PNAS RESEARCH ARTICLE | GENETICS

Genes and evolutionary fates of the amanitin biosynthesis pathway in poisonous mushrooms

Hong Luo[a,b,1], Heather E. Hallen-Adams[c,2], Yunjiao Lüli[a,d,2], R. Michael Sgambelluri[e,f], Xuan Li[g], Miranda Smith[f], Zhu L. Yang[a,b], and Francis M. Martin[h,i,1]

Edited by Louise Chow, University of Alabama at Birmingham, Birmingham, AL; received January 26, 2022; accepted March 22, 2022

The deadly toxin α-amanitin is a bicyclic octapeptide biosynthesized on ribosomes. A phylogenetically disjunct group of mushrooms in Agaricales (*Amanita*, *Lepiota*, and *Galerina*) synthesizes α-amanitin. This distribution of the toxin biosynthetic pathway is possibly related to the horizontal transfer of metabolic gene clusters among taxonomically unrelated mushrooms with overlapping habitats. Here, our work confirms that two biosynthetic genes, *P450-29* and *FMO1*, are oxygenases important for amanitin biosynthesis. Phylogenetic and genetic analyses of these genes strongly support their origin through horizontal transfer, as is the case for the previously characterized biosynthetic genes *MSDIN* and *POPB*. Our analysis of multiple genomes showed that the evolution of the α-amanitin biosynthetic pathways in the poisonous agarics in the *Amanita*, *Lepiota*, and *Galerina* clades entailed distinct evolutionary pathways including gene family expansion, biosynthetic genes, and genomic rearrangements. Unrelated poisonous fungi produce the same deadly amanitin toxins using variations of the same pathway. Furthermore, the evolution of the amanitin biosynthetic pathway(s) in *Amanita* species generates a much wider range of toxic cyclic peptides. The results reported here expand our understanding of the genetics, diversity, and evolution of the toxin biosynthetic pathway in fungi.

Amanita | *Galerina* | *Lepiota* | genome | gene cluster

Significance

Why do unrelated poisonous mushrooms (*Amanita*, *Galerina*, and *Lepiota*) make the same deadly toxin, α-amanitin? One of the most effective and fast strategies for organisms to acquire new abilities is through horizontal gene transfer (HGT). With the help of genome sequencing and the finding of two genes for the amanitin biosynthetic pathway, we demonstrate that the pathway distribution resulted from HGT probably through an unknown ancestral fungal donor. In *Amanita* mushrooms, the pathway evolved, through a series of gene manipulations, to produce very high levels of toxins, generating "the deadliest mushroom known to mankind."

The deadliest mushrooms, such as the Death Cap and the Destroying Angel, belong to the genus *Amanita* sect. *Phalloideae*, yet equally poisonous mushrooms are present in the distantly related genera *Galerina* and *Lepiota* (1, 2). For centuries, stories and incidents associated with these lethal mushrooms have been widespread (1, 3). These three genera of Agaricales are phylogenetically disjunct, but they produce the same deadly toxin, α-amanitin. This bicyclic octapeptide is a ribosomally encoded posttranslationally modified peptide. The toxin has a human LD_{50} (the amount fatal to half of a tested population) of 0.1 mg per kg, and one mature Death Cap fruiting body can contain a lethal dose of 10 to 12 mg (3). α-Amanitin and other related cyclic peptides are nonetheless highly valuable, not only as molecular tools (4) but also as potential treatment for diseases like cancers (5), and they have only been chemically synthesized very recently (6, 7). The cytotoxicity found in amanitins is the result of the inhibition of RNA polymerases that precludes mRNA synthesis in the liver (8).

Amanitin-producing fungi rely on the same set of biosynthetic genes for the production of α-amanitin. The proproteins are encoded by a family of genes called the MSDIN family for the first five conserved amino acids in the precursor peptides (9). The initial posttranslational processing step of the α-amanitin precursor peptide is catalyzed by POPB, a specialized member of the prolyl oligopeptidase family of serine proteases (10). The disjunct phylogenetic distribution of α-amanitin in agarics suggests that the *MSDIN* genes involved in biosynthesis have been dispersed by horizontal gene transfer (HGT) (11, 12). Recent findings in psilocybin-producing mushrooms, including *Psilocybe cyanescens*, *Gymnopilus dilepis*, and *Panaeolus cyanescens*, showed that the biosynthetic pathway for psilocybin is encoded by a gene cluster, which has been transferred throughout phylogenetic-disjunct dung and wood decay mushrooms via HGT (13). The genes coding for psilocybin biosynthesis are located in a gene cluster in *Ps. cyanescens*, *Gy. dilepis*, and *Pa. Cyanescens*, facilitating the HGT process (14, 15). However, such a gene cluster for amanitin biosynthesis has not been found in the *Amanita* or *Lepiota* genomes (11, 16).

The α-amanitin–producing *Amanita*, *Galerina*, and *Lepiota* mushrooms are phylogenetically and ecologically disjunct. The amanitin-producing *Amanita* belongs to the section *Phalloideae*, and they form a mutualistic ectomycorrhizal symbiosis with trees (17). In contrast, members of the deadly *Galerina* species, such as *Galerina marginata*, are white-rot decayers, able to decompose wood by using a large arsenal of ligninolytic

Author contributions: H.L., Z.L.Y., and F.M.M. designed research; H.E.H.-A. and Z.L.Y. contributed new reagents/analytic tools; and H.L., Y.L., R.M.S., X.L., M.S., and F.M.M. analyzed data and wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: francis.martin@inrae.fr or luohong@mail.kib.ac.cn.

[2]H.E.H.-A. and Y.L. contributed equally to this work.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2201113119/-/DCSupplemental.

Published May 9, 2022.

oxidoreductases: dye-decolorizing peroxidase and heme-thiolate peroxidase (18). Finally, members of the lethal *Lepiota* species are soil-saprotrophic species. They have lost genes coding for lignin oxidoreductases, but they have an expanded set of carbohydrate-active enzymes involved in cellulose, hemicellulose, and pectin degradation.

In the present study, we investigated previously unknown amanitin biosynthetic genes and the evolution of the biosynthesis pathway in the *Amanita*, *Galerina*, and *Lepiota* clades. By analyzing independent genome assemblies for the species of these genera, we recaptured the homologous multigene amanitin biosynthesis locus in each poisonous lineage. The enzymatic functions of genes within this locus were confirmed by combined genetic and biochemical approaches. Two genes were confirmed to be important for the production of mature α-amanitin. Based on in-depth phylogenetic, distance, and synteny analyses, the origin of the amanitin biosynthetic pathway was carefully assessed, and three distinct genus-specific evolutionary outcomes were documented. These data support the hypothesis that the amanitin biosynthesis locus follows independent evolutionary pathways in the three deadly mushroom clades and suggest that amanitin production may be part of a larger adaptive strategy to soil niches, which harbor abundant invertebrates that eat or compete with fungi.

Results

Main Features of the Genomes from Amanitin-Producing Mushrooms. We compared 15 genomes from amanitin-producing mushrooms, including previously unobtained genomes from one *Amanita* species, two *Galerina* species, and three *Lepiota* strains (from two species) (see *Materials and Methods* and *SI Appendix*). Whole Genome Shotgun assemblies have been deposited at Data Bank of Japan (DDBJ)/European Nucleotide Archive (ENA)/GenBank under the accession number JAEBUT000000000. We combined these genomes with the genome of *G. marginata* CBS 339.88, available from the Joint Genome Institute (JGI) MycoCosm database (19), and the *Amanita* and *Lepiota* genomes previously published (9, 11, 16, 20, 21). The completeness of the gene repertoires of 13 sequenced genomes, based on Benchmarking Universal Single-Copy Orthologs (BUSCO) and CEGMA (Core Eukaryotic Genes Mapping Approach) analyses, are provided in *SI Appendix*, Table S1. All the sequenced genomes are dikaryon, except for *G. marginata* CBS 339.88, which is a monokaryon strain. The high quality of the genomes allowed a detailed analysis of the amanitin biosynthesis loci. In *Amanita*, the genome size ranges from 45 to 71 Mbp, while in *Galerina*, it varies from 59 to 101 Mbp and in *Lepiota* from 37 to 55 Mbp. Notably, *Galerina* species had on average the largest genome sizes and the lowest content of repetitive sequences (76.89 Mbp and 15.77%, respectively). Further, *Galerina* genomes have the highest number of predicted genes compared to *Lepiota* and *Amanita*, with *Amanita* having the lowest (*SI Appendix*, Table S1). The genomes of the *Amanita* species were the most distinctive, sharing significantly less synteny within the genus and with either *Galerina* or *Lepiota* than was shared between the two saprobes (*SI Appendix*, Fig. S1). This limited genome synteny is possibly due to the high repetitive content of ectomycorrhizal *Amanita* genomes. In contrast, *Galerina* and *Lepiota* species presented significantly higher levels of synteny in the respective genera and between the two groups. Amanitin-producing *Amanita* and *Lepiota* species consistently shared more genes than did *Galerina* and *Lepiota*, and the *Galerina* species had a large number of species-specific genes (*SI Appendix*, Fig. S2).

The Amanitin Biosynthetic Genes and Loci in *Galerina*. In contrast to the *Amanita* species, a single α-amanitin–encoding *MSDIN* gene, *GmAMA1* (with two copies), was identified in the genome of *G. marginata* CBS 339.88. The functional copy, *GmAMA1-1*, is located near a *POPB* gene, *GmPOPB* (22).

