Article, Discoveries

Unmatched level of molecular convergence among deeply divergent complex multicellular fungi

Zsolt Merényi[1], Arun N. Prasanna[1&], Zheng Wang[2], Károly Kovács[5,1], Botond Hegedüs[1], Balázs Bálint[1], Balázs Papp[5,1], Jeffrey P. Townsend[2,3,4], László G. Nagy[1]*

Affiliations

[1]*Synthetic and Systems Biology Unit, Institute of Biochemistry, Biological Research Center, HAS, Szeged, 6726, Hungary*

[2]*Department of Biostatistics, Yale University, New Haven, CT, United States of America,*

[3]*Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, United States of America,*

[4]*Program in Computational Biology and Bioinformatics, Yale University, New Haven, CT, United States of America*

[5]*HCEMM-BRC Metabolic Systems Biology Lab*

*Corresponding author: László G. Nagy lnagy@fungenomelab.com

[&]Current address: Red Sea Science and Engineering Research Center, 4700 King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia

Keywords

Convergent evolution, complex multicellularity, fruiting-body development, developmental transcriptome, fungi

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 developmental transcriptome data to understand the independent evolution of complex multicellularity in the two largest clades of fungi—the Agarico- and Pezizomycotina. Despite >650 million years of divergence between these clades, we find that

very similar sets of genes have repeatedly 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 gamily 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 hypthesis that the coding capacity of ancestral fungal genomes might have promoted the repeated evolution of complex multicellularity. We interpret this 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.

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), which is starting to represent an enigma in evolutionary biology. The independent emergence of similar phenotypes is called convergent evolution, which happens 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 (Stern and Orgogozo 2009; Losos 2011; Blount et al. 2018; Xie et al. 2019). Convergence is widespread in nature (e.g. (Rundle et al. 2000; Muschick et al. 2012; D. Luke et al. 2013; van Velzen et al. 2018)) and its pervasive occurrence suggested that the evolution of certain traits may be predictable (Pankey et al. 2014) and deterministic (Losos 1998; D. Luke et al. 2013; Blount et al. 2018) under some circumstances, although 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 et al. 2018; Nagy 2018), and reached complexity levels that range from simple cell aggregates to the most complex macroscopic organisms (Sebé-Pedrós 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 Szathmary 1995; Knoll 2011).

The highest level of multicellular organization is referred to as complex multicellularity (CM), which, unlike unicells and simple multi-celled aggregates (e.g. filaments, colonies, 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, at first, find developmentally regulated genes shared by the Agarico- and Pezizomycotina, then analyze the evolution of these genes families to uncover genetic mechanisms of convergence in fungal CM.

Results

We assembled two whole-genome datasets, one comprising 19 and another containing 90 species. For simplicity, we discuss the 19-species dataset in the main text and present results for the 90-species dataset in the Supplement. To study the evolution of complex multicellularity in fungi we first inferred ancestral cellularity levels across the phylogenetic tree of 90 species (supplementary table S1) using ancestral character state reconstructions.

This analysis strongly suggests that the two most diverse CM clades, Agarico- and Pezizomycotina, acquired fruiting body formation independently (supplementary note 1, fig. 1/a, supplementary fig. S1). Proportional likelihoods indicate that the most recent common ancestor (MRCA) of the Dikarya 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}: 0.0), followed by two independent acquisitions of CM in the MRCA of the Agaricomycetes 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}: 0.0), which implies a third independent origin of fruiting body production in *Neolecta* (Taphrinomycotina) (Nagy 2017; Nguyen et al. 2017; Nagy et al. 2018) (fig. 1/a, supplementary fig. S1). Consistent with reconstructed ancestral states, the best-fit evolutionary model implied that gains of CM occurred at non-zero rate values (0.21 and 6.09 in ML and stochastic mapping, respectively), whereas the rate of the loss of CM was zero in both analyses of 90-species dataset. These model parameters strongly suggest that the evolution of CM is the preferred direction and in evolution and is non-reversible. Consistent with this, 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 mainy vertically(Mikheyev et al. 2006). Analyses of the 90-species dataset showed identical results (supplementary note 1, fig S1). 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)).

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; Z. Wang 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 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), 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, Aspergillus* and *Coprinopsis* (see supplementary note 3,

supplementary table S2), 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 dataset of 19 species (see 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. 1/b). We identified 1,026 gene families that were developmentally regulated in ≥75% of the species in either or both clades. These comprised 314, 439, and 273 families that showed a developmental expression in ≥7 of 9 Dikarya, ≥3 of 4 Agaricomycotina and ≥4 of 5 Pezizomycotina species, respectively (supplementary table S3). 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.

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 ≥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. Looking at individual species, we found that 11-87% of meiotic genes were developmentally regulated, the highest number observed in *Coprinopsis* and *Fusarium verticillioides* and the lowest in *Hypsizigus* and *Pyronema* (supplementary table S4, supplementary fig. S3). Because environmental stress is a common trigger of sexual development, we examined the contribution of stress-related genes to shared developmental expression patterns, to make sure shared patterns are not driven by stress-related genes. We identifed homologs of Core Environmental Stress Response (CESR) genes and members of the stress activated MAPK pathway (Gómez-Gil et al. 2019) in our species and examined their expression. Very few of the stress-related genes were developmentally regulated in the examined transcriptomes and none of them were so in 7-9 species (supplementary table S5), 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 are not or are 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 S3). Based on currently available information, it is impossible to decide if these regulate sexual reproductive processes, accompanying CM morphogenetic events or both, leaving some uncertainty on the contribution of conserved sexual reproductive genes to our data with regard to the mentioned regulatory families. We also found 3 orthogroups containing Pumilio-family proteins, which regulate cellular processes by sequence-specifically binding RNA molecules. Pumilio proteins have been implicated in morphogenetic processes that precede sexual reproduction in *Cryptococcus neoformans*, a species derived from primitive fruiting body forming Basidiomycota(Wang et al. 2014). Argonaute-family proteins were developmentally regulated in all nine species, suggesting that RNA interference is a key regulatory mechanism of fruiting body formation, as suggested recently for specific species(Lau et al. 2018; Shao et al. 2019).

