Genomic and transcriptomic analyses of the medicinal fungus *Antrodia cinnamomea* for its metabolite biosynthesis and sexual development

Mei-Yeh Jade Lu[a], Wen-Lang Fan[a], Woei-Fuh Wang[a], Tingchun Chen[a], Yi-Ching Tang[a], Fang-Hua Chu[b], Tun-Tschu Chang[c], Sheng-Yang Wang[d,e,f], Meng-yun Li[a], Yi-Hua Chen[a], Ze-Shiang Lin[a], Kai-Jung Yang[a], Shih-May Chen[a], Yu-Chuan Teng[a], Yan-Liang Lin[b], Jei-Fu Shaw[g], Ting-Fang Wang[h], and Wen-Hsiung Li[a,i,1]

[a]Biodiversity Research Center, Academia Sinica, Taipei 11529, Taiwan; [b]School of Forestry and Resource Conservation, National Taiwan University, Taipei 10617, Taiwan; [c]Taiwan Forestry Research Institute, Taipei 10066, Taiwan; [d]Department of Forestry, National Chung-Hsing University, Taichung 402, Taiwan; [e]Agricultural Biotechnology Research Center, Academia Sinica, Taipei 11529, Taiwan; [f]Agricultural Biotechnology Center, National Chung-Hsing University, Taichung 402, Taiwan; [g]Department of Biological Science and Technology, I-Shou University, Kaohsiung 84001, Taiwan; [h]Institute of Molecular Biology, Academia Sinica, Taipei 11529, Taiwan; and [i]Department of Ecology and Evolution, University of Chicago, Chicago, IL 60637

Contributed by Wen-Hsiung Li, September 13, 2014 (sent for review June 10, 2014; reviewed by Robert Friedman, Kenneth Wollfe, and Takashi Gojobori)

Antrodia cinnamomea, a polyporus mushroom of Taiwan, has long been used as a remedy for cancer, hypertension, and hangover, with an annual market of over $100 million (US) in Taiwan. We obtained a 32.15-Mb genome draft containing 9,254 genes. Genome ontology enrichment and pathway analyses shed light on sexual development and the biosynthesis of sesquiterpenoids, triterpenoids, ergostanes, antroquinonol, and antrocamphin. We identified genes differentially expressed between mycelium and fruiting body and 242 proteins in the mevalonate pathway, terpenoid pathways, cytochrome P450s, and polyketide synthases, which may contribute to the production of medicinal secondary metabolites. Genes of secondary metabolite biosynthetic pathways showed expression enrichment for tissue-specific compounds, including 14-α-demethylase (CYP51F1) in fruiting body for converting lanostane to ergostane triterpenoids, coenzymes Q (COQ) for antroquinonol biosynthesis in mycelium, and polyketide synthase for antrocamphin biosynthesis in fruiting body. Our data will be useful for developing a strategy to increase the production of useful metabolites.

medicinal fungus | fruiting body | triterpenes | meiosis | P450

An endemic mushroom (Polyporales, Basidiomycota) of Taiwan, *Antrodia cinnamomea* (syn. *Antrodia camphorate* and *Taiwanofungus camphorates*) is restricted to the endemic aromatic tree *Cinnamomum kanehirai* Hayata (1–3). A phylogenetic analysis using the nuclear large subunit (LSU) and internal transcribed spacer (ITS) sequences showed that the genus *Antrodia* is not monophyletic (3). *A. cinnamomea* (*T. camphorates*) is in a clade that includes *Taiwanofungus, Auriporia, Sarcoporia, Dacryobolus,* and *Amylocystis*. This clade is more closely related to other genera in the order Polyporales such as *Postia* than to the rest of the genus *Antrodia*. No genomes from this clade have yet been sequenced.

A. cinnamomea is rich in polysaccharides (4) and produces specific ergostane triterpenoids, including antcins C and K (5), zhankuic acids A, B, and C (6), and antrocamphin (7). *A. cinnamomea* has long been the most valued medicinal fungus in Taiwan, as a remedy for various diseases, including hypertension, abdominal pains, hepatitis B, and cancers (8–11). Also, the extract of *A. cinnamomea* has long been used to ameliorate the effects of alcohol intoxication. The total market value of *A. cinnamomea* products in Taiwan is estimated to be over 100 million US dollars per year (www.businesstoday.com.tw/article-content-80392-93594). For these reasons, numerous studies of its pharmacological properties have been conducted (12). Also, others have investigated the composition of polysaccharides and lipopolysaccharides and their differences between basidium and mycelium and among strains (11, 13).

The main components isolated from the basidium of *A. cinnamomea* are classified into three categories: triterpenoids (including steroids), phenolic compounds, and polyacetylenes. The lanostane-type triterpenoids are produced in both mycelium and basidium, whereas ergostanes are related to basidiomatal formation and are produced only in basidium (6, 14). Significantly, the ergostane production of *A. cinnamomea* growing on *C. kanehirai* Hayata is of higher efficiency and shows more compound variety than that growing on other trees (14, 15). As to the metabolites of mycelium, lanostane-type triterpenoids and the derivatives of maleic acid, succinic acid and ubiquinone, have been reported (15). Despite its medicinal importance, however, there has been little research on its genetics.

Currently, little is known about the mating system and the mechanism of meiotic control in *A. cinnamomea*. In Basidiomycota, the mating process is controlled by homeodomain-encoding genes (*HD*) and G protein-coupled pheromone/receptor (*P/R*) genes. The involvement of these genes in mating-type determination differs

Significance

Antrodia cinnamomea, a mushroom, has long been used as a remedy for cancer, hypertension, and hangover. However, the molecular basis of its medicinal effects is unclear and its genome has not been studied. We obtained a genome draft and conducted gene annotation. Genome ontology enrichment and pathway analyses shed light on sexual development and metabolite biosynthesis. We identified genes differentially expressed between mycelium and fruiting body and also proteins in the mevalonate pathway, terpenoid pathways, cytochrome P450s, and polyketide synthases, which may contribute to production of medicinal metabolites. Genes of metabolite biosynthesis pathways showed expression enrichment for tissue-specific compounds in mycelium and in fruiting body. Our data will be useful for developing a strategy to increase the production of valuable metabolites.

between the bipolar (one locus) and tetrapolar (two loci) mating systems (16). For example, in the tetrapolar mushroom *Coprinus cinereus*, the *A* genes in the *HD* locus promote the formation of clamp connections and regulate conjugate division of the two nuclei from each monokaryotic mating partner, whereas the *B* genes in the *P/R* locus control the cellular fusions that complete clamp connections during growth and the nuclear migration required for dikaryosis. The clamp connection is a characteristic structure found only in fertile dikaryons with compatible *A*- and *B*-gene alleles (17). Although a report indicated that members in the Polyporale order display either bipolar or tetrapolar mating systems (16), the mating system of *A. cinnamomea* has not been characterized.

This study had three purposes. First, we characterized the genome of *A. cinnamomea* to understand its gene content and genome structure. Second, we sequenced the transcriptomes of the mycelium and fruiting body to study the expression differences between mycelium and fruiting body. Third, we identified the pathways for sexual development (mating systems, clamp connection, mushroom morphogenesis, and meiosis) and for important metabolites, particularly for terpenoid biosynthesis, and also identified the genes that may contribute to the differential terpenoid biosynthesis between different strains or between tissue types or culture conditions.

Results

Genomic DNA and RNA Collection. We obtained the single-nucleated genomic DNA sample of a potent strain, called S27, for genome sequencing and functional annotation. For transcriptome analysis, we collected four samples of total RNA from two monokaryons (S27 and S32), a dikaryon mycelium (AM), and a wood-grown fruiting body (AT).

Genome Sequencing and Assembly. We characterized the genome of a single-nucleated *A. cinnamomea* strain, S27, with whole-genome shotgun sequencing using both Roche 454 and Illumina platforms (details in *Materials and Methods* and *SI Appendix*). We obtained 243.9 millions of reads and reached a genome assembly of 32.15 Mb from ~878 × coverage, comprising 360 scaffolds with 99.64% of the genome assembled in scaffolds exceeding 1 kb in length. The N50 length was 1,035 kb, the N90 length was 260 kb, and there were 11 scaffolds exceeding 1 Mb (2.22, 2.09, 1.67, 1.62, 1.62, 1.57, 1.29, 1.23, 1.22, 117, and 1.03 Mb) (Table 1 and *SI Appendix*, Table S1).

Transcriptome Sequencing and Assembly. To characterize the expressed sequences of *A. cinnamomea*, we conducted de novo transcriptome sequencing of the haploid S27 and S32 strains using Roche 454 (*SI Appendix*, Table S2) and recovered 7,503 and 7,541

isogroups from 11,622 and 19,478 isotig sequences, respectively (*SI Appendix*, Tables S2 and S3). An isotig represents an individual transcript (exons and possible UTRs), whereas an isogroup represents a primary transcript (gene) containing a group of isotig sequences (isoforms of assembled transcripts). We examined the cDNA assembly by searching the National Center for Biotechnolgy Information (NCBI) database and found significant hits to the protein-coding sequences of the polyporus fungus *Fibroporia radiculosa*, *Ceriporiopsis subvermispora*, *Trametes versicolor*, *P. placenta*, and *Dichomitus squalens*. This suggests a high quality of the transcript assembly, which then enhanced the quality of our gene predictions.