In order to find the genes responsible for amanitin biosynthesis, a search of the functional annotations was performed for each of the relevant types of enzymatic activity (cyclization, hydroxylation, sulfoxidation, formation of the tryptathionine bridge, and epimerization) using specific protein families predicted to act on chemically similar substrates to those found in the amanitin biosynthetic pathway. A set of genes possibly involved in α-amanitin biosynthesis was found in the vicinity of *GmAMA1-1* (Fig. 1*A*). By coupling the gene disruption of these candidate genes to the chemical profiling of amanitin and amanitin-related metabolites, we characterized the cluster of genes involved in amanitin biosynthesis. Two biosynthetic genes, *GmFMO1* and *GmP450-29*, involved in the cyclic peptide biosynthesis were identified. *GmFMO1* is predicted to encode a member of the flavin mono-oxygenase (FMO) family of enzymes. Disruption of *GmFMO1* abolished the formation of amanitin and gave rise to three compounds with monoisotopic masses of 902.4, 886.4, and 870.4, corresponding to α-amanitin minus one, two, or three oxygen atoms, respectively (Fig. 1*B*). In addition, there were four genes predicted to encode cytochrome P450s (*CYP450*s) near *GmAMA1-1*. Based on Southern blot hybridization, all four genes were found present in three amanitin-producing species and absent from the amanitin-nonproducing species (*SI Appendix*, Fig. S3). The sequence of one of them, *GmP450-29*, was successfully disrupted. The *GmP450-29* deletion mutant did not biosynthesize α-amanitin, and instead a major product (named θ-amanitin, see below) appeared (Fig. 1*C*). This compound was isolated and identified by mass spectrometry and nuclear magnetic resonance (NMR) (*SI Appendix*, Figs. S4 and S5 and Table S4) as an amatoxin lacking hydroxylation at the C-4 position of proline and the C-5 position of isoleucine, indicating that GmP450-29 catalyzes one or both of these hydroxylations. These two positions are major active sites interacting with RNA polymerase II (RNAP II), and the absence of those hydroxylations caused a drastic decrease in the activity by up to 1,000-fold (C-4 position of proline). This particular amatoxin had not previously been observed and has been named θ-amanitin.

In *G. marginata*, the two previously known genes (*GmAMA1* and *GmPOPB*) that are required for, and dedicated to, amanitin biosynthesis were separated in the genome by a single gene (Fig. 1*A*). This suggested that other biosynthetic genes might also be genomically linked to these two genes. Genes upstream and downstream of *GmAMA1* and *GmPOPB* were therefore annotated and searched for their presence or absence in amanitin-producing and nonproducing species of *Galerina*, and their role was confirmed by targeted gene disruption (*SI Appendix*, Table S2).

Comparison of Southern blot patterns between poisonous and nonpoisonous species (*SI Appendix*, Fig. S3) suggested the presence of six additional candidate genes involved in amanitin biosynthesis (Fig. 1, light-blue arrowheads). Consistent with the Southern blotting results, disruption of three of the six genes led to mycelial phenotypes with abolishment or significant reduction of α-amanitin (*SI Appendix*, Figs. S6–S8). They are located in a gene cluster spanning 111 kbp. To confirm this spatial distribution, a different strain of *G. marginata* and another species, *Galerina sulciceps*, both collected from China, were subsequently

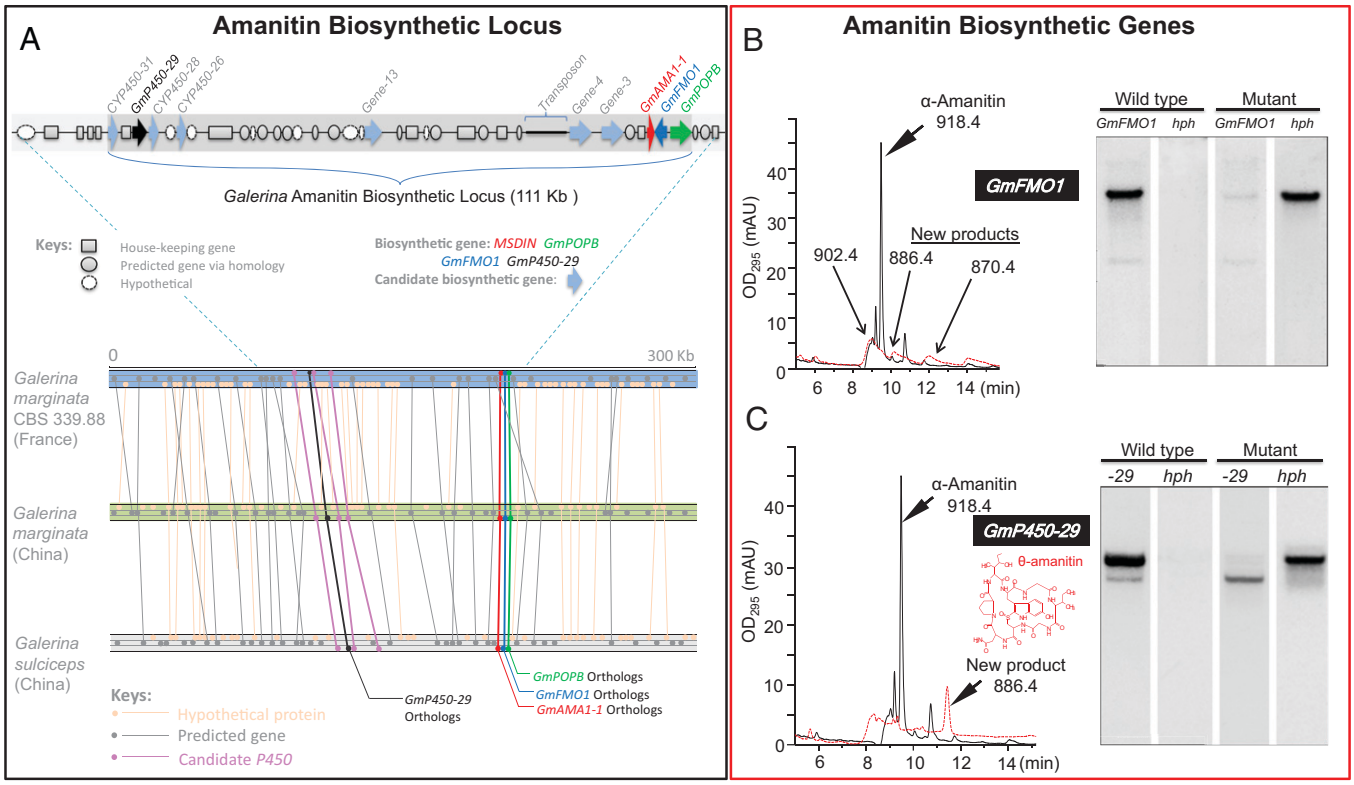

Fig. 1. The amanitin biosynthesis locus and biosynthetic genes in *Galerina* species. (*A*) Amanitin biosynthetic genes are color-coded (red, green, blue, and black). Candidate biosynthetic genes are marked as light blue arrowheads. The amanitin biosynthetic loci were aligned for three *Galerina* genomes. Gene number suffix: position upstream (-) of GmAMA1-1; the two named downstream genes are not numbered based on position. Lines connect the genes (dots) with the highest homology within the genomes. (*B*) Deletion of *GmFMO1* in *G. marginata* results in termination of α-amanitin production. In liquid chromatography–mass spectrometry (LC-MS) analysis: solid line denotes OD_{295} for wild type; red dashed line denotes OD_{295} for *GmFMO1* mutant. In the DNA blot analysis, the wild type and the mutant are shown, with *hph* for the hygromycin B marker. (*C*) Deletion of *GmP450-29* in *G. marginata* results in termination of α-amanitin production. In the LC-MS analysis: solid line denotes OD_{295} for wild type; red dashed line denotes OD_{295} for *GmP450-29* mutant. In the DNA blot analysis, the wild type and the mutant are shown, with *hph* for the hygromycin B marker.

sequenced. Highly similar gene arrangements were found in all three specimens with the amanitin biosynthetic genes in nearly identical placement (Fig. 1*A*); the tight arrangement was used throughout this report for locus comparison instead of using the entire harboring scaffold or contig. Based on the above results, gene homologs of *AMA1*, *POPB*, *FMO1*, and *P450-29* were selected as markers for the amanitin biosynthetic pathway.

Phylogenetic and Genetic Distance Analyses for *rpb2*, *GmP450-29*, and *GmFMO1*. With tBLASTn search (GmP450-29 and GmFMO1), the genomic DNAs (gDNAs) of representative species of *Amanita*, *Galerina*, and *Lepiota* were obtained from the respective genomes. Complementary DNAs (cDNAs) were acquired by aligning gDNAs with their BLASTn hits in their respective transcriptome. Gene trees were constructed and showed the same structure as those for *MSDIN* and *POPB* (11, 12). Topology conflicts of both gene tree vs. species tree are displayed in *SI Appendix*, Fig. S9, in which the normal species tree showed *Amanita* at the base and *Lepiota* at the terminal, while both gene trees showed that *Lepiota* was basal and *Amanita* was terminal. These conflicts, as in our previous reports on *MSDIN* and *POPB* genes, provided strong support to the hypothesis of HGT while rejecting the hypothesis of massive gene loss (11, 12). The gene trees were then analyzed using PAML software, which further supported the HGT hypothesis (*SI Appendix*, Table S5). The species tree, in accordance with that by Matheny et al. (23), represented the phylogeny of the housekeeping gene *rpb2*, while the gene trees represented the evolutionary relationships of the *FMO1* and *P450-29* genes. The results showed that there were significant differences in distances, synonymous rates (dS), and nonsynonymous rates (dN) among the three amanitin-producing genera. The distances of *rpb2* were significantly higher than those of *FMO1* and *P450-29* (*SI Appendix*, Table S5). For *MSDIN*, *POPB*, and *FMO1*, the dN/dS values were below 1, while the value for *P450-29* was above 1. In all cases, the dS values of the biosynthetic genes were significantly smaller that of *rpb2*. These data are consistent with the HGT scenario as smaller distances and lower rates are expected when compared to housekeeping genes.

