



# Unmatched Level of Molecular Convergence among Deeply Divergent Complex Multicellular Fungi

Zsolt Merényi,<sup>1</sup> Arun N. Prasanna,<sup>1,†</sup> Zheng Wang,<sup>2</sup> Károly Kovács,<sup>1,3</sup> Botond Hegedüs,<sup>1</sup> Balázs Bálint,<sup>1</sup> Balázs Papp,<sup>1,3</sup> Jeffrey P. Townsend ,<sup>2,4,5</sup> and László G. Nagy \*<sup>1</sup>

<sup>1</sup>Synthetic and Systems Biology Unit, Institute of Biochemistry, Biological Research Center, Szeged, Hungary

<sup>2</sup>Department of Biostatistics, Yale University, New Haven, CT

<sup>3</sup>Hungarian Centre of Excellence for Molecular Medicine, Metabolic Systems Biology Lab, Szeged, Hungary

<sup>4</sup>Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT

<sup>5</sup>Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT

<sup>†</sup>Present address: Red Sea Science and Engineering Research Center, 4700 King Abdullah University of Science and Technology (KAUST), Thuwal, Kingdom of Saudi Arabia

\*Corresponding author: E-mail: lnagy@fun genomelab.com.

Associate editor: Tal Pupko

## Abstract

Convergent evolution is pervasive in nature, but it is poorly understood how various constraints and natural selection limit the diversity of evolvable phenotypes. Here, we analyze the transcriptome across fruiting body development to understand the independent evolution of complex multicellularity in the two largest clades of fungi—the Agarico- and Pezizomycotina. Despite >650 My of divergence between these clades, we find that very similar sets of genes have convergently been co-opted for complex multicellularity, followed by expansions of their gene families by duplications. Over 82% of shared multicellularity-related gene families were expanding in both clades, indicating a high prevalence of convergence also at the gene family level. This convergence is coupled with a rich inferred repertoire of multicellularity-related genes in the most recent common ancestor of the Agarico- and Pezizomycotina, consistent with the hypothesis that the coding capacity of ancestral fungal genomes might have promoted the repeated evolution of complex multicellularity. We interpret this repertoire as an indication of evolutionary predisposition of fungal ancestors for evolving complex multicellular fruiting bodies. Our work suggests that evolutionary convergence may happen not only when organisms are closely related or are under similar selection pressures, but also when ancestral genomic repertoires render certain evolutionary trajectories more likely than others, even across large phylogenetic distances.

**Key words:** convergent evolution, complex multicellularity, fruiting body development, transcriptome, fungi.

## Introduction

Darwin suggested that organisms can evolve an endless variety of forms (Darwin and Mayr 1995; McGhee 2011). Contrary to his concept of “endless forms,” it is now clear that evolution follows similar paths more often than classic models of genetic change would predict (Shubin et al. 2009; Stern and Orgogozo 2009; McGhee 2011; Blount et al. 2018). The independent emergence of similar phenotypes is called convergent evolution, a phenomenon which still remains an enigma in evolutionary biology. Convergent evolution is known to occur in response to similar selection pressures, bias in the mutational origin of phenotypic variation (cf., Rice and Townsend 2012; Park et al. 2017), or both (Yampolsky and Stoltzfus 2001; Stern and Orgogozo 2009; Braendle et al. 2010; Losos 2011; Blount et al. 2018; Xie et al. 2019; Stoltzfus and Mccandlish 2017). Convergence is widespread in nature (e.g., Rundle et al. 2000; Muschick et al. 2012; Mahler et al. 2013; van Velzen et al. 2018) and its pervasive occurrence suggests that the evolution of certain traits may be predictable (Pankey et al. 2014) and deterministic (Losos

1998; Mahler et al. 2013; Blount et al. 2018) under some circumstances.

What drives divergent lineages to evolve similar phenotypes is poorly understood. A fascinating example of convergent phenotypes is multicellularity: It has evolved at least 25–30 times across the pro- and eukaryotes (Grosberg and Strathmann 2007; Ruiz-Trillo et al. 2007; Knoll 2011; Claessen et al. 2014; Nagy 2018; Nagy et al. 2018) and has reached complexity levels ranging from simple cell aggregates to the most complex macroscopic organisms (Sebé-Pedrés et al. 2017). Instances of the evolution of multicellularity are considered major transitions in evolution—a conceptual label that is difficult to reconcile with repeated origins (Grosberg and Strathmann 2007; Knoll 2011; Nagy 2017). This difficulty follows from the assumption that major transitions are limited by big genetic hurdles and thus should occur rarely during evolution (Smith and Szathmáry 1995; Knoll 2011).

The highest level of multicellular organization is referred to as complex multicellularity (CM), which, unlike unicells and simple multicelled aggregates (e.g., filaments, colonies,

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

Open Access

biofilms, etc.), is characterized by a three-dimensional organization, sophisticated mechanisms for cell–cell adhesion and communication, and extensive cellular differentiation (Knoll 2011; Lord and Read 2011; Sebé-Pedrós et al. 2017; Nagy 2018). CM occurs in metazoans, embryophyte plants, and fungi as well as red and brown algae. In fungi, CM usually refers to sexual fruiting bodies, which are found in 8–11 disparate fungal clades and show clear signs of convergent origins (Nagy et al. 2018). Although they originated independently, CM fungal clades are phylogenetically close enough to provide a tractable system for studying the genetics of major evolutionary transitions in complexity. Fruiting bodies in fungal lineages can be developmentally and morphologically highly distinct—yet there is strong evidence that they repeatedly evolved for the same general purpose: to enclose sexual reproductive structures in a protective environment and to facilitate spore dispersal (Hibbett 2007; Trail et al. 2017; Nagy 2018; Pöggeler et al. 2018; Krizsán et al. 2019). Here, we seek to explain the convergent evolution of fungal fruiting bodies by analyzing the fate of multicellularity-related gene families across the two largest clades of CM fungi: the Agaricomycotina (mushroom-forming fungi, Basidiomycota) and the Pezizomycotina (Ascomycota). We use a comparative transcriptomic and phylogenomic approach to first find developmentally regulated genes shared by the Agarico- and Pezizomycotina, then to analyze the evolution of these gene families to uncover genetic mechanisms of convergence in fungal CM.

## Results

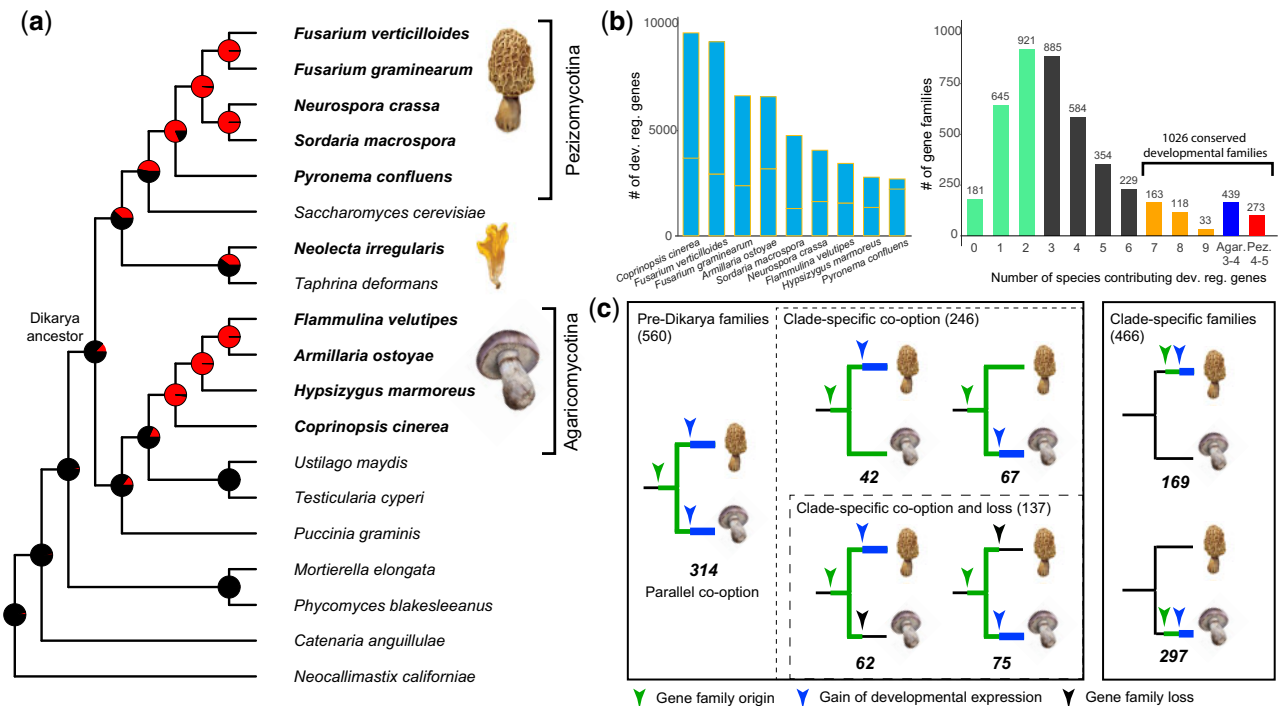
### Multiple Origins of CM in Fungi

We assembled two whole-genome data sets, one comprising 19 and another containing 90 species (supplementary table S1ab). For simplicity, we discuss in detail the 19-species data set in the main text and present results for the 90-species data set in the Supplementary Material online. To study the evolution of CM in fungi, we first inferred ancestral cellularity levels across the phylogenetic tree of 90 species (Supplementary Material online) using ancestral character state reconstruction. This analysis strongly suggests that the most recent common ancestor of the two most diverse CM clades, Agarico- and Pezizomycotina, did not form fruiting bodies (marginal probability of non- $CM_{19sp}$ : 0.87,  $CM_{19sp}$ : 0.13 and non- $CM_{90sp}$ : 1.0,  $CM_{90sp}$ : 0 for trees of 19 and 90 species, respectively; supplementary note 1 and fig. S1, Supplementary Material online, and fig. 1a), followed by two independent acquisitions of CM in the most recent common ancestor (MRCA) of the Agaricomycotina and that of the Pezizomycotina. This analysis also suggests that the MRCA of the Ascomycota did not produce fruiting bodies (state non- $CM_{19sp}$ : 0.616, state  $CM_{19sp}$ : 0.384 and state non- $CM_{90sp}$ : 1.0, state  $CM_{90sp}$ : 0), which implies a third origin of fruiting body production in *Neolecta* (Taphrinomycotina; Nagy 2017; Nguyen et al. 2017; Nagy et al. 2018; fig. 1a and supplementary fig. S1, Supplementary Material online). Consistent with reconstructed ancestral states, the best-fit evolutionary model implied that gains of CM occurred at