In parallel, we applied Illumina RNA-seq to analyze the quantitative transcriptome profiles of four *A. cinnamomea* tissues (stats of sequencing output, assembly, and read preprocessing are summarized in *SI Appendix*, Tables S2–S4). The highest mapping ratio was for reads from S27, followed by S32, AM (binucleated mycelium), and AT (wood-grown fruiting body).

Gene Prediction and Integration. To increase the accuracy of functional annotation, we used multiple gene prediction tools to obtain independent predictions (*Materials and Methods*). The S27 transcriptome isotigs (*SI Appendix*, Table S3) were used to integrate gene models from different predictors by weighing the alignment between isotigs and gene models using the integration pipeline shown in *SI Appendix*, Fig. S1 *B–D*. The final integration led to a collection of an annotated dataset containing 9,254 protein-coding gene models (Table 1) with gene sequence length ranging from 63 to 13,968 bp, resulting in a grand total of 13,927,215 bp, and giving rise to a gene density of 43.28% in the assembled 32,177,404 bp of S27. Among the 9,254 gene models, 9,252 contain a gene sequence >100 bp, with 8,162 models > 500 bp and 5,729 models > 1,000 bp. All annotated *Antrodia* genes are stored in the ACg database (Dataset S1), and each gene ID starts with the prefix ACg.

For the noncoding gene, one single rRNA cluster was found on scaffold 10 (Scaf10) of the S27 assembly, with the genes of 8S-18S-28S clustered in a region of 10,461 bp (*SI Appendix*, Table S19). There were 134 tRNA genes and 14 pseudogenes predicted, corresponding to 4.36% of the ACg genome assembly length (*SI Appendix*, Fig. S5 and Table S18).

Genome-Wide Functional Annotation. To conduct functional annotation of the S27 gene models, we first used the blastp (18) search of the putative protein-coding sequences against the NCBI nr (nonredundant) database.

Table 1. Summary of genome assembly and annotation of *A. cinnamomea*

Genome		Value
Length of genome assembly, bp		32,155,604
No. of scaffolds		360
Length of the largest scaffold, bp		2,218,769
Length of the smallest scaffold, bp		201
N50, bp		1,034,879
Percentage of assembly, %	Scaffolds ≥ 500 bp	99.79
	Scaffolds ≥ 1 kb	99.64
GC content, %		50.60
No. of gaps		442
Total length of gaps, bp		574,579
No. of protein-coding genes		9,254
Average protein length, aa		500.73
Length of largest protein-coding gene, bp		13,968
Length of smallest protein-coding gene, bp		63
Average exon size, bp		245.48
Average no. of exons per gene		6.13
Average intron size, bp		81.97
Average no. of introns per gene		5.13

We sorted the blast hits in *E*-value rank order. *SI Appendix*, Fig. S2, shows these results and homology-based ranking for the *Antrodia* proteins, in which the top five ranked species are all Polyporales (*SI Appendix*, Fig. S2A). We focused on the top 13 ranked species and included another medicinal polyporus fungus *Ganoderma lucidum* for comparison (*SI Appendix*, Fig. S2B). The blastP results indicate that *A. cinnamomea* proteins share high sequence homology with those of *F. radiculosa*, but it is unknown whether this is evidence of taxonomic relatedness or a common origin of their wood-degrading capability (Dataset S1).

In summary, 8,911 of the 9,254 ACg genes have BLASTp hits (96.3%) with an *E*-value $<10^{-5}$ and 8,717 genes (94.2%) with an *E*-value $<10^{-10}$ (*SI Appendix*, Table S5), suggesting the high quality of our genome assembly and gene prediction. We then analyzed the gene ontology using AmiGO for functional enrichment. There were 9,233, 11,000, and 13,762 annotations obtained for the three main categories of cellular component, molecular function, and biological process, distributed among a total of 6,396 gene models (69.1%) of the S27 dataset (*SI Appendix*, Table S5).

Mating Systems, Clamp Connection, Mushroom Morphogenesis, and the Core Meiotic Genes.

The mating system of *A. cinnamomea* has not been characterized, but a recent report indicated that members in the Polyporale order display either bipolar or tetrapolar mating systems. For example, *P. placenta* has a bipolar mating system (19), whereas *F. radiculosa* (20), *Sarcoporia*, *Dacryobolus*, and *Amylocystis* (3) all display a tetrapolar mating system. We found that in the *A. cinnamomea* genome there is a putative *HD* locus on Scaf3 with four homeodomain genes (ACg001308–10 and ACg001316) and a *P/R* locus on Scaf6 with four Mfa (mating pheromone a factor)-like genes (ACg003186, ACg003200, and ACg003207–8) and six Ste3-like pheromone receptor genes (ACg003202–6 and ACg003212) (16), indicating that *A. cinnamomea* may display a tetrapolar mating system with multiple alleles of the mating-type genes.

Pcc1 (*pseudoclamp connection*, ACg003357) and *clp1* (*clampless1*, ACg005835) are two genes involved in the formation of clamp connections in dikaryon. In *C. cinereus*, the *pcc1* gene encodes a high mobility group (HMG) box transcriptional factor that represses *A*-regulated clamp connection genes in the monokaryon (21). The *clp1* gene encodes a Basidiomycota-specific protein that represses the *pcc1* expression in the dikaryon (22).

Like other higher basidiomycetes, *A. cinnamomea*'s sexual development undergoes a highly organized process to generate fruiting bodies (mushrooms). Our Blast search also identified orthologs of *Coprinopsis cinereus* mushroom morphogenesis genes, including *cfs1* (*cyclopropane fatty acid synthase 1*, ACg007212), *dst1* (*WC1 blue-light receptor*, ACg004402), *dst2* (*WC2 blue-light receptor*, ACg005051), *eln2* (*elongationless 2*, ACg006982), *eln3* (*elongationless 3*, ACg007250), and *exp1* (*expansionless1*, ACg007250) (23). These results suggest that *A. cinnamomea* (a polyporus mushroom) and *C. cinereus* (an agaric or gilled mushroom) apparently share a highly conserved mechanism in fruiting body development.

We also surveyed the *A. cinnamomea* genome for the core meiotic genes (24) for the proteins that constitute the conserved meiotic machinery in eukaryotic organisms (particularly *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*). Of the 74 *S. cerevisiae* and/or *S. pombe* meiotic genes surveyed, 57 were found to have homologs in the *A. cinnamomea* genome (*SI Appendix*, Table S6 and Dataset S1), including those that are "meiosis-specific" and hypothesized to be present only in organisms with sexual reproduction or with sexual ancestry. The *A. cinnamomea* genome encodes several evolutionarily conserved proteins required for assembling the synaptonemal complex, e.g., *mer3* (ACg003801), *msh4* (ACg005375), *msh5* (ACg002475), and *zip4/spo22* (ACg005454) (*SI Appendix*, Table S6 and Dataset S1), suggesting that *A. cinnamomea*, like most sexually reproductive organisms, forms the synaptonemal complex to mediate chromosomal synapsis, allowing interacting homologous chromosomes to complete their crossover activities during meiosis (25).

Terpenoid Pathway Gene Search and Analysis.

It has been reported that the mycelium and fruiting body of *A. cinnamomea* produce common secondary metabolites as well as unique medicinal compounds (11). For example, antrocamphins are abundant in mycelium, and lanostane-type triterpenoids are present in both tissues, whereas ergostanes-type triterpenoids are unique to the fruiting body. Analysis of enzymatic pathways in the S27 genome using the KEGG (Kyoto Encyclopedia of Genes and Genomes) map shows that the highest number of genes in KEGG pathways are involved in purine metabolism (Fig. 1), for which the gene and enzyme counts are much higher than those in the other categories (*SI Appendix*, Fig. S3). In addition, we found many hits in the mevalonate branch of the terpenoid backbone biosynthesis pathway. This result is similar to that of *G. lucidum* (26, 27), except for some enzymes downstream of the geranyl diphosphate synthase (GPP) and farnesyl diphosphate synthase (FPP) biosynthesis steps. We also found genes encoding components in the biosynthesis pathways of terpenoid backbone, sesquiterpenoid, ubiquinone, and other terpenoid quinones (*SI Appendix*, Table S7 and Dataset S1). Significantly, we identified many P450-related enzymes in the cytochrome P450 xenobiotic and drug metabolism pathways. Altogether, we found 24, 22, and 20 enzymes of the xenobiotic P450, drug metabolism P450, and terpenoid backbone biosynthesis pathways, including 7 of the *O*-methyltransferase enzymes (*SI Appendix*, Table S7).