***Amanita* Amanitin Biosynthetic Loci.** In six *Amanita* genomes (*Amanita phalloides* and *Amanita bisporigera* were left out due to low N50s [the shortest contig length that needs to be included for covering 50% of the genome], indicating small contig size), the amanitin biosynthetic genes were scattered on contigs or scaffolds totaling from 9.47 to 31.97 Mbp (19.1 to 53.1% of the respective genome) (Fig. 2*A*). These genomes contain a high number of *MSDIN* gene duplications, whereas the *Galerina* genomes encode a single *MSDIN* gene. A summary of *MSDIN* genes identified in all 15 sequenced amanitin-producing *Amanita*, *Galerina*, and *Lepiota* species is shown in Fig. 3, and the details are shown in *SI Appendix*, Table S6 and Dataset S1. These genes are distributed in a loose, largely random pattern in deadly *Amanita* species, but the *POPB*, *FMO1*, and *P450-29* genes are frequently located near *MSDIN* genes (Fig. 2*A*). In the sequenced *Amanita* genomes, the *MSDIN* gene distribution is patchy. Duplication of *MSDIN* genes happens frequently in *Amanita* species (e.g., *Amanita subjunquillea* and *Amanita rimosa*; *SI Appendix*, Figs. S10 and S11). Copies of identical *MSDIN* genes are often linked, which partly causes the patchy patterns. Recent studies showed that most of the

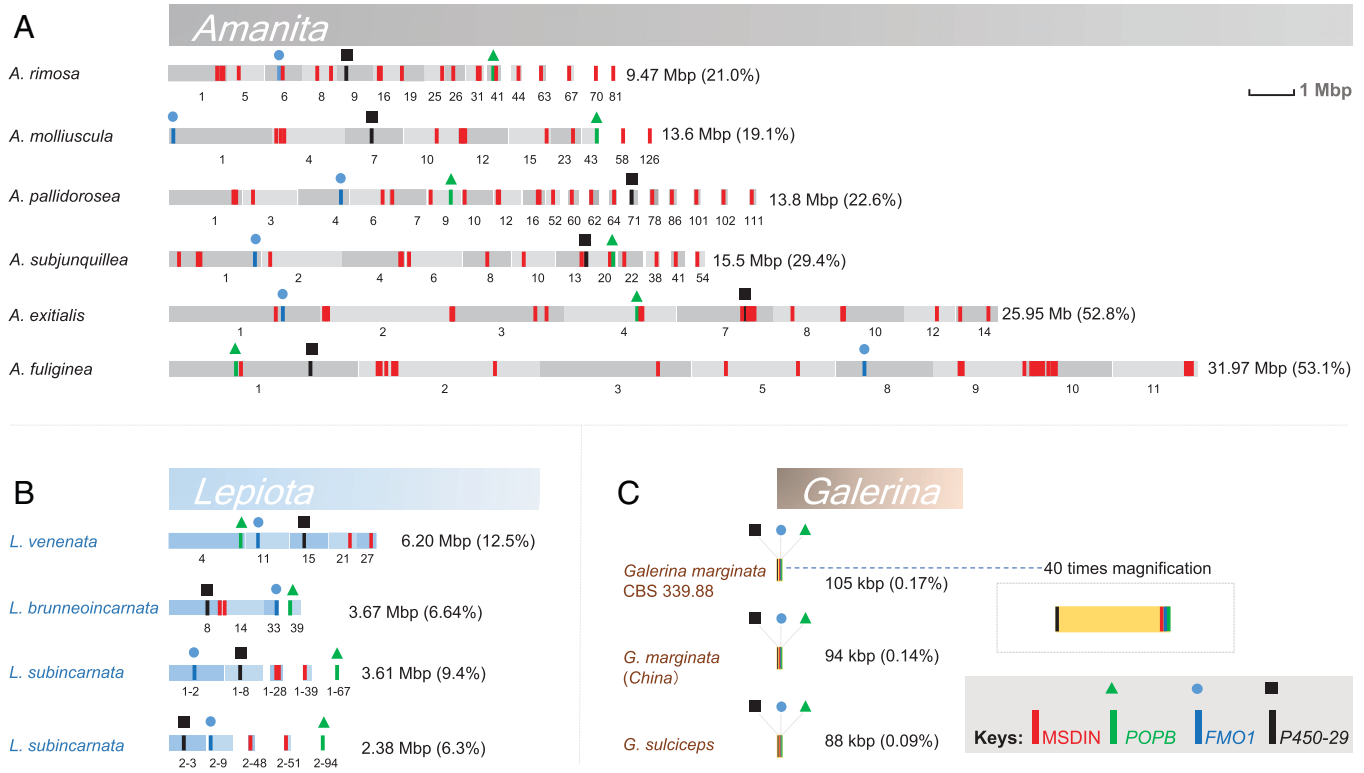

Fig. 2. The amanitin biosynthesis loci. Contigs or scaffolds are indicated as blocks with combined sizes on the right (percentage of genome assemblies in parentheses); the display is in scale, with details in a magnified view for the locus in *Galerina* species. (*A*) Distribution of amanitin biosynthesis genes in the genome of lethal *Amanita* species. (*B*) Distribution of biosynthesis genes in lethal *Lepiota* species. (*C*) Distribution of the biosynthesis genes in lethal *Galerina* species. Homologs of *MSDIN*, *POPB*, *P450-29*, and *FMO1* marked in red, green, blue, and black, respectively.

*MSDIN*s are expressed at the transcription level (16, 20), and through a genome-guided approach we found 14 cyclic peptides in 5 *Amanita* species (20, 21). Known cyclic peptides are shown in Fig. 3 in bold. Quite noticeably, the α-amanitin–encoding *MSDIN* is the only gene shared by all amanitin-producing mushrooms (Fig. 3, red letters), and the ubiquitous presence of α-amanitin in these mushrooms was used for the development of a commercial amanitin detection kit (patent number: ZL 2016 1 0991804.1).

Amanitin Biosynthetic Loci in *Lepiota*. In the sequenced *Lepiota* species, the genes involved in amanitin biosynthesis spanned from 2.38 to 6.2 Mbp on the combined contigs or scaffolds (Fig. 2*B*). Unlike with the *Galerina* species, the four amanitin biosynthetic genes are distributed in a completely random manner in the genome as none of these genes were found in close vicinity, except for a few *MSDIN* genes.

Synteny of Amanitin Biosynthetic Loci across *Amanita*, *Galerina*, and *Lepiota*. With relaxed Synteny Mapping and Analysis Program (SyMAP) settings, the amanitin biosynthetic locus of *G. marginata* was searched for syntenic blocks against the genomes of *A. subjunquillea*, *Amanita pallidorosea*, *Lepiota brunneoincarnata*, and *Lepiota venenata*. In all cases, syntenic regions were found at or near the amanitin biosynthetic genes (*SI Appendix*, Fig. S12), suggesting that the pathways in the three genera possess similar structural regions. The result lends additional support to the HGT origin of the amanitin biosynthesis.

***MSDIN* Family Expansion in *Amanita* and *Lepiota*.** In *Galerina*, there is no expansion of the *MSDIN* gene family (12, 22) (Fig. 3 and *SI Appendix*, Table S6). In contrast, this gene family has undergone a striking expansion in the eight sequenced amanitin-producing *Amanita* species, with 19 to 40 *MSDIN* genes found in each of the sequenced genomes (Fig. 3 and Dataset S1). Most of these sequences displayed the canonical MSDIN amino acid signature (*SI Appendix*, Fig. S13). All known amanitin-producing *Amanita* species are reported to produce multiple cyclic peptides. For example, at least 12 cyclic peptides have been reported in *A. phalloides* (1, 3).

In the *Lepiota* species, we found two to six *MSDIN* genes (Fig. 3 and *SI Appendix*, Table S6), indicating a minor expansion. The canonical MSDIN sequence is often replaced by the M_DAN sequence in most of these species.

Toxin Biosynthetic Genes Specific to *Amanita*. On a lambda gDNA clone of *A. bisporigera*, three *CYP450* genes (*AbP450-1*, *AbP450-2*, and *AbP450-3*) were found near two *MSDIN* genes encoding phallacidin. These *CYP450*s were therefore tested for their presence in amanitin-producing *Galerina*, *Lepiota*, and amanitin-producing and non–amatin-producing *Amanita* species using Southern blotting. Hybridization was detected only in amanitin-producing *Amanita* species (Fig. 4). Consistent with this result, a BLAST search using the three *CYP450* genes as queries returned no convincing hits (National Center for Biotechnology Information [NCBI] BLAST^{+} 2.6.0), except for the amanitin-producing *Amanita* species (sect. *Phalloideae*). These results suggest that these three *P450*s were specific only to the amanitin-producing *Amanita* species. Structures and sequences of the three genes are shown in *SI Appendix*, Fig. S14.

Discussion

Amanitin Biosynthesis Genes. Two types of biosynthetic genes were known to participate in the amanitin biosynthesis, i.e., *MSDIN* and *POPB*. In this study, we characterized two genes,