Several cell-wall active CAZyme families and secreted proteins were found among the 314 shared developmentally regulated families, consistent with active cell wall remodeling during development(Wang and Lin 2012; Krizsán et al. 2019). These included chitin synthases, LysM domain proteins, chitin and glucan active glycoside hydrolase families, lectins, chitooligosaccharide deacethylases, among others (Supplementary Table 3). Laccases were developmentally regulated in eight species, consistent with their fruiting body-specific roles(Thurston 1994; Wang and Ng 2006; Sakamoto et al. 2015). Although cupredoxin domains are often associated with laccases, we found an 8-way developmentally regulated family that harbored only cupredoxin domains. These might be involved in developing pigmentation in fruiting bodies or assist in other copper-dependent processes (e.g. laccase activity)(Lopez et al. 2018). Metallopeptidase M53 (deuterolysin) proteins were developmentally regulated in eight species, which is consistent with previous reports of their upregulation in fruiting bodies of mainly Basidiomycetes (*Grifola, Schizophyllum, Armillaria* (Ohm et al. 2010; Sipos et al. 2017)). Fasciclin homologs were developmentally regulated in 8 species. This family was reported to be involved in the adhesion of pathogenic fungi to host surfaces (Liu et al. 2009), and was shown to be upregulated during primordium formation in *Lentinula edodes* (Miyazaki et al. 2007) and in protoperithecia (fruiting body initials) of *Sordaria macrospora* (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 an plesiomorphic role in adhesion to various surfaces and that they has been independently co-

eukaryote smith 10/24/2019 3:35 PM
Comment [1]: may want to remove this if we want to shorten this chapter.

opted for hypha-hypha adhesion (or cell wall modifications (Teichert et al. 2012)) in fruiting bodies of 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 or the non-CM developmental stages of the examined species, the above-mentioned gene families function in processes ranging from plant-fungal interactions (e.g. LysM proteins and chitooligosaccharide decethylases (Maillet et al. 2011)), the degradation of host extracellular matrix proteins by *Candida* (Rodier et al. 1999), the recognition and adhesion to carbohydrate-covered surfaces (Varrot et al. 2013) (fasciclins, lectins), or simple non-CM developmental processes (chitinases and glucanases). Cell wall remodeling is a key process in non-CM sexual development (e.g. basidium development in *Cryptococcus* (Liu et al. 2018)), asexual development (Beauvais and Latgé 2018) or the development of appressoria in pathogenic fungi (Geoghegan et al. 2017). Taken together, the functions of shared developmentally regulated families in non-CM fungi indicate that genes with diverse roles have been likely co-opted for CM in the Agarico- and Pezizomycotina. While the minority of these families are plesiomorphically linked to sexual development, developmentally regulated families appear to reflect a much broader set of ancestral functions. We explain this by the conserved nature of sexual reproduction (such as meiosis, cell cycle regulation), which, in terms of complexity level did not change in CM fungi, as opposed to morphogenetic events, which became orders of magnitudes more complex in CM Agarico- and Pezizomycotina.

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, while 297 and 169 families are taxonomically restricted to the Agarico- and Pezizomycotina, respectively (fig. 1/c). Similar figures were obtained for the 90-species dataset 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 =0, 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. 1/a), 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

263 carbohydrate metabolism, reminiscent of general functions required for fungal development
264 (supplementary fig. S4). We found that several gene regulatory circuits, including ones
265 involved in sexual reproduction, mating partner recognition, light, nutrient and starvation
266 sensing, fungal cell-wall synthesis and modification, cell-to-cell signaling, and
267 morphogenesis were present in the Dikarya ancestor (and even earlier), suggesting that these
268 might have provided a foundation for the evolution of fruiting bodies.

269 **Gene family expansions correlate with the origins of CM**

270 To obtain a higher-resolution picture of the evolution of developmental gene families, we
271 reconstructed gene duplications and losses in the Agarico- and Pezizomycotina and in other
272 parts of the fungal tree. In both the 19- and the 90-species dataset, we found characteristic
273 expansions of developmentally regulated gene families in CM Agarico- and Pezizomycotina,
274 but no or significantly smaller expansions in other families (fig. 2/a, supplementary fig. S5,
275 see supplementary fig. S6 and supplementary note 2 for results of the 90-species dataset).
276 Across the 314 shared families, we inferred a net expansion (duplication minus loss) of 323
277 and 250 genes in the MRCA of the Agarico- and that of the Pezizomycotina, respectively,
278 indicating that the origin of these CM groups coincided with significant expansions in
279 developmentally regulated gene families. In the MRCA of Dikarya we inferred 442
280 duplications and 45 losses. The observed gene family expansions were driven by increased
281 gene duplication rates, with loss rates remaining approximately constant (supplementary fig.
282 S7). A 6.3 to 8.1-fold higher rate of expansion was found in the 314 shared developmental
283 families compared to other families that were also shared by ≥7 of the nine species (fig 2/a-c),
284 indicating that CM-related gene families are one of the most expanding group in the fungal
285 genomes examined here.
286 Gene families with a developmental expression specific to the Agarico- (439 families)
287 or Pezizomycotina (273 families) show higher duplication rates (1.93–1.95-fold) and
288 substantially expanded in their respective clades, but not in the other CM subphylum or in
289 non-CM species (fig. 2/bc). Of the Agarico- and Pezizomycotina, the latter showed a more
290 gradual expansion of gene families: we reconstructed 104 and 162 net gains (duplications
291 minus losses) in the MRCA of Pezizomycotina and that of Sordariomycetes, respectively.
292 Interestingly, we inferred relatively few (73) duplications along the branch leading to
293 *Pyronema*, a representative of apothecium-forming Pezizomycotina. This species has 287
294 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 to filamentous or CM fungi. We inferred net contractions (19-193 gene duplications and 1977-2406 losses, supplementary fig. S8) 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). While 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 to their common ancestor (supplementary fig S4/b). 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 on fig. 2/d. Of the 314 shared developmental families, 257 (82%) showed parallel expansions in the Agarico- and Pezizomycotina. In contrast, only eight (2.5%) showed no duplications in either class and 49 (16%) showed duplications in only one. The 90-species dataset 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). If we considered families that likewise contained at least seven species, but developmentally regulated genes from up to 2 species

(1747 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. 2/e).

Convergent expansion in the 314 shared developmental families correlates with the number of species represented by developmentally regulated genes in the family (fig. 2/f). Convergence in this group of families is significantly more abundant than in any other combination of gene families we tested ($P < 1.2\times10^{-9}$, Fisher's exact test), including controls for gene family size and the number of developmentally regulated proteins per family, among others (supplementary fig. S10, supplementary table S6/a). In support of this finding, *Aspergillus* gene families with developmental functions verified by knockout studies (Cerqueira et al. 2014) show significantly more convergent expansion than non-developmental families ($P = 1.8\times10^{-6}$, Fisher's exact test, supplementary table S6/c). The accelerated duplication rate in these families differs considerably from that of other families (supplementary table S6), or from that of families with constraint on duplication imposed by a fitness cost of increased dosage (Sopko et al. 2006) (supplementary note 4, fig 2/g), collectively suggesting that the observed convergent gene family expansions were driven by positive selection.

