0.0006*	1.18	*0.1930*
CBM1-CE15	209748	121	280	14774	244	2.31	*0.0009*	122.10	*0.0004*	2.01	*0.0058*
PL1	139877	995	1835	118069	2323	1.84	*0.0206*	118.69	*0.0003*	2.34	*0.0252*
CE5	194321	127	194	14419	516	1.53	*0.0529*	113.85	*0.0002*	4.07	*0.0001*
GH1	194744	98	108	10389	613	1.10	*0.7230*	106.39	*0.0005*	6.28	*0.0010*
CBM1-GH5	191420	267	365	25908	476	1.37	*0.1030*	96.93	*0.0002*	1.78	*0.0036*
GH78	210102	194	387	17845	275	1.99	*0.0233*	92.04	*0.0002*	1.42	*0.0276*
GH28	194940	151	1032	13906	150	6.83	*0.0022*	92.01	*0.0003*	0.99	*0.9720*

CE16	179406	127	147	10087	96	1.16	*0.6270*	79.68	*0.0012*	0.76	*0.1210*
GH3	219902	383	1395	29848	1869	3.64	*0.0119*	77.91	*0.0005*	4.88	*0.0005*
GH61	137828	109	60	8494	442	0.55	*0.0455*	77.58	*0.0671*	4.03	*0.0044*
CE8	151878	16	103	1210	77	6.44	*0.0004*	76.86	*0.0004*	4.87	*0.0008*
GH43	224152	613	763	46615	1031	1.24	*0.0662*	76.06	*0.0001*	1.68	*0.0314*
CE1-CBM1	196213	250	353	17166	843	1.41	*0.1430*	68.79	*0.0010*	3.38	*0.0629*
CBM1	194858	30	58	2089	107	1.93	*0.0821*	68.77	*0.0324*	3.51	*0.0086*
CBM1-GH5	189329	366	357	22438	413	0.98	*0.7650*	61.26	*0.0024*	1.13	*0.1320*
GH43	208425	217	1520	13105	300	7.00	*0.0276*	60.40	*0.0002*	1.39	*0.1330*
GH55	183321	1384	5850	83158	1575	4.23	*0.0521*	60.08	*0.0004*	1.14	*0.3820*
CBM1-GH5	194478	270	354	15454	390	1.31	*0.1950*	57.33	*0.0021*	1.45	*0.0499*
PL3	195581	1027	1155	57520	1355	1.12	*0.5250*	56.02	*0.0004*	1.32	*0.2590*
GH10	191440	363	1029	18571	545	2.83	*0.0198*	51.20	*0.0004*	1.50	*0.0527*
GH74-CBM1	214617	375	620	18092	506	1.65	*0.0849*	48.30	*0.0011*	1.35	*0.0496*
GH5	229390	1124	4694	50706	971	4.18	*0.0001*	45.13	*0.0002*	0.86	*0.3630*
GH61	136180	28	64	1222	6238	2.29	*0.1840*	43.42	*0.0282*	221.63	*0.0000*
GH25	120008	23	5	963	9	0.22	*0.0140*	41.71	*0.0092*	0.40	*0.1380*
GH29	209111	1087	2385	43080	836	2.19	*0.0013*	39.62	*0.0001*	0.77	*0.0338*
CBM1-GH5	192609	678	583	26379	440	0.86	*0.1180*	38.89	*0.0011*	0.65	*0.0253*
CBM1-CE16	192194	232	231	8912	714	1.00	*0.9760*	38.35	*0.0072*	3.07	*0.0001*
GH61-CBM1	192993	147	579	5547	109	3.94	*0.0002*	37.63	*0.0261*	0.74	*0.1940*
GH2	202715	86	233	3037	371	2.71	*0.0012*	35.17	*0.0001*	4.30	*0.0002*
GH115	121649	378	516	12627	904	1.37	*0.4990*	33.43	*0.0004*	2.39	*0.0022*
GH2	188899	312	1121	10087	435	3.59	*0.0252*	32.35	*0.0001*	1.40	*0.2890*
CE5-CBM1	136707	262	297	8054	534	1.13	*0.7390*	30.74	*0.0092*	2.04	*0.0026*

Table S10. Gene contents in wood decay proteins in the genomes of 20 Agaricomycotina.
(11 oxidoreductase and 17 CAZyme families)

Taxonomy		Agaricomycotina																			
		Agaricomycetes																			Dac
		Ab	Abb	Ha	Sh	Ps	Fm	Ad	Tv	Ds	Pc	Pp	Wc	Fp	Gt	Sl	Cp	Lb	Sc	Cc	Da
Ecology		S		WR								BR						ECM	WR	S	BR
												CAZymes									
Genes	***P***																				
GH3	0.000	8	8	11	15	12	7	12	11	7	9	5	7	12	9	9	12	2	11	7	8
GH5	0.003	**19**	**19**	7	6	6	6	8	5	5	5	5	4	5	5	8	8	3	3	6	5
GH6	0.207	1	1	1	1	1	2	2	1	1	1	0	0	0	0	1	2	0	1	5	0
GH7	0.000	1	1	1	3	5	2	6	4	4	8	0	0	0	0	0	2	0	2	6	0
GH10	0.000	2	2	2	6	5	4	4	6	5	6	4	4	2	3	1	3	0	5	6	3
GH11	0.043	2	2	0	1	1	0	3	0	0	1	0	0	0	0	0	0	0	1	6	0
GH12	0.236	2	2	4	5	2	3	1	5	3	2	2	2	2	2	1	4	3	1	1	1
GH28	0.000	6	6	8	17	13	16	10	11	7	4	7	9	13	10	7	13	7	3	3	6
GH61	0.000	11	11	10	16	14	13	19	18	15	15	2	2	4	4	5	10	5	22	35	0
GH74	0.539	1	1	1	2	2	4	1	1	1	4	0	0	0	1	1	0	0	1	1	0
GH43	0.000	4	4	4	10	7	6	26	3	7	4	1	1	7	5	1	6	0	12	4	5
CE1	0.015	2	2	1	1	2	0	3	3	0	4	0	0	0	1	0	0	0	4	3	0
CE16	0.000	10	10	5	10	8	6	29	7	10	2	5	6	11	6	3	6	3	10	5	4
CE5	0.035	6	6	0	1	1	0	3	0	0	0	0	0	0	0	0	1	1	2	6	0
CE8	0.084	2	2	3	4	6	3	3	2	3	2	1	1	2	2	2	2	4	2	0	3
CE12	0.273	3	2	2	3	0	2	1	0	2	0	0	0	0	0	0	0	0	1	1	0
CE15	0.011	0	0	1	1	2	1	6	2	2	2	1	1	1	1	0	0	0	2	8	1
												Oxidoreductases									
POD	0.000	2	2	8	6	11	17	19	26	12	16	1	1	1	0	0	0	1	0	1	0
MCO	0.000	12	12	17	20	13	11	10	10	13	5	5	5	7	4	6	8	11	6	17	5
CRO	0.000	9	9	5	8	9	4	9	9	9	7	3	4	4	2	3	6	11	2	6	3
CDH	0.791	1	1	1	1	1	1	1	1	1	1	0	0	0	1	2	2	0	1	1	0
Cytb562	0.922	0	0	1	1	0	0	0	1	1	1	0	0	0	0	2	3	0	2	0	0
ODC	0.754	2	2	3	3	2	3	3	5	5	7	5	4	5	4	3	2	1	5	1	2
GLP	0.000	2	3	1	11	6	8	1	2	6	3	5	10	10	6	3	11	1	7	0	3
QRD	0.885	4	4	2	1	3	3	4	1	1	4	1	1	1	3	2	2	2	4	3	1
DyP	0.000	0	0	1	2	5	3	11	2	1	0	2	0	0	0	0	0	2	0	4	0
HTP	0.000	**24**	**24**	5	10	8	4	16	3	4	3	5	5	4	6	3	2	5	3	8	6
P450	0.000	109	104	144	215	144	130	249	190	187	149	250	206	190	130	164	238	101	115	139	126

Species : Ab, *A. bisporus* var bisporus; Abb, *A.bisporus* var burnettii; Ad, *Auricularia delicata*; Cc, *C. cinerea*; Cp, *C. puteana*; Da, *Dacryopinax* sp.; Ds, *D. squalens;* Fm, *F. mediterranea*; Fp, *F. pinicola*; Gt, *G. trabeum;* Ha, *H. annosum*; Lb, *L. bicolor*; Pc, *P. chrysosporium*; Pp, *P. placenta*; Ps*, P. strigosozonata*; Sc, *S. commune*; Sh, *S. hirsutum*; Sl, S. *lacrymans*; Tv, *T. versicolor*; Wc, *W. cocos*.
Genes : GH, glycoside hydrolases; CE, carbohydrate esterases; POD, class II peroxidases; MCO, multicopper oxidases; CRO, copper radical oxidases; CDH, cellobiose dehydrogenase; Cytb562, cytochrome b562; ODC, oxalate decarboxylases; GLP, Fe(III)-reducing glycopeptides; QRD, quinone reductases; DyP, dye-decolorizing peroxidases; HTP, heme-thiolate peroxidases; P450, cytochromes P450. P values indicate strength of rejection of model of random diversification in CAFE analyses.
Ecologies : WR, white rot; BR, brown rot; ECM, mycorrhiza; S, non-wood decay saprotroph.

Table S11. Comparison of oxidoreductase encoding genes in *A. bisporus* var. *bisporus* and other Agaricomycotina.

Species[1] Ecology[2]	**Ab** HU	**Abb** HU	**Cc** CP	**Ha** WR	**Pc** WR	**Ad** WR	**Pst** WR	**Fm** WR	**Ds** WR	**Tv** WR	**Sh** WR	**Sc** WR	**Sl** BR	**Pp**[4] BR	**Cp** BR	**Gt** BR	**Fp** BR	**Wc** BR	**Da** BR	**Lb** ECM
Gene[3]																				
LiP	0	0	0	0	**10**	0	0	0	0	**10**	0	0	0	0	0	0	0	0	0	0
MnP	2	2	0	8	5	5	10	**16**	9	13	5	0	0	0	0	0	0	0	0	0
VP	0	0	0	0	0	0	0	0	**3**	0	0	0	0	0	0	0	0	0	0	0
DyP	0	0	4	1	0	**11**	5	3	1	2	2	0	0	2	0	0	0	0	0	2
HTP	**24**	**24**	8	5	3	16	8	4	4	3	10	3	3	5	2	6	4	5	6	5
GLX	3	3	0	0	1	2	3	0	**5**	**5**	3	0	0	0	0	0	0	0	0	0
CRO1	2	2	2	1	1	0	0	1	1	1	0	0	0	0	1	0	1	1	2	**6**
CRO2	2	2	2	2	1	**4**	3	1	1	1	1	1	1	1	**4**	1	1	1	0	2
CRO3-5	1	1	0	1	**3**	1	1	1	1	1	1	0	1	1	0	0	1	1	0	2
CRO6	1	1	1	1	1	**2**	**2**	1	1	1	1	1	1	1	1	1	1	1	1	0
MCO	12	12	**17**	13	0	7	12	10	12	7	15	2	4	3	6	4	5	5	0	9
ODC	2	2	1	3	**7**	3	3	3	5	6	3	5	3	5	2	4	5	4	2	1
CDH	1	1	1	1	1	1	1	1	1	1	1	1	**2**	0	1	1	0	0	0	0
P450	109	ND	139	144	152	**249**	144	130	187	190	215	115	164	254	238	130	190	206	126	101

[1]Species designations: Ab and Abb, *A. bisporus* strains; Ha, *H. annosum*; Pc, *P. chrysosporium*; Ad, *Auricularia delicata*; Pst*, P. strigosozonata*; Fm, *F. mediterranea*; DS, *D. squalens*; Tv, *T. versicolor*; Sh, *S. hirsutum*; Sc, *S. commune*; Sl, S. *lacrymans*; Pp, *P. placenta*; Cp, *C. puteana*; Gt, *G. trabeum*, Fp, *F. pinicola*; Wc, *W. cocos*; Da, *Dacryopinax* sp.; Cc, *C. cinerea*; Lb, *L. bicolor*. [2]Ecology: CP, coprophilic; WR, White-rot; BR, Brown-rot; HU, Humicolous; ECM, mycorrhiza. [3]Gene: LiP, lignin peroxidase; MnP, Manganese peroxidase; VP, Versatile peroxidase; DyP; Dye degrading peroxidase; HTP, Heme thiol peroxidases; GLX, copper radical oxidase glyoxal oxidase; CRO, Copper radical oxidase nomenclature based on *P. chrysosporium*; MCO, Multicopper Oxidase; ODC, oxalate decarboxylase; CDH, Cellobiose dehydrogenase; P450, cytochrome P450s. [4]*P.placenta* estimate of CYPs is inflated by allelism.

Table S12. Up-regulation of transcripts coding for oxidoreductases potentially involved in degradation of lignin and related compounds in *A. bisporus* mycelium grown on agar medium (culture) or compost. Genes significantly up-regulated >10-fold with a p-value FDR, modified t test <0.05 are shown.