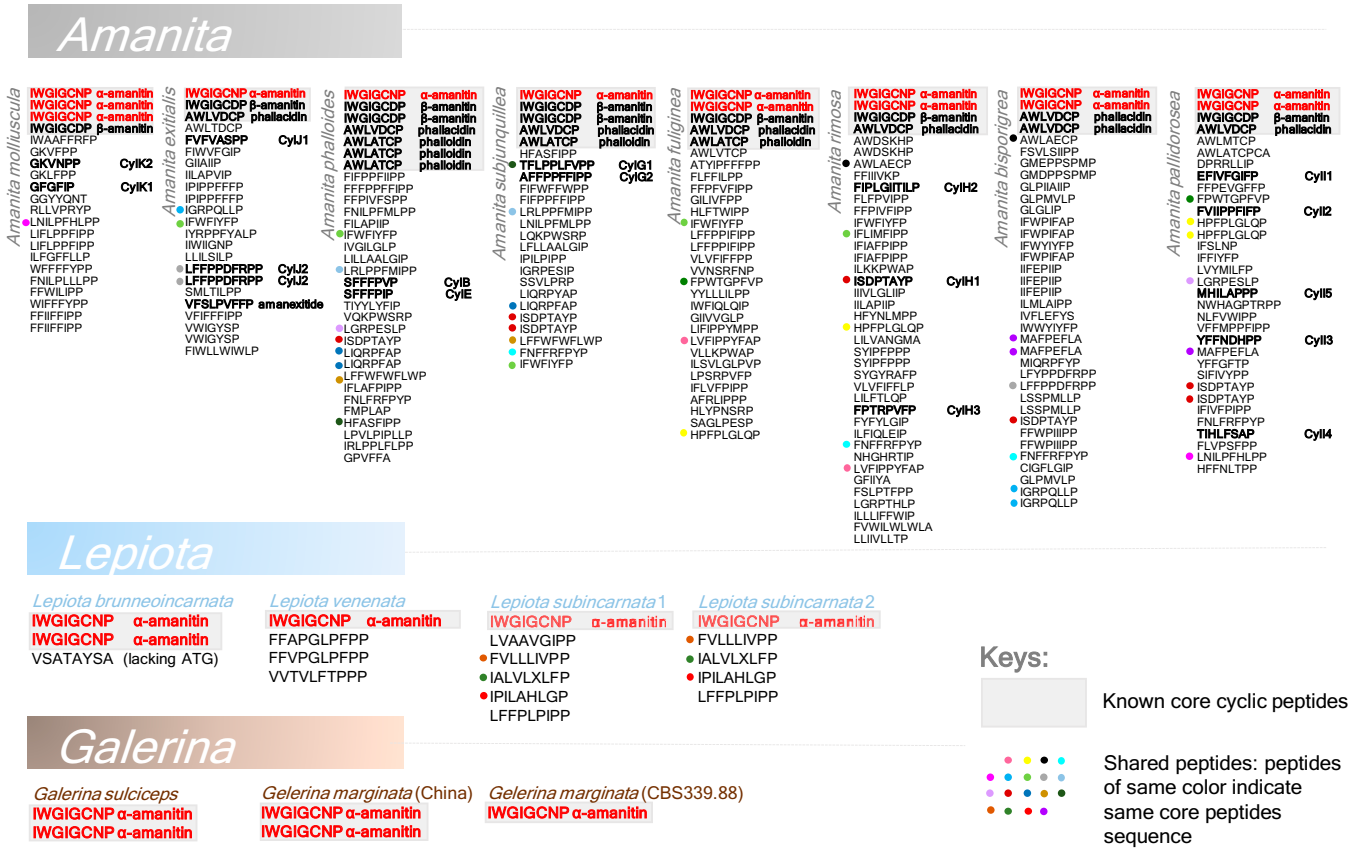

Fig. 3. MSDIN core peptide sequences and known cyclic peptides in lethal *Amanita*, *Galerina*, and *Lepiota*. Core peptides (the amino acid residues in the cyclic peptide products) of known MSDINs across the three genera are listed on top of each column (shaded in gray). The α-amanitin–forming peptide is marked in red, which is the only one shared across the three genera. Shared core peptides across species are indicated with dots of the same colors in front of the peptides.

i.e., *FMO1 and P450-29*, which are involved in posttranslational modifications of α-amanitin. *FMO*s and cytochrome *CYP450*s are two microsomal enzymes involved in metabolizing foreign compounds (xenobiotics) by adding molecular oxygen to their substrates (24). For α-amanitin and the related peptides, oxygenation (hydroxylation and sulfoxidation) is one of the most important posttranslational steps to produce active compounds. Without proper hydroxylation or sulfoxidation, there is a drastic activity decrease (1). The hydroxylation(s) catalyzed by GmP450-29 takes place on C-4 Pro and/or C-5 Ile of α-amanitin, and both are major sites for its affinity to the target RNAP II; e.g., a lack of C-4 Pro hydroxylation causes a 1,000-fold decrease in the activity (1, 25). *GmFMO1* is important for amanitin biosynthesis as well, as not only was α-amanitin production abolished in the deletion mutant but the accumulation of the intermediates, probably the toxin minus one to three oxygen atoms, dropped to trace level. This result indicated that the oxygenation carried out by this enzyme is necessary before other posttranslational modifications can readily occur. Due to the difficulty of working with this

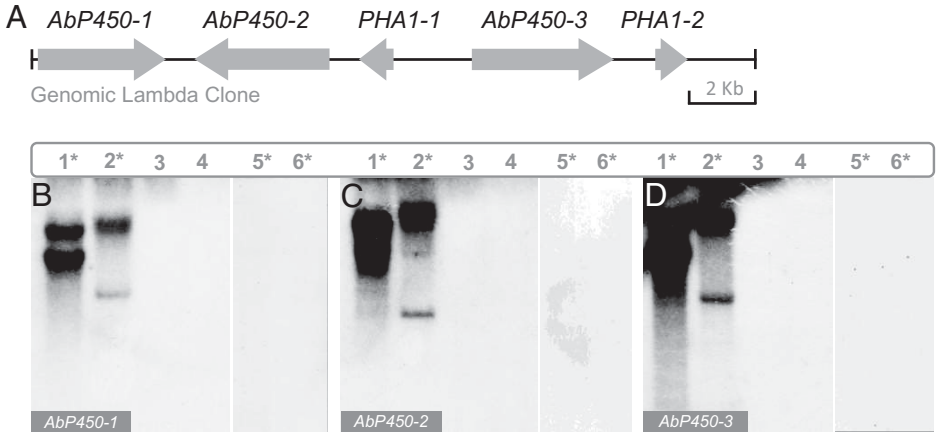

Fig. 4. Southern blotting of amanitin biosynthesis *CYP450* genes (*AbP450-1*, *AbP450-2*, and *AbP450-3*) in *Amanita*, *Lepiota*, and *Galerina* species. (*A*) Map of the genomic lambda clone PA1 from *A. bisporigera*. DNA blots show taxonomic distribution of *AbP450-1* (*B*), *AbP450-2* (*C*), *and AbP450-3* (*D*). Lanes: 1, *A. bisporigera*; 2, *A. phalloides*; 3, *A. porphyria*; 4, *A. franchetii*; 5, *G. marginata*; 6, *L. brunneoincarnata*. Amanitin-producing species are marked with asterisks. *PHA1-1* and *PHA1-2* are the two copies of the phallacidin-encoding *MSDIN* gene.

mutant, identifying the exact site of oxygenation by this enzyme was not possible. The characterization of these two genes greatly improves our understanding of the constitution and function of amanitin biosynthesis.

Diversification of the Amanitin Biosynthetic Pathway. The amanitin biosynthesis in the *Amanita* species is more complex and versatile than the amanitin pathways occurring in *Galerina* and *Lepiota* mushrooms. The basidiocarps of the *Amanita* sect. *Phalloideae* produce the largest set of cyclic peptides as a result of a striking *MSDIN* gene family expansion; 45 cyclic peptides have been documented in deadly *Amanita* species so far, although some of these cyclic peptides, such as antamanide and CylA-D, are not known toxins (1). It is widely believed that many more toxic peptides are produced by *Amanita* mushrooms (1, 16), and our very recent discovery of 14 cyclic peptides in 5 *Amanita* species through a genome-guided approach confirms this hypothesis. Many *MSDIN* genes are expressed at the transcriptional level and the corresponding cyclic peptides, with or without posttranslational modifications, have been detected by mass spectrometry (16). In contrast, the three available sequenced *Galerina* genomes encode a single *MSDIN* gene and the *Galerina* mushrooms synthesize a single cyclic peptide, i.e., α-amanitin. The γ-amanitin peptide is also synthesized by *Galerina*, and α-amanitin is derived from γ-amanitin by posttranslational hydroxylation (26). The presence of a small amount of β-amanitin in *G. marginata* is likely due to the chemical deamination of α-amanitin. In the four sequenced *Lepiota* genomes, α-amanitin is the major amanitin metabolite. Other minor amanitins reported in *Lepiota* (amanin, γ-amanitin, and amaninamide) are likely intermediates of α- or β-amanitin lacking posttranslational hydroxylation(s) (27). The current analysis of the representative genomes of *Amanita*, *Lepiota*, and *Galerina* showed that the α-amanitin–encoding gene is shared by all amatoxin-producing species across the three agaric families. α-Amanitin performed well in our chromogenic test (based on a highly modified Meixner test), becoming the basis of a detection kit for these poisonous mushrooms (http://www.cxbio199.com/sy).

Genes specific to amanitin-producing *Amanita* species have been recruited into the amanitin biosynthetic pathway in this genus. In this report, we showed that three toxin-related *CYP450* genes occur only in amanitin-producing *Amanita* species (sect. *Phalloideae*). They are lacking in the genomes of nonpoisonous *Amanita* species, such as *Amanita rubescens* or *Amanita thiersii* (see genome available at JGI), but also in the genomes of amanitin-producing *Galerina* and *Lepiota* species. These genes can potentially process a larger range of cyclic peptide substrates. Recruitment of new genes into the amanitin biosynthetic pathway is clearly an evolutionary step leading to a wider repertoire of toxic peptides through additional posttranslational steps.

Genomic arrangement of the amanitin biosynthesis genes in the *Amanita*, *Galerina*, and *Lepiota* genera is genus-specific. A gene cluster–like arrangement (ca. 0.1 Mbp) is found in the three *Galerina* species, whereas the amanitin biosynthesis genes are randomly scattered over 2 Mbp in the *Lepiota* species. In *Amanita*, the amanitin biosynthesis genes are distributed over larger genomic regions, up to 30 Mbp. In comparison to the *Lepiota* species, genes are not randomly scattered over the genome. Most of the amanitin biosynthesis genes are found in physical proximity; e.g., the *POPB* and *FMO1* genes are frequently found to be linked to *MSDIN* genes. This approximate gene location likely facilitates the coordinated regulation of their expression. For example, it is well known that the core cyclic peptides, α-amanitin, β-amanitin, phalloidin, and phallacidin, are often expressed in the basidiocarps at similarly high levels (27).

The present genome-wide analysis of the amanitin biosynthesis pathway in three poisonous genera of mushrooms, *Galerina*, *Lepiota*, and *Amanita*, reveals a striking range of genetically encoded biosynthetic capacity in the production of amanitin-related toxins. *Galerina* has the potential for only one MSDIN-family cyclic peptide, while the *Amanita* species can potentially biosynthesize hundreds of toxic cyclic peptides, with 45 biochemically confirmed metabolites. The pathway diversification, involving the expansion of the *MSDIN* gene family and the incorporation of several genes (e.g., *CYP450*) for posttranslational modifications, dramatically increases the outcome of the amanitin biosynthesis pathway in *Amanita* mushrooms. To the best of our knowledge, there is no other secondary metabolic pathway in fungi showing so many innovations.