nonzero rate values (0.21 and 6.09 in maximum likelihood [ML] and stochastic mapping, respectively), whereas the rate of the loss of CM was 0 in both analyses of this 90-species data set. These model parameters strongly suggest that the acquisition of CM has directionality and is not a reversible trait under historical circumstances. Consistent with this directionality, losses of CM are not known or very rare in the Agaricomycotina and the Pezizomycotina, with some ant-associated fungi not known to produce fruiting bodies and may be transmitted mainly vertically (Mikheyev et al. 2006). Analyses of the 90-species data set showed identical results (supplementary note 1 and fig. S1, Supplementary Material online). These results are consistent with the general consensus on the independence of fruiting bodies in the Basidio- and Ascomycota (e.g., Stajich et al. 2009).

### Analyses of RNA-Seq Data

To identify developmentally relevant genes, we used publicly available fruiting body transcriptomes of four Agaricomycotina (Park et al. 2014; Muraguchi et al. 2015; Zhang et al. 2015) and five Pezizomycotina (Sikhakolli et al. 2012; Teichert et al. 2012; Traeger et al. 2013; Wang, Lopez-Giraldez, et al. 2014) species, in which we quantified gene expression across 2–13 developmental stages. We defined developmentally regulated genes as ones showing >4-fold change in expression across development. In the transition from vegetative mycelium to fruiting bodies (the first stage of which is either primordia or protoperithecia), we only considered upregulated genes as developmentally regulated (see Materials and Methods and Krizsán et al. [2019] for a more detailed definition). We detected 2,645–9,444 developmentally regulated genes in the nine species (supplementary fig. S2, Supplementary Material online), corresponding to 20–66% of the proteome (lowest in *Pyronema confluens*, highest in *Coprinopsis cinerea*). The identified developmentally regulated genes contained 27–98%, 5–70%, and 93% of known developmental genes of *Neurospora* (Colot et al. 2006; Trail et al. 2017), *Aspergillus* (Dyer and O’Gorman 2012), and *Coprinopsis* (see supplementary note 3 and table S2, Supplementary Material online), respectively, and their functional annotations were consistent with previous studies of fruiting body development (Sakamoto et al. 2011; Trail et al. 2017; Nagy 2018; Pöggeler et al. 2018; Krizsán et al. 2019). In the 19-species data set (see Materials and Methods), developmentally regulated genes fell into 21,267 families, of which we focused on families that showed developmental regulation in the majority of species (fig. 1b). We identified 1,026 gene families that were developmentally regulated in  $\geq 75\%$  of the species in either or both clades. These comprised 314, 439, and 273 families that showed a developmental expression in  $\geq 7$  of 9 Dikarya,  $\geq 3$  of 4 Agaricomycotina, and  $\geq 4$  of 5 Pezizomycotina species, respectively (supplementary table S3, Supplementary Material online). We hereafter focus on these families because these are most likely to have been developmentally regulated also in the most recent common ancestor of Agaricomycotina and/or that of the Pezizomycotina. We hereafter refer the 314 gene families that are developmentally



**FIG. 1.** The evolution of complex multicellularity in fungi and conserved developmentally regulated gene families. (a) Phylogenetic relationships among 19 species analyzed in this study inferred from 86 conserved, single-copy orthologs. Two independent clades of complex multicellular species are marked, and typical fruiting body morphologies are shown. Pie charts at nodes indicate the proportional likelihoods of CM (red) and non-CM (black) ancestral states reconstructed using ML. Character state coding of extant species is shown as bold (CM) or regular (non-CM) font. (b) The number of developmentally regulated genes detected in each of the nine species (left) and the number of gene families in which these genes belong. The 314 gene families shared by  $\geq 7$  species are highlighted in yellow. Groups of gene families that are developmentally regulated in  $\geq 3$  Agaricomycotina or  $\geq 4$  Pezizomycotina are also shown. (c) Developmentally regulated gene families grouped by evolutionary conservation and history.

regulated in  $\geq 7$  species (both Agarico- and Pezizomycotina) to as shared developmentally regulated gene families.

### Shared Developmentally Regulated Families

We analyzed functional annotations and queried experimentally identified members or orthologs of the 314 gene families that showed a developmental expression in  $\geq 7$  of 9 species. First, as CM fruiting bodies evolved to enclose sexual reproductive processes into a protected environment, we tested whether shared developmentally regulated families comprise meiosis-related gene families. We found that only 4 of the 314 shared developmentally regulated families contained meiotic genes, whereas 78 and 27 meiosis-related families were developmentally regulated in 3–6 and 0–2 species, respectively (supplementary table S4 and fig. S3, Supplementary Material online). Because environmental stress is a common trigger of sexual development, we also examined the contribution of stress-related genes to shared developmental expression patterns. We identified homologs of Core Environmental Stress Response genes and members of the stress activated MAPK pathway (Gómez-Gil et al. 2019) in our species and examined their expression. Very few of these stress-related genes were developmentally regulated in the examined transcriptomes and none of them were so in seven to nine species (supplementary table S5, Supplementary Material online), indicating a low contribution of stress-related genes to the observed developmental expression

patterns. These observations suggest that the shared developmental expression of the 314 gene families is not or is only to a small extent driven by genes related to sexual reproduction or stress.

The 314 shared developmental families included several known regulators of fruiting body development and sexual reproduction, such as the white collar complex, mating pheromone GPCRs, components of the velvet complex, HMG box transcription factors, and a phytochrome red-light photoreceptor, among others (supplementary table S3a, Supplementary Material online). We also found three orthogroups containing Pumilio-family proteins, which regulate cellular processes by sequence-specifically binding mRNA molecules and have been implicated in morphogenetic processes that precede sexual reproduction in *Cryptococcus neoformans*, a species derived from primitive CM Basidiomycota (Wang, Tian, et al. 2014). Argonaute-family proteins were developmentally regulated in all nine species, suggesting RNA interference as a widespread regulatory mechanism of CM morphogenesis, consistent with recent reports on individual species (Lau et al. 2018; Shao et al. 2019). Several cell-wall-active gene families were also developmentally regulated in most species (supplementary note 4, Supplementary Material online). For example, fasciclin homologs were developmentally regulated in eight species. This family is involved in the adhesion of pathogenic fungi to host surfaces (Liu et al. 2009) and was shown to be upregulated in primordial of

*Lentinula* (Miyazaki et al. 2007) and in protoperithecia (fruiting body initials) of *Sordaria* (Teichert et al. 2012). Based on their function in adhesion to host surfaces in filamentous fungi and yeasts, we hypothesize that fasciclin-like proteins have a plesiomorphic role in adhesion to various surfaces and that they have been independently co-opted for hypha–hypha adhesion (or cell wall modifications [Teichert et al. 2012]) in CM Agarico- and Pezizomycotina. More broadly, the family has adhesive functions in unicellular protists, multicellular algae, and animals too (Huber and Sumper 1994; Paulsrud and Lindblad 2002), indicating co-option events not only in fungi but in other multicellular groups as well.

In non-CM fungi and vegetative mycelia of CM species, the shared developmentally regulated gene families function in processes ranging from plant–fungal interactions (Maillet et al. 2011), host protein degradation (Rodier et al. 1999) or appressorium development in pathogenic fungi (Geoghegan et al. 2017), the recognition and adhesion to carbohydrate-covered surfaces (Varrot et al. 2013), or simple non-CM developmental processes. These functions indicate that genes with diverse roles have been likely co-opted for CM in the Agarico- and Pezizomycotina. Although the minority of these families is plesiomorphically linked to sexual development, developmentally regulated families appear to reflect a much broader set of ancestral functions. This breadth of ancestral functions can be explained by the conserved nature of sexual reproduction (such as meiosis, cell cycle regulation): cellular processes that did not change in complexity in CM fungi, as opposed to morphogenetic events, which became an order of magnitude more complex in CM Agarico- and Pezizomycotina based on estimates of cell type diversities (>15 in Pezizomycotina perithecia [Bistis et al. 2003; Lord and Read 2011] and >30 in Agaricomycete fruiting bodies [Kües and Navarro-González 2015] versus up to 3 in vegetative mycelia).