	Average array signal[2]		Ratio[3]	P value[4]			SS[5]
Prot_ID	Compost	Culture	Compost/Culture	Compost/Culture	Broad category[5]	Putative function	Prob
193563	20890	14	1492	*0.0009*	HTP	Aromatic peroxygenase	1
183646[6]	2618	5	523	*0.0021*	GMC OR	Model #141436 preferred	
207786	1421	13	109	*0.0006*	HTP	Aromatic peroxygenase	-
139148	39882	369	108	*0.0173*	MCO	Laccase *sensu stricto*	0.999
146228	36073	381	94	*0.0152*	MCO	Laccase *sensu stricto*	0.999
195432	9776	108	90	*0.0009*	HTP	Aromatic peroxygenase	1
195436	6838	116	58	*0.0011*	HTP	Aromatic peroxygenase	1
211194	3441	64	53	*0.0003*	GMC OR		0.901
193903	18844	395	47	*0.0034*	CRO	GLX-like	0.917
195553	39479	944	41	*0.0020*	GMC OR	Methanol oxidase (probable)	-
217653	10027	276	36	*0.0020*	GMC OR		0.985
181085	8335	236	35	*0.0003*	CRO	CRO3-like	1
226793	11229	321	34	*0.0053*	HTP	Aromatic peroxygenase	0.999
188295	6150	216	28	*0.0132*	GMC OR		0.999
186833	186	7	26	*0.0011*	HTP	Aromatic peroxygenase	0.997
134420	956	40	23	*0.0100*	HTP	Aromatic peroxygenase	1
188178	7034	313	22	*0.0431*	CDH		0.997
179427	3133	149	21	*0.0130*	CRO	GLX-like	0.996
193900	12246	588	20	*0.0099*	CRO	GLX-like	1
186880	12452	709	17	*0.0007*	CYP P450	CYP64	0.84
78998	90995	5251	17	*0.0005*	CYP P450	CYP64	-
186683	7346	430	17	*0.0006*	CYP P450	CYP Not assigned	-
195996	5485	333	16	*0.0006*	CYP P450	CYP64	-
65123	222	14	15	*0.0354*	HTP	Aromatic peroxygenase	-
139842	1311	84	15	*0.0065*	HTP	Aromatic peroxygenase	1
190967	5903	385	15	*0.0272*	HTP	Aromatic peroxygenase	0.875
52638	182	12	15	*0.0021*	HTP CPO	Chloroperoxidase	-
190973	1248	86	14	*0.0606*	HTP	Aromatic peroxygenase	-
194024	23286	1650	14	*0.0002*	CYP P450	CYP64	-
209229	1264	101	12	*0.0039*	MCO	Laccase sensu stricto	1
190985	1978	160	12	*0.0423*	HTP	Aromatic peroxygenase	1
78318	21052	1703	12	*0.0004*	CYP P450	CYP64	0.927
190943	2683	222	12	*0.0184*	HTP	Aromatic peroxygenase	-
64972	751	65	11	*0.0434*	HTP	Aromatic peroxygenase	-

[1] Ranked according to those oxidoreductases with highest array signals on compost relative to culture. Listing includes only those 36 genes exhibiting >10-fold ratios and previously implicated in oxidative degradation of lignin and related aromatic compounds. [2] mean expression values from four biological replicates [3] fold change in transcript levels calculated from the mean expression value of culture and compost samples. [4] p-value FDR, modified t test. [5] Abbreviations: HTP, Heme-thiol peroxidase; GMC OR, Glucose-methanol-choline. [6] 5'-editing needed. Alternative model #14143 is preferred and features a predicted secretion signal.

Table S13. Comparison of genes encoding secreted proteases in *A. bisporus* (*light grey*) and other Agaricomycotina (coprophilic, *dark grey*; brown rots, *brown*; white rots, *white*; ectomycorrhizal symbiont, *green*)

MEROPS ID	Ab	Abb	Cc	Cp	Fp	Wc	Gt	Sl	Dsp	Ad	Tv	Ds	Sh	Ps	Sc	Fm	Pc	Lb
S09X	32	36	51	35	21	20	24	22	22	42	32	31	37	34	38	24	13	26
A01A	11	12	9	33	23	24	21	18	8	20	26	25	22	24	16	16	19	19
S08A	10	11	11	16	15	19	8	13	14	55	27	19	21	9	11	13	13	10
S10	9	8	3	6	8	8	9	9	9	16	19	14	10	13	3	8	7	4
M38	6	3	1	1	2	1	2	0	0	6	1	1	1	0	1	0	0	1
S33	5	6	4	4	7	11	3	5	5	39	5	6	7	12	7	8	6	2
C14B	3	2	0	1	1	1	1	1	1	1	2	0	1	1	2	1	2	3
C26	3	4	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1
C46	3	2	3	0	0	0	1	1	0	0	2	2	1	2	0	1	0	2
M35	3	3	0	0	0	0	2	0	0	4	2	2	0	2	4	0	1	1
M36	3	3	7	3	0	1	2	0	1	8	3	5	3	3	3	2	0	6
C56	2	1	4	1	1	0	1	2	1	5	0	1	2	2	3	5	2	2
M12A	2	2	3	2	4	3	2	3	2	2	3	3	4	3	2	4	2	3
M28X	2	2	1	1	1	1	1	1	1	0	1	0	1	2	1	0	0	1
M43B	2	2	21	0	0	0	0	0	0	5	2	1	1	1	0	0	0	3
S28	2	3	2	5	4	3	2	3	1	1	3	4	2	2	4	2	2	2
S53	2	2	2	16	21	14	9	7	8	13	22	21	14	10	4	5	6	6
C13	1	1	0	1	1	0	1	1	0	1	1	0	1	0	1	0	0	0
M14A	1	1	1	0	0	1	1	1	0	0	1	0	0	0	2	0	0	1
M18	1	1	1	5	2	0	4	1	0	5	1	0	1	1	2	0	1	2
M19	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
M28A	1	1	8	1	2	1	1	1	1	5	1	0	2	3	0	1	0	1
M28E	1	2	4	4	1	1	2	0	1	7	4	3	0	0	3	0	0	1
M76	1	1	1	0	0	1	0	2	0	1	0	0	0	1	1	1	0	0
S01C	1	1	1	0	1	1	1	0	1	0	0	0	0	0	0	0	1	1
S12	1	1	0	0	1	3	1	0	1	2	2	3	1	1	0	0	0	0
S16	1	1	1	1	1	1	1	0	0	0	1	1	0	0	1	0	0	0
S26B	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Species designations: Ab, *A. bisporus*; Abb, *A. bisporus burnettii*; Cc, *Coprinopsis cinerea*; Cp, *Coniophora puteana*; Fp, *Fomitopsis pinicola*; Wc, *Wolfiporia cocos*; Gt, *Gloeophyllum trabeum*; Sl, S*erpula lacrymans*; Dsp, *Dacryopinax* sp.; Ad, *Auricularia delicata*; Tv, *Trametes versicolor*; Ds, *Dichomitus squalens*; Sh, *Stereum hirsutum*; Ps*, Punctularia strigosozonata*; Sc, *Schizophylum commune*; Fm, *Fomitoporia mediterranea*; Pc, *Phanerochaete chrysosporium*; Lb, *Laccaria bicolor*.

Table S14. Up-regulation of transcripts coding for proteases potentially involved in degradation of compost proteins in *A. bisporus* var. *bisporus* mycelium grown on agar medium (Culture), casing layer (Casing), compost or in fruiting body. Genes showing an up-regulation >10-fold on compost are presented. Protease sequences were identified by using the MEROPS database. Non-significant values (p-value FDR, modified t test >0.05) are marked in grey.

		Expression Levels				*Expression ratios*					
SEQ_ID	**MEROPS family**	**Casing**	**Compost**	**Culture**	**Fruiting Body**	**Casing/ Culture**	***P value***	**Compost/ Culture**	***P value***	**Fruiting Body/ Culture**	***P value***
122979	**C56**	1510	32411	56	24	0.047	*0.09*	573.903	*0.000101*	0.418	*0.286*
133541	**S09X**	751	87557	490	782	0.009	*0.02*	178.739	*0.000383*	1.597	*0.0442*
212650	**S08A**	37	15996	101	66	0.002	*0.06*	158.101	*0.000138*	0.657	*0.0584*
68978	**S10**	1380	108057	769	880	0.013	*0.06*	140.572	*0.00008*	1.144	*0.494*
180108	**S12**	45	2331	17	81	0.02	*0.02*	137.523	*0.000335*	4.78	*0.0288*
209748	**S09X**	280	14774	121	244	0.019	*0.00*	122.098	*0.000352*	2.013	*0.00579*
194648	**S08A**	1255	68168	652	76	0.018	*0.56*	104.584	*0.000306*	0.116	*0.000615*
134948	**M43B**	10275	90941	994	386	0.113	*0.09*	91.478	*0.000087*	0.389	*0.0211*
196213	**S09X**	353	17166	250	843	0.021	*0.14*	68.785	*0.00104*	3.379	*0.0629*
188444	**M28X**	3312	15299	229	2362	0.217	*0.00*	66.824	*0.000618*	10.318	*0.00392*
183321	**S08A**	5850	83158	1384	1575	0.07	*0.05*	60.077	*0.000361*	1.137	*0.382*
191440	**S09X**	1029	18571	363	545	0.055	*0.02*	51.199	*0.000421*	1.503	*0.0527*
213137	**S10**	1061	27540	583	697	0.039	*0.26*	47.227	*0.000263*	1.195	*0.514*
196269	**S09X**	136	5682	130	8	0.024	*0.93*	43.737	*0.000737*	0.059	*0.0000136*
211605	**S33**	146	26561	630	41	0.006	*0.00*	42.185	*0.0144*	0.066	*0.000683*
192194	**S33**	231	8912	232	714	0.026	*0.98*	38.35	*0.00715*	3.072	*0.000127*
195776	**S28**	8539	21253	642	820	0.402	*0.02*	33.128	*0.000096*	1.278	*0.161*
120846	**S08A**	110	743	24	84	0.148	*0.00*	31.314	*0.00134*	3.554	*0.00267*
191068	**S09X**	860	13566	445	347	0.063	*0.12*	30.466	*0.00843*	0.779	*0.192*
181785	**S10**	180	1643	55	209	0.11	*0.00*	29.993	*0.0632*	3.824	*0.00213*
193513	**I51**	895	13829	473	1063	0.065	*0.04*	29.222	*0.0258*	2.246	*0.00804*
222358	**S28**	800	8249	293	16	0.097	*0.04*	28.181	*0.000929*	0.054	*0.00823*
74932	**M36**	1244	116532	4564	128	0.011	*0.20*	25.532	*0.000395*	0.028	*0.0000953*
179502	**S09X**	432	4448	194	138	0.097	*0.00*	22.908	*0.00139*	0.709	*0.0903*
226558	**S09X**	620	12651	570	427	0.049	*0.64*	22.178	*0.00089*	0.748	*0.0762*
115977	**S09X**	73	79	4	49	0.923	*0.00*	17.869	*0.00613*	10.965	*0.0276*
200866	**S09X**	928	13725	830	1415	0.068	*0.51*	16.533	*0.00085*	1.705	*0.0493*
194743	**S09X**	966	4725	306	183	0.204	*0.01*	15.416	*0.0000771*	0.596	*0.0228*
120844	**S08A**	773	4573	299	530	0.169	*0.00*	15.271	*0.00188*	1.771	*0.0566*
212363	**A01A**	31	3008	200	64	0.01	*0.04*	15.041	*0.0366*	0.322	*0.0238*
194099	**S33**	158	22856	1577	192	0.007	*0.00*	14.492	*0.000652*	0.121	*0.000981*
116475	**S33**	3672	15701	1106	1389	0.234	*0.04*	14.194	*0.000315*	1.256	*0.474*
118415	**M17**	189	1512	108	75	0.125	*0.03*	13.995	*0.00172*	0.692	*0.462*
222775	**M28E**	8665	68315	4942	4969	0.127	*0.59*	13.822	*0.000147*	1.005	*0.996*
188452	**S33**	364	658	63	167	0.553	*0.01*	10.469	*0.00695*	2.658	*0.147*
189957	**M01**	1420	13443	1288	1609	0.106	*0.65*	10.436	*0.000387*	1.249	*0.387*
142735	**A01A**	14725	7083	683	126	2.079	*0.00*	10.364	*0.000241*	0.184	*0.00704*
115910	**S09X**	1336	5023	490	1306	0.266	*0.00*	10.261	*0.00452*	2.667	*0.000291*
179046	**S09X**	1728	10382	1021	1709	0.166	*0.01*	10.169	*0.000553*	1.674	*0.215*
223428	**S09X**	1822	11559	1139	1933	0.158	*0.03*	10.145	*0.000444*	1.696	*0.229*

Table S15. A. Gene regulation by humic substances in *Agaricus bisporus*.