Evolutionary Fates of the Amanitin Biosynthetic Pathway. In this study, two genes involved in amanitin biosynthesis, *P450-29* and *FMO1*, have been genetically validated through gene disruption and biochemical analysis of the resulting peptides, leading to a better understanding of the function and genomic organization of the amanitin pathway. Our phylogenetic and genetic distance analyses showed topology conflict and shorter genetic distances when compared to the species tree based on the housekeeping gene *rpb2*. In addition, synteny was observed among the amanitin biosynthetic loci across *Amanita*, *Galerina*, and *Lepiota*. The congruence of these data strongly supports that the distribution of the amanitin biosynthetic pathway is a result of HGT. In addition, the divergence time analysis showed a clear radiation of *Lepiota*, *Galerina*, and *Amanita* from a common ancestor, but the amanitin biosynthetic pathway was only found in three subsets of these lineages (*SI Appendix*, Fig. S15). Although the evidence for HGT is strong, surprisingly, the evolutionary outcome of the pathway followed distinctive paths in the three genera. In Fig. 5, we propose a putative scenario for the evolutionary fates of the amanitin biosynthesis pathway in the deadly species of *Amanita*, *Galerina*, and *Lepiota*. Using the genome of *A. subjunquillea* as an example, we showed that the amanitin biosynthesis genes are distributed over a large portion (~15 Mbp) of the mushroom genome, with *POPB*, *FMO1*, and *P450-29* each linked to an *MSDIN* gene. In contrast, the toxin biosynthesis genes are located on a restricted locus of 111 kbp in *G. marginata*. Notably, the expansion of the *MSDIN* gene family only takes place in *Amanita* sect. *Phalloideae*. All sequenced *Galerina* species only have a single *MSDIN* encoding α-amanitin. Finally, the distribution of the amanitin biosynthesis genes in *L. brunneoincarnata* differs from the *Amanita* and *Galerina* species. Unlike the *Amanita* and *Galerina* genomes, the *Lepiota* genome encodes only two *MSDIN* genes with a random genome location.

Based on the current distribution of the genes involved in amanitin biosynthesis, we initially speculated that the pathway may have originated in the *Galerina* clade. The gene cluster would have been transferred to other mushroom species by HGT, as observed for many secondary metabolite gene clusters, and the genes would have been physically unlinked over time. However, our recent phylogenetic and genetic distance analyses (11, 12) and the data in this study do not support this hypothesis as the most parsimonious. Another unlikely scenario is the transfer of the amanitin genes from *Lepiota* to *Galerina* to *Amanita*, but the dispersed loci structure and genetic distances of all four biosynthetic genes would be in conflict (the three genera

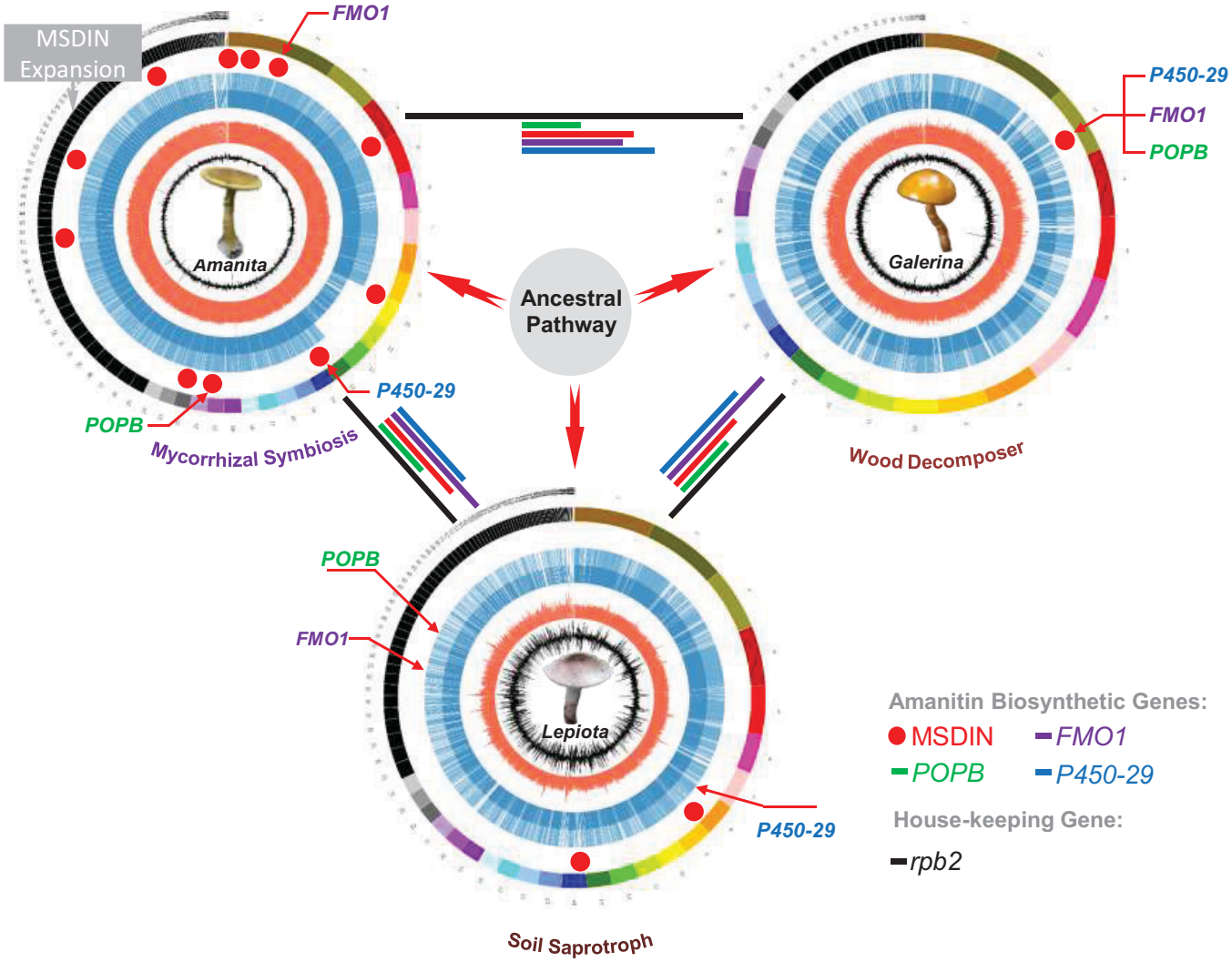

Fig. 5. Evolutionary outcome of the amanitin biosynthesis pathway in mushrooms. The modes of nutrition, mycorrhizal symbiosis, wood decomposition, and soil saprotroph are indicated below the genomes. Within the representative genomes for *Amanita*, *Galerina*, and *Lepiota*, predicted genes are indicated as blue lines, GC content as red lines, and GC skew as black lines. Four toxin biosynthetic genes are color-labeled to show their respective arrangements in the genomes. Genetic distances of the biosynthetic genes are similarly color-coded as lines connecting genome circles (housekeeping *rpb2* indicated as black lines). The thick red arrows represent the best hypothesis from comprehensive evaluations.

have similar distances to one another). In conclusion, our results based on genomic organization, phylogeny, and genetic distance strongly suggest that direct transfer of the pathway among *Amanita*, *Galerina*, and *Lepiota* is highly unlikely. We therefore posit the hypothesis involving an ancestral donor species (Fig. 5). The genome of this ancestral fungal species likely encoded a gene cluster for the amanitin biosynthesis that would have been transferred to other recipient species. Thereafter, amanitin genes dispersed throughout the genome as a result of genome rearrangements and were subsequently maintained by gene sequestration. Clustering of several genes would have been maintained in the *Galerina* and *Amanita* species, retaining some of the traits of the ancestral cluster.

We hypothesize that the amanitin biosynthesis pathway became sequestered within specific lineage genomes because of the loss of the gene cluster organization resulting from transposable element (TE)–driven genome fragmentation and reshuffling. The role of TEs is supported by the fact that TEs are found in the vicinities of most amanitin biosynthesis genes in all sequenced amanitin-producing mushrooms, including *Galerina* (Fig. 1*A*).

Amanitins are well known for their bioactive properties in humans, yet their role in the ecology of poisonous mushrooms is still uncertain. While the prevailing hypothesis that these toxins are used as a defense to mycophagy by invertebrates or vertebrates is plausible, this has not yet been confirmed by chemoecological studies. Is it by chance, or do the three distinct nutritional modes of the investigated mushrooms, i.e., wood decay by *Galerina*, soil saprotrophy by *Lepiota*, and ectomycorrhizal symbiosis by *Amanita*, play a role in shaping the pathways into their current fates? This is certainly a fascinating question to tackle in future research.

Materials and Methods

Samples. Fresh basidiocarps of amanitin-producing mushrooms were collected from the regions of Hunan (*Amanita fuliginea*, *G. marginata*, *G. sulciceps*) and Gansu (*L. brunneoincarnata*) in China. Two fruiting bodies of *Lepiota subincarnata* were collected in Italy. All sampled basidiocarps were immediately wrapped in aluminum foil and snap-frozen on dry ice and stored at −80 °C until further analyses. Most mushrooms were collected at or near sites well known for the presence of poisonous mushrooms as reported by the Chinese Center for Disease Control and Prevention. gDNA extraction and genome sequencing were carried out within one month after sampling to ensure quality.

Genome Sequencing. Morphological identification was done by Zhu L. Yang at the Kunming Institute of Botany. Further, sequencing of the rDNA internal

transcribed spacer (Dataset S2) was carried out before genomic sequencing to confirm the taxonomic affiliations. High-molecular-weight gDNA was then extracted from lyophilized fruiting bodies using the Genomic-tip 100/G kit (Qiagen 10243), according to the manufacturer's instructions. Except for two *Lepiota* strains (see below), genome sequencing was carried out at the Beijing Genomics Institute (BGI) and NextOmics Biosciences following their standard sequencing procedures.