Families with clade-specific developmental expression, on the other hand, did not show signs of convergent expansion (significantly underrepresented according to Fisher's exact test compared to all gene families with $P=3.5\times10^{-9}$ in Basidiomycota and $P=1.4\times10^{-6}$ in Ascomycota, supplementary fig. S11). We note that further convergently expanding families can certainly be found among those containing <7 species, which renders our estimate of convergence conservative. These observations suggest that CM Agarico- and Pezizomycotina have experienced significantly elevated numbers of convergent gene duplication events in gene families that can be linked to fruiting body development based on gene expression patterns.

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; Zou and Zhang 2015b; Storz 2016). We used the 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 is expect 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 fig. S12-S14, supplementary note 5). Developmentally regulated genes are enriched in 28 of these families (supplementary note 5), including genes related to cell division and DNA repair, splicing and ergosterol biosynthesis, among others (see supplementary table S7). However, the extent of convergence in CM clades was overall similar to that observed in other combinations of clades (supplementary note 5), 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 ~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, while 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 million years 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 (Gina L 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 non-adaptive) genetic parallelism (Colosimo et al. 2005; Stern and Orgogozo 2009). Because the probability of repeated recruitment of genes 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 million years ago and their ancestors existed >270 myr 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 8 fungal clades in which CM has evolved repeatedly (Nagy et al. 2018). The large number of independent transitions to CM (of which we here focused on two) in fungi suggests that its evolution may be pre-staged by some widely conserved genetic circuitries, resembling cases of evolutionary determinism documented elswhere(Wake et al. 2011; Elmer et al. 2014). Such genetic circuits perhaps include adhesion and cell-differentiation related genes used by non-CM fungi for various purposes, including perhaps pathogenicity or asexual reproduction, although this remains hard to test experimentally.

In the context of Gould's famous thought experiment (Gould 1990), the repeated emergence of CM in fungi means that if we replayed life's tape, CM would again evolve in fungal clades. Such predictability of evolution is inferred in a growing number of cases and has been attributed to shared genetic variation (Muschick et al. 2012), similar selective regimes (Losos 2011; D. Luke 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 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 to 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 J. B. S 1932). There is probably 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 towards 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 is not only limited 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 (FB) forming species from the Agaricomycotina (*Armillaria ostoyae*, *Coprinopsis cinerea*, *Hypsizygus marmoreus*, *Flammulina velutipes*) and the Pezizomycotina (*Fusarium graminearum*, *F. verticillioides*, *Sordaria macrospora*, *Neurospora crassa*, *Pyronema confluens*, supplementary table S1). 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 pre-filter, an FPKM value less than two was considered insignificant. The homogeneity of biological replicates was checked by constructing MDS plots based on overall expression levels. One 24h sample of *F. verticilloides* 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 four-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).

Comparative genomic approaches

In addition to the nine above mentioned fruiting body forming species, 10 additional species were included in the analysis for comparative purposes (supplementary table S1). This 19 genome dataset 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).

Gene Ontology (GO) enrichment analysis for yeast orthologs was performed using GOrilla ((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, Molecular function) were considered. Experimentally verified gene function from *Aspergillus nidulans* and *Neurospora crassa* 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 dataset containing 90 species (supplementary table S1). From this dataset, gene families were reconstructed using all vs all blastp search, $1e^{-5}$ e-value and an 30% bidirectional coverage filter in blastp output and Markov Cluster (MCL) 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 dataset, 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 2014), 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), maximum likelihood (ML) ancestral state reconstruction and Stochastic character mapping (Huelsenbeck et al. 2003) were performed with the asr_max_parsimony function of castor (Louca and Doebeli 2018), ace (ancestral character estimations) function of the ape (Paradis et al. 2004) and make.simmap function of phytools (Revell) R packages (R Development Core Team, 2018), respectively. The extant character states of species are listed in supplementary table S1. The ARD model was used in both ML reconstruction and stochastic mapping. For stochastic character mapping the estimated ancestral state and a fixed value of transition matrix from the best model for our empirical data was used and 1000 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, i.e. from non-CM to CM and from CM to non-CM).

To infer the ancestral cellularity level of the Dikarya ancestor (MRCA of the two main fruiting body forming clade: Agaricomycotina and Pezizomycotina), we used Maximum Likelihood ancestral state reconstruction in 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 free and two constrained models, where Dikarya was fixed as complex or simple multicellular state based on 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 (9686 gene families). Sequences in each gene family were aligned using the L-INS-I method of MAFFT and trimmed with trim-AL (-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 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: a) a gene family has genes conserved in ≥7 of the 9 Dikarya species, ≥4 of the 5 Pezizomycotina species or ≥3 of the 4 Agaricomycotina species or b) a gene family has developmental expression conserved in ≥7 of the 9 Dikarya species, ≥3 of the 4 Agaricomycotina species or ≥4 of the 5 Pezizomycotina species.

Gene families of the 90-species dataset 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 if 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 to the conserved developmentally regulated gene family set (developmentally expression conserved in ≥7 of the 9 Dikarya species). The first control groups comprised families that similarly contained ≥7 species but only 0-2 (1747 gene families) or 3-6 species (2052 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 non-developmentally 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 (4113) gene families represented by at least 7 out of the 9 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 "Profile Change with One Change" (PCOC) (Rey et al. 2018) (downloaded from https://github.com/CarineRey/pcoc on 2018.11.05). We used reconciled gene trees with branch lengths re-estimated with RAxML (raxmlHPC-PTHREADS-SSE3) as input. Each of the most inclusive clades that contained only CM species were 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 3799 gene families that contained at least 10 proteins and at least one 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, Taphrinomycotina). In control 2, CM Agaricomycotina species were paired with the basal clades of Ascomycota (Saccharomycotina, Taphrinomycotina) while in control 3 CM Pezizomycotina were paired with the basal clades of Basidiomycota (Ustilagomycotina, 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.

Acknowledgements:

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 programme under grant agreement No 739593 to B.P. and National Science Foundation (IOS 1457044) to J.P.T.).

