Report

Current Biology

Genome-level analyses resolve an ancient lineage of symbiotic ascomycetes

Highlights

- Sampling gaps in uncultured non-model fungi are filled using metagenomics
- Numerous putative early-diverging *Ascomycota* lineages resolve in a novel clade
- The clade includes lichen and insect symbionts, endophytes, and mycorrhizae
- Fungi in this clade possess small genomes with reduced metabolic arsenals

Authors

David Díaz-Escandón, Gulnara Tagirdzhanova, Dan Vanderpool, ..., Robert Lücking, Philipp Resl, Toby Spribille

Correspondence

toby.spribille@ualberta.ca

In brief

Genome-scale evolutionary reconstructions of ascomycete fungi have so far relied on sampling well-known groups, but many lesser-known lineages remain unsampled. Using 30 new genomes of putative early-diverging ascomycetes, Díaz-Escandón et al. demonstrate the existence of an ancient symbiont clade with reduced genome sizes.

Díaz-Escandón et al., 2022, Current Biology *32*, 5209–5218
December 5, 2022
https://doi.org/10.1016/j.cub.2022.11.014

CellPress

Current Biology

CellPress
OPEN ACCESS

Report

Genome-level analyses resolve an ancient lineage of symbiotic ascomycetes

David Díaz-Escandón,[1] Gulnara Tagirdzhanova,[1] Dan Vanderpool,[2] Carmen C.G. Allen,[1] André Aptroot,[3] Oluna Češka,[4,14] David L. Hawksworth,[5,6,7] Alejandro Huereca,[1] Kerry Knudsen,[8] Jana Kocourková,[8] Robert Lücking,[9] Philipp Resl,[10,11,12] and Toby Spribille[1,11,12,13,15,*]

[1]Department of Biological Sciences CW405, University of Alberta, Edmonton, AB T6G 2R3, Canada
[2]National Genomics Center for Wildlife and Fish Conservation, Rocky Mountain Research Station, 800 E Beckwith, Missoula, MT 59812, USA
[3]Laboratório de Botânica / Liquenologia, Instituto de Biociências Universidade Federal de Mato Grosso do Sul, Avenida Costa e Silva s/n Bairro Universitário, Campo Grande, Mato Grosso do Sul CEP 79070-900, Brazil
[4]Victoria, BC, Canada
[5]Comparative Fungal Biology, Royal Botanic Gardens, Kew, Surrey TW9 3DS, UK
[6]Department of Life Sciences, The Natural History Museum, Cromwell Road, London SW7 5BD, UK
[7]Jilin Agricultural University, Changchun, Jilin Province 130118, China
[8]Czech University of Life Sciences, Faculty of Environmental Sciences, Department of Ecology, Kamýcká 129, Praha-Suchdol 165 00, Czech Republic
[9]Botanischer Garten, Freie Universität Berlin, Königin-Luise-Straße 6-8, 14195 Berlin, Germany
[10]Institute of Biology, University of Graz, Universitätsplatz 2, 8010 Graz, Austria
[11]These authors contributed equally
[12]Senior author
[13]Twitter: @TobySpribille
[14]In memoriam; deceased November 9, 2022
[15]Lead contact
*Correspondence: toby.spribille@ualberta.ca
https://doi.org/10.1016/j.cub.2022.11.014