Rank	JGI_ID	Ratio humic/non-humic	Motif number	motif lowest *p*-value	Protein description
1	190200	1561	2.00	8.61E-06	cutinase
2	227728	1219	1.00	5.96E-06	glycoside hydrolase, family 12
4	190202	646	1.00	2.03E-06	cutinase
6	183646	524	1.00	1.85E-06	pyranose dehydrogenase
7	195861	497	1.00	1.14E-05	MFS sugar transporter
8	194576	406	1.00	2.03E-06	alpha-N-arabinofuranosidase
9	226235	356	2.00	4.73E-08	cellulose-growth-specific protein
11	190390	306	1.00	9.26E-06	exoglucanase 3
12	146117	238	1.00	4.36E-06	glycoside hydrolase, family 88
13	212650	235	1.00	4.53E-07	serine proteinase
15	192849	208	4.00	2.59E-07	glycoside hydrolase, family 5
16	194280	199	1.00	4.71E-06	endoglucanase
20	149407	140	2.00	4.01E-06	small sub-unit laccase *sslcc*
24	68978	107	1.00	6.87E-07	serine proteinase
25	194648	103	2.00	5.93E-09	serine proteinase *spr1*
27	175634	101	1.00	9.04E-05	hypothetical protein
28	211605	98	1.00	5.21E-05	fumarylacetoacetate hydrolase family
30	79647	94	1.00	7.70E-07	xylanase
31	139148	93	2.00	9.69E-08	laccase *lcc2*
32	211295	89	2.00	7.99E-06	alpha amylase
33	194521	88	1.00	1.40E-05	exoglucanase
40	146228	75	2.00	1.38E-06	laccase *lcc1*
Negative control on promoters of 33 random sequences					
4008	193722	1	3	3.20E-06	hypothetical protein
4018	199045	1	1	8.00E-06	hypothetical protein
4029	65196-190694*	1	1	8.00E-06	hypothetical protein
4024	189364	1	1	1.70E-05	transcription initiation factor TFIID subunit 1
4026	225676	1	1	2.40E-05	prolyl-tRNA synthetase
4011	134340	1	1	2.60E-05	ribose-phosphate pyrophosphokinase
4021	119786	1	1	3.30E-05	short-chain dehydrogenase/reductase SDR
4025	189220	1	1	4.00E-05	enoyl-CoA hydratase
4012	67030	1	1	5.70E-05	ankyrin repeat protein
4020	209944	1	1	7.20E-05	delta-12 fatty acid desaturase
4010	191839	1	1	7.50E-05	deoxyhypusine synthase

Grey – motif *p*-values of humic overexpressed genes lower than the control motif *p*-values. * Protein IDs of the gene model used for expression study and of a better model used to define the promoter boundary.

Microarray derived gene expression data were compared and ranked as ratios of expression in humic/non-humic environments (i.e. expression in compost/ Average expression in culture, casing and as fruit bodies). The table shows the transcript ratios, rankings (highest to lowest) and protein description of 23 genes sharing the promoter motif, TC[CA][TG]G[AT][GTA]A[AC]AATCTC. The motif was discovered with the Motif-based sequence analysis tools (MEME Suite) 4.4 (http://meme.sdsc.edu/) applied on the 33+1 promoters (1566 bp upsteam the ATG) of humic overexpressed genes + 4 promoters of *A. bisporus var. burnetii* genes homologous of laccase and serine proteinase overexpressed genes and 4 negative sequences (promoter of genes having a transcript ratio of 1). MAST tool was used to find the motif among the 33 control sequences. Lowest p-values of the motifs found among 33+1 promoters of humic overexpressed genes compared to those of motifs found among 33 promoters (of genes having a ratios of expression of 1 in humic/non-humic environments).

B. Genomic frequency of the humic-response motif in *A. bisporus* and in *C. cinereus, P. chrysosporium*, selected Agaricomycotina and two Ascomycetes with and without soil niches.

Species	Genome Size (Mb)	Expected Frequency (both directions)	Actual Frequency	Ratio of Actual/Expected Frequencies
Agaricus bisporus var. *burnettii*	32.8	31.24	132	4.2
Agaricus bisporus var. *bisporus*	30.4	28.98	89	3.1
Coprinopsis cinereus	36.2	34.48	69	2.0
Phanerochaete chrysosporium	35.2	33.54	62	1.8
Serpula lacrymans	42.7	40.76	52	1.3
Tuber melanosporum	125	119.18	143	1.2
Laccaria bicolor	59.1	56.54	58	1.0
Cryptococcus neoformans	18.9	18.00	18	1.0
Schizophyllum commune	38.6	36.84	29	0.79
Botrytis cinerea	42.7	40.72	29	0.72

Table S16. A. Comparison of fruiting body (FB) regulated transcripts from *A. bisporus var. bisporus*, *Laccaria bicolor* and *Schizophyllum commune*. Listed are all transcripts with a significant >10 fold up- or down-regulation in Agaricus FB compared to agar medium, as well as significantly regulated homologs from *L. bicolor* and *S. commune* (FB compared to agar medium). Expression data for *A. bisporus* and *L. bicolor* were generated by oligoarrays, *S. commune* data by massively parallel signature sequencing (MPSS). * p-value FDR, modified t test < 0.05.

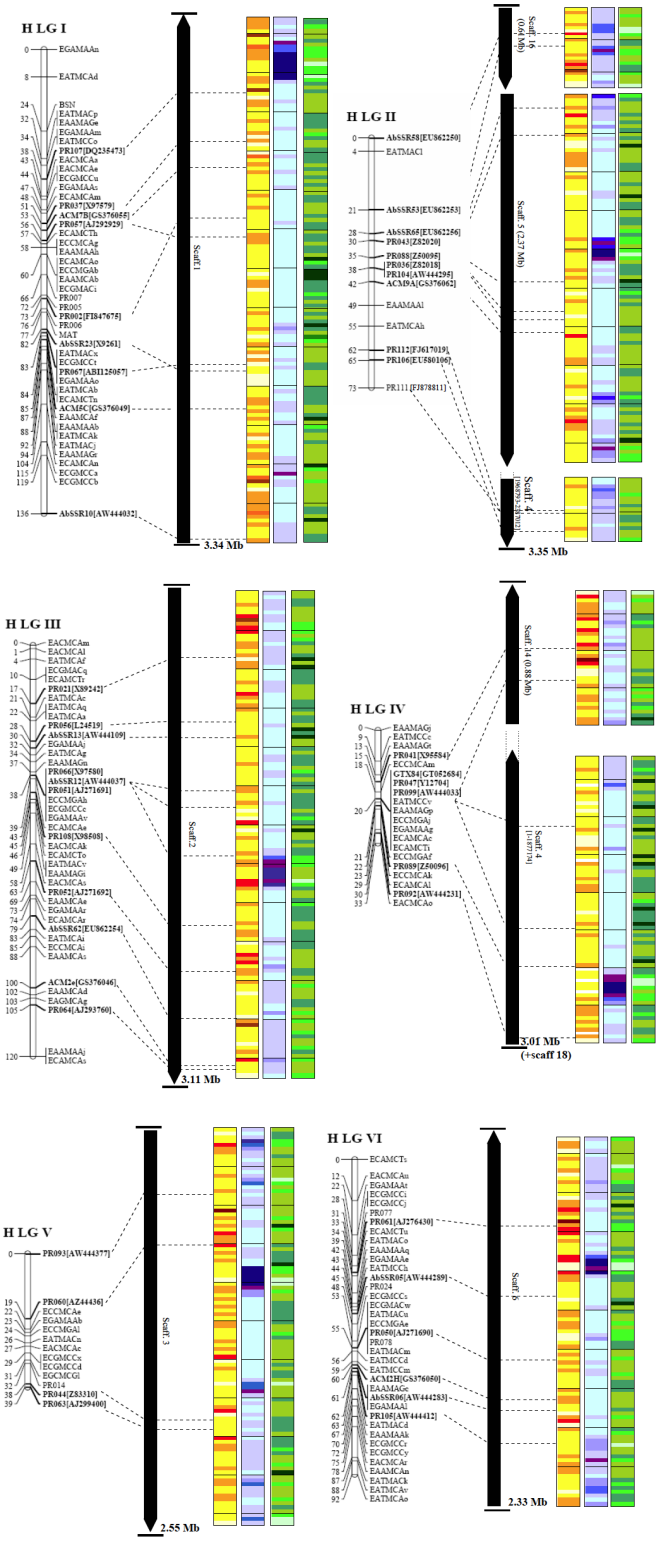
Agaricus bisporus		***Laccaria bicolor***		***Schizophyllum commune***		
JGI Protein ID	ratio FB/Agar medium	JGI Protein ID	ratio FB/Agar medium	JGI Protein ID	ratio FB/Agar medium	Putative function
			significantly* up (>10-fold)			
189424	391.9	301146	5.7	73392	9.4	Hypothetical protein
195454	241.8	298461	4.3	258034	21.1	Aromatic-ring hydroxylase
219810	184.7	309640	5.1	72462	28.5	Hypothetical protein
74558	58.6	236299	99.4	258323	23.4	RNA recognition motif. RNP-1
188626	53.2	245697	27.3	64227	10.6	Glycoside hydrolase 16
136509	49.7	299929	475.3	235154	>3.5	Hypothetical protein
185602	25.7	295954	29	69239	5.1	FAD linked oxidase
114425	21	242884	4.6	13353	76.9	Exonuclease
225837	17	254799	3.2	111512	3.7	Fatty acid desaturase
194141	16.9	300021	6.4	85903	4.5	Hypothetical protein
191073	16.7	306386	5.6	108884	8.4	Hypothetical protein
200889	15.7	312067	3.6	27681	5.5	Hypothetical protein
83052	11.6	306176	5.3	105081	8.2	DNA-binding HORMA domain
			significantly* down (>10-fold)			
190902	0.01	253329	0.16	50449	0.18	Cytochrome P450, E-class. group IV
19143	0.02	308057	0.35	27314	0.27	Hypothetical protein
205540	0.04	305032	0.07	32888	0.07	Glycoside hydrolase 5

B. Expression of *A. bisporus* var. *bisporus* and *Laccaria bicolor* homologs of transcription factor genes that have previously been shown to be involved in mushroom development in *Schizophyllum commune* and *Coprinopsis cinerea*. Homologs were identified using a bi-directional/reciprocal best hit analysis. [1]Expression data for *A. bisporus* and *L. bicolor* were generated by oligoarrays, [2]*S. commune* data by massively parallel signature sequencing (MPSS). Changes in gene expression with p-value FDR, modified t test < 0.05 were considered as significant, non-significant differences in gene expression compared to fruiting bodies are marked in grey. [3] compared to vegetative mycelium, as determined by northern analysis (40). [4] determined by northern analysis (41).

Schizophyllum commune						***A. bisporus var bisporus***				***Laccaria bicolor***			
		Expression level[2]						**Expression level[1]**				**Expression level[1]**	
			Fruiting bodies										
Protein ID	**Name**	**Culture**	**Stage I**	**Stage II**	**mature**	**Protein ID**	**E value**	**Culture**	**Fruiting bodies**	**Protein ID**	**E value**	**Culture**	**Fruiting bodies**
257987	hom2	53	2	75	23	192725	6,72E-28	8	173	293988	1,51E-52	16993	3856
66861	fst4	82	75	257	372	223670	0	2336	4533	308722	0	4698	9218
114363	c2h2	17	14	317	244	230069	5,47E-38	7	766	487295	4,34E-38	2660	3659
257422	fst3	100	35	234	363	191328	3,18E-102	3255	12769	307309	0	3556	24281
255004	gat1	104	22	437	138	195724	1,72E-55	232	149	685209	1,67E-55	2889	3064
257652	hom1	90	50	354	772	192433	3,11E-31	794	3556	324166	2,63E-28	2003	5992
255701	bri1	26	13	82	37	189729	1,92E-114	2139	3372	700295	4,36E-125	13211	5276
Coprinopsis cinerea													
CC1G_01334	exp1	up-regulated in fruiting bodies[3]				188638	1,26E-147	450	319	644689	8,76E-144	4936	9363
CC1G_07392	pcc1	higher expressed in dikaryons than in monokaryons[4]				192530	5,44E-32	4151	6119	386478	5,42E-36	8499	1708

Supplemental Figures

Figure S1. Alignment of *A. bisporus* var. *bisporus* H97 genome scaffolds on *A. bisporus* genetic linkage map. Linkage groups (LG) corresponding to chromosomes are drawn on the left, scaffolds on the middle, and heat maps displaying distribution of single sequence repeats (SSRs) (in yellow), transposable elements (TEs) (in blue) and protein-coding gene models (GMs) (in green) per 25kbp-window on the right. The heat maps describe 98.2 %, 97.3 %, 98.3% of SSRs, TEs and GMs, respectively, found in the H97 genome. Increasing abundance of the different structures is represented by a color scale from light to darker colors.