At BGI, sequencing was done using the Illumina HiSeq 4000 and the Pacific Biosciences (PacBio) RSII with 250 bp, 10 kbp, and 20 kbp DNA libraries. PacBio polymerase reads < 1,000 bp or with a quality score < 80% were removed. Subreads were extracted from polymerase reads and were adapter filtered. Subreads were corrected using Pbdagcon (https://anaconda.org/bioconda/pbdagcon), Falcon (https://github.com/PacificBiosciences/FALCON-integrate), and Proovread (28). The resultant filtered reads were assembled using the Celera Assembler (29) (version 8.3; parameters: doTrim_initialQualityBased = 1, doTrim_finalEvidenceBased = 1, doRemoveSpurReads = 1, doRemoveChimericReads = 1, -d properties -U) or Falcon (version 0.3.0; parameters: -v -dal8 -t32 -h60 -e.96 -l500 -s100 -H3000). Scaffolds were constructed using SSPACE Basic (version 2.0) (30) and gap closing with PBJelly2 (31) (version 15.8.24 with default settings). The GATK (https://gatk.broadinstitute.org/hc/en-us) and SOAP tool packages (SOAP2, SOAPsnp, SOAPindel) (32, 33) were applied for single-base corrections.

At NextOmics Biosciences, high-quality gDNA was extracted as above and used for a 20 kbp library construction. The gDNA was then randomly fragmented using Covaris g-TUBE. Large DNA fragments were enriched and purified by magnetic beads and fragmented DNA was repaired. At the ends of DNA fragments, the stem circular sequencing joints were connected, and unconnected fragments were removed by exonuclease. A 20 kbp library was constructed using a PacBio template prep kit and analyzed using an Agilent 2100 Bioanalyzer for quality control. After the completion of the library, the DNA template and the enzyme complex were transferred to the PacBio Sequel sequencer for real-time single-molecule sequencing. The Illumina HiSeq ×10 platform was used for nucleotide level correction, based on a 350 bp library, and the company's standard method was applied.

The genomes of *L. subincarnata* 1 and *L. subincarnata* 2 strains were sequenced at the Research Technology Support Facility at Michigan State University. The main pipeline included ABySS (34) and ALLPATHS (35). For draft assembly, the two strains were sequenced using the Illumina HiSeq 4000, generating paired-end and mate-pair libraries for both samples. The libraries were cleaned using Trimmomatic (version 0.32; command line options: LEADING:20 TRAILING:20 SLIDINGWINDOW:5:20 MINLEN:85, phred33) (36) and FastQC (version 0.11.3) (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to remove adaptors and low-quality reads. The cleaned reads were then assembled using ABySS (version 1.5.2; parameters: k = 45). New overlapping paired-end Illumina MiSeq data were then received and cleaned using Trimmomatic and FastQC (same parameters as above). This dataset, along with the mate-pair libraries, was assembled using ALLPATHS-LG. For K-mer analysis, the distribution of read lengths was analyzed using Jellyfish (version 2/2.2.0) (37) to estimate genome size. The distribution graph of the reads showed two peaks. To estimate the size for exclusively homozygous reads, the area under the second curve was calculated.

Besides the above genomes sequenced in this study, those of *Amanita molliuscula*, *A. rimosa*, *A. pallidorosea*, *A. subjunquillea*, *Amanita exitialis*, *A. phalloides*, *A. bisporigera*, and *L. venenata* were generated from previous studies (9, 11, 16, 20, 21). In addition, a genome of a monokaryotic strain of *G. marginata* (France, CBS 339.88) was acquired through the Joint Genome Institute (19).

The quality of the genome assemblies, including the ones previously published by our group, was checked by BUSCO (38) and CEGMA (39) analyses (*SI Appendix*, Table S1).

Transcriptome sequencing was based on the Illumina RNAseq platform, and clean data were obtained through NextOmics Biosciences via the company's pipeline. Hisat2 (40) was then used to align transcriptome data with default parameters to produce Sam files. SAMtools (41) was then applied to convert Sam files to binary Bam files. Finally, the assembly was completed using StringTie (42) with default settings.

Genome Synteny Analysis. The synteny analysis was conducted via the SyMAP 4.2 (43). Alignments were computed using default SyMAP parameters except for the comparison of *G. marginata* loci to those of the *Amanita* and *Lepiota* species (Min Dots = 3, Top *n* = 2). Genomes were not repeat-masked prior to analysis. MUMmer (44) and BLAT (45) were used to compute the raw anchors. All scaffolds or contigs for all the genomes were loaded in the analysis.

Gene Prediction. Repeated sequences were identified using RepeatMasker (46), RepeatProteinMasker (47), and Tandem Repeats Finder (48). Repeat sequence prediction was carried out via RepeatModeler (http://www.repeatmasker.org), LTRfinder (49), and PILER (50). Three methods were then applied to predict genes. AUGUSTUS (51) and GENSCAN (52) were used to construct models for de novo prediction. GeneWise (53) was used to annotate homologous protein sequences. Transcriptome data and genome comparison were used to predict gene structure with Exonerate (https://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate). The predicted results were integrated using EVidenceModeler (54). TransposonPSI (http://transposonpsi.sourceforge.net/) was then applied to remove the genes containing transposons, resulting in the final gene set. The Eukaryotic Orthologous Groups, Gene Ontology, Kyoto Encyclopedia of Genes and Genomes, NCBI nonredundant, and UniProt databases were used for the functional annotation of predicted genes. Amanitin biosynthetic genes were manually annotated via homology comparison using BLAST (55).

BLAST Search. NCBI BLAST$^+$ 2.6.0. was used for BLAST searches. Nucleotide and amino acid sequences of amanitin biosynthetic genes (*MSDIN*, *POPB*, *P450-29*, and *FMO1*) from the sequenced genomes of *A. molliuscula*, *A. exitialis*, *A. phalloides*, *A. subjunquillea*, *A. fuliginea*, *A. rimosa*, *A. pallidorosea*, *A. bisporigera*, *G. marginata*, *G. sulciceps*, *L. brunneoincarnata*, *L. venenata*, and *L. subincarnata* were identified and captured through BLAST (NCBI BLAST$^+$ 2.6.0) alignments. The same search method also applied to all the transcriptomes.

Identification of *MSDIN* Genes. The abovementioned search strategy for *MSDIN* genes can miss sequences with a low percentage similarity. In *Lepiota* species, the C-terminal amino acid sequences of several MSDIN proteins showed a weak similarity to *Amanita* sequences. Consequently, we missed three *MSDIN* genes in *L. venenata* using this BLAST-based search. To find additional *MSDIN* sequences in the amanitin-producing mushrooms, we used a modified search strategy. A first round of queries was conducted with the known MSDIN protein and nucleotide sequences from *Amanita*, *Galerina*, and *Lepiota* species. The second search round used the most divergent *MSDIN* gene sequences resulting from the first search round. When no new *MSDIN* gene was detected, the query was terminated. When new *MSDIN* sequences were found, a third search round was performed using the new *MSDIN* sequences as queries. Three search rounds allowed a thorough identification of *MSDIN* genes for the present species. We applied this search strategy to the published genomes of *A. bisporigera* and *A. phalloides* (16), leading to the identification of additional unpublished *MSDIN* genes.

Venn Diagram Construction. Predicted protein sequences of *A. subjunquillea*, *A. molliuscula*, *L. venenata*, *G. marginata*, and *G. sulciceps* were clustered in orthogroups using OrthoFinder (56) with default settings. The output dataset was used to construct Venn diagrams using JVENN (http://jvenn.toulouse.inra.fr/app/example.html).

Construction of Amanitin Biosynthetic Loci in *Galerina* Species. Nucleotide BLAST was used to locate orthologs to *GmAMA1-1*–the α-amanitin–encoding *MSDIN* of *G. marginata*–in the genome assemblies. Approximately 150-kbp upstream and downstream regions of *GmAMA1-1* (300 kbp in total) located in the genomes of *G. marginata* (from China), *G. sulciceps*, and *G. marginata* (CBS 339.88) were chosen for further analysis. The predicted genes situated in the selected regions were then annotated and compared manually. An orthologous gene was determined by finding the best homologous hit for a given gene.

Visualization of Genomes and Amanitin Biosynthetic Genes. Circos (57) was used to map genomes and amanitin biosynthetic genes. Python scripts for obtaining guanine and cytosine (GC) content and GC skew were generated. Genome annotation files were processed mainly through Excel, and the resulting track files were used for building information tracks in the Biopython environment. The tracks were loaded into Circos to produce a genome overview in the Perl environment. Coordinates of amanitin biosynthetic genes were loaded as a

track to show their precise genomic locations. In addition, the coordinates of the biosynthetic genes in each genome were manually labeled in the graph, and the sizes of each contig and scaffold were shown in scale.

Cloning of *CYP450* Genes from *A. bisporigera*. PCR primers unique to *AbP450-1*, *AbP450-2*, and *AbP450-3* were designed and used for isolating the genomic clones of each gene. For *AbP450-1*, primers used were 5′-CTCCAATCCCCCAACCACAAA- 3′ (forward) and 5′-GTCGAACACGGCAACAACAG-3′ (reverse). For *AbP450-2*, the primers were 5′-GAAAACCGAATCTCCAATCCTC-3′ (forward) and 5′-AGCTCACTCGTTGCCACTAA-3′ (reverse). For *AbP450-3*, the forward primers were 5′-TTTAGGGCAGTGATTTCGTGACA-3′ and 5′-AACAGGGAGGCGATTATTCAAC-3′.

gDNA sequences were used for primer design to obtain full-length cDNAs by Rapid Amplification of cDNA Ends (RACE) using the SMART RACE kit (Clontech). A cDNA copy of *AbP450-1* was obtained using the following primers: 5′-CCAACGACAGGCGGGACACG-3′ (5′-RACE) and 5′-GACCTTTTTGCTTTAACATCTACA-3′ (3′-RACE), and for *AbP450-2* with the primers 5′-GTCAACAAGTCCAGGAGACATTCAAC-3′ (5′-RACE) and 5′-ACCGAATCTCCAATCCTCCAACCA-3′0 (3′-RACE). The RACE primers for *AbP450-3* were 5′-CGGCGTTCCAAGGCGATGATAATA-3′ (5′-RACE) and 5′-CATCTCCATCGACCCCTTTTTCAGC-3′ (3′-RACE).

Sequences generated from the RACE reactions were used to assemble full-length cDNAs of all three genes. Alignments of genomic and cDNA copies were done using Spidey (https://www.ncbi.nlm.nih.gov/spidey/) and Splign (https://www.ncbi.nlm.nih.gov/sutils/splign/splign.cgi).