### Widespread Parallel Co-Option of Developmental Families

We analyzed the origin of CM by reconstructing the evolution of developmental gene families along the phylogeny. Of the 1,026 conserved developmental families, 560 predate the origin of CM, consistent with their co-option for multicellularity-related functions, whereas 297 and 169 families are taxonomically restricted to the Agarico- and Pezizomycotina, respectively (fig. 1c). Similar figures were obtained for the 90-species data set too. Of the 560 ancient families, 314 (56%) are developmentally regulated in both the Agarico- and Pezizomycotina, indicating parallel co-option for fruiting body development. This frequency is significantly more common than expected by chance ( $P < 10^{-4}$ , permutation-test). The remaining 246 families can be divided into those that have homologs in only one CM clade (because they were lost in the other, 24%, 137 families) and those that have homologs in both CM clades but are developmentally regulated only in one (19%, 109 families), consistent with clade-specific co-option. The frequency of clade-specific co-option is low, with 42 and 67 families in the Agarico- and Pezizomycotina (7.5% and 12%), respectively. The observation of limited clade-specific but widespread parallel co-option

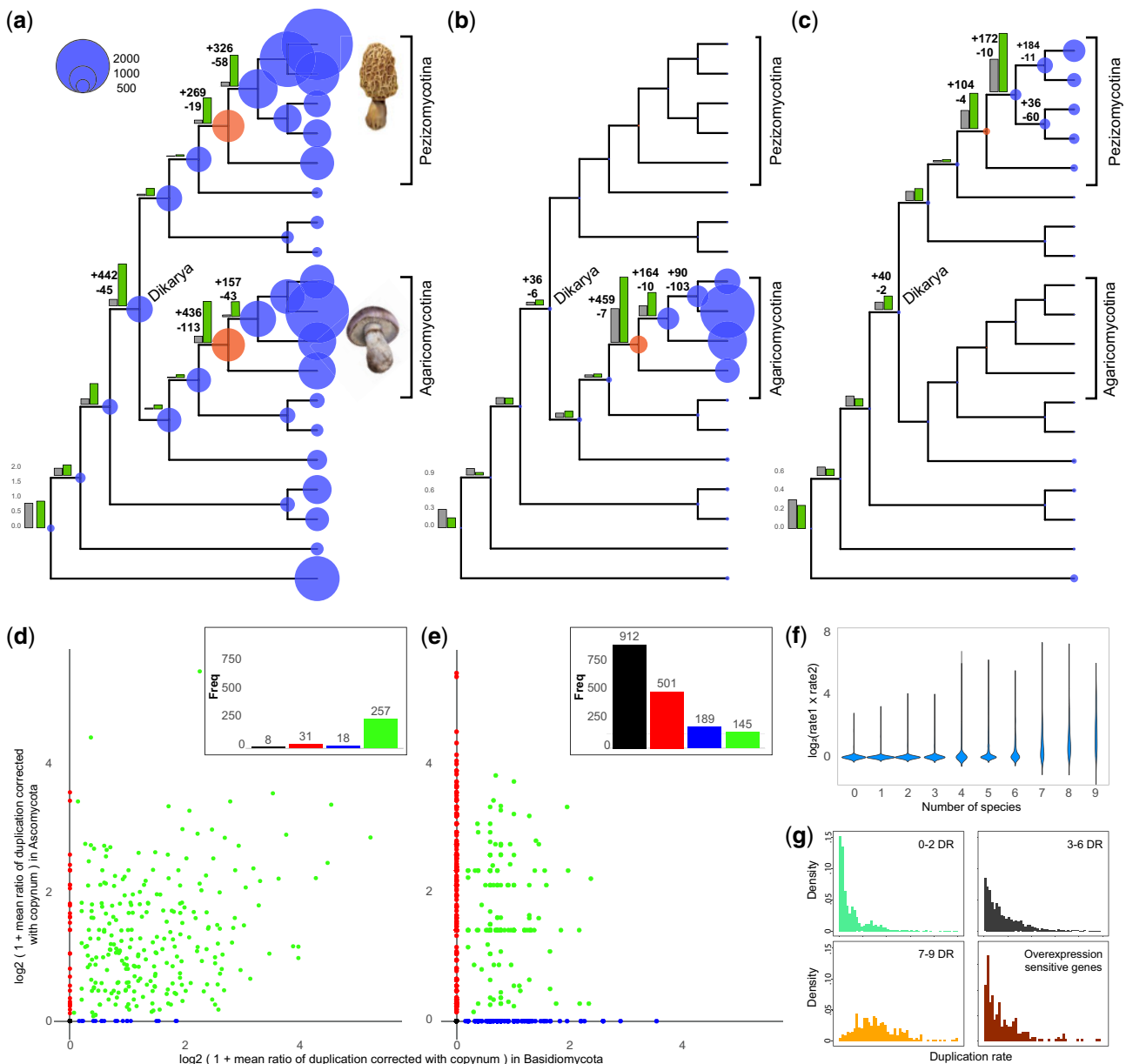
suggests that gene families with suitable properties for CM rarely escaped integration into the genetic toolkit of CM. It also agrees with genes suitable for a given phenotype being rare and thus they mostly end up being recruited under similar selection regimes (Christin et al. 2010).

The observed distribution of developmentally regulated gene families is consistent with two hypotheses. Families with clade-specific taxonomic distribution or clade-specific developmental expression conform to expectations under the simplest model of convergent evolution at the phenotypic level: independent gains of CM in the Agarico- and Pezizomycotina. Similarly, shared developmentally expressed families could have been independently co-opted in the two CM clades. However, this set of families could also encode plesiomorphic functions that were present in the Dikarya ancestor and were independently integrated into CM in the Agarico- and Pezizomycotina. Sexual reproduction is obviously one such function, but our data suggest that genes related to sexual reproduction make up only a minority of the 314 shared developmentally regulated families. Some plesiomorphic functions (e.g., adhesion via fasciclins) may have been present in the ancestor of the Agarico- and Pezizomycotina and proven useful for CM, explaining their parallel co-option. Yet, other functions might have served as precursors to CM (e.g., as morphogenetic modules linked with asexual development [Wang et al. 2009]). Such “precursor traits” could have predisposed lineages for evolving CM by providing stepping stones for evolution, leading to a higher likelihood of phenotypic convergence, as suggested for many cases of convergent evolution (Nagy et al. 2014; Griesmann et al. 2018; Nagy 2018).

To understand what functions, if any, might have predisposed fungal lineages for evolving CM, we looked at the genetic repertoire of the Dikarya ancestor. Although the ancestor of the Dikarya most likely did not have fruiting bodies (see fig. 1a), we reasoned that its ancestral gene complement could reveal whether evolutionary predisposition is a viable hypothesis to explain CM in fungi. We inferred that the Dikarya ancestor had 989 genes in the 314 shared developmental families (fig. 2), which we functionally characterized by examining *Saccharomyces cerevisiae* orthologs. Analyses of Gene Ontology terms revealed an enrichment of genes for the regulation of growth, filamentous growth in response to starvation, transmembrane transport, cell communication, gene expression regulation, and carbohydrate metabolism, reminiscent of general functions required for fungal development (supplementary fig. S4, Supplementary Material online). We found that several gene regulatory circuits, including ones involved in sexual reproduction, mating partner recognition, light, nutrient and starvation sensing, fungal cell wall synthesis and modification, cell-to-cell signaling, and morphogenesis were present in the Dikarya ancestor (and even earlier), suggesting that these might have provided a foundation for the evolution of fruiting bodies.

### Gene Family Expansions Correlate with the Origins of complex multicellularity

To obtain a higher-resolution picture of the evolution of developmental gene families, we reconstructed gene



**FIG. 2.** Convergent expansion of developmentally regulated gene families in independent complex multicellular fungi. (a–c) Reconstructed copy number evolution of 314 shared developmentally regulated gene families (a), 439 families with Agaricomycotina-specific developmental expression and (c) 273 families with Pezizomycotina-specific developmental expression. Bubble size proportional to the number of reconstructed ancestral gene copies across the analyzed families. Numbers next to internal nodes denote the number of inferred duplications and losses. Bar graphs show genome-wide duplication rates (gray) versus duplication rates of the depicted developmental families (green). Inferred gains of CM are indicated by red bubbles. (d, e) Scatterplot of Agarico- and Pezizomycotina duplication rates across 314 shared developmentally regulated gene families (d) and 1,747 families containing  $\leq 2$  developmentally regulated species (e). Black, red, blue, and green denote families with no duplications, Pezizomycotina-specific, Agaricomycotina-specific, and parallel duplications, respectively. Bar diagrams show the number of gene families in each category. (f) Correlation between the extent of convergence and the number of species contributing developmentally regulated genes to a family. (g) The distribution of gene duplication rates across gene families containing developmentally regulated genes from  $\leq 2$ , 3–6, and  $\geq 7$  species (see in fig. 1b) and families in which dosage effects constrain duplications rates (Sopko et al. 2006).

duplications and losses in the Agarico- and Pezizomycotina and in other parts of the fungal tree. In both the 19- and the 90-species data set, we found characteristic expansions of developmentally regulated gene families in CM Agarico- and Pezizomycotina, but no or significantly smaller expansions in other families (fig. 2a and supplementary fig. S5, Supplementary Material online; see supplementary fig. S6

and note 2, Supplementary Material online, for results of the 90-species data set). Across the 314 shared families, we inferred a net expansion (duplication minus loss) of 323 and 250 in the MRCA of the Agarico- and that of the Pezizomycotina, respectively, indicating that the origin of these CM groups coincided with significant expansions in developmentally regulated gene families. In the MRCA of

Dikarya, we inferred 442 duplications and 45 losses. In contrast, we inferred much smaller expansions, or net contractions in other, non-CM branches of the tree (fig. 2a). The observed gene family expansions were driven by increased gene duplication rates, with loss rates remaining approximately constant (supplementary fig. S7, Supplementary Material online). A 6.3- to 8.1-fold higher rate of expansion was found in the 314 shared developmental families compared with other families that were also shared by  $\geq 7$  of the 9 species (fig. 2a–c), indicating that CM-related gene families are one of the most expanding groups in the fungal genomes examined here.

Gene families with a developmental expression specific to the Agarico- (439 families) or Pezizomycotina (273 families) show higher duplication rates (1.93–1.95-fold) and substantially expanded in their respective clades, but not in the other CM subphylum or in non-CM species (fig. 2b and c). Of the Agarico- and Pezizomycotina, the latter showed a more gradual expansion of gene families: We reconstructed 104 and 162 net gains (duplications minus losses) in the MRCA of Pezizomycotina and that of Sordariomycetes, respectively. Interestingly, we inferred relatively few (73) duplications along the branch leading to *Pyronema*, a representative of apothecium-forming Pezizomycotina. This species has 287 genes in the 273 Pezizomycotina-specific families, whereas other species have 401–849 genes. Given that *Pyronema*'s fruiting bodies probably reflect the ancestral morphology (apothecium) in the Pezizomycotina (Liu and Hall 2004; Spatafora et al. 2006; Schoch et al. 2009; Zhang and Wang 2015), these figures could indicate that the developmental gene repertoire of *Pyronema* reflects the ancestral condition in the Pezizomycotina.