SUMMARY

Ascomycota account for about two-thirds of named fungal species.[1] Over 98% of known *Ascomycota* belong to the *Pezizomycotina*, including many economically important species as well as diverse pathogens, decomposers, and mutualistic symbionts.[2] Our understanding of *Pezizomycotina* evolution has until now been based on sampling traditionally well-defined taxonomic classes.[3–5] However, considerable diversity exists in undersampled and uncultured, putatively early-diverging lineages, and the effect of these on evolutionary models has seldom been tested. We obtained genomes from 30 putative early-diverging lineages not included in recent phylogenomic analyses and analyzed these together with 451 genomes covering all available ascomycete genera. We show that 22 of these lineages, collectively representing over 600 species, trace back to a single origin that diverged from the common ancestor of *Eurotiomycetes* and *Lecanoromycetes* over 300 million years BP. The new clade, which we recognize as a more broadly defined *Lichinomycetes*, includes lichen and insect symbionts, endophytes, and putative mycorrhizae and encompasses a range of morphologies so disparate that they have recently been placed in six different taxonomic classes. To test for shared hidden features within this group, we analyzed genome content and compared gene repertoires to related groups in *Ascomycota*. Regardless of their lifestyle, *Lichinomycetes* have smaller genomes than most filamentous *Ascomycota*, with reduced arsenals of carbohydrate-degrading enzymes and secondary metabolite gene clusters. Our expanded genome sample resolves the relationships of numerous "orphan" ascomycetes and establishes the independent evolutionary origins of multiple mutualistic lifestyles within a single, morphologically hyperdiverse clade of fungi.

RESULTS AND DISCUSSION

Our attention was initially drawn to seven recently described classes of *Ascomycota*,[6–11] none of which had been previously resolved with respect to evolutionary relationships, using various taxon and locus samplings.[3,4,7,8,12–20] We obtained genomes from 15 representatives of each of these "orphan" classes by genome sequencing from both culture and metagenome-assembled genomes (MAGs). We augmented this sampling by targeting several other fungal lineages treated as "incertae sedis" or previously undersampled in genome-scale studies. In total, we generated 30 new genomes (Table S1; Data S1A and S1B).

Next, we analyzed the relationships of the newly obtained genomes by assembling a dataset that included at least one genome each from every genus of *Ascomycota* currently

(legend on next page)

represented in public databases (Data S1C), extracting 1,292 single-copy gene orthologs, using both concatenation and coalescent gene trees. Genomes from six of the seven sampled "orphan" classes—*Candelariomycetes*, *Coniocybomycetes*, *Geoglossomycetes*, *Lichinomycetes*, *Sareomycetes*, and *Xylonomycetes*—resolved in a single clade sibling to *Lecanoromycetes* and *Eurotiomycetes* (Figure 1). The newly recognized clade, which we will refer to by the oldest name contained within it, *Lichinomycetes*,[6] was highly supported in both analytical approaches we used (ultrafast bootstrap = 100; local posterior probability = 1). The only representative of a previously proposed class-level taxon that did not resolve in this clade is *Cirrosporium*—one of two genera of *Xylobotryomycetes*[10]—which we recovered as an early-diverging lineage of *Eurotiomycetes* (Figure 1). Of the additional 15 genomes from undersampled lineages, eight were also recovered in the *Lichinomycetes* clade, for a total of 22 lineages. The remaining seven were recovered with high support within the classes *Arthoniomycetes*, *Dothideomycetes*, *Lecanoromycetes*, and *Leotiomycetes*.

To test whether the recovered relationships were sensitive to phylogenetic reconstruction methods, we compared the results of maximum likelihood analyses based on concatenated gene sets and coalescence calculations based on 1,292 individual gene trees. We recovered no major discrepancies between the topology obtained from either concatenation or the species tree (Figure 2A; Data S2A and S2B). Of 959 edges, 43 were incongruent between these two reconstruction methods, almost all representing rearrangements among closely related taxa (Data S1H and S2C). Of the 43 incongruent edges, 6 were in *Lichinomycetes* (Figure 2B). The most notable of these concerns was the placement of *Vezdaea aestivalis*, which clustered with *V. acicularis* in the concatenated phylogeny but in the coalescent phylogeny was sibling to the clade including *Arthoniomycetes*, *Dothideomycetes*, *Lichinomycetes*, *Lecanoromycetes*, and *Eurotiomycetes* (AD+LLE). Consistent with this, some of the lowest gene concordance factors (gCFs) and site concordance factors (sCFs)[21] were found in branches located near *V. aestivalis* as well as *Symbiotaphrina buchneri* and the *Chaenotheca/Sclerophora* clade (Data S1I). The gCF of the *Lichinomycetes* clade was 2.48, indicating that 32 gene trees were concordant with the branch underlying both analyses. The sCF for the ancestral lichinomycete node was 41.1, one of the highest values for the major ascomycotan clades. Of the 1,292 genes, 1,151 (89%) support the concatenated topology based on phylogenetic signal[22] and an approximately unbiased (AU) test,[23] as well as a parsimonious version of the concatenated tree (663 genes, 51%) calculated based on three topologically constrained hypotheses of the relationship between *Lichinomycetes*, *Lecanoromycetes*, and *Eurotiomycetes* (L+LE; Figures 2C and 2D; Data S1J and S1K).