To search for genes involved in terpenoid biosynthesis, we collected terpenoid-related protein sequences from the NCBI database. We ran blastP of this query dataset against the S27 genome annotation file and identified 242 sequences putatively coding for genes involved in the triterpenoid biosynthesis of *A. cinnamomea* (*SI Appendix*, Table S12). We identified genes for the terpenoid backbone synthesis, including the single-copy gene of mevalonate kinase and that of HMG-CoA synthase, the two upstream core enzymes in the pathway. In addition, many genes involved in fatty acid synthesis (particularly, the enoyl-CoA reductase) were identified, suggesting the importance of fatty acid synthesis in this medicinal fungus. Also, we identified candidate genes of the secondary metabolite category, including 24 monoterpene synthases, 7 sesquiterpene cyclases, and 119 cytochrome P450s (*SI Appendix*, Tables S7 and S8). Further details are given below.

Cytochrome P450s.

To find P450 genes, we searched for homologs to the P450 protein sequences in the 27 domain classes in the Conserved Domains Database (www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). Based on domain sequence similarity, 81 putative P450 proteins were identified and classified into 14 types, among which the "cytochrome_P450" domain type included 42 P450 candidates (Fig. 1B). Furthermore, we collected 21 additional gene IDs in the P450-related pathways of "metabolism of xenobiotics by cytochrome P450" and in "drug metabolism-cytochrome P450" in KEGG analyses. In total, we obtained 102 gene IDs for the P450 search.

To broaden the search, we conducted Blast search for CYP proteins using the 2,753 query protein sequences from the Fungal Cytochrome P450 Database (p450.riceblast.snu.ac.kr/index.php?a=view). This led to the identification of 96 proteins, which were classified into 39 CYP families (*SI Appendix*, Tables S9 and S10). The results show the distribution of the candidates in the CYP families characteristic for fungal species, with higher counts in CYP5150, CYP63, and CYP512 (Fungal Cytochrome P450 Database). Combining the putative P450 proteins from KEGG pathways, P450 domain hunting, and CYP family analysis, we found a total of 119 candidate P450 genes (*SI Appendix*, Table S12).

In particular, we identified many monooxygenase candidates of CYP5144 and CYP504 (6 and 10 candidates, respectively). The 10 CYP504 genes (7 monoterpene synthases, 2 enoyl-CoA reductases, and one 1-deoxy-D-xylulose-5-phosphate_synthase) (*SI Appendix*, Table S9) were not reported for *G. lucidum*, *P. placenta*, and *Phanerochaete chrysosporium*. The CYP504 enzymes (phenyl-acetate-oxidizing Cytochrome P450, EC 1.14.13) were reported to catalyze phenylacetate catabolism (28) and may participate in the diverse functions of detoxifying xenobiotic compounds (29), so they
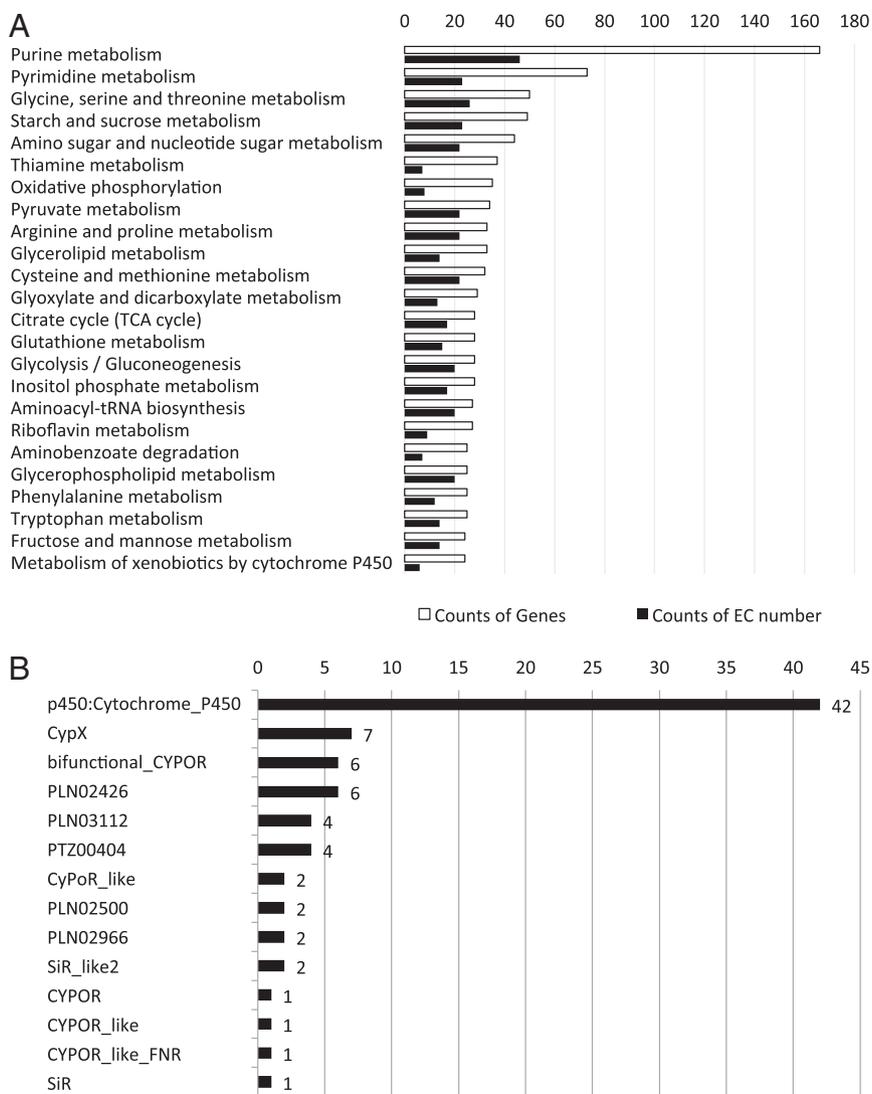
A

☐ Counts of Genes ■ Counts of EC number

B

Fig. 1. Functional categorization of *A. cinnamomea* gene models. (*A*) Distribution of genes in different KEGG categories. White bars indicate counts of EC numbers and solid black bars indicate gene counts. (*B*) Distribution of P450 genes in different cytochrome P450 domain categories.

may likely contribute to the unique growth of *A. cinnamomea* on trees of *C. kanehirai* Hayata, which inhibit the nearby growth of other fungi (30).

A single candidate, ACg0004767, of the CYP6001 family has a high sequence similarity to lactoperoxidase and is likely involved in phenylalanine metabolism and phenylpropanoid biosynthesis for antroquinonol coumpounds. It contains a CypX domain for oxidative degradation of environmental toxins by haem-thiolate proteins and a linoleate diol synthase domain of fungal protein similar to animal heme peroxidases. In addition, extracts of *A. cimmamomea* mycelium were reported to have a protective effect from oxidative stress for hepatic cells (31) and a clinical effect against non-small-cell lung cancer (12).

COQ enzymes for isoprenoid CoQ biosynthesis. To understand the wood-decaying nature and production of antroquinonol compounds of *A. cinnamomea*, we searched for enzymes involved in the biosynthesis of coenzyme Q10 for ubiquinone bioproduction, extending the isoprenoids from C5 (5-carbon compound) to C30 or greater. Using the query sequences of COQ1~COQ9 in Kawamukai (32), we identified 11 putative COQ proteins in the *A. cinnamomea* genome, most of which are encoded by single genes, except there are four putative genes for COQ3, whereas, COQ6 and COQ7 are putatively encoded by the same gene (*SI*

Appendix, Table S11). We verified these candidates by blastp search against the nr database. Significantly, all COQ candidates have high similarities to their counterparts in wood-decaying fungal genomes (*E*-value $<10^{-89}$), particularly to those of *Dichomitus squalens* and *Trametes versicolor* (33). These enzymes are part of the "ubiquinone and other terpenoid-quinone biosynthesis" pathway (KEGG map00130). The expression levels of these COQ genes, except for COQ1, were noticeable in both AM and AT, but were much lower in the fruiting body (AT) than in the mycelial (AM) tissues. This finding is consistent with the much higher content of antroquinonol in agar-cultured mycelium than in wood-grown fruiting body (34).

Terpenoid skeleton and triterpenoid biosynthesis. We identified a putative friedelin synthase [2,3-oxidosqualene cyclase (OSC); EC 5.4.99.50] in the pentacyclic triterpene biosynthesis pathway, which converts 2,3-oxidosqualene into pentacyclic triterpenoid skeletons. The protein sequence is highly similar to the lanosterol synthase (LSS) of *G. lucidum* (26), and significantly, the LSS gene showed the highest expression in the wood-grown fruiting body among the four RNA samples.

A candidate gene, ACg002141, for demethylating lanosterol involved in the ergosterol synthesis pathway, was identified. It belongs to the CYP51F1 (CypX) family and has a high similarity to AcCyp51, which encodes a cytochrome P450 sterol

14-α-demethylase cloned from *A. cinnamomea*. AcCyp51 was shown to have the demethylation activity of lanosterol, converting lanostane-type triterpenoids to ergonstane-type triterpenoids from C30 to C29 compounds by removing one methyl group, and a higher expression level in fruiting body than in other tissues (35). The CYP51F1 homolog of *Candida albicans* has the function of demethylating lanoterol and is capable of binding several azole antifungal agents (36). ACg002141 showed three- to fivefold higher expression in the binucleated mycelium and fruiting bodies relative to the two monokaryotic strains examined. Although there is the possibility of karyon-type–specific expression, it is likely that this CypX candidate plays a role in ergostane biosynthesis and involvement for fruiting body formation.