In contrast to CM clades, we reconstructed considerable gene loss in unicellular yeast-like fungi in the clades Saccharomycotina, Taphrinomycotina, Pucciniomycotina, and Ustilaginomycotina. Species in these clades spend most of their life cycle in a unicellular yeast stage and, although capable of performing multicellularity-related functions (e.g., cell-to-cell communication, biofilm formation) they have severely reduced development compared with filamentous or CM fungi. We inferred net contractions (19–193 gene duplications and 1,977–2,406 losses, supplementary fig. S8, Supplementary Material online) in these clades, consistent with previous reconstructions (Nagy et al. 2014). An enigmatic CM species, *Neolecta irregularis* belongs to the Taphrinomycotina, where most of its relatives have a simple yeast-like morphology (Nagy 2017; Nguyen et al. 2017). Although the Taphrinomycotina generally shows reductions across the 314 shared developmental families (497 losses and 0 duplications), we observed a slight expansion in *Neolecta* (201 duplications and 178 losses), as opposed to its relatives (e.g., *Taphrina deformans* 38 duplications and 133 losses), which showed further reductions in gene number compared with their common ancestor (supplementary fig. S5b, Supplementary Material online). The slight expansions of shared developmental gene families in *Neolecta* provide further evidence for a potential causative link between the expansion of shared developmental gene families and the independent evolution of CM.

## Convergent Expansions in Shared Developmental Families

We next asked whether the expansions observed in the two CM clades happened in the same gene families (i.e., convergent) or in differing ones (i.e., divergent). We calculated subphylum-specific gene duplication rates in the Agarico- and Pezizomycotina, which we plotted against each other, as shown in figure 2d. Of the 314 shared developmental families, 257 (82%) showed parallel expansions in the Agarico- and Pezizomycotina. In contrast, only 8 (2.5%) showed no duplications in either class and 49 (16%) showed duplications in only one. The 90-species data set showed similar patterns: 92% of shared developmentally regulated gene families showed parallel expansion, 1% showed no duplications, and 7% showed clade-specific duplications (supplementary fig. S9, Supplementary Material online). If we considered families that likewise contained at least seven species, but developmentally regulated genes from up to two species (1,747 families), the pattern was the opposite: only 145 (8.3%) showed parallel duplications in the Agarico- and Pezizomycotina, 1602 (92%) showed no duplications, or in only one of the subphyla (fig. 2e).

Convergence in the 314 developmental families is significantly more abundant than in any other combination of gene families we tested ( $P = 1.2 \times 10^{-9}$  to  $2.7 \times 10^{-84}$ , Fisher's exact test), including controls for gene family size and the number of developmentally regulated proteins per family, among others (supplementary fig. S10 and table S6a and b, Supplementary Material online). Convergence remains significantly more frequent in the 314 families even if we restrict the comparison to one node (the most possible emergence of CMC, fig. 1a) in each of the Agarico- and Pezizomycotina ( $P < 2.95 \times 10^{-39}$ , Fisher's exact test). However, we prefer to consider all (nonterminal) CM nodes in these comparisons, because the evolution and elaboration of CM in these clades might have been a gradual process that spanned several adjacent speciation events (e.g., see *Pyronema*, above). In support of the convergent expansion of developmental families, an independent set of gene families that we assembled based on *Aspergillus* genes with verified developmental functions (by knockout studies) (Cerqueira et al. 2014) also show significantly more convergent expansion than nondevelopmental families ( $P = 5.3 \times 10^{-1}$  to  $6.7 \times 10^{-7}$ , Fisher's exact test; supplementary table S6c, Supplementary Material online). The accelerated duplication rate in the 314 families differs considerably from that of control and nondevelopmental families (supplementary table S6a, Supplementary Material online), or from that of families with constraint on duplication imposed by a fitness cost of increased dosage (Sopko et al. 2006; supplementary note 5, Supplementary Material online, and fig. 2g), collectively suggesting that the observed convergent gene family expansions were driven by positive selection.

Convergent expansion in the 314 shared developmental families correlates with the number of species represented by developmentally regulated genes: Families that contain developmentally regulated genes from seven to nine species show

a much higher incidence of convergent expansion in CM clades (fig. 2f). To check that this is not caused by an intrinsically higher duplication rate of the 314 families, we performed two control analyses. First, we compared family-wise duplication rates in the 314 gene families with that of a control set of genes families that closely resembled the 314 in terms of size (which primarily determines the number of duplications), but contained developmentally regulated genes from  $\leq 6$  species (supplementary fig. S11b, Supplementary Material online). Duplication rates of the 314 and the control families matched almost perfectly, indicating that the higher frequency of convergent duplication (80% vs. 38% in the control set) in the developmental gene family set is not caused by the higher overall duplicability of the former. Second, we compared the frequency of convergent duplications in pairs of nodes in CM clades (excluding terminals) with that observed in other pairs of nodes. This showed that across the 314 families, nodes in which CM likely evolved show a much higher incidence of convergent expansion than other pairs of nodes. The three node pairs in which the highest number of families was convergently expanding contain the MRCA of CM Agaricomycotina (node 28 in supplementary fig. S11a, Supplementary Material online) and one of three nodes from CM Pezizomycotina (nodes 34, 35, or 37 in supplementary fig. S11a, Supplementary Material online), indicating that the 314 developmental families show convergent expansion only in parts of the tree where CM was inferred to have evolved.

Families with clade-specific developmental expression, on the other hand, did not show signs of convergent expansion (convergent duplications were significantly underrepresented,  $P = 3.5 \times 10^{-9}$  in Basidiomycota and  $P = 1.4 \times 10^{-6}$  in Ascomycota, Fisher's exact test; supplementary fig. S12, Supplementary Material online). We note that further convergently expanding families can certainly be found among those containing  $< 7$  species, which renders our estimate of convergence conservative.

Collectively, these analyses indicate that convergent expansion in the 314 developmental families in CM clades is significantly more frequent than in other gene families and that CM Agarico- and Pezizomycotina experienced significantly elevated numbers of convergent gene duplication events in gene families that can be linked to fruiting body development. From these observations, we infer that the observed expansions in developmental families were probably linked to the emergence and elaboration of CM in fungi.

### Limited Evidence for Convergence at the Protein Sequence Level

We also examined the extent of convergence in amino acid sites among CM Agarico- and Pezizomycotina, using approaches that incorporate null models (Rey et al. 2018) proposed in response to criticisms of previous measures of amino acid convergence (Zou and Zhang 2015a, 2015b; Storz 2016). To detect amino acid convergence in multigene families, we applied the Profile Change with One Change (PCOC) model, for detecting amino acid convergence in multigene families, which scans for similar (not necessarily identical)

substitutions in clades with convergently evolved phenotypes (Rey et al. 2018). We required each of the clades in the gene tree made up of proteins of CM species to converge on biochemically similar amino acids—a comparatively strict requirement, that should be expected to yield a conservative estimate of amino acid convergence. We found 129 families in which convergent shifts in amino acid preference are significantly enriched relative to control analyses (supplementary figs. S13–S15 and note 6, Supplementary Material online). Developmentally regulated genes are enriched in 28 of these families (supplementary note 6, Supplementary Material online), including genes related to cell division and DNA repair, splicing and ergosterol biosynthesis, among others (see supplementary table S7, Supplementary Material online). However, the extent of convergence in CM clades was overall similar to that observed in other combinations of clades (supplementary note 6, Supplementary Material online), which could indicate that the extent of amino acid convergence is either not outstanding in CM fungi or that other, unknown traits drove convergence also in non-CM clades. This observation holds under a variety of thresholds for detecting convergent amino acid shifts (0.8, 0.9, 0.95, results only shown for 0.8).

### Discussion

Our genome-wide analyses revealed extensive parallel co-option of ancient genes and convergent gene family expansions in two complex multicellular clades of fungi. We observed molecular convergence in hundreds of gene families, with  $\sim 82\%$  of shared developmentally regulated families showing convergent expansions and others showing convergent amino acid substitutions. Several recent studies suggested that molecular convergence may be widespread in nature (Rokas and Carroll 2008; Shen et al. 2012; Zou and Zhang 2015a). However, although most previous examples were restricted to a few genes (Hughes and Friedman 2003; Shen et al. 2012; Emms et al. 2016; Wirthlin et al. 2018) or to closely related species (Colosimo et al. 2005; Elmer et al. 2014), our results suggest that molecular convergence can be pervasive in clades separated by  $> 650$  My of evolution and can affect hundreds of gene families.

The observed extent of convergence in gene family co-option and expansion exceeds expectations based on previous predictions (Conte et al. 2012) or observations (Hughes and Friedman 2003; Rokas and Carroll 2008; Castoe et al. 2009; Shen et al. 2012; Zhen et al. 2012; Parker et al. 2013) at this phylogenetic scale. In closely related populations of the same species or in sister species, evolution works with nearly the same standing genetic variation, providing the opportunity for a higher incidence of (potentially even nonadaptive) genetic parallelism (Colosimo et al. 2005; Stern and Orgogozo 2009). Because the probability of convergence declines with phylogenetic distance, much less convergence is expected among distantly related clades that have diverged in the architecture of gene regulatory networks—even if the genes themselves are conserved. Molecular clock estimates suggest that the Agarico- and Pezizomycotina diverged  $> 650$  Ma and

their ancestors existed  $>270$  My after the Dikarya ancestor (Floudas et al. 2012; Kohler et al. 2015). Based on the regression model of Conte et al (2012), this deep divergence among CM fungi implies a very low probability of repeated co-option due to neutral processes or phylogenetic proximity. Thus, we conclude that the extent of parallel co-option and convergent diversification of developmental families in CM Agarico- and Pezizomycotina is not explainable by phylogenetic proximity or neutral processes alone and was probably driven by selection.