Lichinomycete fungi are disproportionately symbionts

The lichinomycete clade includes a disproportionate number of fungal symbionts with diverse morphological outcomes (Figure 3). The largest lifestyle group represented by our sampling is lichen symbionts, accounting for about 75% of estimated species richness, followed by putative mycorrhizae, endophytes, and yeast-like insect symbionts. Altogether, the 22 lineages are representatives of higher-ranked taxa that include over 600 species (Data S1U).

Lichinomycete fungi have small genomes

To test for shared traits among members of this clade, we annotated and compared 29 of our new genomes to a selected subset of 86 genomes, representing all classes and lifestyles in *Pezizomycotina*. *Lichinomycetes* have smaller genomes (25.1 Mb; sd = 4.7 Mb) than any of the major clades of *Pezizomycotina* (~27–80 Mb). This reduction comes with a major decrease in average gene numbers (7,749 genes; sd = 1,035.2) compared to other classes of *Pezizomycotina* (Table S2). Gene functional comparisons demonstrate conspicuous differences in carbohydrate degradation, secondary metabolite gene clusters, and protein family-based functional annotations between *Lichinomycetes* and other major ascomycotan groups (Figure 1). For carbohydrate-active enzymes (CAZymes), *Lichinomycetes* exhibit some of the smallest inventories in all compared *Ascomycota*, with reduced numbers in 4 out of 5 major CAZyme classes and near-absence of polysaccharide lyases. The number of CAZymes was 44% lower than in the *Lecanoromycetes-Eurotiomycetes* (LE) clade, largely accounted for by a functional loss in the putative ability to degrade cellulose and pectin. In the case of genes associated with lignin degradation, the repertoire resembles that of early divergent lineages such as *Orbiliomycetes* and *Pezizomycetes* (Figure S1).

The reduction in gene numbers is echoed in secondary metabolite biosynthetic gene clusters (BGCs). The lichinomycete clade includes far fewer type-I polyketide synthases and non-ribosomal peptide synthetases than, for example, *Lecanoromycetes* and relative reductions in most other BGC subcategories compared to other *Ascomycota* (Figure 1; Data S2F). The low number of BGCs in *Lichinomycetes* resembles the early divergent lineages *Pezizomycetes* and *Orbiliomycetes* (Figures 4C and 4D); all other classes have undergone BGC expansions. *Lecanoromycetes* and *Eurotiomycetes* possess the most BGCs, followed by *Sordariomycetes* (Table S2).

Finally, *Lichinomycetes* have the lowest average annotated Pfam number per genome in *Pezizomycotina* (8,644 Pfam annotations; sd = 1,088.9; comparisons in Table S2). By extension, their Pfam compositional profiles diverge from those of other *Ascomycota*, notably also their closest relatives,

Figure 1. Position of newly sampled early-diverging lineages in a maximum likelihood reconstruction of 115 *Ascomycota* genomes based on 1,292 concatenated loci

Dating based on the program LSD2 with six fossils, showing confidence interval bars (see Star Methods). Newly generated genomes are indicated by bold text; underlying highlighting indicates ascomycotan classes in which these lineages have been placed until now. Single asterisk = lineages classified in "orphan" classes; double asterisk = additionally sampled putatively early-diverging lineages. Genome size, GC content, number of genes, tRNAs, CAZy classes, Pfams, and BGCs based on *de novo* annotations of all genomes. Voucher data on newly sequenced genomes are in Table S1; expanded background data on genome assemblies and annotations are found in Table S2 and Data S1L and S1M and on CAZyme and BGC annotations in Data S1O and S1P. An underlying tree with 481 taxa is in Data S2B and S2D, an expanded CAZyme heat map with node dating is in Figure S1, and underlying CAZyme and BGC data are in Data S2E and S2F.