In addition, we identified a putative lignin peroxidase gene (EC 1.11.1.14), *ACLnP*, which was reported for its extracellular enzymatic function in biosynthesis of antcins in *A. cinnamomea* (37). We also identified a putative manganese peroxidase (EC 1.11.1.13) (*SI Appendix*, Table S7). These two enzymes may contribute to the wood degradation and colonization capacity of the fungus in its unique habitat.

Polyacetylene biosynthesis. The biosynthesis of antrocamphins (which are polyacetylenes) in *A. cinnamomea* requires polyketide synthases (PKS). By homology search we identified 14 putative PKS-encoding genes, which can be largely separated into two groups by protein lengths. One group contains four PKS enzymes that are over 2,000 aa in length and consist of domains characteristics of the multimodular type I PKS, such as PKS (cd00833), PKS_AT, PKS_KS, and acyl-carrier protein domains (FabD, hot_dog superfamily, acpP domain). The other group contains 10 smaller proteins of lengths ranging from 316 to 411 aa, consisting of domains such as PKS_ER (enoylreductase), MDR, and ADH_zinc, but no acyl-carrier protein domains (*SI Appendix*, Table S8 and Fig. S4*B*). Interestingly, a recent study of the polyketide synthase gene *pks4* found it essential for sexual development and regulating fruiting body morphology in *Sordaria macrospora*. Mutants overexpressing the *pks4* developed enlarged, malformed fruiting body (38).

The latter step of the enzymatic reaction by polyketide cyclase is potentially carried out by the single gene for a putative olivetolic acid cyclase (OAC; EC 4.4.1.26) encoded by a single gene, ACg002677. OAC has been reported for polyketide biosynthesis in plants and is similar to the olivetolic acid cyclase for cannabinoid biosynthesis in *Cannabis sativa* (39).

Differential Gene Expression Between Transcriptomes.

We studied expression differences between transcriptomes of the single-nucleated (monokaryon) mycelia of S27 and S32, a dikaryon mycelium (AM), and a wood grown fruiting body (AT). We first confirmed that only AM, but not S27 or S32, formed a clamp connection, a bulgelike hyphal outgrowth structure for maintaining the binucleate stage (*SI Appendix*, Fig. S7*A*). The *C. cinereus clp1* gene is specifically expressed in a dikaryon, resulting from mating of two compatible monokaryons, and the Clp1 protein regulates dikaryon-specific clamp development (22). Here, the results of de novo transcriptome sequencing experiments further confirmed that the *A. cinnamomea clp1* gene (ACg005835) was expressed at least 10-fold higher in AM and AT than in the two bona fide monokaryons (S27 and S32) (*SI Appendix*, Fig. S7*B*).

SI Appendix, Fig. S3*A*, shows the variance scatter plot for pairwise comparisons. The expression variation was the smallest between the S27 and S32 monokaryon mycelia and, as expected, was the greatest between the monokaryons and the wood-grown fruiting body. Similarly, the highest expression correlation was between the two monokaryons, S27 and S32 (Fig. 2*A*).

The genes expressed across four samples are shown in Fig. 2*B*. Among the 9,254 gene models of *A. cinnamomea*, 8,079 were expressed in at least one sample and 7,741 were expressed in all tissues, indicating that our sequencing was sufficiently deep for gene expression detection (Fig. 2*B*). A total of 25 genes were expressed only in AM or AT, with 12 genes expressed in both AM and AT. In comparison, 98 genes were expressed only in the monokaryon mycelia. There are 9 genes that were expressed only

in the wood-grown fruiting bodies (ACg006161, ACg002837, ACg000902, ACg004753, ACg008058, ACg000033, ACg004031, ACg004725, and ACg005279), and it will be of interest to determine their functions in the fruiting body.

We further looked at the function of the 25 genes uniquely expressed in AT or AM (*SI Appendix*, Table S13). There are 23 genes having homologous proteins from 14 species in the NCBI nr database. Eleven of the 14 species are wood-grown basidiomycetes. The top hit species is *T. versicolor* (*Polyporus versicolor*), whose mushroom extract reportedly has antitumor effects (40). Also, for the 98 genes uniquely expressed in S27 or S32, most of the hit species are also wood-grown basidiomycetes.

Among the genes expressed uniquely in the wood-grown fruiting body, a putative transcription factor coded by the ACg004725 gene model is predicted to contain two zinc-finger BTB domains (pfam00651) located at the N terminus and midbody of the protein, respectively. The BTB/POZ domain has been reported in mediation of protein–protein dimerization, regulation of cytoskeleton dynamics, and transcriptional regulation (41–43). The mutations at the dimer interface were found to result in misfolded proteins (the single-pocket mutation R49Q repressed transcription, whereas the double-mutant D35N/R49Q could not) (44). The *A. cinnamomea* BTB-containing protein has homologs in many fungi grown on decayed wood, including *P. placenta* and *T. versicolor* which is also a medicinal fungus full of triterpenoids. In contrast to the report that BTB is present as a single-copy domain in the 17 nonfungal eukaryotes surveyed, ACg004725 in this study and its homologs from *P. placenta* and *T. versicolor* have the multiple BTB domains (e.g., accession ID XP_002471622 and EIW58581.1).

Most of the genes in the terpenoid backbone synthesis pathway were differentially expressed (17 of 20) (Fig. 3). We classified the triterpenoid biosynthesis-related genes that were differentially expressed between any two of the four transcriptomes into four clusters of expression patterns (Fig. 4). Cluster C1 (34 genes) represents genes with enriched expression in AM, C2 (36 genes) in AT, C3 (67 genes) in S32, and C4 (48 genes) in S27.

In addition, we studied the clustering of putative P450 genes and the set of all differentially expressed genes (DEGs) (*SI Appendix*, Fig. S3*B* and Table S14). Clustering by Z-scores showed categorization of the 92 differentially expressed P450 genes into four clusters, as in the case of all 4,898 DEGs (*SI Appendix*, Fig. S3*B*), suggesting the unique combinations of DEGs in these four transcriptomes required for their genome natures (single vs. binucleated, S27 vs. S32, mycelium vs. fruiting body, agar culture vs. wood grown).

Genome Ontology Enrichment for DEGs in Different Transcriptomes.

The genome ontology (GO) enrichment test was applied to the clustered DEGs of the triterpenoid pathway, using two criteria: (*i*) the FDR *P* value is $\leq 1 \times 10^{-4}$ (45) and (*ii*) at least two genes in a cluster have the same GO assignment. The analysis classified these genes into 16 GO categories, among which the wood-grown fruiting body (AT) showed the broadest representation, and the binucleated mycelium (AM) showed the fewest number of GO categories (Fig. 4).

Among the enriched GOs, the categories of electron carrier activity, heme binding, iron ion binding, and oxidoreductase activity showed enrichment in all four transcriptomes, indicative of their basic functions in *A. cinnamomea*. The GO groups of secondary metabolism and trichodiene synthase activity are enriched in the fruiting body and the binucleated sample. In comparison, several gene clusters showed unique enrichments in specific transcriptomes. Among them, the C3 cluster (S32 mycelium) is enriched for glutathione transferase, whereas C4 (S27 mycelium) is enriched for the glutamate biosynthetic process and flavin-related oxidoreductase activity (Fig. 4 and *SI Appendix*, Tables S15 and S16). Interestingly, C2 (AT fruiting body) is enriched for the DIM/DIP cell-wall layer assembly (*SI Appendix*, Table S15), indicating the unique cell-wall organization of the wood-grown fruiting body on *C. kanehirai* trees for special growth and metabolism,
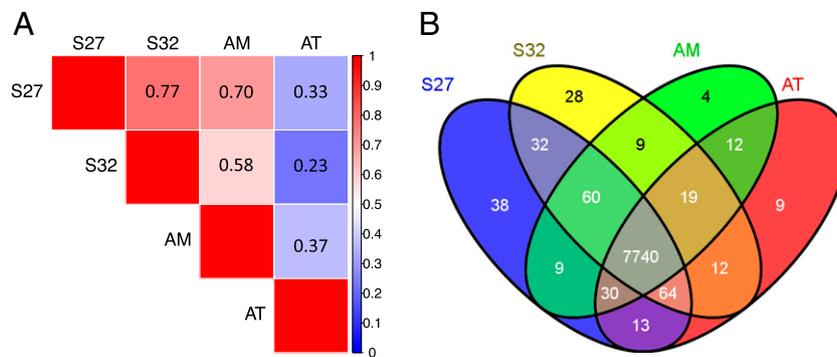
Fig. 2. Relationships between the four transcriptomes of *A. cinnamomea*. (*A*) Pairwise correlation of normalized RPKMs between RNA samples. The Pearson correlation coefficient ranges from no correlation (blue) to perfect correlation (red). S27 and S32, single-nucleated mycelia; AM, binucleated mycelium; AT, wood-grown fruiting body. (*B*) Venn diagram of 8,079 expressed genes in four *Antrodia* transcriptomes.

and C3 (AM mycelium) is enriched for the ergosterol biosynthetic process and the sesquiterpene biosynthetic process. Moreover, similar to the fruiting body AT, S27 (but not S32) mycelium is enriched for steroid metabolism, a correlation worthy of a deeper investigation.