Genomic Lambda Library of *A. bisporigera*. A lambda gDNA library of *A. bisporigera* was prepared using the λBlueSTAR Vector System (Novagen 69242) according to the manufacturer's instructions. Screening of the library was performed using a standard Southern hybridization process, using *AMA1* and *PHA1* (the phallacidin-encoding *MSDIN*) as the probes. Positive clones were sequenced using the standard primer walking technique (58).

DNA Isolation from *G. marginata*. Mycelium of *G. marginata* was cultured in liquid medium (HSV-2C) for 15 to 25 d with rotary shaking at 120 rpm at 23 °C. The medium (per liter) contained 1 g yeast extract, 2 g glucose, 0.1 g NH_4Cl, 0.1 g $CaSO_4{\cdot}5H_2O$, 1 mg thiamine·HCl, and 0.1 mg biotin, pH 5.2 (59). DNA extraction was performed using lyophilized fruiting bodies or cultures with the DNeasy Plant Mini Kit (Qiagen 69106) (for DNA blot hybridization) and the Genomic-tip 100/G (Qiagen 10243) (for constructing genomic libraries and genome sequencing), following the manufacturer's protocols.

Gene Knockout in *G. marginata*. *Agrobacterium*-mediated transformation of *G. marginata* was performed as described earlier (10). For the selected genes, ∼1.5-kbp upstream and downstream sequences were used for homologous recombination (Dataset S3). The fragments were cloned into the pHg vector (60) and transformed into *Agrobacterium tumefaciens* strain LBA1100 (61). The vectors were linearized before transformation to facilitate crossing over both upstream and downstream of the gene of interest and hence stable gene deletion. Ectopic integration was estimated at ∼10% of all transformants. All the transformants for each gene were analyzed, among which at least four were chemically analyzed. Conclusions about the transformant phenotype were drawn only if all independent deletion transformants displayed the same phenotype.

DNA Blot Hybridization. Probe labeling, DNA blotting, and filter hybridization followed standard protocols (62). DNA for blotting was cut with *Pst*I and electrophoresed in 0.7% agarose. Hybridizations were performed overnight at 65 °C in 4× SET buffer plus 0.1% sodium pyrophosphate, 0.2% SDS, 10% dextran sulfate, and 625 μg/mL heparin. SET buffer (20×) contains 3 M NaCl, 0.6 M Tris, and 0.04 M EDTA, pH 7.4. For probe preparation, gene fragments (∼500 bp) generated by PCR for each candidate gene in *A. bisporigera* or *G. marginata* were used. For *G. marginata*, genes selected for blotting were located upstream and downstream of *GmAMA1-1* and *GmPOPB* (*SI Appendix*, Table S2). For the *Amanita* species, three *CYP450* genes were chosen based on the lambda clone results.

Phylogenetic and Genetic Distance Analyses of *rpb2*, *FMO1*, and *P450-29*. The method adopted in the present study is similar to that in our previous reports (11, 12). Coding sequences (CDS) of *rpb2*, *FMO1*, and *P450-29* from the genomes used in this study were identified by tBLASTn (NCBI BLAST$^+$ 2.12.0) and amino acid sequences of GmP450-29 and GmFMO1 were applied as query sequences. Accurate CDS sequences were acquired using respective transcriptomes as the reference. The CDSs were aligned by MAFFT v7.304b (63) with default settings. The taxa included representative species from *Amanita*, *Galerina*, and *Lepiota*. For the alignment, GTR + I + G was selected as the best model for the CDSs of *rpb2*, *FMO1*, and *P450-29* genes, using MrModeltest v2.3 (64) under Akaike Information Criterion. To maintain the correct topology of the *rbp2* species tree, more species were selected than the biosynthesis genes (nonorthologous) in the gene trees (*FMO1* and *P450-29*). Maximum likelihood analyses and bootstrapping (1,000 replicates) were performed using RAxML v7 (65). The Codeml program in PAML v4.9 (66) was used for postphylogenetic analysis, including genetic distance and divergence rate calculations.

Characterization of Amanitin Peptides by Liquid Chromatography–Mass Spectrometry. We analyzed the amanitin-related peptides by using the Agilent Model 1200 high-performance liquid chromatography (HPLC) system coupled to an ultraviolet detector (monitoring at 280, 295, and 305 nm) and an Agilent 6120 mass spectrometer. The HPLC elution solution A was 0.02 M ammonium acetate:acetonitrile (90:10, vol/vol), adjusted to pH 5 with glacial acetic acid, and solution B was 0.02 M aqueous ammonium acetate:acetonitrile (76:24, vol/vol), pH 5 (38).

θ-Amanitin was purified from the *P450-29* mutant in two steps. The first separation was performed on a semipreparative C18 column (25 cm × 10 mm, 5 mm; Supelcosil LC-18, Supelco). The flow rate was 2 mL/min with a stepwise gradient profile of 100% A for 3 min, 43% A for 7 min, and 0% A for 9 min. Fractions containing θ-amanitin were pooled, lyophilized, and then redissolved in H_2O. The second separation was performed on a 250 × 4.6 mm C18 column (Higgins Analytical, http://www.higanalyt.com). The flow rate was 1 mL/min with a gradient of 100% solution A to 100% solution B in 15 min. The fractions containing θ-amanitin were collected, dried under vacuum, and redissolved in H_2O.

NMR. θ-Amanitin (820.6 μg) was dissolved in 200 μL DMSO-d6 to a final concentration of 4.63 mM. 4,4-Dimethyl-4-silapentane-1-sulfonic acid (10 mM) (Sigma-Aldrich) was included as the chemical shift reference. Then, 1-D and 2-D spectra were recorded at 25 °C on a Bruker Avance 900 MHz spectrometer using a TCI cryoprobe at the Max T. Rogers NMR Facility, Michigan State University.

Divergence Estimation of the Genera *Amanita*, *Galerina*, and *Lepiota*. Given that fossil records of fungi are limited, it has been difficult to choose a reliable calibration point to estimate the divergence time for any fungal groups. Therefore, an extensive sampling of outgroup species for which fossils were available were selected to estimate the divergence time of those species producing cyclic peptide toxins. The split between Ascomycota and Basidiomycota inferred from the fossil *Paleopyrenomycites devonicus* was used as the calibration. From that point, a normal distribution with a mean of 582.5 Mya and an SD of 50.15 Mya was applied (67). Three gene fragments, nrLSU, *rpb2*, and *tef1-α*, were concatenated for molecular dating using the phylogenetic framework described in James et al. (68). Nucleotide sequences were retrieved from GenBank, the Assembling the Fungal Tree of Life database, JGI (*SI Appendix*, Table S3), and the genomes of this study. All introns within *rpb2* and *tef1-α* were removed due to the difficulty in alignment when large numbers of less closely related taxa were included. MrModeltest version 2.3 (64) was used to select the best models of evolution using the hierarchical likelihood ratio test. The divergence time was estimated in BEAST version 1.6.1 (69), with the molecular clock and substitutions models unlinked but the trees linked for each gene partition. The BEAST input files were constructed using BEAUti (within BEAST), in which the GTR + G + I model (MrModeltest output) was selected. The lognormal relaxed molecular clock model and the Yule speciation prior set were used to estimate the divergence times and the corresponding credibility intervals. The posterior distributions of parameters were obtained using Markov chain Monte Carlo analysis for 100 million generations with a burn-in percentage of 25%. The convergence of the chains was checked using Tracer version 1.5 (http://tree.bio.ed.ac.uk/software/tracer/2013) to confirm that the analysis reached a stationary distribution. Samples from the posterior distributions were summarized on a maximum clade credibility tree with the maximum sum of posterior probabilities on its internal nodes using TreeAnnotator version 1.8.1 with the posterior probabilities limit set to 0.5 to summarize the mean node heights. FigTree version 1.4.4 (http://

tree.bio.ed.ac.uk/software/figtree/) was used to visualize the resulting tree and to obtain the means and 95% higher posterior densities.

Data Availability. The Whole Genome Shotgun assemblies are available at NCBI (BioProject PRJNA679796). All additional study data are included in the article and/or supporting information.

ACKNOWLEDGMENTS. This research was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDB31010000), the National Natural Science Foundation of China (No. 31972477, 31772377), the International Partnership Program of the Chinese Academy of Sciences (No. 151853KYSB20170026), and the Scientific Research Foundation of the Kunming Institute of Botany, Chinese Academy of Sciences. F.M.M.'s research is supported by the Laboratory of Excellence ARBRE (ANR-11-LABX-0002-01) and the Beijing Advanced Innovation Center for Tree Breeding by Molecular Design. Dr. Jonathan D. Walton has substantially contributed to this work over the past eight years. His saddening passing prevented him from approving the final version of this manuscript. At the request of his family, his name was not associated with the present paper, but we would like to acknowledge his outstanding contribution to this work and thank his family for everlasting support. In addition, we thank Dr. Qing Cai for her kind assistance with the divergence time analysis.