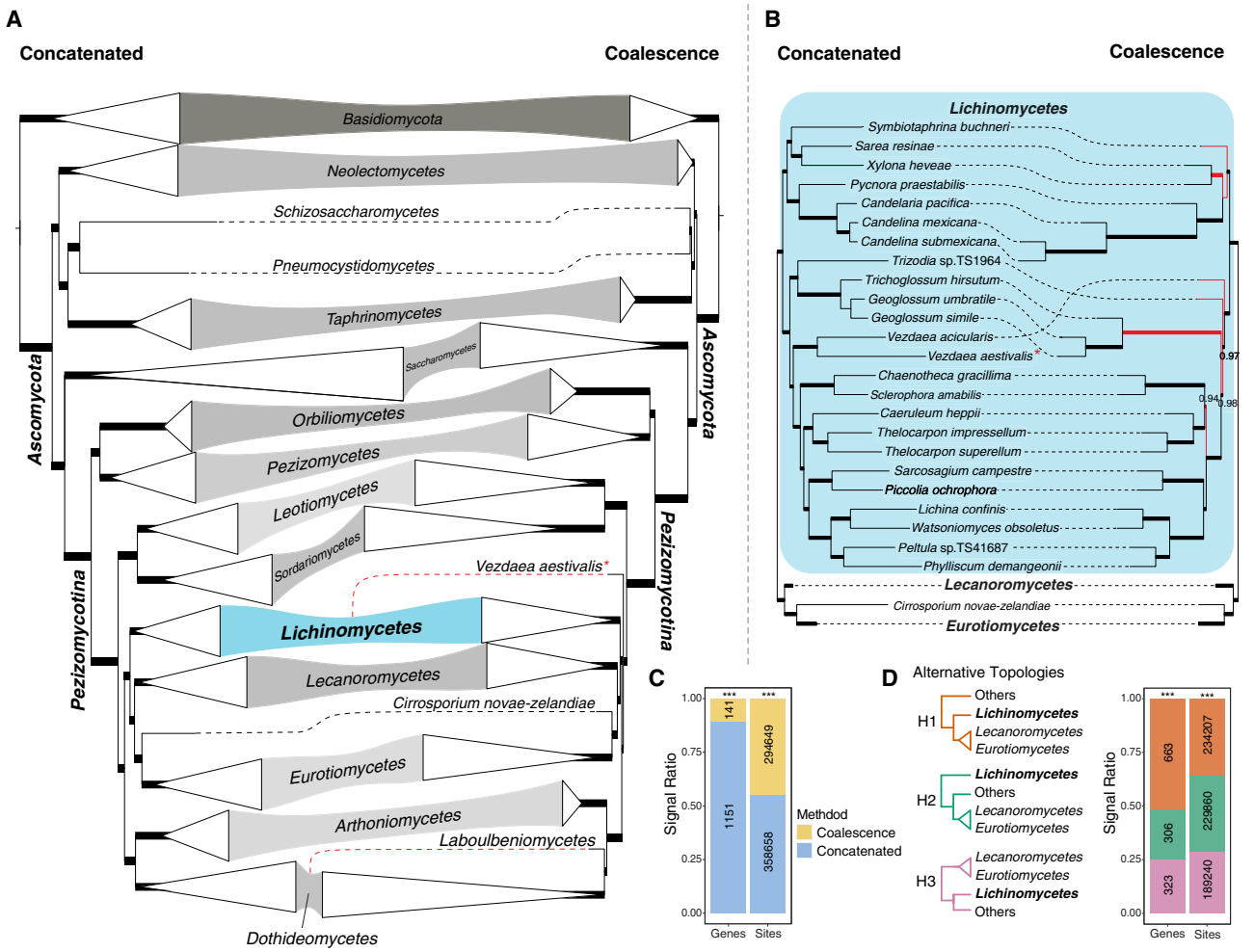

Figure 2. Stability of phylogenomic topologies using analyses with concatenated alignments and gene tree coalescence

(A) Congruence of major class-level clades across 481 genomes including one sample for every available genus of *Ascomycota* and representatives of *Basidiomycota* (compared trees are in Data S2A, S2B, and S2C; see Star Methods).