Physical Clustering of P450 Genes in the Genome. We examined the physical clustering of P450 genes and triterpenoid pathway-related genes in the genome and found nine gene clusters, which are distributed on six scaffolds (*SI Appendix*, Fig. S4*B*). In particular, two clusters (each consisting of four genes) are on Scaf2, and a cluster of six genes is on Scaf14. Some genes lie within a short region, such as the cluster of five genes on Scaf4 within a 26.6-kb region and another one on Scaf4 within a region of only 12.7 kb. Among the 31 genes in the nine physical clusters, 28 genes were differentially expressed (~90.3%), in contrast to the proportion of 60.6% (4,898 DEGs) in the 8,079 expressed genes.

Some of the genes in the same cluster showed a high correlation of expression pattern ($pcc > 0.9$ or $pcc < -0.9$). For example, on Scaf2, $pcc = -0.91$ between ACg000810 and ACg000822 and $pcc = 0.91$ between ACg001140 and ACg001143. Also, on Saf14 ACg005636 is highly correlated to both ACg005615 ($pcc = 0.93$) and ACg005625 ($pcc = 0.96$). The correlation between ACg001953 and ACg001959 (on Scaf4) is the highest ($pcc = 0.99$) [RPKMs (reads per kilobase of transcript per million reads mapped) are shown in *SI Appendix*, Table S23]. A negative correlation may mean strong expression differentiation between tissues or media, whereas a positive correlation may mean coexpression.

Repeat Analysis of the S27 Genome Scaffolds. The S27 genome contains 17.7% (~5.7 Mb) repetitive sequences, among which the LTR sequences account for 6.98% (*SI Appendix*, Table S20). The high repeat content rendered de novo genome assembly challenging. Because of high repeat content, some regions showed a very high level of read coverage (>2,000× compared with the average 200×). Many regions with low GC% show a very high level of coverage.

Genome Variation Between Two Mycelium Strains. To characterize the differences between the S27 and S32 genomes, we carried out variant detection analyses by mapping the preprocessed reads from S32 with different library insert lengths (average of 237 and 481 bp) on the S27 genome scaffolds and used variant calling and annotation algorithms to assess functional impact among these sites. Almost 60% of the reads remained after the preprocessing for both of the single-nucleated datasets, leaving 201× and 347× of the genome coverage fold for the two samples, respectively (*SI Appendix*, Table S21). More than 93% of the S32 reads and 98% of S27 reads were mapped onto the S27 genome scaffolds in a properly paired fashion, indicating the high quality of the data and assembly accuracy.

The genomic differences between S27 and S32 include 98,930 SNPs and 7,715 insertion/deletions (INDELs). The variants are categorized according to their positions on the genomic region and classified for their functional impacts (*SI Appendix*, Fig. S6). The snpEFF analysis showed that 33.91, 15.93, and 50.15% of the SNP variants and 8.58, 15.66, and 75.76% of the INDEL variants occur on exonic, intronic, and intergenic regions, respectively. The functional analysis shows that 44.36% of the 33,511 exonic SNPs located in the genic regions potentially have a deleterious impact on protein functions (43.70 and 0.66% for missense and nonsense, respectively). About 50.15% of exonic INDELs would cause a frameshift (*SI Appendix*, Fig. S6).

Our SNP density analysis indicated that the average number of SNPs per kilobase between S27 and S32 is 2.41 in the exonic regions, 3.58 for the nonexonic regions, and 3.07 for the whole genome (*SI Appendix*, Table S22). Thus, the genetic difference between S27 and S32 is about three times higher than that between two humans. In addition, 7,715 INDELs were found between the two genomes. The difference between the two genomes is not high, but may be sufficient to cause substantial gene expression differences between the two strains.

Discussion

In this study, we carried out de novo genome sequencing, assembly, and annotation of *A. cinnamommea*, a polyporus mushroom. The genomes of two other mushrooms, *P. placenta* and *F. radiculosa*, have been sequenced (46, 47). These two brown rot fungi are commonly responsible for the destructive decay of wood in buildings and timbers in the field (46, 47). Moreover, both fungi secrete high levels of oxalate and thus show high resistance to copper-based wood preservatives (47). In contrast, *A. cinnamommea* is an endemic mushroom that is restricted to the endemic aromatic tree *Cinnamomum kanehirai* Hayata in Taiwan. The economic value of *A. cinnamomea* is mainly due to its medicinal use (8–11). According to Ortiz-Santana et al. (3), *A. cinnamomea* (*T. camphorates*) belongs to a clade that includes *Taiwanofungus*, *Dacryobolus*, and other related genera. This clade is distinct from that of *P. placenta*, which is a diverse clade including *Rhodonia*, *Amyloporia*, and the rest of *Antrodia* and is even more divergent from *F. radiculosa*. The lack of a closely related genome and the existence of many short exons were obstacles in the functional annotation of this genome. The intron prediction accuracy was improved by using the 454 transcriptome assembly during integration of gene prediction datasets.

We conducted GO enrichment and pathway analyses, providing insights into sexual development and the secondary metabolite biosynthesis pathways and the wood-degrading nature of *A. cinnamommea* on its unique tree habitat. Our results reveal that *A. cinnamomea* and *C. cinereus* might share a mechanism in sexual development, although their mushrooms display polyporus and gilled structures, respectively. Comparative functional
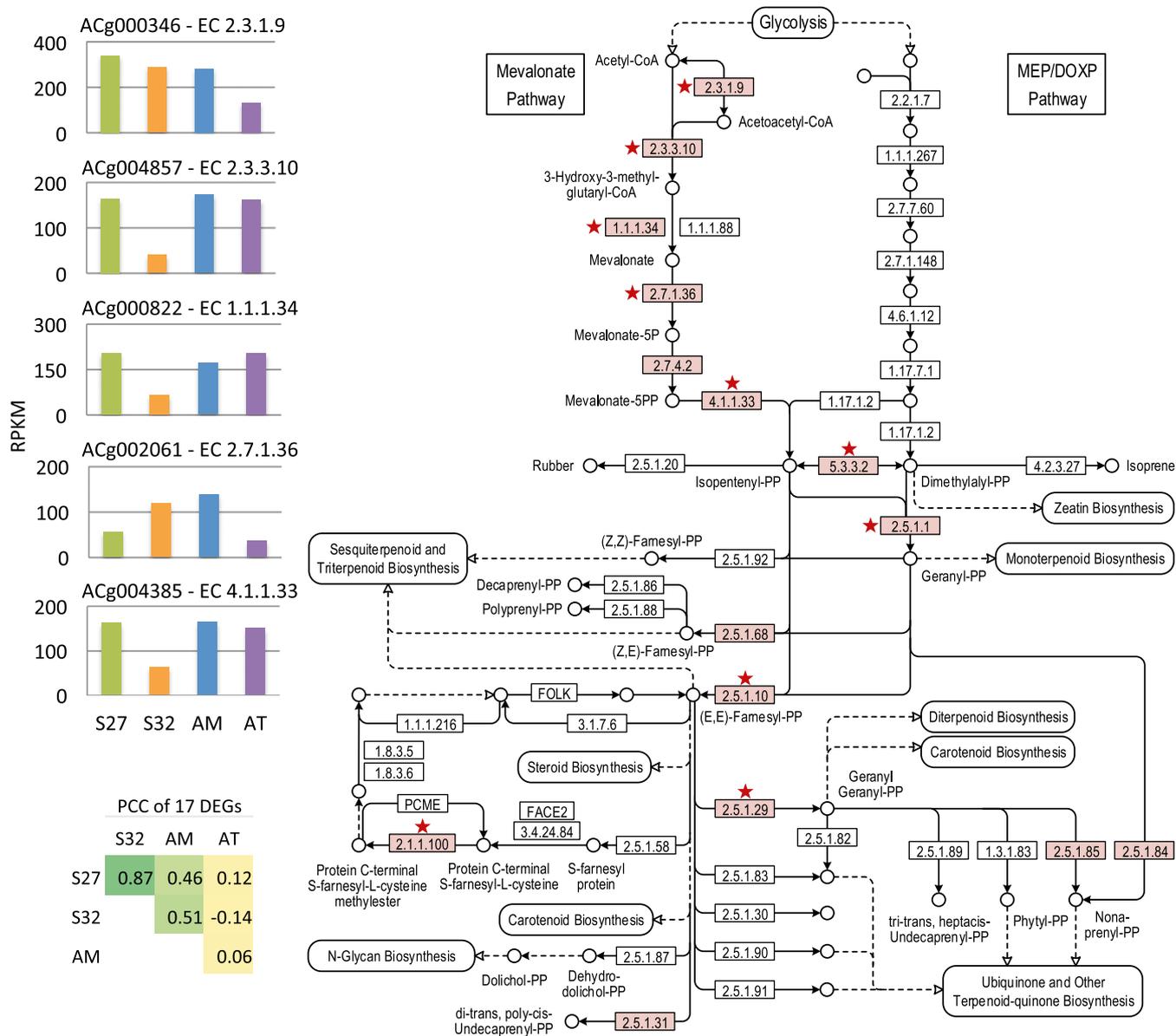
Lu et al.