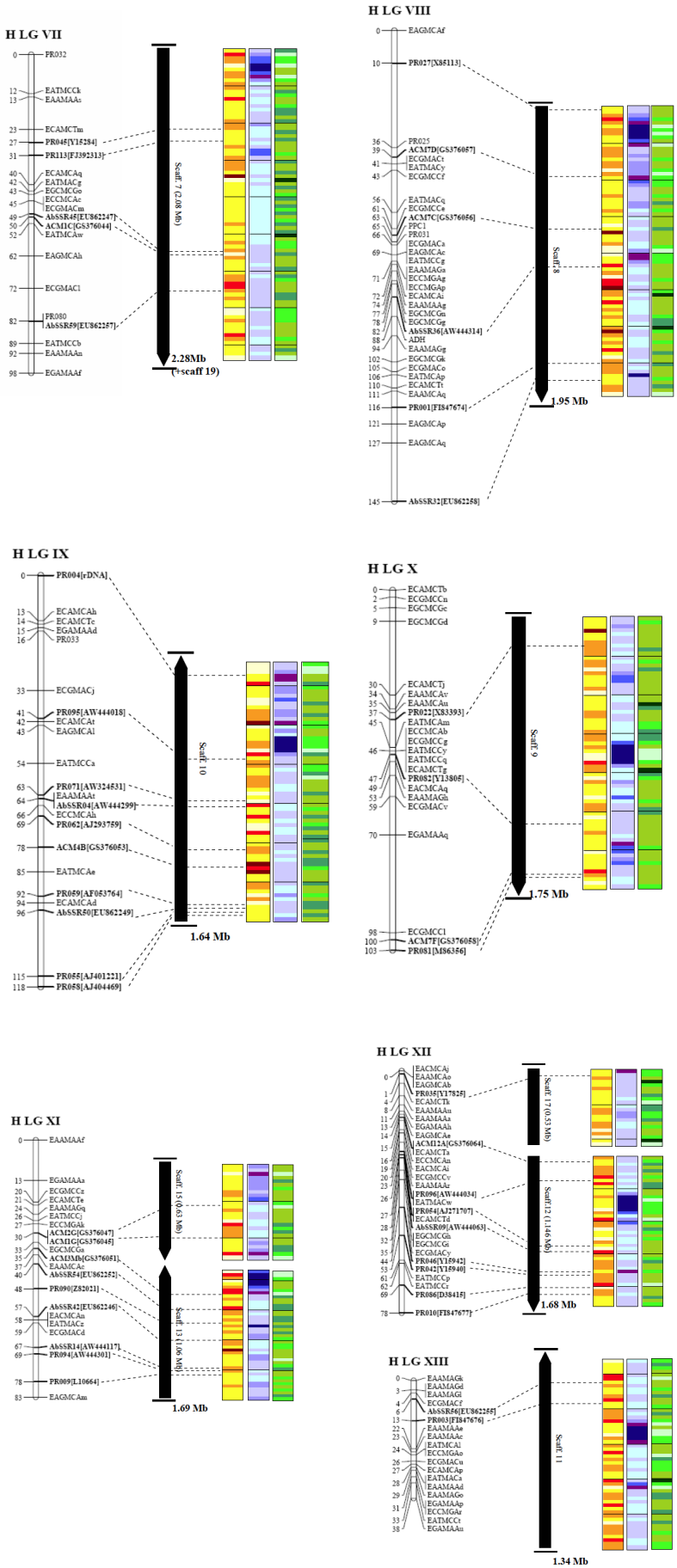

H LG VII
H LG VIII
H LG IX
H LG X
H LG XI
H LG XII
H LG XIII
Scaff 7 (2.08 Mb)
2.28Mb
(+scaff 19)
Scaff 8
1.95 Mb
Scaff 10
1.64 Mb
Scaff 9
1.75 Mb
Scaff 14 (0.63 Mb)
Scaff 13 (1.06 Mb)
1.69 Mb
Scaff 17 (0.53 Mb)
Scaff 12 (1.146 Mb)
1.68 Mb
Scaff 11
1.34 Mb

Figure S2. Macro synteny between *Agaricus bisporus* H97 scaffolds (right panel) and the 134 largest scaffolds of *A. bisporus* var *burnetti* JB137-S8 scaffolds (left panel). *A. bisporus* scaffolds are depicted by the colored blocks and *A. bisporus* var *burnetti* scaffolds are represented by gray blocks. Only regions larger than 5,000 bp are connected with links of colors matching those used for coloring *A. bisporus* scaffolds. Comparison between the two genomes sequences was performed with VISTA (http://genome.lbl.gov/vista/).

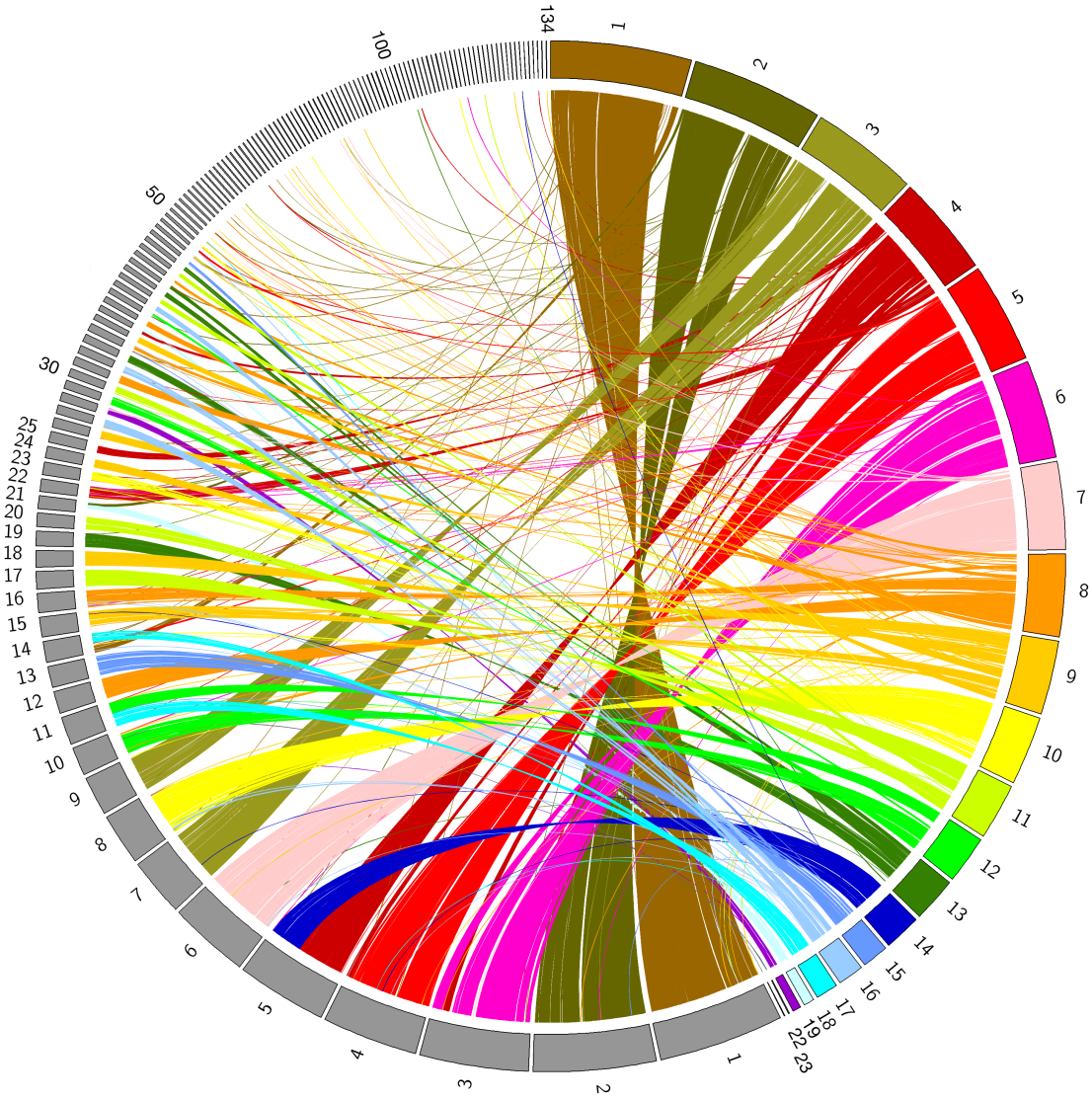


Figure S3. The diversity and distribution of class I and class II transposable elements in *A. bisporus* var. *bisporus* H97. **A.** Distribution in various TE families, number of copies and % genome assembly coverage. **B**. Average insertion age of full-length copies of LTR retrotransposons.

A.

Agaricus bisporus H97 Transposable elements diversity

	Number of families	Number of copies	% Genome assembly coverage
Class 1 LTR Gypsy-like	6	574	3.5
Class 1 LTR Copia-like	6	453	1.96
Class 1 Other LTR	6	84	0.16
Class 1 Non LTR	14	285	1.39
Class 2 TIR	12	220	0.4
Class 2 Helitron	1	5	0.01
Not Categorized	171	2008	3.76

B.

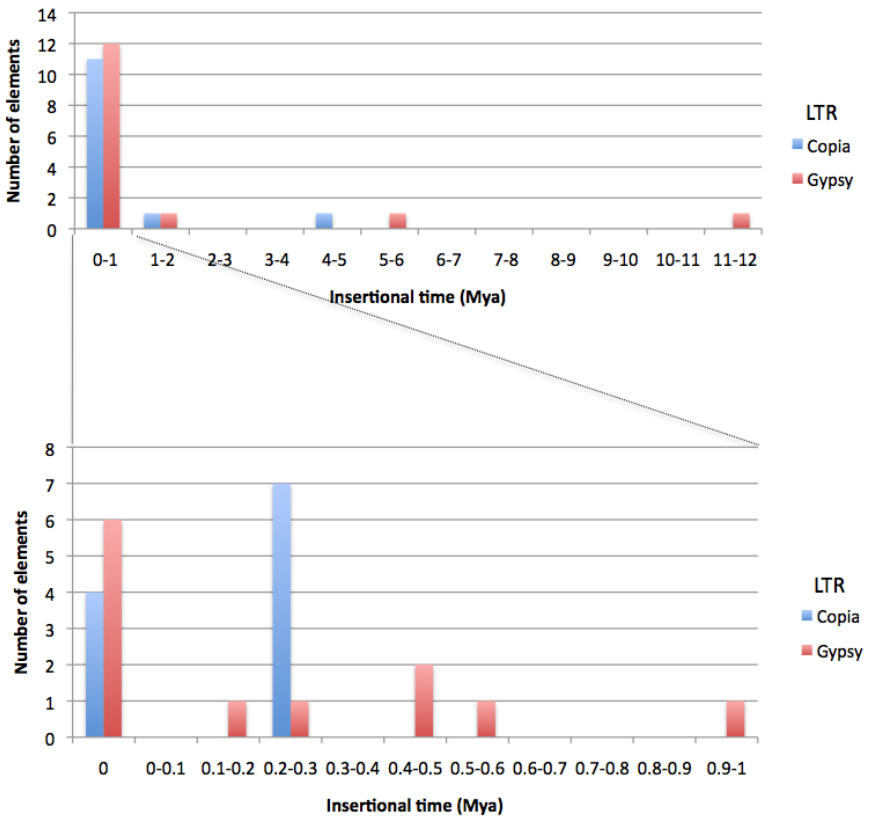


Figure S4. Organismal phylogeny produced with RAxML from a 36 genes dataset. Each branch is labeled with the confidence value from 1,000 bootstrapped topologies. Posterior probabilities are >0.95 for each branches. The separation time (in Myears) of some node was estimated using Floudas *et al.* (2012) and indicated in parentheses.

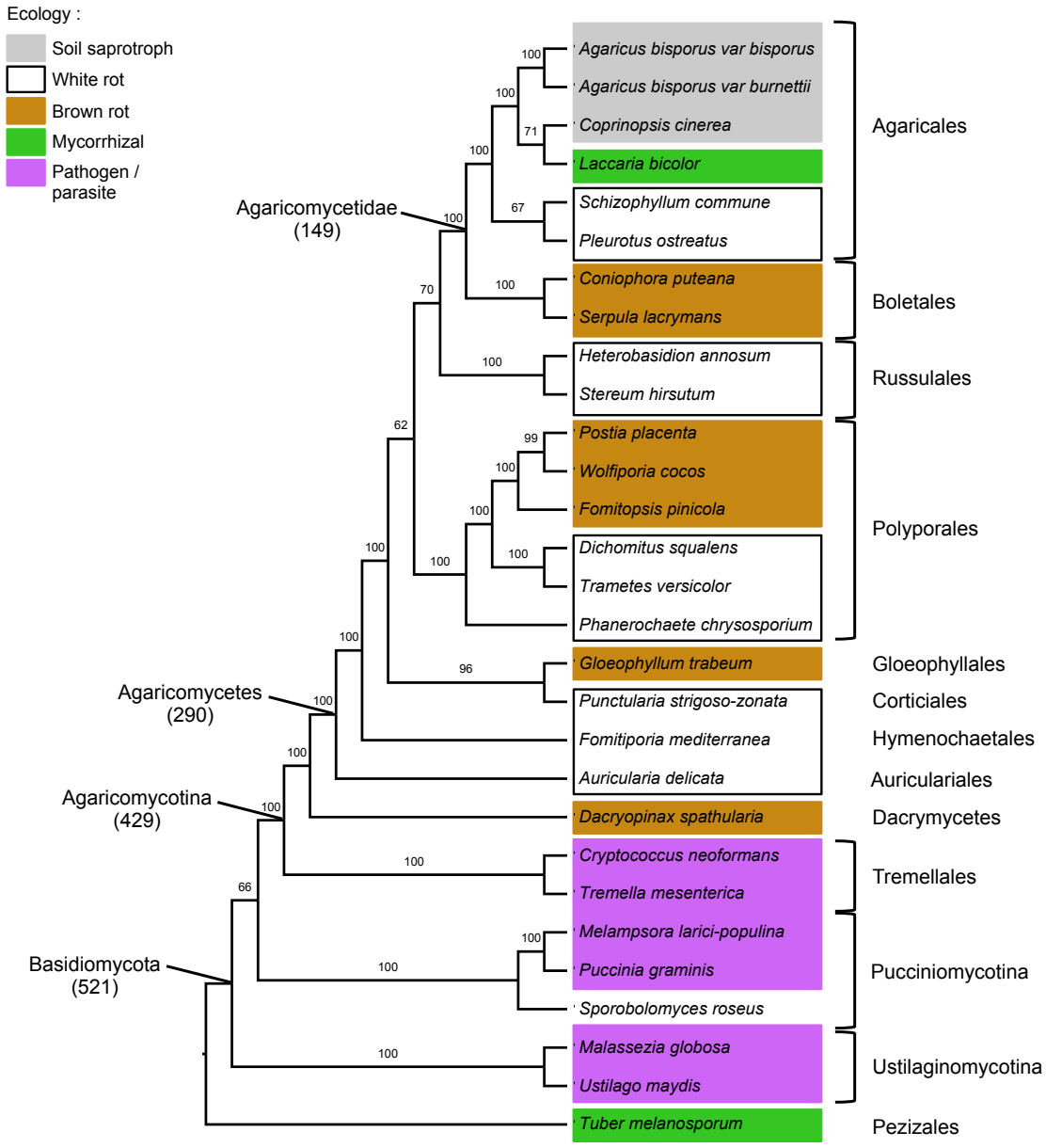


Figure S5. Genome redundancy in the *Agaricus bisporus* var. *bisporus* genome. The figure represents the total number of gene families in each species or node. The numerals on branches and pie charts at each branch terminus show the proportion of expanded (red), unchanged (black/grey) or contracted (blue) gene families along lineages by comparison to the putative pan-genome. CAFE analysis, p-value <0.001. MRCA, most recent common ancestor.