(B) Congruence of sampled lineages within *Lichinomycetes* based on the same two methods. Red lines indicate edges incongruent between the two methods. A complete comparison of edges is in Data S1H.

(C) Gene and site phylogenetic signal for the two methodological approaches. Related to Data S1K.

(D) Phylogenetic signal of three topological hypotheses of the relationship of *Lichinomycetes*, *Lecanoromycetes*, and *Eurotiomycetes*. Statistical differences from G test (with ratios 1:1 and 1:1:1 respectively) is expressed as *** $p < 0.001$. Related to Data S1K.

Lecanoromycetes and *Eurotiomycetes* (Figures 4E and 4F). We also identified qualitative and quantitative Pfam occurrence differences compared to other groups of *Ascomycota* using two additional analyses (Data S1Q). In the first, we set out to identify Pfams with fidelity to individual clades. We constructed a correlation matrix consisting of Pfam statistical association with 10 *a priori* defined phylogenetic clades and calculated fidelity scores for each Pfam. We recovered sets of high-fidelity Pfams for every clade except *Lichinomycetes* (plotted in Data S2G). In the second analysis, we tested all 5,385 annotated Pfams for differences in gene numbers between *Lichinomycetes* and non-*Lichinomycetes* (Wilcoxon test). Of the 881 Pfams with significant differences between the two groups ($p < 0.01$), a total of 43 were statistically overrepresented in *Lichinomycetes*, but none were unique (Data S1R). Finally, since *Lichinomycetes* lack representation in databases, we tested if the low Pfam numbers were a result of annotation bias. We found no differences (Kruskal-Wallis test: $p > 0.05$) in the ratio of successfully predicted proteins used in annotations between *Lichinomycetes* and non-*Lichinomycetes*.

All three of the main annotation groups we analyzed—CAZymes, BGCs, and Pfams—exhibit distinct trends for genomes of *Lichinomycetes* compared to the other sampled *Ascomycota* genomes. In principal components analyses (PCAs) of CAZymes and Pfams both without phylogenetic correction and with phylogenetic correction (pPCA), *Lichinomycetes* form clusters distinct from *Lecanoromycetes*, the largest class of lichen fungal symbionts, as well as its sibling group, *Eurotiomycetes*. The pPCA results based on BGCs, which are anomalously underrepresented in *Lichinomycetes*, exhibit a conspicuous