We here focused on the evolution of CM fungi in two clades, the Agarico- and Pezizomycotina, out of at least eight fungal clades in which CM has evolved repeatedly (Nagy et al. 2018). The large number of independent transitions to CM in fungi suggests that its evolution may be prestaged by some widely conserved genetic circuitries, resembling cases of evolutionary bias that have been documented elsewhere (Wake et al. 2011; Elmer et al. 2014). Such genetic circuits perhaps include adhesion and cell-differentiation-related genes that are used by non-CM fungi for pathogenicity or asexual reproduction, although these hypotheses remain challenging to test experimentally.

In the context of Gould's (1990) famous thought experiment, the repeated emergence of CM in fungi argues that if we replayed life's tape, CM would again evolve in fungal clades. Evidence of such predictability of evolution has been supplied in a growing number of cases and has been attributed to shared genetic variation (Muschick et al. 2012), similar selective regimes (Losos 2011; Mahler et al. 2013; Blount et al. 2018), or constraints on the array of acceptable changes (Losos 2011) and on how novelty arises (Losos 2011; McGhee 2011). A special case of bias in the emergence of novelty is when proteins or gene regulatory networks with suitable biochemical properties are available in ancestral species and can easily be co-opted for the same new functionalities (e.g., due to the modularity in genetic circuits). Our results are compatible with the scenario that ancient fungi have been predisposed for evolving CM by a rich repertoire of genes in the Dikarya ancestor that are used by extant species for CM-related functions. We speculate that these included simple mechanisms for cell-to-cell communication, adhesion, cell differentiation, or other processes that existed in the (non-CM) Dikarya ancestor. Compared with non-CM fungi, these might have become more sophisticated in CM lineages through gene duplications and the evolution of developmentally dynamic expression patterns. Predisposed lineages are more likely to show phenotypic convergence (Losos 2011; Nagy et al. 2014; Agrawal 2017; Nagy 2018), simply because of the availability of genetic tools that can be repeatedly co-opted for the same functions. It follows that if predisposition indeed happened, then the repeated evolution of CM is not as surprising as it may seem under the more widely accepted model that necessitates hundreds of independent events to happen convergently.

Haldane speculated that similar phenotypes emerge not only as a result of similar selection pressures but also as a result of shared genetic biases (Haldane 1932). There are a finite number of genes that can potentially fulfill CM-

associated functions, which explains why the same gene families were co-opted and started diversifying in complex multicellular Agarico- and Pezizomycotina. Our study provides an example of how the genomic repertoire may channel phenotypic evolution toward similar solutions and how this can lead to extensive genetic convergence even at large phylogenetic scales. Such genetic biases on phenotypic evolution suggest that "endless forms" created by the tireless tinkering of evolution are limited not only by the environment but also by the genetic ingredients at hand.

## Materials and Methods

### Bioinformatic Analysis of RNA-Seq Data

We downloaded publicly available transcriptome data related to different developmental stages of nine fruiting body forming species from the Agaricomycotina (*Armillaria ostoyae*, *Coprinopsis cinerea*, *Hypsizygus marmoreus*, and *Flammulina velutipes*) and the Pezizomycotina (*Fusarium graminearum*, *F. verticillioides*, *Sordaria macrospora*, *Neurospora crassa*, and *Pyronema confluens*, supplementary table S1a, Supplementary Material online). We ran a quality check on raw fastq files using fastqc v0.11.5 (Andrews 2010) and trimmed the adapter sequences and low-quality bases with trimmomatic v0.36 (Bolger et al. 2014). Next, we used kallisto v0.43.1 (Bray et al. 2016) to quantify the abundance of transcripts for each stage. Specifically, we utilized the estimated counts from abundance data to calculate Fragments Per Kilobase Million (FPKM) and used it as the quantification metric. As a prefilter, an FPKM value  $<2$  was considered insignificant. The homogeneity of biological replicates was checked by constructing multidimensional scaling (MDS) plots based on overall expression levels. One 24-h sample of *F. verticillioides* was excluded from the identification of developmentally regulated genes, because it appeared as an outlier on MDS plots.

### Identification of Developmentally Regulated Genes

The RNA-Seq data comprised 2–13 stages, which was used to identify developmentally regulated genes: those that show at least 4-fold change in expression between any two fruiting body stages or tissue types and that show an expression level FPKM  $>4$ . Fold change values were calculated for all biologically relevant pairwise comparisons (supplementary fig S2, Supplementary Material online).

### Identification of Meiosis, Stress-Related Genes, and CAZymes

Best DELTA-BLAST (Boratyn et al. 2012) hits were identified as orthologs using meiosis-related genes (Toome et al. 2014; Knapp et al. 2018) and stress-related genes (Gómez-Gil et al. 2019) as the query against the nine proteomes. CAZy classification was downloaded from the JGI-MycoCosm database (Grigoriev et al. 2014).

### Comparative Genomic Approaches

In addition to the nine above-mentioned fruiting body forming species, ten additional species were included in the



analysis for comparative purposes (supplementary table S1a, Supplementary Material online). This 19 genome data set was clustered into gene families using OrthoFinder v1.1.8 (Emms and Kelly 2015) with the default inflation parameter of 1.5 to facilitate interspecies comparison.

For functional annotation of genes and gene families, InterProScan search was performed with InterProScan version 5.28-67.0 (Jones et al. 2014) across the 19 fungal proteomes (supplementary table S3, Supplementary Material online).

Gene Ontology enrichment analysis for yeast orthologs was performed using GOzilla (Eden et al. 2009; <http://cbl-gorilla.cs.technion.ac.il/>) with *Saccharomyces cerevisiae* as the reference organism, a  $10^{-3}$  *P* value threshold and false discovery rate correction for multiple testing. Terms in all three ontologies (Biological process, Cellular component, and Molecular function) were considered. Experimentally verified gene function from *Aspergillus nidulans* (Dyer and O’Gorman 2012) and *Neurospora crassa* (Colot et al. 2006; Trail et al. 2017) were also considered during the functional annotation of developmentally regulated gene families. We also used known developmentally regulated gene set of *Coprinopsis* from literature to verify the efficiency of our designation.

To get a more reliable estimation of ancestral state reconstruction and evolution of gene families, we used a broader data set containing 90 species (supplementary table S1b, Supplementary Material online). From this data set, gene families were reconstructed using all versus all BlastP search,  $1e^{-5}$  *e*-value and a 30% bidirectional coverage filter in BlastP output and Markov Cluster algorithm (Dongen 2000) with the inflation parameter 1.5 to identify clusters.

### Phylogenetic Analyses and Ancestral State Reconstructions

Orthofinder clustering identified 86 single-copy gene families (without any paralogs) shared by all 19 species. These clusters were used to reconstruct a species tree. After multiple sequence alignment using the L-INS-I algorithm of MAFFT (Katoh and Standley 2013) and trimming with trimAL (−gt 0.6) (Capella-Gutierrez et al. 2009), sequences were concatenated into a supermatrix and used for phylogenetic reconstruction in raxmlHPC-PTHREADS-SSE3 (Stamatakis 2014). The supermatrix was partitioned by gene and the PROTGAMMAWAG model was used with 100 rapid bootstrap replicates, to estimate branch support.

For the broader data set, species tree reconstruction was based on 301 single-copy gene families. To find single-copy gene families, those containing 45–180 proteins were aligned using L-INS-I algorithm of MAFFT and trimmed with trimAL (−gt 0.4). Gene families either with no or only terminal duplications were allowed; of terminally duplicated genes (inparalogs) only those proteins were retained, which showed the lowest amino acid distance to other members of the family. In the chosen gene families, at least 50% of species had to be represented, and the trimmed alignment had to be >60 amino acids long, resulting in a final set of 301 gene families. These families were aligned with PRANK (Löytynoja 2013), trimmed with trimAL (−strict), and concatenated into a supermatrix. Phylogenetic

reconstruction was performed in RaxmlHPC-PTHREADS-AVX2 version 8.2.10. The supermatrix was partitioned by gene and the PROTGAMMAWAG model was used with 200 rapid bootstrap replicates in the CIPRES Science Gateway (Miller et al. 2010).

To reconstruct the evolution of fruiting body formation, maximum parsimony (MP), ML ancestral state reconstruction, and stochastic character mapping (Huelsenbeck et al. 2003) were performed. For MP, we used the *asr\_max\_parsimony* function of castor (Louca and Doebeli 2018), for ML we used the *ace* (ancestral character estimations) function of the ape (Paradis et al. 2004), and for stochastic mapping we used the *make.simmmap* function of phytools (Revell 2012) R packages (R Core Team 2019), respectively. The extant character states of species are listed in supplementary table S1b, Supplementary Material online. The ARD (all rates different) model was used in both ML reconstruction and stochastic mapping. For stochastic character mapping, the estimated ancestral state and a fixed transition matrix from the best model for the data were used and 1,000 stochastic character maps were summarized. In the case of MP, multiple transition costs (all\_equal, sequential, proportional, and exponential) were tested. We combined each transition cost option with both possible transitions, that is, from non-CM to CM and from CM to non-CM.

To infer the ancestral cellularity level of the Dikarya ancestor (MRCA of the Agaricomycotina and Pezizomycotina), we executed ML ancestral state reconstruction using BayesTraits (Pagel and Meade 2007). Analyses were run using 200 trees from the bootstrap analysis of ML species tree reconstruction. We compared the likelihoods of the unconstrained and two constrained models, in which the ancestral state of the Dikarya MRCA was fixed either as complex or simple multicellular. We compared these models based on log-likelihood values, with a difference of two log-likelihood units taken as evidence for the better-fitting model (Pagel 1999).