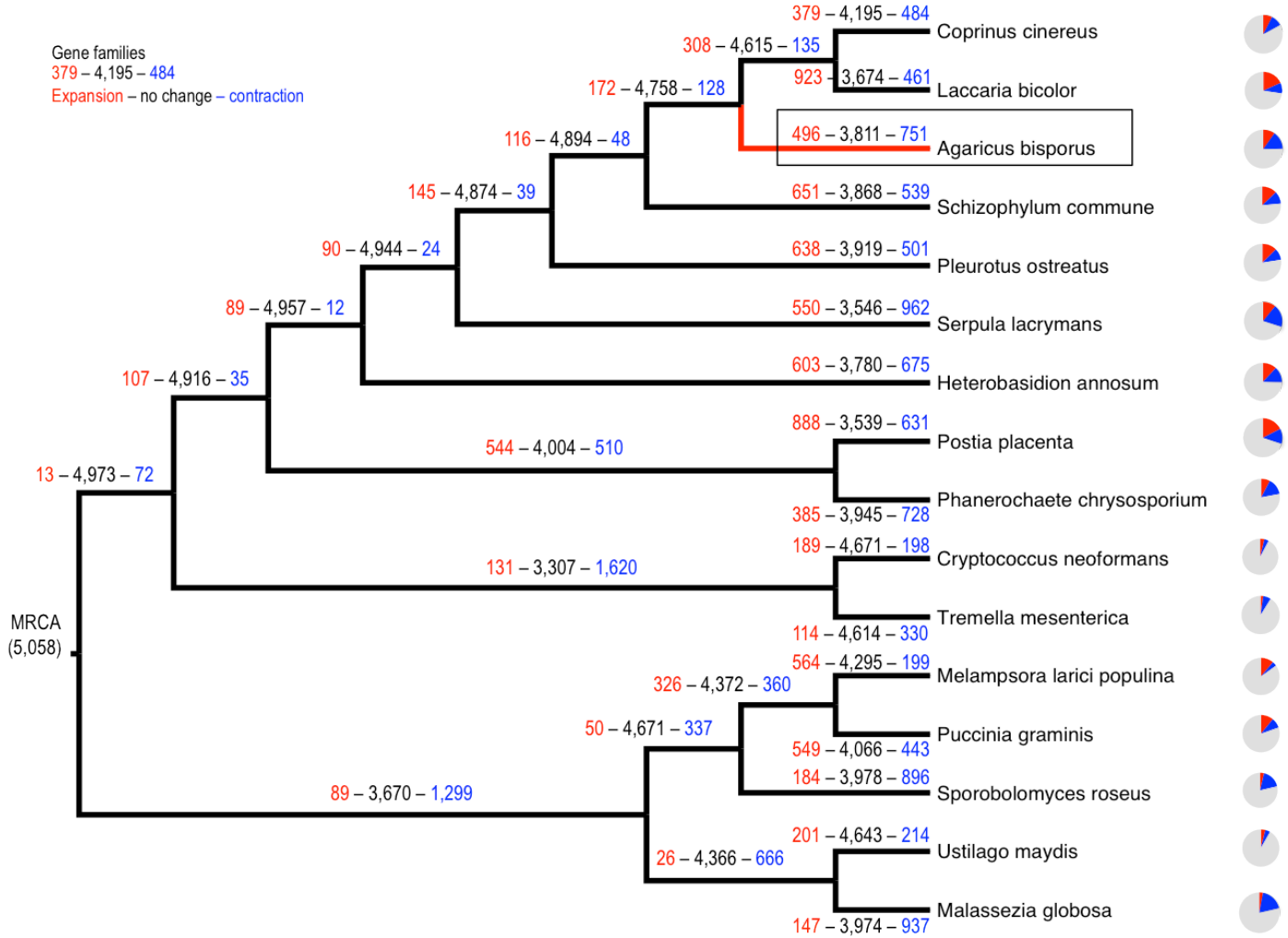
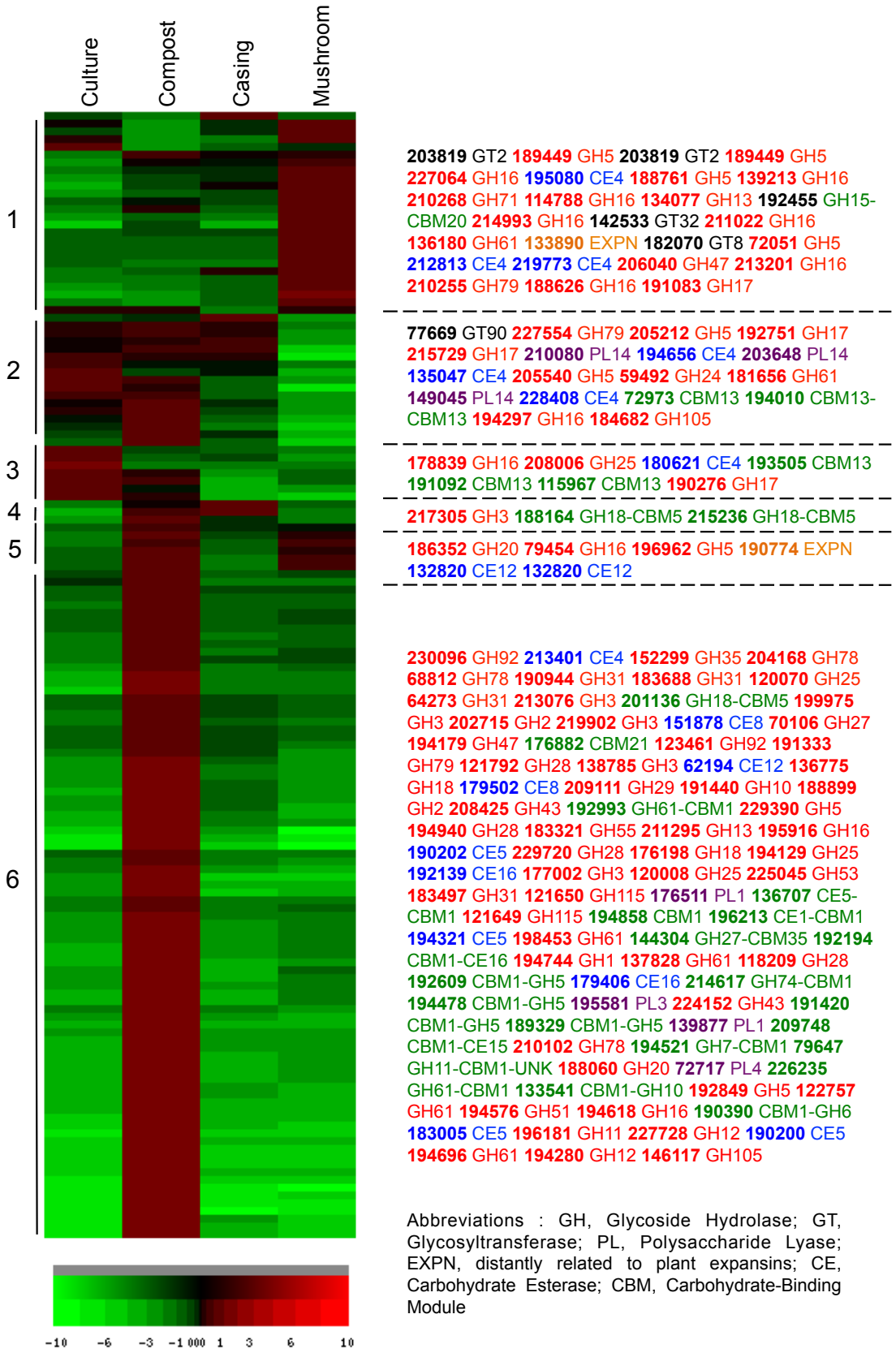


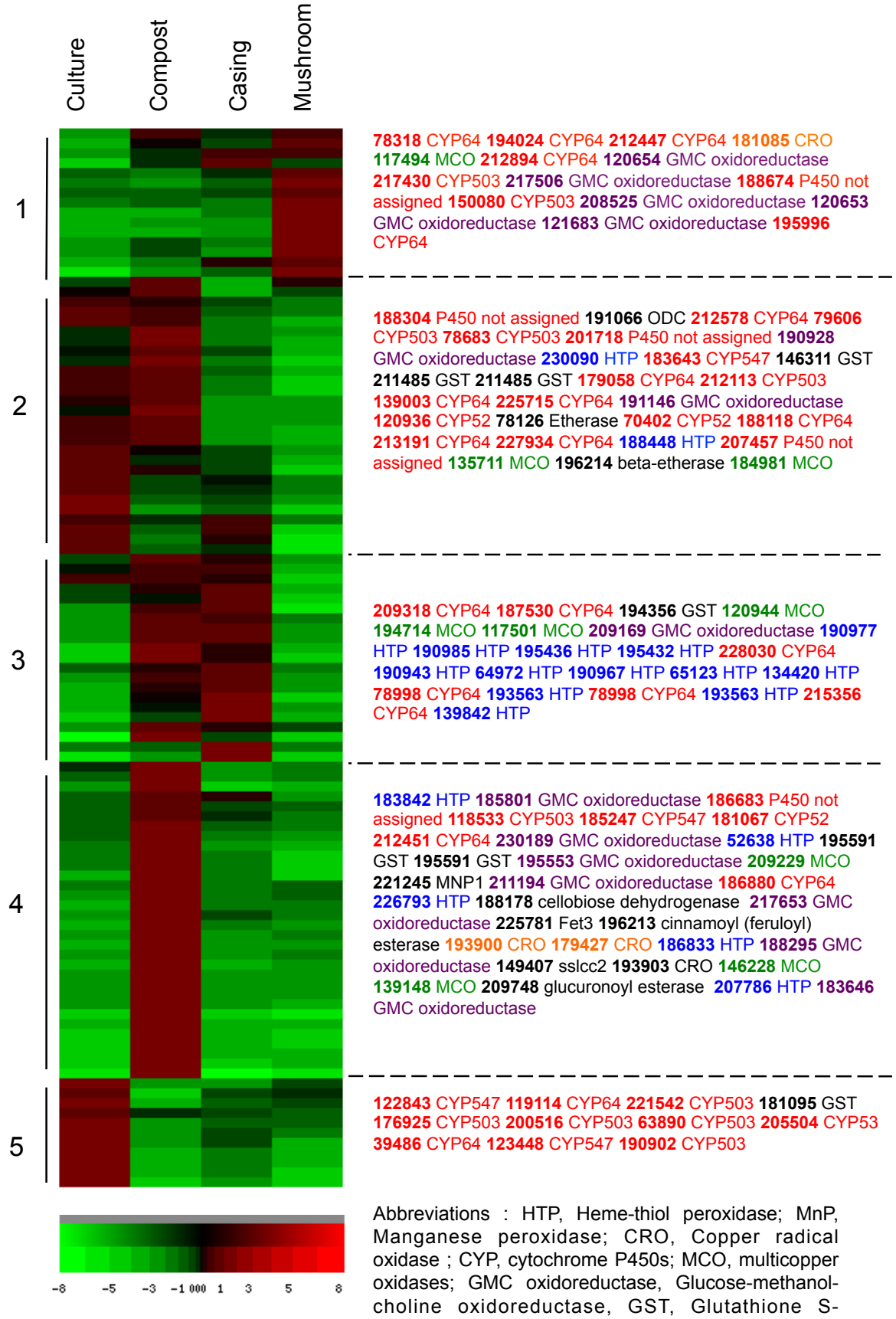
Figure S6. Hierarchical clustering of regulated transcripts from *A. bisporus* var. *bisporus* either grown on defined agar-medium [Culture], compost [Compost] or casing layer [Casing] or in fruiting bodies [Mushroom].