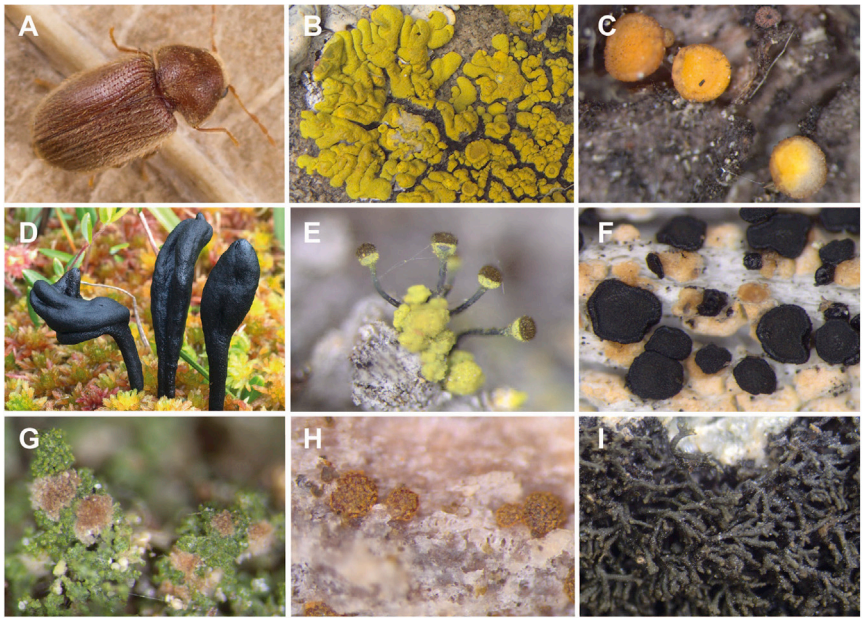

Figure 3. Lifestyle diversity of *Lichinomycetes*

(A) *Stegobium paniceum* beetle host of *Symbiotaphrinabuchneri* (photo by Nikolai Vladimirov).
(B) *Candelina submexicana*, a lichen from rocks in Mexico.
(C) *Sarea resinae*, a fungus on conifer resin in boreal forests.
(D) *Geoglossum glabrum*, an earth-tongue of *Sphagnum* bogs (photo by Jozef Pavlík).
(E) *Chaenotheca chrysocephala*, a "pin lichen" of ancient forests.
(F) *Pycnora praestabilis*, a lichen on wood in the Alps.
(G) *Vezdaea aestivalis*, a lichen on plant detritus in the Rocky Mountains.
(H) *Piccolia ochrophora*, a lichen on poplar bark in British Columbia, Canada.
(I) *Lichina confinis*, a lichen on seashore rocks in western Scotland.
Photos by the authors except where indicated.

lichinomycete cluster, close to *Orbiliomycetes* and *Pezizomycetes* (Figure 4), though overlapping with the some of the overspread classes.

Ancestral node reconstruction

Since the effect of the *Lichinomycetes* on the timing of evolutionary events and ancestral states is unexplored, we reconstructed ancestral CAZyme machinery for multiple deep nodes as a proxy for lifestyle. We recovered a hypothetical ancestor of *Lichinomycetes* diverging 305 million years ago (mya) (CI 340:272), with a putatively already reduced ability to degrade cellulose, lignin, and pectin (Figures 4A and S1). The other ancestral reconstructions showed a stepwise reduction in CAZyme numbers from the most recent common ancestor (MRCA) of AD+LLE (396 mya; CI 437:359) to LLE (344 mya; CI 365:285) to *Lichinomycetes*. In other clades (e.g., AD and LE), later expansion events occurred, with the largest occurring in *Dothideomycetes*. The numbers of CAZymes recovered in this group that putatively act on plant cell walls (cellulose, lignin, or pectin) are the highest for any lichen fungal symbiont to date.[24]

A morphologically hyperdiverse group with a common origin

Various combinations of taxon and locus sampling have captured patterns that foreshadowed the evolutionary relationships we report here. The L+LE topology was first recovered by Beimforde et al.[16] using representatives of three of the seven lineages we recovered as *Lichinomycetes*. Hashimoto et al.[25] recovered all seven lineages in the topology we recover here but focused on the position of *Xylonomycetes* without exploring the implications of a monophyletic group sibling to LE. In the only two genome-scale studies to include relevant taxa, Shen et al.[5] and Li et al.[26] likewise recovered *Xylona* and *Symbiotaphrina* (*Xylonomycetes*) as siblings to LE. In some studies, two or three lineages were recovered as monophyletic but part of a different topology,[3,11,19] while in others, *Lichinomycetes* were polyphyletic.[4,7–10,12,13,15,18–20,27–34]

Given that lineages recovered here as part of the lichinomycete clade have been widely sampled, how did it go overlooked for so long? Both the morphology and evolutionary history of the lineages may have played a role. Historically, morphology-based systematics has not anticipated or predicted this grouping, likely owing to the lack of shared morphological traits. Another possible reason is that extant *Lichinomycetes* are remnants of a clade that experienced more extinctions than other fungi, leaving branching points between extant lineages too deep to resolve with shallow sampling. Testing this using, for example, speciation-extinction modeling is not trivial. Such models are sensitive to assumptions concerning sampling completeness, about which we lack certainty.