Author contributions:

ZM and LGN conceived the study. ZM, ZW, ANP, JPT, BH and BB analyzed data. ANP analyzed transcriptomic data, ZW, JPT and ZM evaluated developmentally regulated genes, ZM, BH and BB reconstructed gene family evolution. ZM, KK and BP evaluated adaptivity of gene family expansions. ZM, BP and LGN 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. [Internet] 190:S1–S12. Available from: http://www.journals.uchicago.edu/t-and-c

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

Beauvais A, Latgé JP. 2018. Special issue: Fungal cell wall. J. Fungi 4.

Blount ZD, Lenski RE, Losos JB. 2018. Contingency and determinism in evolution: Replaying life's tape. Science [Internet] 362:eaam5979. Available from: http://science.sciencemag.org/

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics [Internet] 30:2114–2120. Available from:

https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btu170

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. [Internet] 34:525–527. Available from: http://www.nature.com/articles/nbt.3519

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics [Internet] 25:1972–1973. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19505945

Castoe TA, Jason De Koning AP, 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. PNAS [Internet] 106:8986–8991. Available from: www.pnas.org/cgi/content/full/

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. [Internet] 42:D705–D710. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt1029

Christin P-A, Weinreich DM, Besnard G. 2010. Causes and evolutionary significance of genetic convergence. Trends Genet. [Internet] 26:400–405. Available from: https://ac.els-cdn.com/S0168952510001289/1-s2.0-S0168952510001289-main.pdf?_tid=12c40414-f1da-4774-a9ba-a1fe149c4998&acdnat=1548084251_a8fb97f2a75053ec166c4d971b4665db

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. [Internet] 12:115–124. Available from: http://www.nature.com/articles/nrmicro3178

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 [Internet] 307:1928–1933. Available from: http://science.sciencemag.org/

Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic parallelism and convergence in natural populations. Proc. R. Soc. B Biol. Sci. [Internet] 279:5039–5047. Available from: http://rspb.royalsocietypublishing.org.

Conte GL, Arnegard ME, Peichel CL, Schluter D. 2012. The probability of genetic

690 parallelism and convergence in natural populations. Proc. R. Soc. B Biol. Sci. [Internet]
691 279:5039–5047. Available from:
692 https://royalsocietypublishing.org/doi/10.1098/rspb.2012.2146
693 D. Luke M, Ingram T, Revell LJ, Losos JB. 2013. Exceptional Convergence on the
694 Macroevolutionary Landscape in Island Lizard Radiations. Science [Internet]. Available
695 from: http://science.sciencemag.org/
696 Darby CA, Stolzer M, Ropp PJ, Barker D, Durand D. 2016. Xenolog classification.
697 Bioinformatics [Internet] 33:btw686. Available from:
698 https://academic.oup.com/bioinformatics/article-
699 lookup/doi/10.1093/bioinformatics/btw686
700 Darwin C, Mayr E. 1995. On the origin of species. Harvard University Press Available from:
701 http://www.hup.harvard.edu/catalog.php?isbn=9780674637528
702 Dongen SM van. 2000. Graph clustering by flow simulation. Available from:
703 https://dspace.library.uu.nl/handle/1874/848
704 Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. 2009. GOrilla: a tool for discovery and
705 visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics [Internet]
706 10:48. Available from: http://www.biomedcentral.com/1471-2105/10/48
707 Elmer KR, Fan S, Kusche H, Luise Spreitzer M, Kautt AF, Franchini P, Meyer A. 2014.
708 Parallel evolution of Nicaraguan crater lake cichlid fishes via non-parallel routes. Nat.
709 Commun. [Internet] 5:5168. Available from:
710 http://www.nature.com/articles/ncomms6168
711 Emms DM, Covshoff S, Hibberd JM, Kelly S. 2016. Independent and Parallel Evolution of
712 New Genes by Gene Duplication in Two Origins of C4 Photosynthesis Provides New
713 Insight into the Mechanism of Phloem Loading in C4 Species. Mol. Biol. Evol.
714 [Internet] 33:1796–1806. Available from: http://creativecommons.
715 Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome
716 comparisons dramatically improves orthogroup inference accuracy. Genome Biol.
717 [Internet] 16:157. Available from: http://genomebiology.com/2015/16/1/157
718 Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R,
719 Spatafora JW, Yadav JS, et al. 2012. The Paleozoic Origin of Enzymatic Lignin
720 Decomposition Reconstructed from 31 Fungal Genomes. Science [Internet] 336:1715–
721 1719. Available from: http://science.sciencemag.org/content/336/6089/1715/tab-pdf
722 Geoghegan I, Steinberg G, Gurr S. 2017. The Role of the Fungal Cell Wall in the Infection of
723 Plants. Trends Microbiol. 25:957–967.

724 Gómez-Gil E, Franco A, Madrid M, Vázquez-Marín B, Gacto M, Fernández-Breis J, Vicente-
725 Soler J, Soto T, Cansado J. 2019. Correction: Quorum sensing and stress-activated
726 MAPK signaling repress yeast to hypha transition in the fission yeast
727 Schizosaccharomyces japonicus. PLOS Genet. [Internet] 15:e1008282. Available from:
728 http://dx.plos.org/10.1371/journal.pgen.1008282
729 Gould SJ. 1990. Wonderful life : the Burgess Shale and the nature of history. New York:
730 WW Norton & Company Available from: https://www.amazon.com/Wonderful-Life-
731 Burgess-Nature-History/dp/039330700X
732 Griesmann M, Chang Y, Liu X, Song Y, Haberer G, Crook MB, Billault-Penneteau B,
733 Lauressergues D, Keller J, Imanishi L, et al. 2018. Phylogenomics reveals multiple
734 losses of nitrogen-fixing root nodule symbiosis. Science [Internet] 361:eaat1743.
735 Available from: http://www.ncbi.nlm.nih.gov/pubmed/29794220
736 Grosberg RK, Strathmann RR. 2007. The Evolution of Multicellularity: A Minor Major
737 Transition? Annu. Rev. Ecol. Evol. Syst. [Internet] 38:621–654. Available from:
738 http://www.annualreviews.org/doi/10.1146/annurev.ecolsys.36.102403.114735
739 Haldane J. B. S. 1932. The causes of evolution. London: Harper and Brothers
740 Hibbett DS. 2007. After the gold rush, or before the flood? Evolutionary morphology of
741 mushroom-forming fungi (Agaricomycetes) in the early 21st century. Mycol. Res.
742 [Internet] 111:1001–1018. Available from:
743 http://www.ncbi.nlm.nih.gov/pubmed/17964768
744 Huber O, Sumper M. 1994. Algal-CAMs: isoforms of a cell adhesion molecule in embryos of
745 the alga Volvox with homology to Drosophila fasciclin I. EMBO J. 13:4212–4222.
746 Huelsenbeck JP, Nielsen R, Bollback JP. 2003. Stochastic Mapping of Morphological
747 Characters.Schultz T, editor. Syst. Biol. [Internet] 52:131–158. Available from:
748 http://academic.oup.com/sysbio/article/52/2/131/1634311
749 Hughes AL, Friedman R. 2003. Parallel evolution by gene duplication in the genomes of two
750 unicellular fungi. Genome Res. [Internet] 13:794–799. Available from:
751 http://www.ncbi.nlm.nih.gov/pubmed/12727899
752 Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J,
753 Mitchell A, Nuka G, et al. 2014. InterProScan 5: genome-scale protein function
754 classification. Bioinformatics [Internet] 30:1236–1240. Available from:
755 http://www.ncbi.nlm.nih.gov/pubmed/24451626
756 Katoh K, Standley DM. 2013. MAFFT Multiple Sequence Alignment Software Version 7:
757 Improvements in Performance and Usability. Mol. Biol. Evol. [Internet] 30:772–780.