Fig. 3. KEGG mapping of terpenoid backbone biosynthesis pathway and differential expression of the enzymes identified in *A. cinnamomea*. (*Right*) KEGG map 00900. Red stars indicate the hits of differentially expressed genes in this map. (*Upper Left*) The expression levels (RPKM values) of mapped genes (EC 2.3.1.9, EC 2.3.3.10, EC 1.1.1.34, EC 2.7.1.36, EC 4.1.1.33) of different strains. (*Lower Left*) Pearson correlation coefficients of the 17 mapped genes in the terpenoid backbone biosynthesis pathway.

analyses between these higher basidiomycotes may provide insights to improve cultivation of *A. cinnamomea* fruiting bodies.

We identified many genes encoding protein candidates for the biosynthesis of bioactive compounds in *A. cinnamommea* (Fig. 5 and *SI Appendix*, Fig. S8). These include the important triterpenoids (especially ergostanes, C-29 type), which are synthesized via the mevalonate pathway followed by cyclizing squalene into the lanostane triterpenoid skeleton (C-31) and subsequent demethylation into ergostanes and modifications into various triterpenoids, including antcins. Monoterpene synthases and sesquiterpene synthases, the enzymes converting the mevalonate pathway intermediate compounds to monoterpenes and sesquiterpenes, were encoded by multiple gene copies in the genome. A search for the enzymes responsible for converting isoprenes into ubiquinonol and ubiquinones led to identification of genes for the COQ enzymes in the ubiquinone biosynthesis pathway. In addition, we found several type III polyketide synthases and one

gene for polyketide cyclase putatively involved in the biosynthesis of antcamphins.

Some of the above key enzymes showed expression preference for the compounds they synthesize or their biological role. For example, 14-α-demethylase showed high expression in the dikaryon mycelium and fruiting body, suggesting its role in basidiomatal formation (35) and correlation with the high content of ergostanes-triterpenoids in fruiting body. Uniquely, the five genes encoding CYP512 P450 enzymes and the three genes encoding CYP5140 enzymes all showed highest expression in fruiting body, so they are possibly involved in the modification of ergostane compounds.

All COQ protein candidates (except COQ1) have much higher expression in mycelium than in fruiting body, consistent with the much higher content of antroquinonol in mycelium (31). In contrast, the three PKS III genes showed high expression in the fruiting body, leading to the high content of antcamphin (7).

In summary, our study provides a comprehensive collection of key enzymes in the production of the main bioactive metabolites
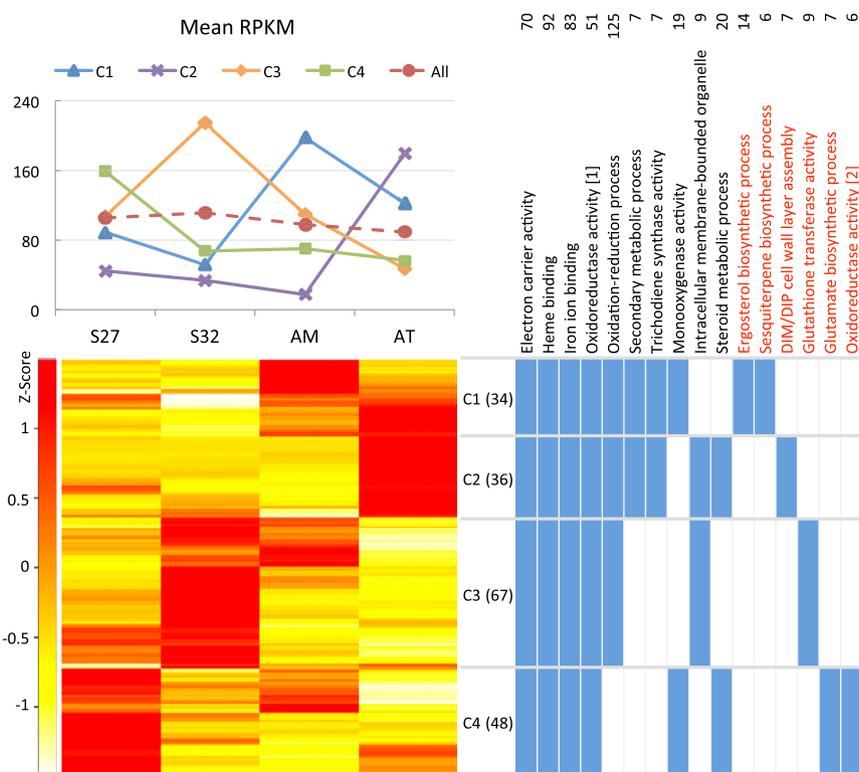
Lu et al.

Mean RPKM

C1   C2   C3   C4   All

70 92 83 51 125 7 7 19 9 20 14 6 7 9 7 6

Electron carrier activity
Heme binding
Iron ion binding
Oxidoreductase activity [1]
Oxidation-reduction process
Secondary metabolic process
Trichodiene synthase activity
Monooxygenase activity
Intracellular membrane-bounded organelle
Steroid metabolic process
Ergosterol biosynthetic process
Sesquiterpene biosynthetic process
DIM/DIP cell wall layer assembly
Glutathione transferase activity
Glutamate biosynthetic process
Oxidoreductase activity [2]

C1 (34)
C2 (36)
C3 (67)
C4 (48)

Fig. 4. Expression profile clustering of triterpenoid biosynthesis genes in *A. cinnamomea*. The four RNA samples were from the mycelia of S27 and S32 (two single-nucleated isolates), the mycelium of AM (a binucleated isolate), and the fruiting body of AT (a wood-grown isolate). A total of 185 differentially expressed genes are clustered into four groups: C1, C2, C3, and C4. (*Lower Left*) Clustering of expressed gene models (with RPKMs transformed to Z-values) by R script. (*Upper Left*) Median RPKM values of the genes in the four clusters. (*Upper Right*) GO categories enriched in the genome-wide annotation of S27. The two oxidoreductase activity categories are the following: (*i*) oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, and (*ii*) oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen. (*Lower Right*) Clusters showing expression enrichment of the GO categories are shown in blue.

of *A. cinnamomea*. The data will be valuable for the biotech industry in producing and packaging these metabolites for commercial applications.

Materials and Methods

Methods for DNA/RNA extraction, methods for transcriptome analysis and identification of differential gene expression, and some other methods are described in *SI Appendix*.

Origins of Strains and Culture Conditions. A fruiting body of *A. cinnamomea* was collected from a native *C. kanehrai* tree in Kaohsiung, Taiwan. Basidiospores of different karyotypes that resulted from meiotic recombination were isolated and subjected to a mating test. Among the single-basidiospore isolates, strains S27 and S32 (pure monokaryons) were mated and became the dikaryon strain B496, which was incapable of undergoing nuclear fusion. This phenotype is consistent with our parallel finding of the tetrapolar mating system of *A. cinnamomea*. The natural fruiting bodies (AT) were obtained directly from rotten wood. Vegetative monokaryon mycelia S27 and S32 and dikaryon mycelium B479 (AM, which can undergo basidio-matal formation) were produced by inoculation on potato dextrose agar (Bacto) medium and incubated at 24 °C in darkness for 1 mo. The mycelial layer at the top of culture was then scrapped off the petri dish. All four collected strains were frozen in liquid nitrogen and stored at −80 °C.

High-Throughput Sequencing. Two types of NGS platforms (Roche 454 and Illumina) were used. All libraries were prepared using the manufacturer's reagents and protocols. The data from the two platforms were used in a complementary manner. For de novo genome sequencing, Roche 454 GS FLX+ was used to generate long reads (400–1,000 bp) for primary de novo assembly and for long-distance paired-end reads for long-range scaffolding. Illumina sequencers (GAIIx, HiSeq2000, and MiSeq) were used to generate massive amounts of shorter reads for base correction and contig-end extension. For bridging over gaps with distances ranging from 2 to 15 kb,

mate-pair data were generated with a stepwise increment of the jump distance at 2–4, 6–4, 6–10, and 10–15 kb using the Illumina platform. For transcriptome sequencing, long reads of RNA-seq data were generated by Roche 454 for primary cDNA assembly (454 titanium for S27 and 454-Plus for S32), and large amounts of shorter reads (75–150 bp) were obtained using Illumina for quantitative profiling. For quantitative transcriptome, all four RNA samples were subjected to RNA-seq using Illumina sequencing on HiSeq2000 by PE2*100nt.

Genome and Transcriptome Assembly. The 454 shotgun and paired-end reads were assembled by the GS De novo Assembler version 2.6 (Newbler, Roche). Newbler contig base-calling errors and INDELS/homopolymer errors were corrected by Illumina paired-end reads. Possible PCR duplicates in Illumina mate-pair reads were removed by MarkDuplicates in Picard tools (picard.sourceforge.net), and possible paired-end read contamination in Illumina mate-paired reads was removed by BWA v6.2 (48) mapping result. Subsequent scaffolding of the 4,018 Newbler contigs was conducted by SSPACE v2.0 (www.baseclear.com/landingpages/basetools-a-wide-range-of-bioinformatics-solutions/sspacev12/) using the Illumina paired-end, mate-paired (3.67- and 5.42-kb fragments) and 454 paired-end (20-kb fragments) data. Gaps inside the scaffolds were closed with Illumina paired-end and mate-paired data using GapCloser (soap.genomics.org.cn/about.html). To overcome potential assembly redundancy arising from the gap closure process, sequences that were aligned to another sequence over 97% of the length by BLAT were removed from the assembly. Transcriptome assembly was carried out using the Roche 454 Newbler software to assemble the RNA-seq reads into isotigs (cDNA).