Author affiliations: [a]CAS Key Laboratory for Plant Diversity and Biogeography of East Asia, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China; [b]Yunnan Key Laboratory for Fungal Diversity and Green Development, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China; [c]Department of Food Science and Technology, University of Nebraska-Lincoln, Lincoln, NE 68588; [d]College of Life Sciences, University of the Chinese Academy of Sciences, Beijing 100049, China; [e]Department of Energy Plant Research Laboratory, Michigan State University, East Lansing, MI 48824; [f]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824; [g]Department of Environmental Science and Engineering, Kunming University of Science and Technology, Kunming 650091, China; [h]Beijing Advanced Innovation Centre for Tree Breeding by Molecular Design, Institute of Microbiology, Beijing Forestry University, Beijing 100083,China; and [i]Université de Lorraine, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, Unité Mixte de Recherche Interactions Arbres/Microorganismes, 54280 Champenoux, France

1. J. D. Walton, *The Cyclic Peptide Toxins of Amanita and Other Poisonous Mushrooms* (Springer International Publishing AG, Cham, Switzerland, 2018).
2. Z. L. Yang, *Atlas of the Chinese Species of Amanitaceae* (Science Press, Beijing, China, 2015).
3. T. Wieland, *Peptides of Poisonous Amanita Mushrooms* (Springer-Verlag New York Inc., New York, NY, 1986).
4. I. B. Buchwalow, W. Böcker, *Immunohistochemistry: Basics and Methods* (Springer-Verlag Berlin Heidelberg, Heidelberg, Germany, 2010).
5. Y. Liu *et al.*, TP53 loss creates therapeutic vulnerability in colorectal cancer. *Nature* **520**, 697–701 (2015).
6. K. Matinkhoo, A. Pryyma, M. Todorovic, B. O. Patrick, D. M. Perrin, Synthesis of the death-cap mushroom toxin alpha-amanitin. *J. Am. Chem. Soc.* **140**, 6513–6517 (2018).
7. M. J. Siegert, C. H. Knittel, R. D. Süssmuth, A convergent total synthesis of the death cap toxin alpha-amanitin. *Angew. Chem. Int. Ed. Engl.* **59**, 5500–5504 (2020).
8. D. A. Bushnell, P. Cramer, R. D. Kornberg, Structural basis of transcription: Alpha-amanitin-RNA polymerase II cocrystal at 2.8 A resolution. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 1218–1222 (2002).
9. H. E. Hallen, H. Luo, J. S. Scott-Craig, J. D. Walton, Gene family encoding the major toxins of lethal *Amanita* mushrooms. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19097–19101 (2007).
10. H. Luo *et al.*, Peptide macrocyclization catalyzed by a prolyl oligopeptidase involved in α-amanitin biosynthesis. *Chem. Biol.* **21**, 1610–1617 (2014).
11. Y. Lüli *et al.*, Genome of lethal *Lepiota venenata* and insights into the evolution of toxin-biosynthetic genes. *BMC Genomics* **20**, 198 (2019).
12. H. Luo *et al.*, The MSDIN family in amanitin-producing mushrooms and evolution of the prolyl oligopeptidase genes. *IMA Fungus* **9**, 225–242 (2018).
13. H. T. Reynolds *et al.*, Horizontal gene cluster transfer increased hallucinogenic mushroom diversity. *Evol. Lett.* **2**, 88–101 (2018).
14. E. Gluck-Thaler, J. C. Slot, Dimensions of horizontal gene transfer in eukaryotic microbial pathogens. *PLoS Pathog.* **11**, e1005156 (2015).
15. J. C. Slot, Fungal gene cluster diversity and evolution. *Adv. Genet.* **100**, 141–178 (2017).
16. J. A. Pulman, K. L. Childs, R. M. Sgambelluri, J. D. Walton, Expansion and diversification of the MSDIN family of cyclic peptide genes in the poisonous agarics *Amanita phalloides* and *A. bisporigera*. *BMC Genomics* **17**, 1038 (2016).
17. Y. Y. Cui, Q. Cai, L. P. Tang, J. W. Liu, Z. L. Yang, The family Amanitaceae: Molecular phylogeny, higher-rank taxonomy and the species in China. *Fungal Divers.* **91**, 5–230 (2018).
18. F. Martin, A. Kohler, C. Murat, C. Veneault-Fourrey, D. S. Hibbett, Unearthing the roots of ectomycorrhizal symbioses. *Nat. Rev. Microbiol.* **14**, 760–773 (2016).
19. R. Riley *et al.*, Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9923–9928 (2014).
20. S. Zhou *et al.*, Novel cyclic peptides from lethal *Amanita* mushrooms through a genome-guided approach. *J. Fungi (Basel)* **7**, 204 (2021).
21. Y. Lüli *et al.*, Differential expression of amanitin biosynthetic genes and novel cyclic peptides in *Amanita molliuscula*. *J. Fungi (Basel)* **7**, 384 (2021).
22. H. Luo, H. E. Hallen-Adams, J. S. Scott-Craig, J. D. Walton, Ribosomal biosynthesis of α-amanitin in *Galerina marginata*. *Fungal Genet. Biol.* **49**, 123–129 (2012).
23. P. B. Matheny, J. M. Curtis, V. Hofstetter, Major clades of Agaricales: A multi-locus phylogenetic overview. *Mycologia* **98**, 984–997 (2007).
24. S. Eswaramoorthy, J. B. Bonanno, S. K. Burley, S. Swaminathan, Mechanism of action of a flavin-containing monooxygenase. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 9832–9837 (2006).
25. T. Wieland, Chemical and toxicological studies with cyclopeptides of *Amanita phalloides*. *Pure Appl. Chem.* **6**, 309–350 (1963).
26. S. Muraoka, N. Fukamachi, K. Mizumoto, T. Shinozawa, Detection and identification of amanitins in the wood-rotting fungi *Galerina fasciculata* and *Galerina helvoliceps*. *Appl. Environ. Microbiol.* **65**, 4207–4210 (1999).
27. R. M. Sgambelluri *et al.*, Profiling of amatoxins and phallotoxins in the genus *Lepiota* by liquid chromatography combined with UV absorbance and mass spectrometry. *Toxins (Basel)* **6**, 2336–2347 (2014).
28. T. Hackl, R. Hedrich, J. Schultz, F. Förster, proovread: Large-scale high-accuracy PacBio correction through iterative short read consensus. *Bioinformatics* **30**, 3004–3011 (2014).
29. E. W. Myers *et al.*, A whole-genome assembly of *Drosophila*. *Science* **287**, 2196–2204 (2000).
30. M. Boetzer, W. Pirovano, SSPACE-longread: Scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics* **15**, 211 (2014).
31. A. C. English *et al.*, Mind the gap: Upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS One* **7**, e47768 (2012).
32. R. Li *et al.*, SNP detection for massively parallel whole-genome resequencing. *Genome Res.* **19**, 1124–1132 (2009).
33. R. Li *et al.*, SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966–1967 (2009).
34. I. Birol *et al.*, De novo transcriptome assembly with ABySS. *Bioinformatics* **25**, 2872–2877 (2009).
35. J. Butler *et al.*, ALLPATHS: De novo assembly of whole-genome shotgun microreads. *Genome Res.* **18**, 810–820 (2008).
36. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
37. G. Marçais, C. Kingsford, A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764–770 (2011).
38. M. Seppey, M. Manni, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness. *Methods Mol. Biol.* **1962**, 227–245 (2019).
39. G. Parra, K. Bradnam, I. Korf, CEGMA: A pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **23**, 1061–1067 (2007).
40. D. Kim, B. Langmead, S. L. Salzberg, HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).
41. H. Li *et al.*; 1000 Genome Project Data Processing Subgroup, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
42. M. Pertea *et al.*, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
43. C. Soderlund, M. Bomhoff, W. M. Nelson, SyMAP v3.4: A turnkey synteny system with application to plant genomes. *Nucleic Acids Res.* **39**, e68 (2011).
44. A. L. Delcher, S. L. Salzberg, A. M. Phillippy, Using MUMmer to identify similar regions in large sequence sets. *Curr. Protoc. Bioinformatics* **Chapter 10**, Unit 10.3 (2003).
45. W. J. Kent, BLAT–The BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
46. A. F. Smit, The origin of interspersed repeats in the human genome. *Curr. Opin. Genet. Dev.* **6**, 743–748 (1996).
47. M. Tarailo-Graovac, N. Chen, Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics* **Chapter 4**, Unit 4.10 (2009).
48. G. Benson, Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
49. Z. Xu, H. Wang, LTR_FINDER: An efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265-8 (2007).
50. R. C. Edgar, PILER-CR: Fast and accurate identification of CRISPR repeats. *BMC Bioinformatics* **8**, 18 (2007).
51. M. Stanke, R. Steinkamp, S. Waack, B. Morgenstern, AUGUSTUS: A web server for gene finding in eukaryotes. *Nucleic Acids Res.* **32**, W309-12 (2004).
52. C. Burge, S. Karlin, Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**, 78–94 (1997).
53. E. Birney, M. Clamp, R. Durbin, GeneWise and genomewise. *Genome Res.* **14**, 988–995 (2004).
54. B. J. Haas *et al.*, Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* **9**, R7 (2008).
55. S. F. Altschul *et al.*, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
56. D. M. Emms, S. Kelly, OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157 (2015).
57. M. Krzywinski *et al.*, Circos: An information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).
58. F. Sterky, J. Lundeberg, Sequence analysis of genes and genomes. *J. Biotechnol.* **76**, 1–31 (2000).
59. S. Muraoka, T. Shinozawa, Effective production of amanitins by two-step cultivation of the basidiomycete, *Galerina fasciculata* GF-060. *J. Biosci. Bioeng.* **89**, 73–76 (2000).
60. M. Kemppainen, A. Circosta, D. Tagu, F. Martin, A. G. Pardo, Agrobacterium-mediated transformation of the ectomycorrhizal symbiont *Laccaria bicolor* S238N. *Mycorrhiza* **16**, 19–22 (2005).
61. M. J. Kemppainen, A. G. Pardo, pHg/pSILBAγ vector system for efficient gene silencing in homobasidiomycetes: Optimization of ihpRNA-triggering in the mycorrhizal fungus *Laccaria bicolor*. *Microb. Biotechnol.* **3**, 178–200 (2010).
62. J. S. Scott-Craig, D. G. Panaccione, F. Cervone, J. D. Walton, Endopolygalacturonase is not required for pathogenicity of *Cochliobolus carbonum* on maize. *Plant Cell* **2**, 1191–1200 (1990).

63. K. Katoh, D. M. Standley, A simple method to control over-alignment in the MAFFT multiple sequence alignment program. *Bioinformatics* **32**, 1933–1942 (2016).
64. J. A. A. Nylander, *MrModeltest v2.2 Uppsala: Evolutionary Biology Centre* (Uppsala University, 2004).
65. A. Stamatakis, RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688–2690 (2006).
66. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
67. R. Lücking, S. Huhndorf, D. H. Pfister, E. R. Plata, H. T. Lumbsch, Fungi evolved right on track. *Mycologia* **101**, 810–822 (2009).
68. T. Y. James *et al.*, Reconstructing the early evolution of fungi using a six-gene phylogeny. *Nature* **443**, 818–822 (2006).
69. A. J. Drummond, M. A. Suchard, D. Xie, A. Rambaut, Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).