How, then, can we be sure that the lichinomycete clade is stable and reflects a single origin? We addressed this in two ways. First, we stress-tested the 481-genome phylogenomic analysis using analytical approaches that generate topologies based on data concatenation and gene trees, respectively, echoing approaches recently applied to large-scale fungal phylogenies.[5,26] We also tested for the gene-level source of signal. Only one lineage, *Vezdaea aestivalis*, is not supported as part of a single origin under both of these approaches (Figure 2A). Its behavior in the species tree would seem to lend new evidence to a decades-old hypothesis, based on ascus and ascomatal characters, that *Vezdaea* may be "primitive".[35] Second, we annotated all lichinomycete genomes and analyzed key features compared to a subset of *Pezizomycotina* genomes. This analysis revealed an unexpected degree of uniformity in lichinomycete genomes, notwithstanding their morphological and lifestyle diversity. In analyses of the *Pezizomycotina* subset, *Lichinomycetes* have, on average, some of the smallest genomes, the lowest number of annotated Pfams, the lowest number of CAZymes, and the smallest secondary metabolite arsenal. Only *Herpomyces*—if it is ultimately maintained at class level (*Laboulbeniomycetes*)—would have a smaller

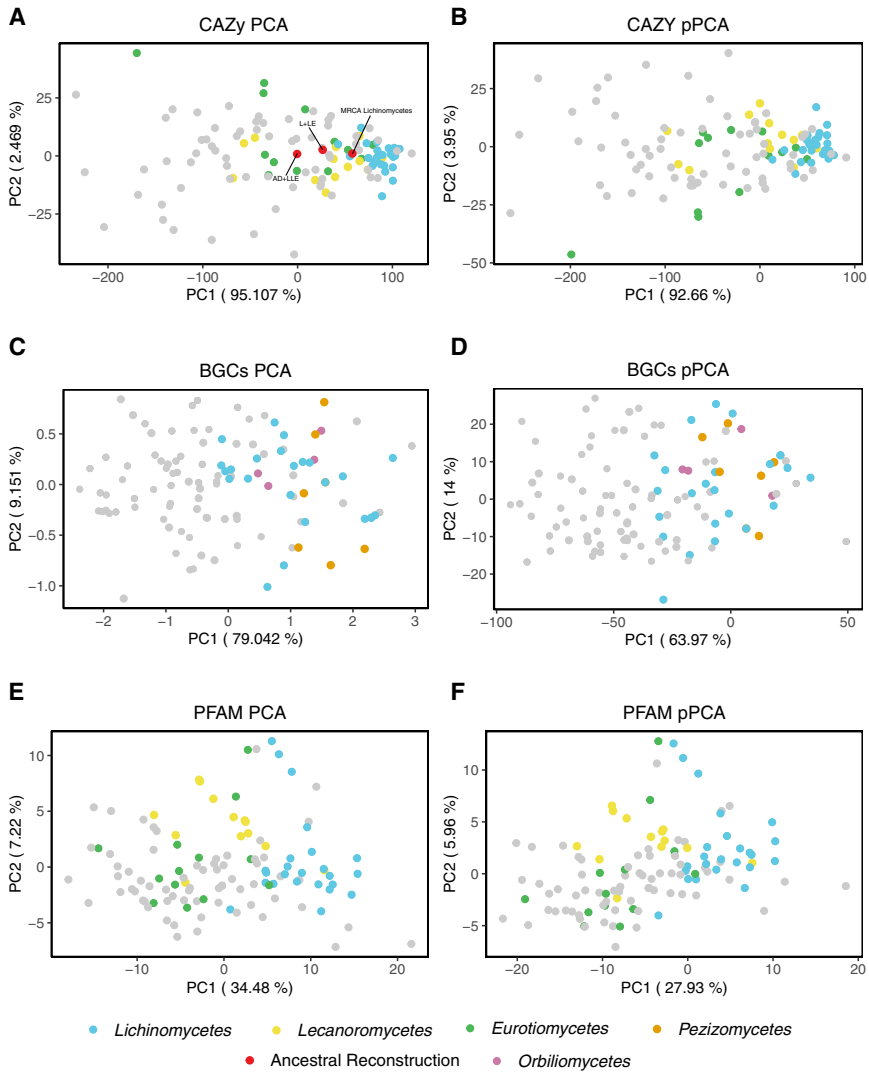

Figure 4. PCAs based on the 115 genomes used in Figure 1, both with and without phylogenetic correction