### Evolutionary History of Gene Families

To reconstruct the duplication and loss events of gene families across the species tree, the COMPARE pipeline (Nagy et al. 2014) was used. For this analysis, gene trees were reconstructed for gene families containing at least four proteins (9,686 gene families). Sequences in each gene family were aligned using the L-INS-I method of MAFFT and trimmed with trimAL (−gt 0.2). Gene tree reconstructions were performed in RAXML under the PROTGAMMAWAG model with 100 rapid bootstrap replicates. Gene trees were rerooted and reconciled with the species tree using Notung 2.9 (Darby et al. 2016) with 80% bootstrap support as the edge-weight threshold for topological rearrangements. After ortholog coding, duplications and losses for each orthogroup were mapped onto the species tree using Dollo parsimony (Nagy et al. 2014). The visualization of reconstructed duplication/loss histories and further statistical analyses (Fisher’s exact test) were performed with custom R scripts (available from the authors upon request).

To quantify convergent gene family expansions, we filtered families with the following criteria: 1) A gene family has genes

conserved in  $\geq 7$  of the 9 Dikarya species,  $\geq 4$  of the 5 Pezizomycotina species, or  $\geq 3$  of the 4 Agaricomycotina species or 2) a gene family has developmental expression conserved in  $\geq 7$  of the 9 Dikarya species,  $\geq 3$  of the 4 Agaricomycotina species, or  $\geq 4$  of the 5 Pezizomycotina species.

Gene families of the 90-species data set were processed with the same analysis pipeline as the smaller set, except that the gene tree support was approximated by the SH-like support in RAXML instead of the more computationally intensive bootstrap analysis.

### Convergence in Gene Family Expansions

To quantify the level of convergent gene family expansions, subphylum-specific gene duplication rates were compared between the Agarico- and Pezizomycotina. Rates were calculated by normalizing the raw number of inferred duplications for a given node by both the length of the preceding branch and by gene family size, because along longer branches and in larger gene families the probability of duplications is naturally higher. After this correction step, duplication rates were averaged across nodes (gene duplication rate for Agarico- and Pezizomycotina) and plotted using custom R scripts. The numbers of four possible events were recorded: duplications in only one (Agarico- or Pezizomycotina), both or none of the CM clades.

To assess whether developmental gene families show more or less convergence than expected by chance, different control groups of gene families were generated and compared using Fisher's exact test. Control families were always compared with the conserved developmentally regulated gene family set (developmentally expression conserved in  $\geq 7$  of the 9 Dikarya species). The first control groups comprised families that similarly contained  $\geq 7$  species but only 0–2 species (1,747 gene families) or 3–6 species (2,052 gene families) with developmental expression. Next, to test if gene family size (i.e., number of proteins) impacts convergence, we also generated control groups with similar gene family size distribution but containing less developmentally regulated genes than the 314 shared developmental gene families. A custom R script was used to find a nondevelopmentally regulated gene family for each of the developmentally regulated gene family with a matching size one by one. If it was not possible to find a gene family with similar size (permitted maximum difference of 10%), the gene family was excluded from the comparison. If there were more than one gene families with the same size, the one with most similar species composition and least developmentally regulated genes (in terms of the number of species represented by developmentally regulated genes) was chosen.

We tested whether gene families with known developmental phenotype show more parallel gene family expansion than expected by chance. For this, we collected genes of *Aspergillus nidulans* and *A. fumigatus* using the “phenotypes” tools of the *Aspergillus* genome portal (<http://www.aspgd.org/>). We chose genes related to conidiation, sporulation, and formation of anatomical structures (e.g., cleistothecium, conidiophore, metula, and Hulle cell) while omitted all genes

related to hyphal structures. Gene families were identified with BLAST search ( $e$ -value  $< 1e^{-6}$ , alignment coverage  $> 0.5$ ) keeping only the best hit per CM genome for each annotated *Aspergillus* query sequence. When a gene family contained at least 60% of the best hits for a given query sequence, its annotation was transferred to the family. Fisher's exact test was performed for the comparison of parallel duplication across Ascomycota and Basidiomycota between the developmental (205) and all (4,113) gene families represented by at least seven out of the nine species.

### Detecting Convergent Amino Acid Changes

In order to gain insights into amino acid convergence between the Agaricomycotina and Pezizomycotina, we followed Rey et al.'s approach (Rey et al. 2018) to identify convergent shifts in amino acid preference at a given site. Convergence is defined not only as changes to identical amino acids from different ancestral states but also as changes to amino acids with similar biochemical properties (referred to as convergent shifts in amino acid composition). We identified such shifts across all gene families in the 19 species' genomes using the model PCOC (Rey et al. 2018; downloaded from <https://github.com/CarineRey/pcoc> on November 5, 2018). We used reconciled gene trees with branch lengths reestimated with RAXML (raxmlHPC-PTHREADS-SSE3) as input. Each of the most inclusive clades that contained only CM species was designated as phenotypically convergent clade. Automated designation of convergent clades was done using a custom R script, followed by execution of PCOC with default settings. For considering a site as convergent, we chose the PCOC model with a posterior probability threshold of 0.8. We performed the analysis on 3,799 gene families that contained at least 10 proteins and at least 1 protein from both the Agaricomycotina and the Pezizomycotina. Three sets of control analyses were run to assess the amount of convergence caused by chance events. In control 1, the basal lineages of Ascomycota and Basidiomycota were designated as convergent clades (Ustilagomycotina, Pucciniomycotina, Saccharomycotina, and Taphrinomycotina). In control 2, CM Agaricomycotina species were paired with the basal clades of Ascomycota (Saccharomycotina and Taphrinomycotina), whereas in control 3 CM Pezizomycotina were paired with the basal clades of Basidiomycota (Ustilagomycotina and Pucciniomycotina). This resulted in three control analyses in which CM is not shared by clades designated as convergent (note however, that other traits might be). The numbers of detected amino acid sites showing convergent shifts in each gene family were recorded and correlations between CM and control groups were evaluated with a Pearson correlation test. We also compared these values after correction by branch lengths between or in the designated clades to avoid the effect of divergence (i.e., branch length) on the amount of amino acid changes.

We assumed that gene families which contain more convergent amino acid sites in CM lineages than in non-CM clades might be involved in the shaping of convergent phenotypes. For identification of these gene families, a linear

model was fit to predict the number of convergent sites between CM clades from the corresponding values of non-CM clades (control 1). Gene families with more convergent sites than the upper limit of 95% prediction interval of the linear model were considered as displaying significant number of convergent sites in CM clades.

## Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

## Acknowledgments

We acknowledge inspiring discussions of this topic in the Fungal Genomics and Evolution Laboratory (Szeged, Hungary). We are thankful to Marin Talbot Brewer, Gunther Doehlemann, and Jon Magnuson for permission to utilize unpublished genomic data of *Exobasidium maculosum*, *Exobasidium vaccinii*, and *Melanotaenium endogenum*, respectively. This work was supported by the “Momentum” program of the Hungarian Academy of Sciences (contract no. LP2014/12 to L.G.N.) and the European Research Council (grant no. 758161 to L.G.N.); GINOP-2.3.2-15-2016-00052 (to L.G.N.), “Frontline” Research Excellence Programme of the National Research, Development and Innovation Office, Hungary (KKP-129814 to B.P.), GINOP-2.3.2-15-2016-00026 (iChamber to B.P.), The European Union’s Horizon 2020 research and innovation program under Grant Agreement No. 739593 (to B.P.), and National Science Foundation IOS 1457044 and IOS 1916137 (to J.P.T.).

## Author Contributions

Z.M. and L.G.N. conceived the study. Z.M., Z.W., A.N.P., J.P.T., B.H., and B.B. analyzed data. A.N.P. analyzed transcriptomic data. Z.W., J.P.T., and Z.M. evaluated developmentally regulated genes. Z.M., B.H., and B.B. reconstructed gene family evolution. Z.M., K.K., and B.P. evaluated adaptivity of gene family expansions. Z.M., B.P., and L.G.N. wrote the paper. All authors have read and commented on the manuscript.

## References

Agrawal AA. 2017. Toward a predictive framework for convergent evolution: integrating natural history, genetic mechanisms, and consequences for the diversity of life. *Am Nat.* 190(S1):S1–S12.

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data [Internet]. Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.

Bistis GN, Perkins DD, Read ND. 2003. Different cell types in *Neurospora crassa*. *Fungal Genet Rep.* 50(1):17–19.

Blount ZD, Lenski RE, Losos JB. 2018. Contingency and determinism in evolution: replaying life’s tape. *Science* 362(6415):eaam5979.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.

Boratyn GM, Schäffer AA, Agarwala R, Altschul SF, Lipman DJ, Madden TL. 2012. Domain enhanced lookup time accelerated BLAST. *Biol Direct* 7:12.

Braendle C, Baer CF, Félix MA. 2010. Bias and evolution of the mutationally accessible phenotypic space in a developmental system. *PLoS Genet.* 6(3):e1000877.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol.* 34(5):525–527.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25(15):1972–1973.

Castoe TA, de Koning APJ, Kim H-M, Gu W, Noonan BP, Naylor G, Jiang ZJ, Parkinson CL, Pollock DD. 2009. Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc Natl Acad Sci U S A.* 106(22):8986–8991.

Cerqueira GC, Arnaud MB, Inglis DO, Skrzypek MS, Binkley G, Simison M, Miyasato SR, Binkley J, Orvis J, Shah P, et al. 2014. The *Aspergillus* Genome Database: multispecies curation and incorporation of RNA-Seq data to improve structural gene annotations. *Nucleic Acids Res.* 42(D1):D705–D710.

Christin P-A, Weinreich DM, Besnard G. 2010. Causes and evolutionary significance of genetic convergence. *Trends Genet.* 26(9):400–405.