Gene Prediction and Integration. Whole-genome annotation was carried out in two phases: (*i*) the overlapping gene sets were subjected to the integration pipeline (see details below), and (*ii*) the gene models uniquely predicted by a single tool were selected by blast hits to the NCBI nr database (*E*-value < 1×10^{-5}).
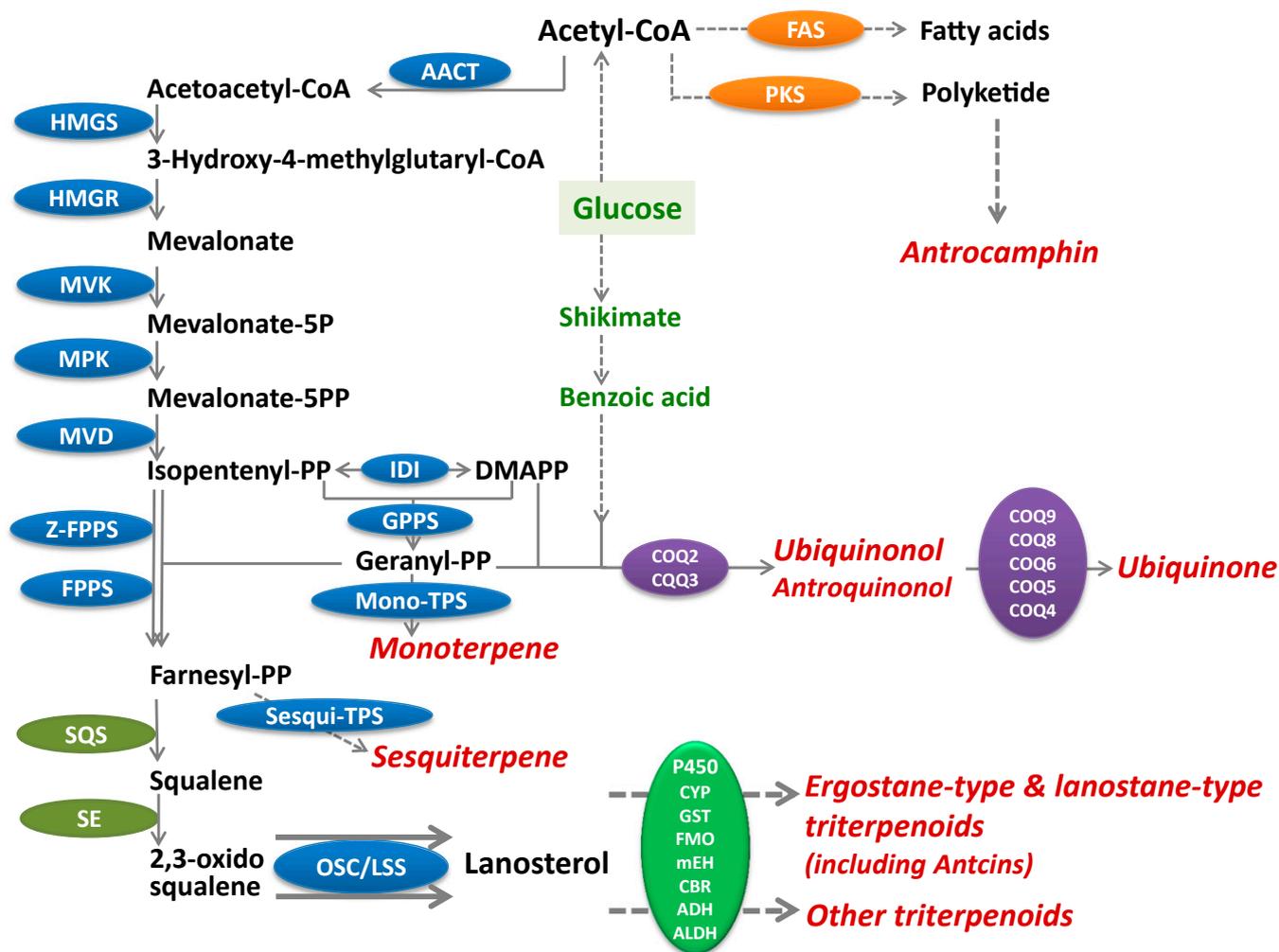
Fig. 5. Representative secondary metabolites and biosynthesis pathways of triterpenoids and antrocamphin of *A. cinnamomea*. The enzymes are differentially colored by their corresponding pathways: blue, KEGG map00900; olive, map00909; purple, map00130; green, map00980 and map00982; orange, putative antrocamphin pathway; red, various modifying enzymes converting triterpenoid skeletons into different bioactive compounds. Three enzymes, AACT, IDI, and GPPS, all showed significantly higher expression in the S27 mycelium than in the AT fruiting body. OSC/LSS showed the highest expression in AT among all four transcriptomes.

Phase I. Genes were predicted by applying two de novo prediction tools GeneMark-ES v2.3e (49) and FgeneSH of MolQuest v2.4.3.1111 (www.molquest.com/molquest.phtml; MolQues –Bioinformatics Toolbox for analysis of biomedical data) to the *A. cinnamomea* S27 genome scaffolds. Gene models were categorized according to their coordinates on the S27 genome scaffolds, with the indication of whether or not overlapping occurred between the predictions from the different algorithms. Next, they were separated into identical (same start/stop position and exon–intron boundaries), overlapping (>5% overlapping of the shorter gene model), and unique gene sets (all except the above). From these categories, all identical genes were selected, whereas in overlapping genes, we used a filter score to integrate the two gene sets from GeneMark and FgeneSH (adopting the weighting formula from Joint Genome Institute (genome.jgi.doe.gov/programs/fungi/FungalGenomeAnnotationSOP.pdf). If the correlation coefficient (calculated from the result of alignment to overlapping isotigs) was undetermined, we selected the blast results only. In the unique gene set, any gene that had either RPKM > 1 or a significant blastP hit (*E*-value <1 × 10^{-5}) was selected (*SI Appendix,* Fig. S1 *B* and *C*).

Phase II. First, we predicted genes with MAKER (50), using the Augustus, Fgenesh, and GeneMark modules without the SNAP module, but based on EST evidence (454.isotig.fna). SNAP is an HMM (Hidden Markov model)-based ab initio gene predictor that attempts to be more adaptable to different organisms, addressing problems related to using a gene finder on a genome sequence that it was not trained against. In this study, SNAP was trained by using the prediction of MAKER without SNAP to seed the first HMM for SNAP and then by repeating the training step using the resultant HMM. Second, the predictions were used to

seed the first HMM used by SNAP, as the MAKER tutorial recommended. After two cycles of SNAP self-training, the protein collections from MAKER output were integrated into the phase I integration set (*SI Appendix,* Fig. S1 *B–D*). Genes were classified into the overlapping and unique sets by comparing gene positions with the cutoff overlapping region >10% of the shorter gene. In the overlapping set, for genes with same blast hits within an overlapping pair, MAKER genes were selected; otherwise, we selected both gene models. In the unique set, genes with blast hits were selected and put into the phase II integration gene set. We considered a gene redundant if the overlapping region shared >10% with others in the phase II integrated gene set. For each redundant case, we selected the one that has the better blast coverage, which was calculated by multiplication of query coverage and hit coverage (*SI Appendix,* Fig. S1 *B* and *D*).

Functional Annotation and Pathway Analysis. The functional annotation of S27 was performed with BLASTp using ACg protein sequences against the NCBI nr protein database (*E*-value ≤1 × 10^{-5}) to find the homologous protein species. Gene ontology was identified by AmiGO v1.8 (51, 52) (*E*-value ≤1 × 10^{-5}). We mapped ACg proteins to InterPro domains and KEGG pathways using Blast2GO v2.7.0 (53, 54).

For terpenoid pathway gene identification, we collected proteins related to the triterpenoid biosynthesis from the NCBI domain database to construct a pathway gene query database, and blast the S27 integrated genes against the database (*E*-value ≤1 × 10^{-3}). We also identified the genomic locations of putative P450 genes in the S27 genome to explore the pathway gene clustering.

Functional Annotation of Wood-Decaying Enzymes. The EC numbers from the output of Blast2GO and InterProScan were mapped into the CaZy database, and the unique EC numbers and enzyme families were collected from the matched records. The families of GH, GT, PL, CE, CBM, and 2 AA families are confirmed from EC mapping, and other AA families are inferred from homology search (*SI Appendix*, Table S17).