A. Transcripts coding for carbohydrate-active enzymes (CAZymes) potentially involved in degradation of polysaccharides.



Abbreviations : GH, Glycoside Hydrolase; GT, Glycosyltransferase; PL, Polysaccharide Lyase; EXPN, distantly related to plant expansins; CE, Carbohydrate Esterase; CBM, Carbohydrate-Binding Module

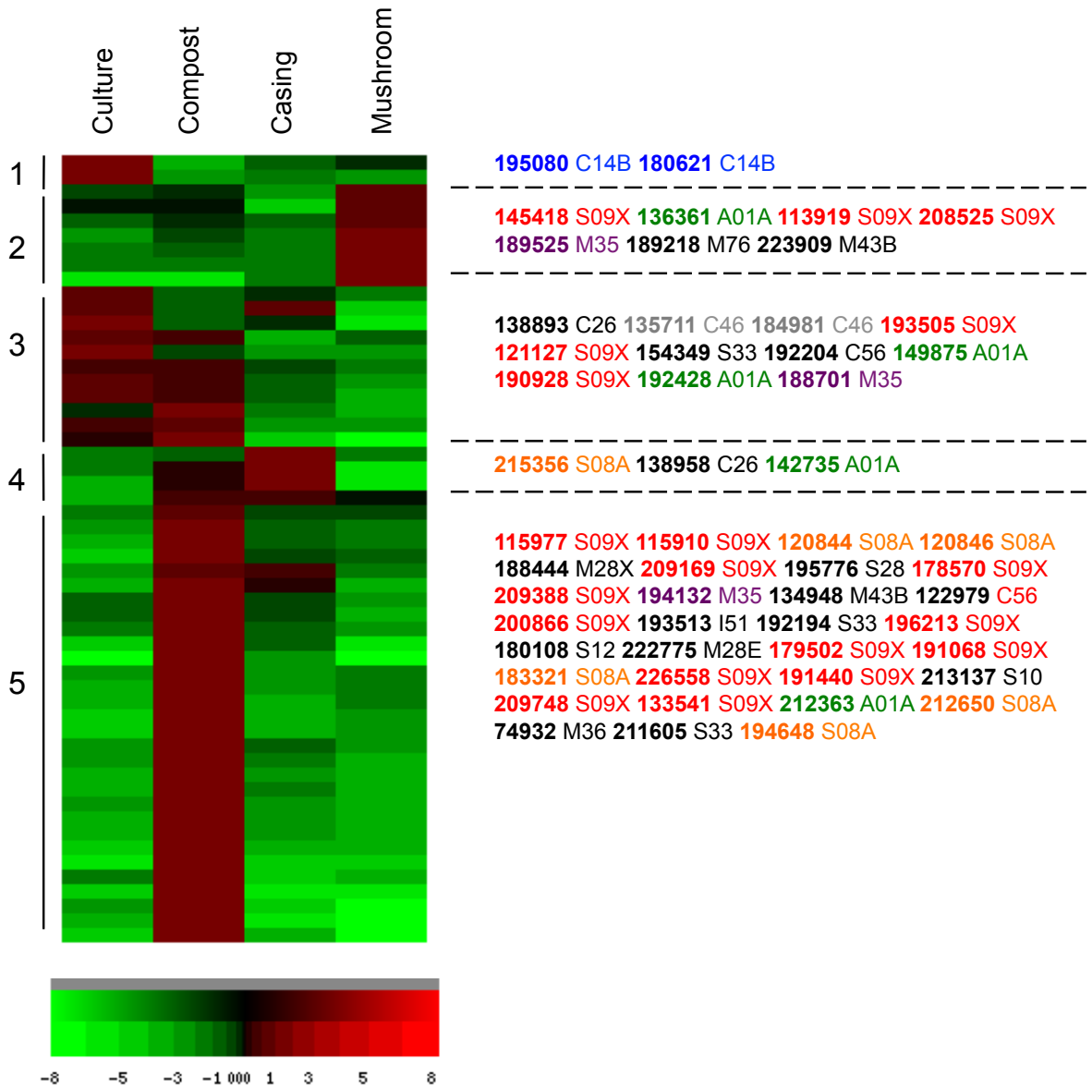
B. Transcripts coding for oxidoreductases potentially involved in degradation of lignin and related compounds



Abbreviations : HTP, Heme-thiol peroxidase; MnP, Manganese peroxidase; CRO, Copper radical oxidase ; CYP, cytochrome P450s; MCO, multicopper oxidases; GMC oxidoreductase, Glucose-methanol-choline oxidoreductase, GST, Glutathione S-transferase; ODC, Oxalate decarboxylase.

C. Transcripts coding for secreted proteases potentially involved in degradation of proteins.

MEROPS families are given on the right panel.



Transcripts significantly regulated (>5-fold; FDR, modified t-test <0.05) in minimum one of the conditions were included in the analysis. Relative expression indexes (REI) were calculated for the dataset. For each gene, a mean expression level was calculated from the four conditions, and the REI corresponds to the ratio between the expression level measured for a given condition and the mean reference. Log2 transformed data were subjected to EPCLUST software (http://www.bioinf.ebc.ee/EP/EP/EPCLUST/) using correlation measure based distance (uncentered) and average linkage (UPGMA). In the boxes on the right of the cluster, the protein ID and the gene family of each transcript are presented. Transcripts from the same gene family are highlighted by the same color.

Figure S7. Growth of *A. bisporus* monokaryons (H39, H97) and dikaryon (U1) on various carbon sources with and without the addition of humic extracts. Growth was quantified by eye independently by two researchers and averaged. Differences between duplicate plates were less than 1 mm.

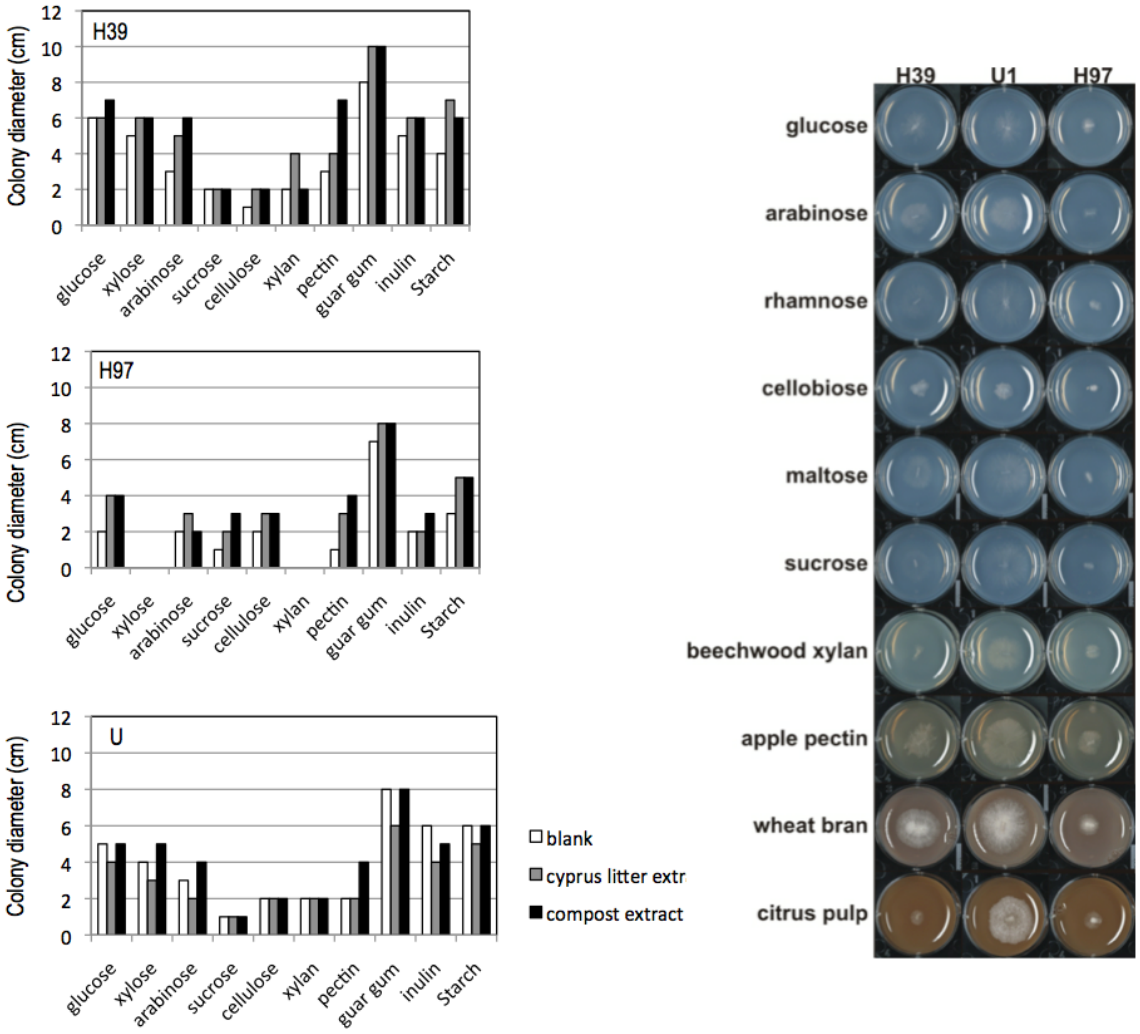


Figure S8. Double clustering of the CAZyme families from representative fungal genomes.
Top tree: fungal species. Left tree: the enzyme families are represented by their class (GH, glycoside hydrolase; PL, polysaccharide lyase) and family number according to the CAZymes database (34). Right side: known substrate for the CAZy families (most common forms in brackets): BPG, bacterial peptidoglycan; BEPS, bacterial exopolysaccharides; CW, cell walls; E, energy storage and recovery; FCW, fungal cell walls; PCW, plant cell walls; PG, protein glycosylation; U, undetermined; a-gluc, a-glucans (including starch/glycogen); bglyc, b-glycans; b-1,3-gluc, b-1,3-glucans; cell, cellulose; chit, chitin/chitosans; dext, dextrans; hemi, hemicelluloses; inul, inulin; N-glyc, N-glycans; N-/O-glyc, N-/O-glycans; pect, pectins; sucr, sucrose; and tre, trehalose. Abundance of the different enzymes within a family is represented by a color scale from 0 (white) to 44 occurrences (red) per species. In the top tree, black circles, plant and fungal pathogens; red squares, saprotrophs; white squares, white rots; orange circles, brown-rots; and green triangle, ectomycorrhizal symbiont.