Claessen D, Rozen DE, Kuipers OP, Søgaard-Andersen L, van Wezel GP. 2014. Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. *Nat Rev Microbiol.* 12(2):115–124.

Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* 307:1928–1933.

Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. 2006. A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *Proc Natl Acad Sci U S A.* 103(27):10352–10357.

Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic parallelism and convergence in natural populations. *Proc R Soc B* 279(1749):5039–5047.

Darby CA, Stolzer M, Ropp PJ, Barker D, Durand D. 2016. Xenolog classification. *Bioinformatics* 33:btw686.

Darwin C, Mayr E. 1995. On the origin of species. Cambridge (MA): Harvard University Press. Available from: <http://www.hup.harvard.edu/catalog.php?isbn=9780674637528>.

Dongen SV. 2000. Graph clustering by flow simulation (Doctoral dissertation). Available from: <https://dspace.library.uu.nl/handle/1874/848>.

Dyer PS, O’Gorman CM. 2012. Sexual development and cryptic sexuality in fungi: insights from *Aspergillus* species. *FEMS Microbiol Rev.* 36(1):165–192.

Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. 2009. GOrrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics* 10(1):48.

Elmer KR, Fan S, Kusche H, Luise Spreitzer M, Kautt AF, Franchini P, Meyer A. 2014. Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. *Nat Commun.* 5:5168.

Emms DM, Covshoff S, Hibberd JM, Kelly S. 2016. Independent and parallel evolution of new genes by gene duplication in two origins of C4 photosynthesis provides new insight into the mechanism of phloem loading in C4 species. *Mol Biol Evol.* 33(7):1796–1806.

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16(1):157.

Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS, et al. 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. *Science* 336(6089):1715–1719.

Geoghegan I, Steinberg G, Gurr S. 2017. The role of the fungal cell wall in the infection of plants. *Trends Microbiol.* 25(12):957–967.

Gómez-Gil E, Franco A, Madrid M, Vázquez-Marín B, Gacto M, Fernández-Breis J, Vicente-Soler J, Soto T, Cansado J. 2019. Correction: quorum sensing and stress-activated MAPK signaling repress yeast to hypha transition in the fission yeast *Schizosaccharomyces japonicus*. *PLoS Genet.* 15(7):e1008282.

Gould SJ. 1990. Wonderful life: the Burgess Shale and the nature of history. New York: WW Norton & Company.

Griesmann M, Chang Y, Liu X, Song Y, Haberer G, Crook MB, Billault-Penneteau B, Laussergues D, Keller J, Imanishi L. 2018. Phylogenomics reveals multiple losses of nitrogen-fixing root nodule symbiosis. *Science* 361:eaat1743.

- Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otilar R, Riley R, Salamov A, Zhao X, Korzeniewski F, et al. 2014. MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res.* 42(Database issue): D699–D704.
- Grosberg RK, Strathmann RR. 2007. The evolution of multicellularity: a minor major transition? *Annu Rev Ecol Syst.* 38(1):621–654.
- Haldane JBS. 1932. The causes of evolution. London: Harper and Brothers.
- Hibbett DS. 2007. After the gold rush, or before the flood? Evolutionary morphology of mushroom-forming fungi (*Agaricomycetes*) in the early 21st century. *Mycol Res.* 111(9):1001–1018.
- Huber O, Sumper M. 1994. Algal-CAMs: isoforms of a cell adhesion molecule in embryos of the alga *Volvox* with homology to *Drosophila fasciclin* I. *EMBO J.* 13(18):4212–4222.
- Huelsenbeck JP, Nielsen R, Bollback JP. 2003. Stochastic mapping of morphological characters. *Syst Biol.* 52(2):131–158.
- Hughes AL, Friedman R. 2003. Parallel evolution by gene duplication in the genomes of two unicellular fungi. *Genome Res.* 13(5):794–799.
- Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9):1236–1240.
- Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
- Knapp DG, Németh JB, Barry K, Hainaut M, Henrissat B, Johnson J, Kuo A, Lim JHP, Lipzen A, Nolan M, et al. 2018. Comparative genomics provides insights into the lifestyle and reveals functional heterogeneity of dark septate endophytic fungi. *Sci Rep.* 8:6321.
- Knoll AH. 2011. The multiple origins of complex multicellularity. *Annu Rev Earth Planet Sci.* 39(1):217–239.
- Kohler A, Kuo A, Nagy LG, Morin E, Barry KW, Buscot F, Canbäck B, Choi C, Cichocki N, Clum A, et al. 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nat Genet.* 47(4):410–415.
- Krizsán K, Almási É, Merényi Z, Sahu N, Virágh M, Kószó T, Mondo S, Kiss B, Bálint B, Kües U, et al. 2019. Transcriptomic atlas of mushroom development reveals conserved genes behind complex multicellularity in fungi. *Proc Natl Acad Sci U S A.* 116(15):7409–7418.
- Kües U, Navarro-González M. 2015. How do *Agaricomycetes* shape their fruiting bodies? 1. Morphological aspects of development. *Fungal Biol Rev.* 29(2):63–97.
- Lau AYT, Cheng X, Cheng CK, Nong W, Cheung MK, Chan R-F, Hui JHL, Kwan HS. 2018. Discovery of microRNA-like RNAs during early fruiting body development in the model mushroom *Coprinopsis cinerea*. *PLoS One* 13(9):e0198234.
- Liu TB, Chen GQ, Min H, Lin FC. 2009. MoFLP1, encoding a novel fungal fasciclin-like protein, is involved in conidiation and pathogenicity in *Magnaporthe oryzae*. *J Zhejiang Univ Sci B* 10(6):434–444.
- Liu YJ, Hall BD. 2004. Body plan evolution of ascomycetes, as inferred from an RNA polymerase II phylogeny. *Proc Natl Acad Sci U S A.* 101(13):4507–4512.
- Lord KM, Read ND. 2011. Perithecial morphogenesis in *Sordaria macrospora*. *Fungal Genet Biol.* 48(4):388–399.
- Losos JB. 1998. Contingency and determinism in replicated adaptive radiations of island lizards. *Science* 279(5359):2115–2118.
- Losos JB. 2011. Convergence, adaptation, and constraint. *Evolution (N Y)*. 65(7):1827–1840.
- Louca S, Doebeli M. 2018. Efficient comparative phylogenetics on large trees. *Bioinformatics* 34(6):1053–1055.
- Löytynoja A. 2013. Phylogeny-aware alignment with PRANK. In: Clifton NJ, editor. *Methods in molecular biology*. Vol. 1079. Totowa (NJ): Humana Press. p. 155–170.
- Mahler DL, Ingram T, Revell LJ, Losos JB. 2013. Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* 341(6143):292–295.
- Maillet F, Poinsoit V, André O, Puech-Pagés V, Haouy A, Gueunier M, Cromer L, Giraudet D, Formey D, Niebel A, et al. 2011. Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. *Nature* 469(7328):58–64.
- McGhee GR. 2011. Convergent evolution: limited forms most beautiful. Cambridge: MIT Press.
- Mikheyev AS, Mueller UG, Abbot P. 2006. Cryptic sex and many-to-one coevolution in the fungus-growing ant symbiosis. *Proc Natl Acad Sci U S A.* 103(28): 10702–10706.
- Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. Proceedings of the Gateway Computing Environments Workshop (GCE), 14 November 2010; New Orleans, LA. p. 1–8. Available from: [http://www.phylo.org/sub\\_sections/portal/sc2010\\_paper.pdf](http://www.phylo.org/sub_sections/portal/sc2010_paper.pdf).
- Miyazaki Y, Kaneko S, Sunagawa M, Shishido K, Yamazaki T, Nakamura M, Babasaki K. 2007. The fruiting-specific *Le.flp1* gene, encoding a novel fungal fasciclin-like protein, of the basidiomycetous mushroom *Lentinula edodes*. *Curr Genet.* 51(6):367–375.
- Muraguchi H, Umezawa K, Niikura M, Yoshida M, Kozaki T, Ishii K, Sakai K, Shimizu M, Nakahori K, Sakamoto Y, et al. 2015. Strand-specific RNA-Seq analyses of fruiting body development in *Coprinopsis cinerea*. *PLoS One* 10(10):e0141586.
- Muschick M, Indermaur A, Salzburger W. 2012. Convergent evolution within an adaptive radiation of cichlid fishes. *Curr Biol.* 22(24):2362–2368.
- Nagy LG. 2017. Evolution: complex multicellular life with 5,500 genes. *Curr Biol.* 27(12):R609–R612.
- Nagy LG. 2018. Many roads to convergence. *Science* 361:125–126.
- Nagy LG, Kovács GM, Krizsán K. 2018. Complex multicellularity in fungi: evolutionary convergence, single origin, or both? *Biol Rev.* 93(4):1778–1794.
- Nagy LG, Ohm RA, Kovács GM, Floudas D, Riley R, Gácsér A, Sipiczki M, Davis JM, Doty SL, de Hoog GS, et al. 2014. Latent homology and convergent regulatory evolution underlies the repeated emergence of yeasts. *Nat Commun.* 5:4471.
- Nguyen TA, Cissé OH, Yun Wong J, Zheng P, Hewitt D, Nowrousian M, Stajich JE, Jedd G. 2017. Innovation and constraint leading to complex multicellularity in the Ascomycota. *Nat Commun.* 8:14444.
- Pagel M. 1999. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Syst Biol.* 48(3):612–622.
- Pagel M, Meade A. 2007. BayesTraits. Computer program and documentation. (Reading Evolutionary Biology Group, 2007). Available from: <http://www.evolution.rdg.ac.uk/BayesTraits.html>.
- Pankey MS, Minin VN, Imholte GC, Suchard MA, Oakley TH. 2014. Predictable transcriptome evolution in the convergent and complex bioluminescent organs of squid. *Proc Natl Acad Sci U S A.* 111(44):E4736–E4742.
- Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20(2):289–290.
- Park B, Rutter MT, Fenster CB, Symonds VV, Ungerer MC, Townsend JP. 2017. Distributions of mutational effects and the estimation of directional selection in divergent lineages of *Arabidopsis thaliana*. *Genetics* 206(4):2105–2117.
- Park Y-J, Baek JH, Lee S, Kim C, Rhee H, Kim H, Seo J-S, Park H-R, Yoon D-E, Nam J-Y, et al. 2014. Whole genome and global gene expression analyses of the model mushroom *Flammulina velutipes* reveal a high capacity for lignocellulose degradation. *PLoS One* 9(4):e93560.
- Parker J, Tsagkogeorga G, Cotton JA, Liu Y, Provero P, Stupka E, Rossiter SJ. 2013. Genome-wide signatures of convergent evolution in echolocating mammals. *Nature* 502(7470):228–231.
- Paulsrud P, Lindblad P. 2002. Fasciclin domain proteins are present in *Nostoc* symbionts of lichens. *Appl Environ Microbiol.* 68(4):2036–2039.
- Pöggeler S, Nowrousian M, Teichert I, Beier A, Kück U. 2018. Fruiting-body development in ascomycetes. In: *Physiology and Genetics. Cham (Switzerland): Springer International Publishing.* p. 1–56.
- R Core Team. 2019. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available from: <https://www.R-project.org/>.
- Revell LJ. 2012. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol.* 3:217–223.