Available from: https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/mst010
Knoll AH. 2011. The Multiple Origins of Complex Multicellularity. Annu. Rev. Earth Planet. Sci. [Internet] 39:217–239. Available from: http://www.annualreviews.org/doi/10.1146/annurev.earth.031208.100209
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. [Internet] 47:410–415. Available from: http://www.nature.com/articles/ng.3223
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. [Internet] 116:7409–7418. Available from: https://www.biorxiv.org/content/10.1101/349894v1
Lau AYT, Cheng X, Cheng CK, Nong W, Cheung MK, Chan RH-F, Hui JHL, Kwan HS. 2018. Discovery of microRNA-like RNAs during early fruiting body development in the model mushroom Coprinopsis cinerea. PLoS One [Internet] 13:e0198234. Available from: http://www.ncbi.nlm.nih.gov/pubmed/30231028
Liu L, He GJ, Chen L, Zheng J, Chen Y, Shen L, Tian X, Li E, Yang E, Liao G, et al. 2018. Genetic basis for coordination of meiosis and sexual structure maturation in cryptococcus neoformans. Elife 7.
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 10:434–444.
Liu YJ, Hall BD. 2004. Body plan evolution of ascomycetes, as inferred from an RNA polymerase II phylogeny. Proc. Natl. Acad. Sci. [Internet] 101:4507–4512. Available from: www.pnas.orgcgidoi10.1073pnas.0400938101
Lopez D, Ribeiro S, Label P, Fumanal B, Venisse JS, Kohler A, de Oliveira RR, Labutti K, Lipzen A, Lail K, et al. 2018. Genome-wide analysis of Corynespora cassiicola leaf fall disease putative effectors. Front. Microbiol. 9.
Lord KM, Read ND. 2011. Perithecium morphogenesis in Sordaria macrospora. Fungal Genet. Biol. [Internet] 48:388–399. Available from: https://www.sciencedirect.com/science/article/pii/S1087184510002227?via%3Dihub
Losos JB. 1998. Contingency and Determinism in Replicated Adaptive Radiations of Island Lizards. Science [Internet] 279:2115–2118. Available from:

http://www.sciencemag.org/cgi/doi/10.1126/science.279.5359.2115
Losos JB. 2011. Convergence, adaptation, and constraint. Evolution (N. Y). [Internet] 65:1827–1840. Available from: http://doi.wiley.com/10.1111/j.1558-5646.2011.01289.x
Louca S, Doebeli M. 2018. Efficient comparative phylogenetics on large trees.Valencia A, editor. Bioinformatics [Internet] 34:1053–1055. Available from: https://academic.oup.com/bioinformatics/article-abstract/34/6/1053/4582279
Löytynoja A. 2014. Phylogeny-aware alignment with PRANK. In: Methods in molecular biology (Clifton, N.J.). Vol. 1079. p. 155–170. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24170401
Maillet F, Poinsot 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 469:58–64.
McGhee GR. 2011. Convergent evolution : limited forms most beautiful. MIT Press
Mikheyev AS, Mueller UG, Abbot P. 2006. Cryptic sex and many-to-one coevolution in the fungus-growing ant symbiosis. Available from: www.pnas.orgcgidoi10.1073pnas.0601441103
Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. Available from: 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. 51:367–375.
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.Kothe E, editor. PLoS One [Internet] 10:e0141586. Available from: https://dx.plos.org/10.1371/journal.pone.0141586
Muschick M, Indermaur A, Salzburger W. 2012. Convergent Evolution within an Adaptive Radiation of Cichlid Fishes. Curr. Biol. [Internet] 22:2362–2368. Available from: http://dx.doi.org/10.1016/j.cub.2012.10.048
Nagy LG. 2017. Evolution: Complex Multicellular Life with 5,500 Genes. Curr. Biol. [Internet] 27:R609–R612. Available from: https://www.sciencedirect.com/science/article/pii/S0960982217304761
Nagy LG. 2018. Many roads to convergence. Science [Internet] 361:125–126. Available from: http://science.sciencemag.org/content/361/6398/125#BIBL