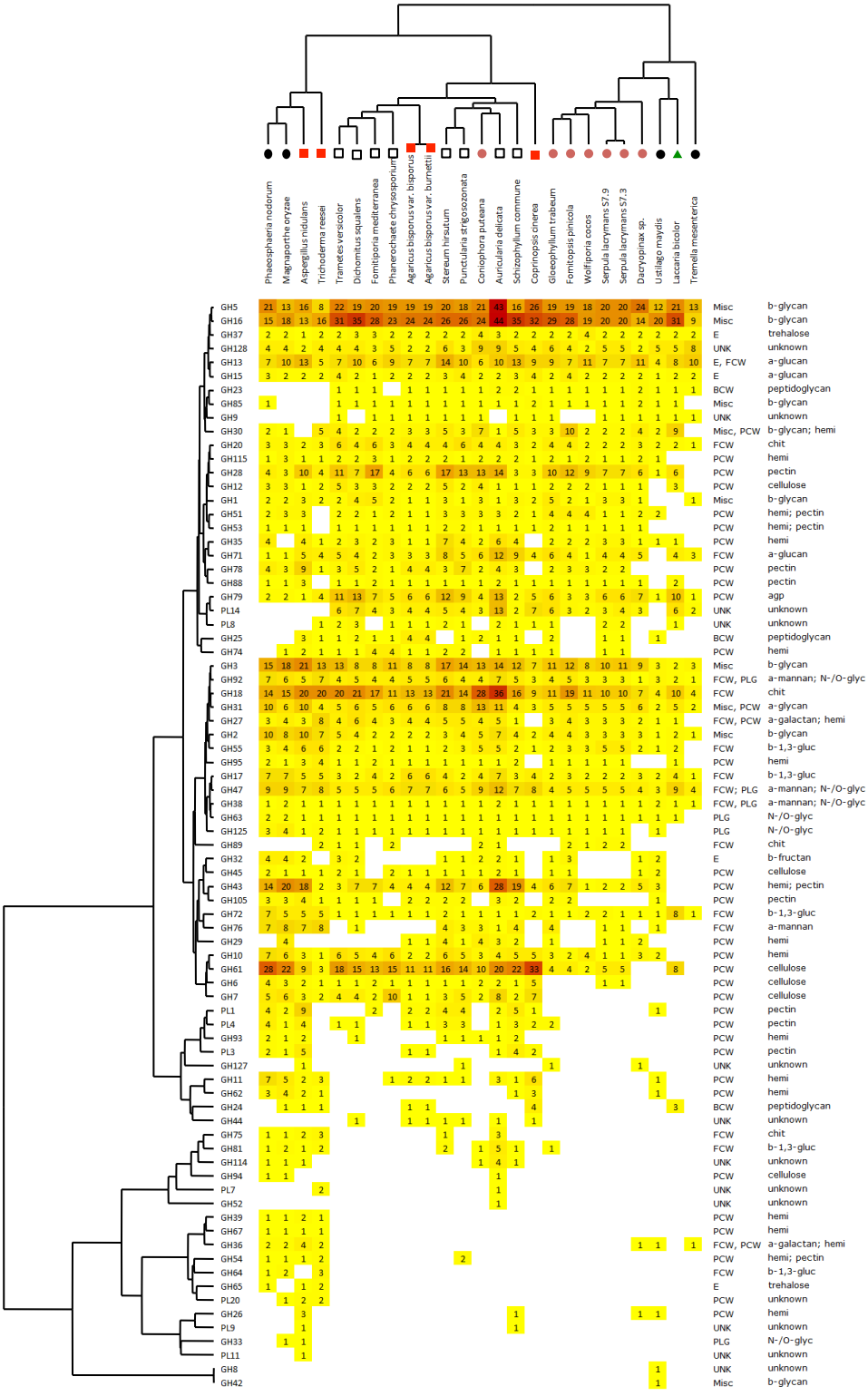
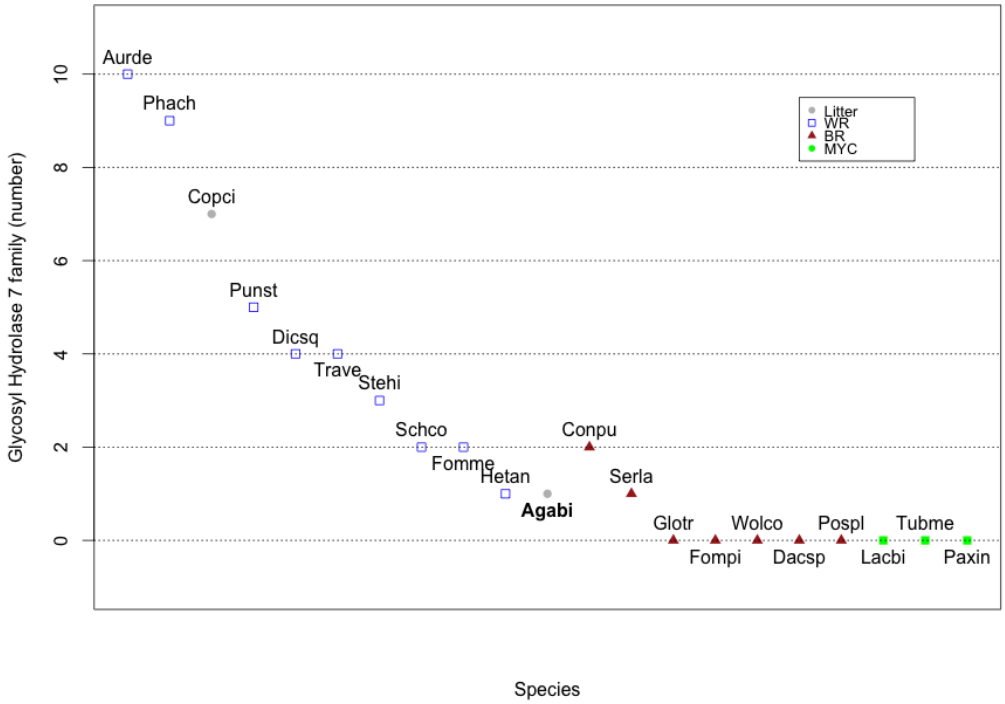
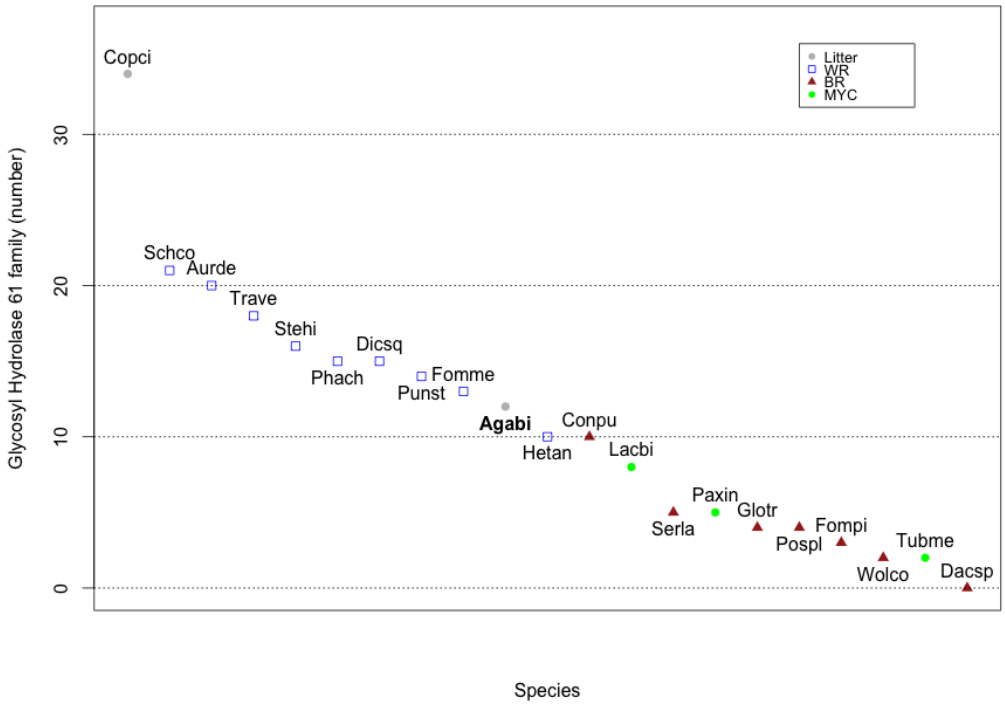


Figure S9. Distribution of genes coding for cellulose-degrading enzymes GH7 and GH61 in *Agaricus bisporus* var. *bisporus,* the coprophilous *C. cinerea*, white-rotters (WR), brown-rotters (BR), and ectomycorrhizal (MYC) fungi. **A**, Content in GH7 cellulase and **B**, GH61 cellulose 'oxidoreductase'.

A.

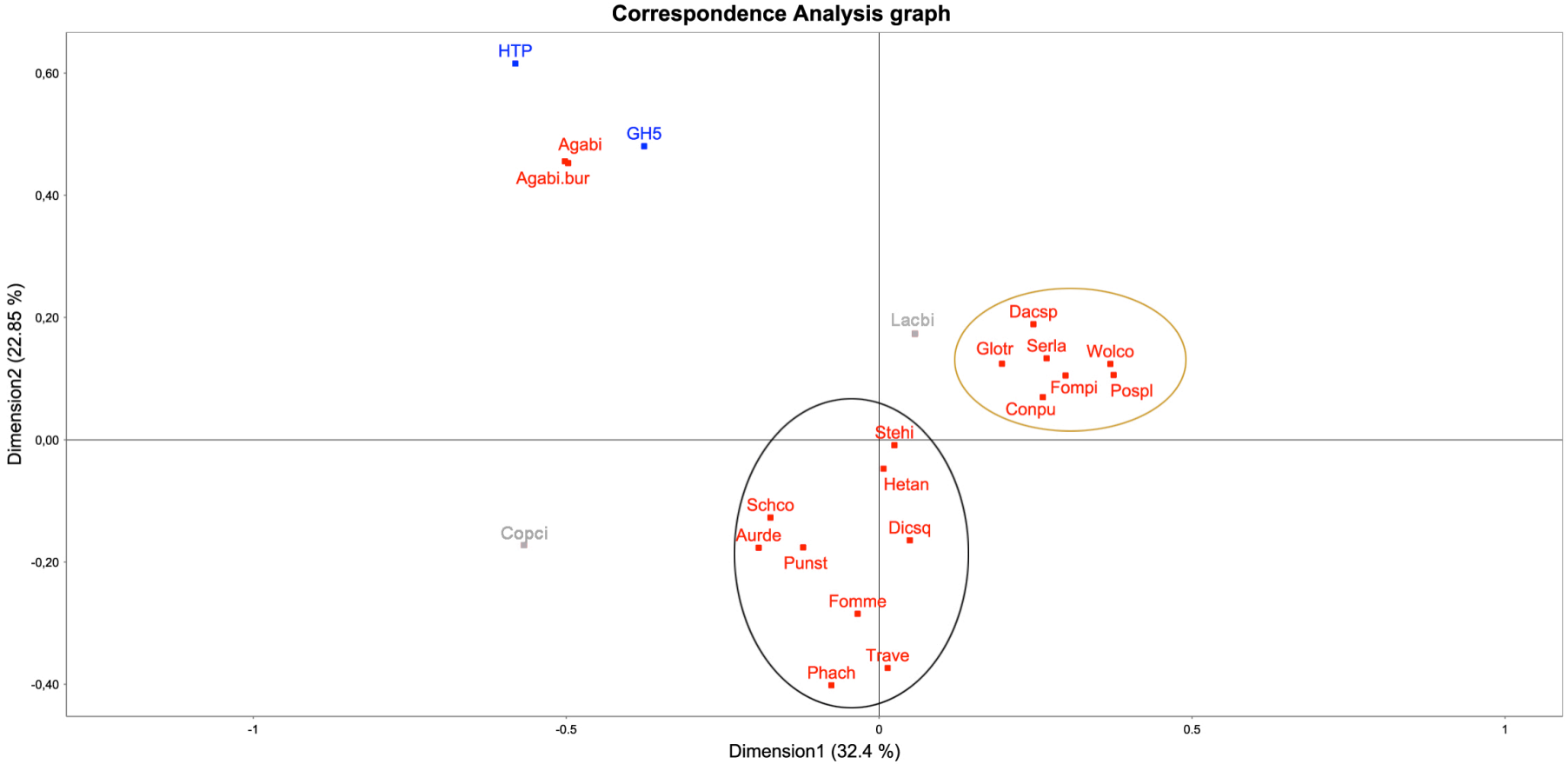


B.



Abbreviations: AGABI, *A. bisporus*; AURDE, *A. delicata;* CONPU, *C. puteana;* COPCI, *C. cinerea*; DACSP, *D. spathularia*; DICSQ, *D. squalens*; FOMME, *F. mediterranea*; FOMPI, *F. pinicola*; GLOTR, *G. trabeum*; HETAN, *H. annosum*; LACBI, *Laccaria bicolor;* PAXIN, *P. involutus*; PHACH, *P.chrysosporium*; POSPL, *P. postia*; PUNST, *P. strigoso-zonata*; SCHCO, *S. commune*; SERLA, *S. lacrymans;* STEHI, S. hirsutum; TRAVE, *T. versicolor*; TUBME, *T. melanosporum*; Wolco, *W. cocos*.

Figure S10. Correspondance Analysis on wood decay gene families. Dataset is built from 27 genes encoding oxidoreductases and carbohydrate active enzymes (CAZymes) known to degrade wood (Table S10). The black circle delimitate white-rotters, the brown one the brown-rotters. The blue dots show the proteins over represented in *A. bisporus* strains.



Abbreviations: Agabi, *A. bisporus var bisporus*; Agabi.bur, *A.bisporus var burnettii*; Aurde, *Auricularia delicata*; Copci, *C. cinerea*; Conpu, *C. puteana*; Dacsp, *Dacryopinax sp.*; Dicsq, *D. squalens*; Fomme, *F. mediterranea*; Fompi, *F. pinicola*; Glotr, *G. trabeum*; Hetan, *H. annosum*; Lacbi, *L. bicolor*; Phach, *P. chrysosporium*; Pospl, *P. placenta*; Punst, *P. strigosozonata*; Schco, *S. commune*; Stehi, *S. hirsutum*; Serla, *S. lacrymans*; Trave, *T. versicolor*; Wolco, *W. cocos*.

Fig S11. The phylogenetic distribution of Class-II-peroxidases in the Agaricus genome. We observed no significant expansion or contraction in Agaricus as compared to other litter-decomposing species (A). Numbers above branches refer to ancestral copy numbers as reconstructed by Notung under two edge-weight thresholds (separated by slash). Numbers following species names show the copy numbers of fPOX-s found in the genomes of these species. (B) Both varieties of *Agaricus bisporus* possess two copies of fPOX-s, which form two well-supported clades within the phylogeny of fPOX-s in Agaricomycetes, as shown on a maximum clade credibility chronogram inferred in BEAST 1.7.2. Thickened branches denote nodes with >0.95 posterior probabilities.