- Rey C, Guéguen L, Sémon M, Boussau B. 2018. Accurate detection of convergent amino-acid evolution with PCOC. *Mol Biol Evol.* 35(9):2296–2306.
- Rice DP, Townsend JP. 2012. A test for selection employing quantitative trait locus and mutation accumulation data. *Genetics* 190(4):1533–1545.
- Rodier M-H, El Moudni B, Kauffmann-Lacroix C, Daniault G, Jacquemin J-L. 1999. A *Candida albicans* metallopeptidase degrades constitutive proteins of extracellular matrix. *FEMS Microbiol Lett.* 177(2):205–210.
- Rokas A, Carroll SB. 2008. Frequent and widespread parallel evolution of protein sequences. *Mol Biol Evol.* 25(9):1943–1953.
- Ruiz-Trillo I, Burger G, Holland PWH, King N, Lang BF, Roger AJ, Gray MW. 2007. The origins of multicellularity: a multi-taxon genome initiative. *Trends Genet.* 23(3):113–118.
- Rundle HD, Nagel L, Boughman JW, Schluter D. 2000. Natural selection and parallel speciation in sympatric sticklebacks. *Science* 287(5451):306–308.
- Sakamoto Y, Nakade K, Konno N. 2011. Endo- $\beta$ -1,3-glucanase GLU1, from the fruiting body of *Lentinula edodes*, belongs to a new glycoside hydrolase family. *Appl Environ Microbiol.* 77(23):8350–8354.
- Schoch CL, Sung G-H, Pez-Gir Ldez FL, Townsend JP, Miadlikowska J, Hofstetter VR, Robbertse B, Matheny PB, Kauff F, Wang Z, et al. 2009. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. *Syst Biol* 58(2):224–239.
- Sebé-Pedrós A, Degnan BM, Ruiz-Trillo I. 2017. The origin of Metazoa: a unicellular perspective. *Nat Rev Genet.* 18(8):498–512.
- Shao Y, Tang J, Chen S, Wu Y, Wang K, Ma B, Zhou Q, Chen A, Wang Y. 2019. miR4 and miR16 mediated fruiting body development in the medicinal fungus *Cordyceps militaris*. *Front Microbiol.* 10:83.
- Shen Y-Y, Liang L, Li G-S, Murphy RW, Zhang Y-P. 2012. Parallel evolution of auditory genes for echolocation in bats and toothed whales. *PLoS Genet.* 8(6):e1002788.
- Shubin N, Tabin C, Carroll S. 2009. Deep homology and the origins of evolutionary novelty. *Nature* 457(7231):818–823.
- Sikhakolli UR, López-Giráldez F, Li N, Common R, Townsend JP, Trail F. 2012. Transcriptome analyses during fruiting body formation in *Fusarium graminearum* and *Fusarium verticillioides* reflect species life history and ecology. *Fungal Genet Biol.* 49(8):663–673.
- Smith JM, Szathmari E. 1995. The major transitions in evolution. New York: W.H. Freeman and Company.
- Sopko R, Huang D, Preston N, Chua G, Papp B, Kafadar K, Snyder M, Oliver SG, Cyert M, Hughes TR, et al. 2006. Mapping pathways and phenotypes by systematic gene overexpression. *Mol Cell* 21(3):319–330.
- Spatafora JW, Sung G-H, Johnson D, Hesse C, O’rourke B, Serdani M, Spotts R, Lutzoni F, Hofstetter V, Miadlikowska J, et al. 2006. A five-gene phylogeny of Pezizomycotina. *Mycologia* 98(6):1018–1028.
- Stajich JE, Berbee ML, Blackwell M, Hibbett DS, James TY, Spatafora JW, Taylor JW. 2009. The fungi. *Curr Biol.* 19(18):R840–R845.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.
- Stern DL, Orgogozo V. 2009. Is genetic evolution predictable? *Science* 323:746–751.
- Stoltzfus A, Mccandlish DM. 2017. Mutational biases influence parallel adaptation. *Mol Biol Evol.* 34(9):2163–2172.
- Storz JF. 2016. Causes of molecular convergence and parallelism in protein evolution. *Nat Rev Genet.* 17(4):239–250.
- Teichert I, Wolff G, Kück U, Nowrousian M. 2012. Combining laser microdissection and RNA-seq to chart the transcriptional landscape of fungal development. *BMC Genomics.* 13(1):511.
- Toome M, Ohm RA, Riley RW, James TY, Lazarus KL, Henrissat B, Albu S, Boyd A, Chow J, Clum A, et al. 2014. Genome sequencing provides insight into the reproductive biology, nutritional mode and ploidy of the fern pathogen *Mixia osmundae*. *New Phytol.* 202(2):554–564.
- Traeger S, Altegoer F, Freitag M, Gabaldon T, Kempken F, Kumar A, Marcet-Houben M, Pöggeler S, Stajich JE, Nowrousian M. 2013. The genome and development-dependent transcriptomes of *Pyronema confluens*: a window into fungal evolution. *PLoS Genet.* 9(9):e1003820.
- Trail F, Wang Z, Stefanko K, Cubba C, Townsend JP. 2017. The ancestral levels of transcription and the evolution of sexual phenotypes in filamentous fungi. *PLoS Genet.* 13(7):e1006867.
- van Velzen R, Holmer R, Bu F, Rutten L, van Zeijl A, Liu W, Santuari L, Cao Q, Sharma T, Shen D, et al. 2018. Comparative genomics of the nonlegume *Parasponia* reveals insights into evolution of nitrogen-fixing rhizobium symbioses. *Proc Natl Acad Sci U S A.* 115(20):E4700–E4709.
- Varrot A, Basheer SM, Imberty A. 2013. Fungal lectins: structure, function and potential applications. *Curr Opin Struct Biol.* 23(5):678–685.
- Wake DB, Wake MH, Specht CD. 2011. Homoplasy: from detecting pattern to determining process in evolution, but with a secondary role for morphology? *Zootaxa* 331(6020):1032–1068.
- Wang L, Tian X, Gyawali R, Upadhyay S, Foyle D, Wang G, Cai JJ, Lin X. 2014. Morphotype transition and sexual reproduction are genetically associated in a ubiquitous environmental pathogen. *PLoS Pathog.* 10(6):e1004185.
- Wang Z, Johnston PR, Yang ZL, Townsend JP. 2009. Evolution of reproductive morphology in leaf endophytes. *PLoS One* 4(1):e4246.
- Wang Z, Lopez-Giraldez F, Lehr N, Farré M, Common R, Trail F, Townsend JP. 2014. Global gene expression and focused knockout analysis reveals genes associated with fungal fruiting body development in *Neurospora crassa*. *Eukaryot Cell* 13(1):154–169.
- Wirthlin M, Lima NCB, Guedes RLM, Soares AER, Almeida LGP, Cavaleiro NP, Loss de Moraes G, Chaves AV, Howard JT, Teixeira M. D M, et al. 2018. Parrot genomes and the evolution of heightened longevity and cognition. *Curr Biol.* 28(24):4001–4008.e7.
- Xie KT, Wang G, Thompson AC, Wucherpfennig JJ, Reimchen TE, MacColl ADC, Schluter D, Bell MA, Vasquez KM, Kingsley DM. 2019. DNA fragility in the parallel evolution of pelvic reduction in stickleback fish. *Science* 363(6422):81–84.
- Yampolsky LY, Stoltzfus A. 2001. Bias in the introduction of variation as an orienting factor in evolution. *Evol Dev.* 3(2):73–83.
- Zhang J, Ren A, Chen H, Zhao M, Shi L, Chen M, Wang H, Feng Z. 2015. Transcriptome analysis and its application in identifying genes associated with fruiting body development in basidiomycete *Hypsizygus marmoreus*. *PLoS One* 10(4):e0123025.
- Zhang N, Wang Z. 2015. Pezizomycotina: sordariomycetes and leotiomycetes. In: McLaughlin D, Spatafora JW, editors. *The Mycota*. 2nd ed. Heidelberg: Springer. p. 57–88.
- Zhen Y, Aardema ML, Medina EM, Schumer M, Andolfatto P. 2012. Parallel molecular evolution in an herbivore community. *Science* 337(6102):1634–1637.
- Zou Z, Zhang J. 2015a. No genome-wide protein sequence convergence for echolocation. *Mol Biol Evol.* 32(5):1237–1241.
- Zou Z, Zhang J. 2015b. Are convergent and parallel amino acid substitutions in protein evolution more prevalent than neutral expectations? *Mol Biol Evol.* 32(8):2085–2096.