826 Nagy LG, Kovács GM, Krizsán K. 2018. Complex multicellularity in fungi: evolutionary
827 convergence, single origin, or both? Biol. Rev. [Internet] 93:1778–1794. Available
828 from: http://doi.wiley.com/10.1111/brv.12418
829 Nagy LG, Ohm RA, Kovács GM, Floudas D, Riley R, Gácser A, Sipiczki M, Davis JM, Doty
830 SL, de Hoog GS, et al. 2014. Latent homology and convergent regulatory evolution
831 underlies the repeated emergence of yeasts. Nat. Commun. [Internet] 5:4471. Available
832 from: www.nature.com/naturecommunications
833 Nguyen TA, Cissé OH, Yun Wong J, Zheng P, Hewitt D, Nowrousian M, Stajich JE, Jedd G.
834 2017. Innovation and constraint leading to complex multicellularity in the Ascomycota.
835 Nat. Commun. [Internet] 8:14444. Available from:
836 http://www.nature.com/doifinder/10.1038/ncomms14444
837 Ohm RA, De Jong JF, Lugones LG, Aerts A, Kothe E, Stajich JE, De Vries RP, Record E,
838 Levasseur A, Baker SE, et al. 2010. Genome sequence of the model mushroom
839 Schizophyllum commune. Nat. Biotechnol. 28:957–963.
840 Pagel M. 1999. The Maximum Likelihood Approach to Reconstructing Ancestral Character
841 States of Discrete Characters on Phylogenies. Available from:
842 https://academic.oup.com/sysbio/article-abstract/48/3/612/1641545
843 Pagel M, Meade A. 2007. BayesTraits. Computer program and documentation available at
844 http://www.evolution.rdg.ac.uk/BayesTraits.html. :1216–1223.
845 Pankey MS, Minin VN, Imholte GC, Suchard MA, Oakley TH. 2014. Predictable
846 transcriptome evolution in the convergent and complex bioluminescent organs of squid.
847 Proc. Natl. Acad. Sci. U. S. A. [Internet] 111:E4736-42. Available from:
848 http://www.ncbi.nlm.nih.gov/pubmed/25336755
849 Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of Phylogenetics and Evolution in R
850 language. Bioinformatics [Internet] 20:289–290. Available from:
851 https://academic.oup.com/bioinformatics/article-
852 lookup/doi/10.1093/bioinformatics/btg412
853 Park B, Rutter MT, Fenster CB, Symonds VV, Ungerer MC, Townsend JP. 2017.
854 Distributions of Mutational Effects and the Estimation of Directional Selection in
855 Divergent Lineages of Arabidopsis thaliana. Genetics [Internet] 206:2105–2117.
856 Available from: http://www.genetics.org/content/genetics/206/4/2105.full.pdf
857 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,
858 et al. 2014. Whole Genome and Global Gene Expression Analyses of the Model
859 Mushroom Flammulina velutipes Reveal a High Capacity for Lignocellulose

Degradation.Schmoll M, editor. PLoS One [Internet] 9:e93560. Available from: http://dx.plos.org/10.1371/journal.pone.0093560
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 [Internet] 502:228–231. Available from: http://www.nature.com/articles/nature12511
Paulsrud P, Lindblad P. 2002. Fasciclin domain proteins are present in Nostoc symbionts of lichens. Appl. Environ. Microbiol. 68:2036–2039.
Pöggeler S, Nowrousian M, Teichert I, Beier A, Kück U. 2018. Fruiting-Body Development in Ascomycetes. In: Physiology and Genetics. Cham: Springer International Publishing. p. 1–56. Available from: http://link.springer.com/10.1007/978-3-319-71740-1_1
Revell LJ. APPLICATION phytools: an R package for phylogenetic comparative biology (and other things). Available from: http://www.respond2articles.com/MEE/
Rey C, Guéguen L, Sémon M, Boussau B. 2018. Accurate Detection of Convergent Amino-Acid Evolution with PCOC.Pupko T, editor. Mol. Biol. Evol. [Internet] 35:2296–2306. Available from: https://academic.oup.com/mbe/article-abstract/35/9/2296/5050468
Rice DP, Townsend JP. 2012. A Test for Selection Employing Quantitative Trait Locus and Mutation Accumulation Data. Genetics [Internet] 190:1533–1545. Available from: http://www.genetics.org/content/
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. [Internet] 177:205–210. Available from: https://academic.oup.com/femsle/article-lookup/doi/10.1111/j.1574-6968.1999.tb13733.x
Rokas A, Carroll SB. 2008. Frequent and Widespread Parallel Evolution of Protein Sequences. Mol. Biol. Evol. [Internet] 25:1943–1953. Available from: https://academic.oup.com/mbe/article-abstract/25/9/1943/1301656
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. [Internet] 23:113–118. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17275133
Rundle HD, Nagel L, Boughman JW, Schluter D. 2000. Natural Selection and Parallel Speciation in Sympatric Sticklebacks. Science [Internet] 287:306–308. Available from: http://science.sciencemag.org/
Sakamoto Y, Nakade K, Konno N. 2011. Endo-β-1,3-Glucanase GLU1, from the Fruiting Body of Lentinula edodes, Belongs to a New Glycoside Hydrolase Family. Appl.

894 Environ. Microbiol. [Internet] 77:8350–8354. Available from:
895 https://aem.asm.org/content/77/23/8350
896 Sakamoto Y, Nakade K, Yoshida K, Natsume S, Miyazaki K, Sato S, van Peer AF, Konno N.
897 2015. Grouping of multicopper oxidases in Lentinula edodes by sequence similarities
898 and expression patterns. AMB Express 5.
899 Schoch CL, Sung G-H, Pez-Gir Ldez FL, Townsend JP, Miadlikowska J, Hofstetter VR,
900 Robbertse B, Matheny PB, Kauff F, Wang Z, et al. 2009. The Ascomycota Tree of Life:
901 A Phylum-wide Phylogeny Clarifies the Origin and Evolution of Fundamental
902 Reproductive and Ecological Traits. Syst. Biol [Internet] 58:224–239. Available from:
903 http://www.homepage.mac.com/
904 Sebé-Pedrós A, Degnan BM, Ruiz-Trillo I. 2017. The origin of Metazoa: a unicellular
905 perspective. Nat. Rev. Genet. [Internet] 18:498–512. Available from:
906 http://www.ncbi.nlm.nih.gov/pubmed/28479598
907 Shao Y, Tang J, Chen S, Wu Y, Wang K, Ma B, Zhou Q, Chen A, Wang Y. 2019. MilR4 and
908 milR16 mediated fruiting body development in the medicinal fungus cordyceps
909 militaris. Front. Microbiol. 10.
910 Shen Y-Y, Liang L, Li G-S, Murphy RW, Zhang Y-P. 2012. Parallel Evolution of Auditory
911 Genes for Echolocation in Bats and Toothed Whales.Zhang J, editor. PLoS Genet.
912 [Internet] 8:e1002788. Available from:
913 https://dx.plos.org/10.1371/journal.pgen.1002788
914 Shubin N, Tabin C, Carroll S. 2009. Deep homology and the origins of evolutionary novelty.
915 Nature [Internet] 457:818–823. Available from:
916 http://www.nature.com/articles/nature07891
917 Sikhakolli UR, López-Giráldez F, Li N, Common R, Townsend JP, Trail F. 2012.
918 Transcriptome analyses during fruiting body formation in Fusarium graminearum and
919 Fusarium verticillioides reflect species life history and ecology. Fungal Genet. Biol.
920 [Internet] 49:663–673. Available from:
921 https://www.sciencedirect.com/science/article/pii/S1087184512000977?via%3Dihub
922 Sipos G, Prasanna AN, Walter MC, O'Connor E, Bálint B, Krizsán K, Kiss B, Hess J, Varga
923 T, Slot J, et al. 2017. Genome expansion and lineage-specific genetic innovations in the
924 forest pathogenic fungi Armillaria. Nat. Ecol. Evol. [Internet] 1:1931–1941. Available
925 from: http://www.nature.com/articles/s41559-017-0347-8
926 Smith JM, Szathmary E. 1995. The Major Transitions in Evolution. New York:W. H.
927 Freeman and Company