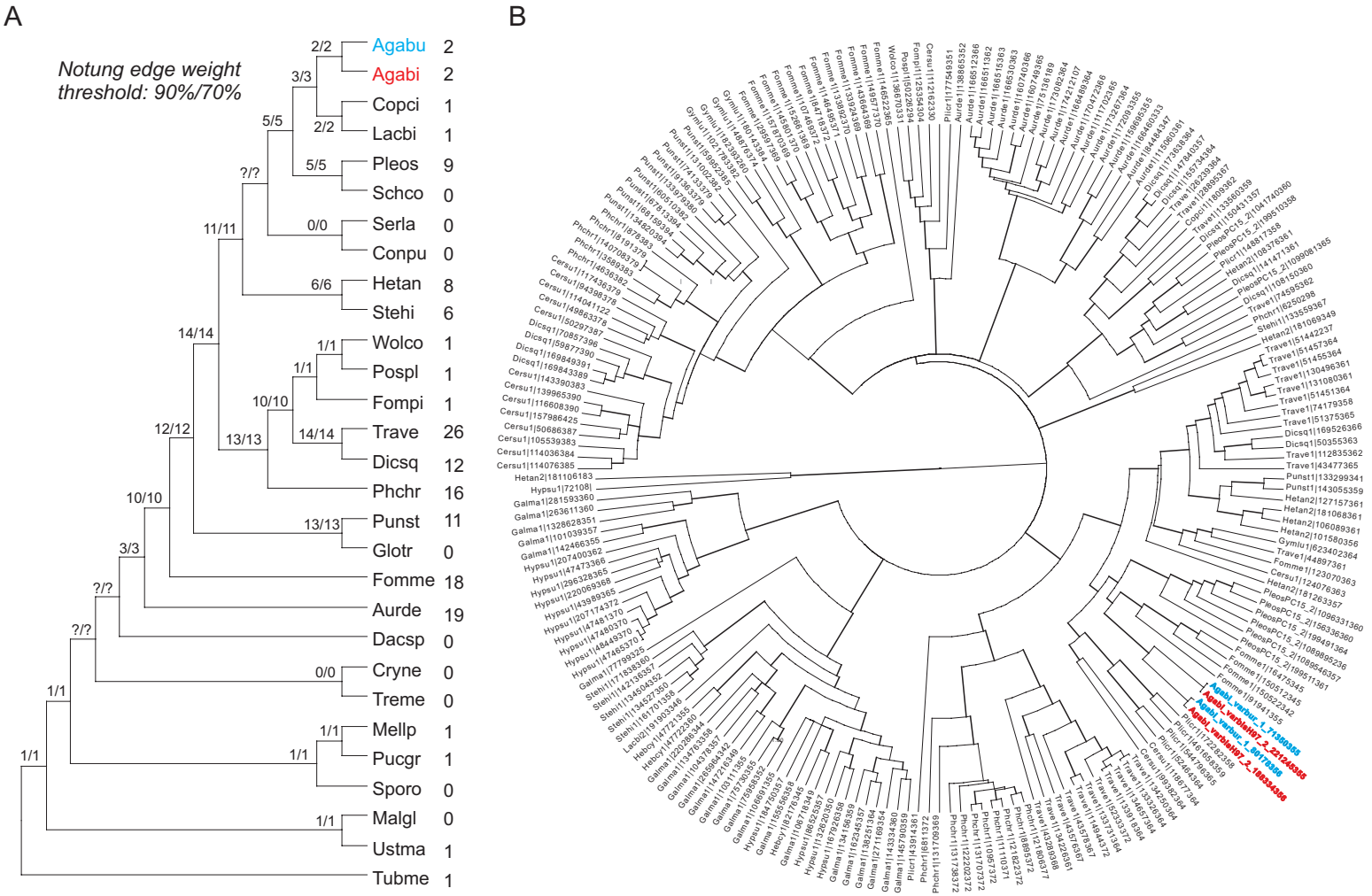
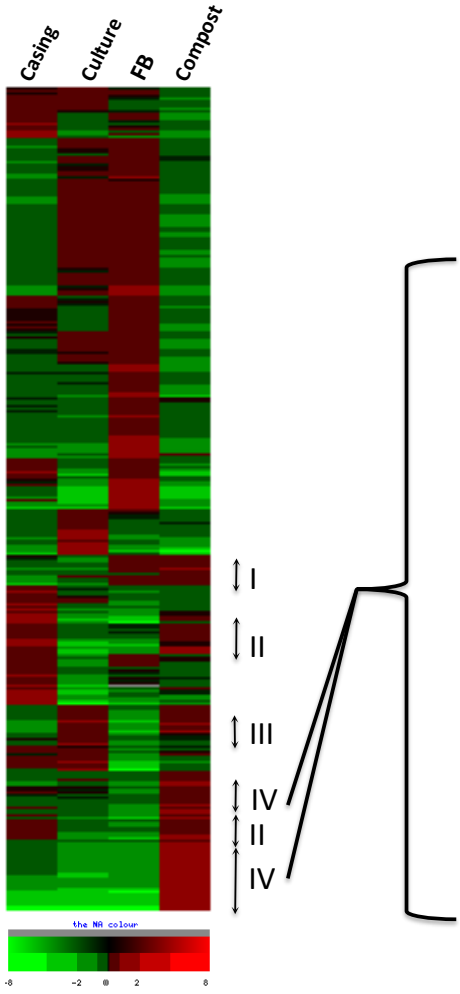


Figure S12. Regulation of genes coding for membrane transporters in *A. bisporus*. Heat map with hierarchical clustering of 333 gene models encoding transporters from *Agaricus bisporus* var. *bisporus*. Gene expression profiles of compost up-regulated genes encoding transporters show four clusters of peak expression. Cluster I contains transporters with gene expression up-regulated in mycelium cultured on compost or from fruiting body (FB). Cluster II, divided in two subgroups, contains transporters up-regulated in mycelium cultured on compost and casing medium. Cluster III comprises transporters up-regulated in mycelium cultured on compost and culture medium. For cluster IV which contains transporters transcripts specifically up-regulated in mycelium cultured on compost medium, a detailed overview of the proteins ID, the putative function and the expression in the different treatments are given in the table to the right of the heat map. Expression levels corresponding to the different treatments (casing, culture, fruiting body (FB) and compost) were converted to relative expression index (REI, expression level at a given treatment divided by the mean level across the treatments series) to help direct comparison of expression profiles. EPCLUST (http://www.bioinf.ebc.ee/EP/EP/EPCLUST/) and Log2-REI were used for the hierarchical clustering. REI levels range from pale to saturated colors (red for induction; green for repression). Black indicates no change in gene expression.



Prot ID	Putative function (TC family)	Cazing	Compost	Culture	FB
175913	Related to Yeast nicotinic acid transporter TNA1 (MFS; ACS)	663	1420	744	701
194062	MFS transporter (MFS)	2267	4926	2542	2459
73149	ABC transporter (ABC)	216	1555	533	210
200580	Related to yeast quinidine transporter QDR1 (MFS; DHA1)	708	7710	2296	456
184138	MFS transporter (MFS)	4558	6395	4199	3333
74076	Uncharacterized MFS transporter (MFS; DHA1)	2996	4732	3138	2198
200631	Related to amino acid exporter ScAQR1 (MFS; DHA1)	474	828	465	328
118249	Related to yeast quinidine transporter QDR (MFS; DHA1)	458	934	441	293
187594	Uncharacterized MFS transporter (MFS)	591	2591	665	291
202707	Uncharacterized transporter (MFS; ACS)	5472	12639	3755	2263
62308	Oligopeptide transporter (OPT)	7543	29410	4233	1159
187542	Myo-inositol Transporter (MFS; SP)	1127	2606	916	802
201241	Multidrug resistance-associated protein (ABC)	995	2932	658	558
118762	Related to nicotinic acid transporter SpSPAC11D3.18c (MFS; ACS)	1378	5612	891	589
177416	Quinate permease (MFS; SP)	875	2203	431	404
214100	Related to yeast Multidrug transporter HOl1 (MFS; DHA1)	1313	4973	667	638
66982	Related to nicotinic acid transporter SpSPAC11D3.18c (MFS; ACS)	1800	6719	791	651
183682	Related to yeast quinidine transporter QDR1 (MFS; DHA1)	488	1864	199	171
175915	Hypothetical Transporter (MFS family)	6597	27121	2589	2249
195761	Related to Cellodextrin/Cellobiose transporter (MFS; SP)	6737	18990	1992	2835
195794	Related to Cellodextrin/cellobiose transporter (MFS; SP)	8317	27574	2014	2269
208266	Oligopeptide transporter (OPT)	4631	16125	1208	2115
117146	Allantoate permease (MFS; ACS)	1057	4510	235	538
187070	Oligopeptide transporter (OPT)	1791	10590	541	1139
191173	Probable glucose transporter (MFS; SP)	2449	16354	2061	2518
179046	Related to nicotinic acid transporter SpSPAC11D3.18c (MFS; ACS)	1728	10382	1021	1709
223428	Related to nicotinic acid transporter SpSPAC11D3.18c (MFS; ACS)	1822	11559	1139	1933
141940	Sugar transporter (MFS; SP)	4174	39062	3460	2050
213473	Monosaccharide tranporter (MFS; SP)	2303	23653	899	512
188055	Related to allantoate-ureidosuccinate permease	1009	26875	941	171
201865	Related to Cellodextrin/Cellobiose transporter (MFS; SP)	223	6414	114	260
194194	Related to Cellodextrin/Cellobiose transporter (MFS; SP)	644	44557	621	493
195861	Related to Cellodextrin/Cellobiose transporter (MFS; SP)	101	31311	64	24

Figure S13. Gene regulation by humic substances. Microarray derived gene expression data were compared and ranked as ratios of expression in humic/non-humic environments (i.e. expression in compost/ Average expression in culture, casing and as fruitbodies) (see Table S15). The majority of compost-induced genes shared the promoter motif, `TC[CA][TG]G[AT][GTA]A[AC]AATCTC`; 31 occurrences of this motif was detected in 20 5'-regions of the top 33 up-regulated genes of *A. bisporus* var. *bisporus* grown on compost. The motif was discovered with the Motif-based sequence analysis tools (MEME Suite) 4.4 (http://meme.sdsc.edu/). It was also detected twice in the 5'-region of the laccase *lcc1* gene.

Name	Strand	Start	*p*-value	Sites
Prom_Abu_130896_spr1	+	298	5.93e-09	TGAAATCTCG TCCTGTGAAAATCTC GGTTCCACCT
prom_194648_SPR1	+	311	5.93e-09	TGAAATCTCG TCCTGTGAAAATCTC GGTTCCACCT
Prom_226235	-	1522	4.73e-08	GAGGGAAGAC GCCTGATAAAATCTC TTCAGCTTTG
Prom_139148_lcc2	+	1419	9.69e-08	ACCCGCGAAC TCCTGAAAAAATTTC GGGCCATACA
Prom_192849	+	1377	2.59e-07	GGTCCTCAAT TCCTGATACAATGTC TTCTCGAGCT
Prom_Abu_79685_homol212650(H97)_serine	-	1019	4.53e-07	GAAAAAGTAT TCATGATACACTCTC TAGAAAGGTA
prom_212650_protase	-	1008	4.53e-07	GAAAAACTAT TCATGATACACTCTC TAGAAGGTAA
Prom_Abu_37296_homol68978_serine	+	1157	6.87e-07	CACAGATGGA GCATGAGACAATATC GCGCTTTTAC
prom_68978_protase	+	1157	6.87e-07	CACAGATGGA GCATGAGACAATATC GCGTTTTTAC
Prom_79647	-	675	7.70e-07	AGGGTGTGAT TCCGGATAAAATCGC GGGTGAATGG
prom_146228_lcc1	+	1427	1.38e-06	GCGCGTAAAC TCCTGAAAAATTTTC GTGGATTGTG
Prom_183646	-	902	1.85e-06	TTTTGTTGCT TCATGTGAAAGGCTC TCCGTCCGCT
Prom_194576	+	1029	2.03e-06	TATATGTGGA TCAGGAAGAAATCTC GGGGAATCTT
Prom_190202_Cutinase	+	257	2.03e-06	GTCTGACTTA TCCTGTTAGAATGTC GGAGAAGCAG
Prom_Abu_85532_sslcc2	+	1148	4.01e-06	TCGTCTCGTC TTATGTGAAAATCTG CAATGTCACA
prom_149407_sslcc2	+	1145	4.01e-06	TCGTCTCGTC TTATGTGAAAATCTG CAATGTCACA
Prom_146117	+	862	4.36e-06	GGTTGGTGAA TCCTGTGCAAATCGC TTCCCTGTTG
Prom_194280_endocellulase	+	1451	4.71e-06	GTTCACAAAC TCCGGACAAAAGCTC CACAACTTTT
Prom_227728	+	1130	5.93e-06	GACCCTACCT TCTTGGTGAAATCTC GGTTTAGGTG
prom_211295_alpha-amylase	-	102	7.99e-06	GGCTCAATTA TCTTGAAACAGTATC ACGCGCGCCA
Prom_190200	+	1461	8.61e-06	TCTTCCATGA GCCGGATACAGTTTC CACATGAATA
Prom_190390	-	1446	9.26e-06	CTGTGAAGTC TCCGGTGAAAATCCG TCAAGGGAGC
Prom_195861	-	1070	1.14e-05	ACGGCGGTGT TACTGGAAAAATATC GGATGATGTC
Prom_194521	+	658	1.40e-05	ACGTACGACA TCATGGGAGATTTTC GAGACTTTTT

Figure S14. The expansion of HTP-s in *Agaricus bisporus*. Maximum Clade credibility chronogram of heme-thiolate peroxidases of the Agaricomycetes showing six independent lineages of HTPs in *Agaricus bisporus*. Branch lengths correspond to relative time. Thickened branches are strongly supported by Maximum Likelihood bootstrap and Bayesian posterior probabilities.