1. Chang TT, Chou WN (1995) Antrodia cinnamomea sp. nov. on Cinnamomum kanehirai in Taiwan. *Mycol Res* 99(6):756–758.
2. Wu S-H, et al. (2004) Taiwanofungus, a polypore new genus. *Fungal Sci* 19(3-4):109–116.
3. Ortiz-Santana B, Lindner DL, Miettinen O, Justo A, Hibbett DS (2013) A phylogenetic overview of the antrodia clade (Basidiomycota, Polyporales). *Mycologia* 105(6):1391–1411.
4. Lee IH, et al. (2002) Antrodia camphorata polysaccharides exhibit anti-hepatitis B virus effects. *FEMS Microbiol Lett* 209(1):63–67.
5. Yu YL, et al. (2009) A triterpenoid methyl antcinate K isolated from Antrodia cinnamomea promotes dendritic cell activation and Th2 differentiation. *Eur J Immunol* 39(9):2482–2491.
6. Chang TT, Wang WR, Chou CJ (2011) Differentiation of Mycelia and Basidiomes of Antrodia cinnamomea using certain chemical components. *Taiwan J For Sci* 26(2):125–133.
7. Hsieh YH, et al. (2010) Antrocamphin A, an anti-inflammatory principal from the fruiting body of Taiwanofungus camphoratus, and its mechanisms. *J Agric Food Chem* 58(5):3153–3158.
8. Huang NK, Cheng JJ, Lai WL, Lu MK (2005) Antrodia camphorata prevents rat pheochromocytoma cells from serum deprivation-induced apoptosis. *FEMS Microbiol Lett* 244(1):213–219.
9. Huang CH, et al. (2010) Fruiting body of Niuchangchih (Antrodia camphorata) protects livers against chronic alcohol consumption damage. *J Agric Food Chem* 58(6):3859–3866.
10. Huang GJ, et al. (2010) Analgesic effects and the mechanisms of anti-inflammation of ergostatrien-3beta-ol from Antrodia camphorata submerged whole broth in mice. *J Agric Food Chem* 58(12):7445–7452.
11. GeethangiliMTzengYM (2011) Review of pharmacological effects of Antrodia camphorata and its bioactive compounds. *Evid Based Complement Alternat Med* 2011:212641.
12. Lu MC, et al. (2013) Recent research and development of Antrodia cinnamomea. *Pharmacol Ther* 139(2):124–156.
13. Ao ZH, et al. (2009) Niuchangchih (Antrodia camphorata) and its potential in treating liver diseases. *J Ethnopharmacol* 121(2):194–212.
14. Lin TY, et al. (2011) Metabolite profiles for Antrodia cinnamomea fruiting bodies harvested at different culture ages and from different wood substrates. *J Agric Food Chem* 59(14):7626–7635.
15. Wang HC, et al. (2013) Establishment of the metabolite profile for an Antrodia cinnamomea health food product and investigation of its chemoprevention activity. *J Agric Food Chem* 61(36):8556–8564.
16. James TY, et al. (2013) Polyporales genomes reveal the genetic architecture underlying tetrapolar and bipolar mating systems. *Mycologia* 105(6):1374–1390.
17. Kamada T (2002) Molecular genetics of sexual development in the mushroom Coprinus cinereus. *BioEssays* 24(5):449–459.
18. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215(3):403–410.
19. Nobles MK (1943) A contribution toward a clarification of the Trametes serialis complex. *Can J Res* 21(7):211–234.
20. Lombard FF (1990) A cultural study of several species of Antrodia (Polyporaceae, Aphyllophorales). *Mycologia* 82(2):185–191.
21. 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(4):1753–1761.
22. Inada K, Morimoto Y, Arima T, Murata Y, Kamada T (2001) The clp1 gene of the mushroom Coprinus cinereus is essential for A-regulated sexual development. *Genetics* 157(1):133–140.
23. Stajich JE, et al. (2010) Insights into evolution of multicellular fungi from the assembled chromosomes of the mushroom Coprinopsis cinerea (Coprinus cinereus). *Proc Natl Acad Sci USA* 107(26):11889–11894.
24. Villeneuve AM, Hillers KJ (2001) Whence meiosis? *Cell* 106(6):647–650.
25. Sugawara H, et al. (2009) Coprinus cinereus Mer3 is required for synaptonemal complex formation during meiosis. *Chromosoma* 118(1):127–139.
26. Liu D, et al. (2012) The genome of Ganoderma lucidum provides insights into triterpenes biosynthesis and wood degradation [corrected]. *PLoS ONE* 7(5):e36146.
27. Chen S, et al. (2012) Genome sequence of the model medicinal mushroom Ganoderma lucidum. *Nat Commun* 3:913.
28. Rodríguez-Sáiz M, et al. (2001) Reduced function of a phenylacetate-oxidizing cytochrome p450 caused strong genetic improvement in early phylogeny of penicillin-producing strains. *J Bacteriol* 183(19):5465–5471.
29. Porter TD, Coon MJ (1991) Cytochrome P-450. Multiplicity of isoforms, substrates, and catalytic and regulatory mechanisms. *J Biol Chem* 266(21):13469–13472.
30. Chang TT, Wu WR (2008) The role of four essential oils on mycelial growth and basidiomatal formation of Antrodia cinnamomea. *Taiwan J For Sci* 23(1):105–110.
31. Kumar KJ, et al. (2011) Antroquinonol from ethanolic extract of mycelium of Antrodia cinnamomea protects hepatic cells from ethanol-induced oxidative stress through Nrf-2 activation. *J Ethnopharmacol* 136(1):168–177.
32. Kawamukai M (2009) Biosynthesis and bioproduction of coenzyme Q10 by yeasts and other organisms. *Biotechnol Appl Biochem* 53(Pt 4):217–226.
33. Floudas D, et al. (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336(6089):1715–1719.
34. Huang RL, Huang Q, Chen CF, Chang TT, Chou CJ (2003) Anti-viral effects of active compounds from Antrodia camphorata on wild-type and lamivudine-resistant mutant HBV. *Chinese Pharm. J* 55(5):371–379.
35. Lee CH, et al. (2010) Cloning and characterization of the lanosterol 14alpha-demethylase gene from Antrodia cinnamomea. *J Agric Food Chem* 58(8):4800–4807.
36. Park HG, et al. (2011) Heterologous expression and characterization of the sterol 14α-demethylase CYP51F1 from Candida albicans. *Arch Biochem Biophys* 509(1):9–15.
37. Huang ST, Tzean SS, Tsai BY, Hsieh HJ (2009) Cloning and heterologous expression of a novel ligninolytic peroxidase gene from poroid brown-rot fungus Antrodia cinnamomea. *Microbiology* 155(Pt 2):424–433.
38. Schindler D, Nowrousian M (2014) The polyketide synthase gene pks4 is essential for sexual development and regulates fruiting body morphology in Sordaria macrospora. *Fungal Genet Biol* 68:48–59.
39. Gagne SJ, et al. (2012) Identification of olivetolic acid cyclase from Cannabis sativa reveals a unique catalytic route to plant polyketides. *Proc Natl Acad Sci USA* 109(31):12811–12816.
40. Kobayashi H, Matsunaga K, Oguchi Y (1995) Antimetastatic effects of PSK (Krestin), a protein-bound polysaccharide obtained from basidiomycetes: An overview. *Cancer Epidemiol Biomarkers Prev* 4(3):275–281.
41. Bardwell VJ, Treisman R (1994) The POZ domain: A conserved protein-protein interaction motif. *Genes Dev* 8(14):1664–1677.
42. Melnick A, et al. (2000) In-depth mutational analysis of the promyelocytic leukemia zinc finger BTB/POZ domain reveals motifs and residues required for biological and transcriptional functions. *Mol Cell Biol* 20(17):6550–6567.
43. Zollman S, Godt D, Privé GG, Couderc JL, Laski FA (1994) The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in Drosophila. *Proc Natl Acad Sci USA* 91(22):10717–10721.
44. Bomont P, et al. (2000) The gene encoding gigaxonin, a new member of the cytoskeletal BTB/kelch repeat family, is mutated in giant axonal neuropathy. *Nat Genet* 26(3):370–374.
45. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc B Met* 57(1):289–300.
46. Martinez D, et al. (2009) Genome, transcriptome, and secretome analysis of wood decay fungus Postia placenta supports unique mechanisms of lignocellulose conversion. *Proc Natl Acad Sci USA* 106(6):1954–1959.
47. Tang JD, et al. (2012) Short-read sequencing for genomic analysis of the brown rot fungus Fibroporia radiculosa. *Appl Environ Microbiol* 78(7):2272–2281.
48. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754–1760.
49. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M (2008) Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res* 18(12):1979–1990.
50. Holt C, Yandell M (2011) MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12:491.
51. Carbon S, et al.; AmiGO Hub; Web Presence Working Group (2009) AmiGO: Online access to ontology and annotation data. *Bioinformatics* 25(2):288–289.
52. Ashburner M, et al.; The Gene Ontology Consortium (2000) Gene ontology: Tool for the unification of biology. *Nat Genet* 25(1):25–29.
53. Götz S, et al. (2008) High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res* 36(10):3420–3435.
54. Götz S, et al. (2011) B2G-FAR, a species-centered GO annotation repository. *Bioinformatics* 27(7):919–924.