928 Sopko R, Huang D, Preston N, Chua G, Papp B, Kafadar K, Snyder M, Oliver SG, Cyert M,
929 Hughes TR, et al. 2006. Mapping Pathways and Phenotypes by Systematic Gene
930 Overexpression. Mol. Cell [Internet] 21:319–330. Available from:
931 https://www.sciencedirect.com/science/article/pii/S1097276505018538#fig5
932 Spatafora JW, Sung G-H, Johnson D, Hesse C, O'rourke B, Serdani M, Spotts R, Lutzoni F,
933 Hofstetter V, Miadlikowska J, et al. 2006. A five-gene phylogeny of Pezizomycotina. #
934 2006 by The Mycological Society of America Available from:
935 http://www.fieldmuseum.org/myconet
936 Stajich JE, Berbee ML, Blackwell M, Hibbett DS, James TY, Spatafora JW, Taylor JW.
937 2009. The fungi. Curr. Biol. [Internet] 19:R840-5. Available from:
938 http://www.ncbi.nlm.nih.gov/pubmed/19788875
939 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
940 large phylogenies. Bioinformatics [Internet] 30:1312–1313. Available from:
941 http://www.ncbi.nlm.nih.gov/pubmed/24451623
942 Stern DL, Orgogozo V. 2009. Is Genetic Evolution Predictable? Science [Internet] 323:746–
943 751. Available from: http://science.sciencemag.org/content/323/5915/746/tab-pdf
944 Storz JF. 2016. Causes of molecular convergence and parallelism in protein evolution. Nat.
945 Rev. Genet. [Internet] 17:239–250. Available from: www.nature.com/nrg
946 Teichert I, Wolff G, Kück U, Nowrousian M. 2012. Combining laser microdissection and
947 RNA-seq to chart the transcriptional landscape of fungal development. BMC Genomics
948 [Internet] 13:511. Available from: http://www.biomedcentral.com/1471-2164/13/511
949 Thurston CF. 1994. The structure and function of fungal laccases. Microbiology 140:19–26.
950 Traeger S, Altegoer F, Freitag M, Gabaldon T, Kempken F, Kumar A, Marcet-Houben M,
951 Pöggeler S, Stajich JE, Nowrousian M. 2013. The Genome and Development-Dependent
952 Transcriptomes of Pyronema confluens: A Window into Fungal Evolution.Richardson
953 PM, editor. PLoS Genet. [Internet] 9:e1003820. Available from:
954 https://dx.plos.org/10.1371/journal.pgen.1003820
955 Trail F, Wang Z, Stefanko K, Cubba C, Townsend JP. 2017. The ancestral levels of
956 transcription and the evolution of sexual phenotypes in filamentous fungi.Fraser HB,
957 editor. PLOS Genet. [Internet] 13:e1006867. Available from:
958 https://dx.plos.org/10.1371/journal.pgen.1006867
959 Varrot A, Basheer SM, Imberty A. 2013. Fungal lectins: structure, function and potential
960 applications. Curr. Opin. Struct. Biol. [Internet] 23:678–685. Available from:
961 http://www.ncbi.nlm.nih.gov/pubmed/23920351

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. [Internet] 115:E4700–E4709. Available from: http://www.ncbi.nlm.nih.gov/pubmed/29717040

Wake DB, Wake MH, Specht CD. 2011. Homoplasy: From detecting pattern to determining process in evolution, but with a secondary role for morphology? Zootaxa 1032:67–68.

Wang HX, Ng TB. 2006. Purification of a laccase from fruiting bodies of the mushroom Pleurotus eryngii. Appl. Microbiol. Biotechnol. [Internet] 69:521–525. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16075291

Wang L, Lin X. 2012. Morphogenesis in Fungal Pathogenicity: Shape, Size, and Surface.Heitman J, editor. PLoS Pathog. [Internet] 8:e1003027. Available from: https://dx.plos.org/10.1371/journal.ppat.1003027

Wang, 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.Sil A, editor. PLoS Pathog. [Internet] 10:e1004185. Available from: https://dx.plos.org/10.1371/journal.ppat.1004185

Wang Z, Johnston PR, Yang ZL, Townsend JP. 2009. Evolution of Reproductive Morphology in Leaf Endophytes. PLoS One [Internet] 4:4246. Available from: www.plosone.org

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 [Internet] 13:154–169. Available from: https://ec.asm.org/content/13/1/154.long

Wirthlin M, Lima NCB, Muniz RL, Tereza A, Vasconcelos R, Prosdocimi F, Mello Correspondence C V. 2018. Parrot Genomes and the Evolution of Heightened Longevity and Cognition. Curr. Biol. [Internet] 28:4001–4008.e7. Available from: https://doi.org/10.1016/j.cub.2018.10.050

Xie KT, Wang G, Thompson AC, Wucherpfennig JI, 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 363:81–84.

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.Nowrousian M, editor. PLoS

One [Internet] 10:e0123025. Available from: http://dx.plos.org/10.1371/journal.pone.0123025

Zhang N, Wang Z. 2015. Pezizomycotina: Sordariomycetes and Leotiomycetes. In: McLaughlin, David, Spatafora JW, editor. Second. Springer. p. 57–88.

Zhen Y, Aardema ML, Medina EM, Schumer M, Andolfatto P. 2012. Parallel Molecular Evolution in an Herbivore Community. Science [Internet] 337:1634–1637. Available from: http://flybase.

Zou Z, Zhang J. 2015a. No Genome-Wide Protein Sequence Convergence for Echolocation. Mol. Biol. Evol. [Internet] 32:1237–1241. Available from: https://academic.oup.com/mbe/article-abstract/32/5/1237/1126808

Zou Z, Zhang J. 2015b. Are Convergent and Parallel Amino Acid Substitutions in Protein Evolution More Prevalent Than Neutral Expectations? Mol. Biol. Evol. [Internet] 32:2085–2096. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4833076/pdf/msv091.pdf

Fig. titles and legends

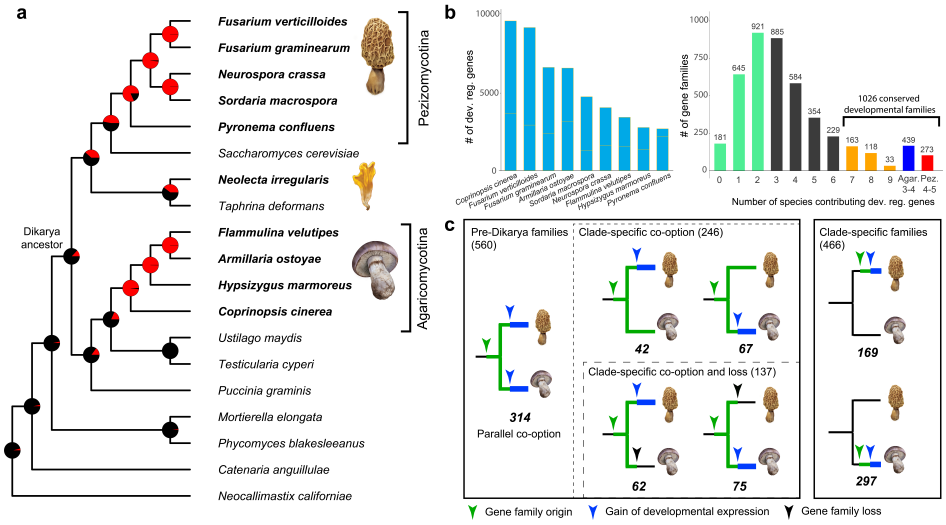

Figure 1. The evolution of complex multicellularity in fungi and conserved developmentally regulated gene families. **(a)** phylogenetic relationships among 19 species analyzed in this study inferred from 86 conserved, single-copy orthologs. Two independent clades of complex multicellular species are 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 Maximum Likelihood. Character state coding of extant species are 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 ≥7 species are highlighted in yellow. Groups of gene families that are developmentally regulated in ≥3 Agaricomycotina or ≥4 Pezizomycotina are also shown. **(c)** developmentally regulated gene families grouped by evolutionary conservation and history.

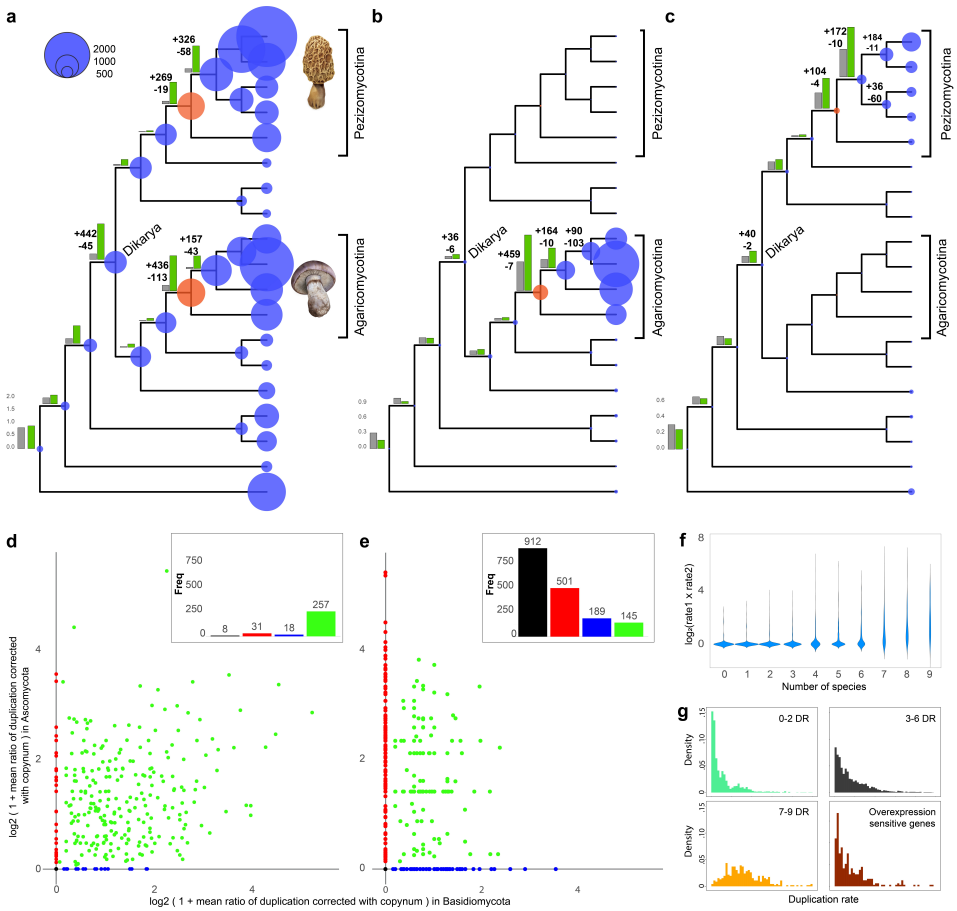

Figure 2. Convergent expansion of developmentally regulated gene families in independent complex multicellular fungi. **(a–c)** Reconstructed copy number evolution of 314 shared developmentally regulated gene families (a), 439 families with Agaricomycotina-specific developmental expression and (c) 273 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 (grey) 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 1747 families containing ≤2 developmentally regulated species (e). Black, red, blue and green denote families with no duplications, Pezizomycotina specific-, Agarimycotina 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 ≤2, 3-6 and ≥7 species (see in fig1/b) and families in which dosage effects constrain duplications rates (Sopko et al. 2006).

Captions for Supplementary Materials

Supplementary notes 1-5 (Merenyi_et_al_Supplement_0809.docx)

Supplementary Figs. S1 to S3 and S5 to S14 (Merenyi_et_al_Supplement_0809.docx)

Supplementary fig. S4. (Suppl_Fig_4.pdf) Gene Ontology terms of 314 developmental gene families.

Supplementary table S1. (Suppl_Table_1.xlsx) List of 19 species used in comparative genomic analyses

Supplementary table S2. (Suppl_Table_2.xlsx) Comparisons developmentally regulated genes to homologues with known developmental roles in model organisms of fungal CM.

Supplementary table S3. (Suppl_Table_3.xlsx) List of 1026 gene families with conserved developmental expression

Supplementary table S4. (Suppl_Table_4.xlsx) Meiosis related gene families.

Supplementary table S5. (Suppl_Table_5.xlsx) Stress related gene families.

Supplementary table S6. (Suppl_Table_6.xlsx) Statistical comparisons of the amount of parallel duplication in developmental versus control gene families.

Supplementary table S7. (Suppl_Table_7.xlsx) Gene families with significantly more amino acid shifts than expected by chance.