(A and B) Data based on all major CAZyme classes and carbohydrate-binding modules. Ancestral reconstructed nodes are shown as red dots (AD+LLE, L+LE, MRCA *Lichinomycetes*; see text). *Eurotiomycetes* and *Lecanoromycetes* are colored as green and yellow, respectively, to highlight differences between those two classes and *Lichinomycetes* (blue); genomes from all other classes shown in grey. Related to Figure S1 and Data S1O.

(C and D) Data based on BGCs, analyzed as four groups (the three most abundant types: NRPS, NRPS-like, type-1 PKS, plus all others as one category). *Orbiliomycetes* and *Pezizomycetes* are shown in pink and orange, respectively, to highlight clustering trends compared to *Lichinomycetes* (in blue); genomes from all other classes shown in grey. Related to Data S1P and S2F.

(E and F) PCAs based on Pfams. Only Pfams were included which had a standard deviation >1 based on raw per genome counts averaged across all annotated genomes. *Eurotiomycetes* and *Lecanoromycetes* are colored as green and yellow, respectively, to highlight differences between those two classes and *Lichinomycetes* (blue); genomes from all other classes shown in grey.

arsenal, but how representative this genome is of its class requires sampling beyond the scope of this study. While future gene characterization will likely identify lichinomycete-specific genes and gene expansions, current data are consistent with a pattern of genome size reduction at the origin of the lichinomycete clade.

Despite the morphological disparities in *Lichinomycetes*, it may be premature to rule out the existence of shared traits. Although this trait is widespread in *Ascomycota*, *Lichinomycetes* have an unusual concentration of species with polysporous asci. By targeting unclassified lineages with this trait, we found that several lineages previously treated as *incertae sedis* in *Ascomycota* (*Caeruleum*, *Piccolia*, *Sarcosagium*, and *Thelocarpon*) resolved inside the lichinomycete clade. Traits such as cell wall composition, which diverges from that of *Lecanoromycetes*,[36] also deserve further study.

Genome annotations challenge lifestyle assumptions

Our annotations challenge assumptions about the lifestyles of several of the included fungal groups. First, lichen symbionts in *Lichinomycetes* deviate in multiple ways from others in *Ascomycota*. Lichen symbionts are often noted for their diversity of secondary metabolites, numbering over 1,000 named substances,[37] about 90% of them unique to lichens.[38] Lichinomycete lichen symbionts, however, have far smaller predicted BGC arsenals than lichen symbionts in *Lecanoromycetes*, *Arthoniomycetes*, *Dothideomycetes*, and *Eurotiomycetes*. In fact, lichinomycete BGC numbers are similar to those of the ancestral clades *Pezizomycetes* and *Orbiliomycetes*. The low BGC numbers correspond to the paucity of reported secondary metabolites in lichinomycete lichens (about 16 named substances, mostly derived from the shikimic acid pathway; Data S1T). Lichinomycete lichen symbionts also possess smaller CAZyme inventories than those in *Lecanoromycetes*[24] and far fewer than lichen symbionts in *Dothideomycetes*; the CAZyme toolkits in the latter group are revealed in this study to be the largest found to date in lichen symbionts, consistent with past predictions that they are facultative saprotrophs.[39] The reduced CAZyme machinery in some lichen fungi, but not others, underscores questions about whether transferred photosymbiont sugar alcohols serve the same function in all lichen systems.[40]

Second, our genome annotations do not appear to be consistent with the assumption that earthtongues (*Geoglossales*) are saprotrophs.[41,42] The recovered CAZyme inventories in all three *Geoglossales* genomes are much smaller than those of well-known saprotrophs; they possess the lowest number of predicted GH5s, active on cellulose, in *Lichinomycetes*, fewer than in any lichen symbiont. They also possess no GH12s, active

on cellulose; no GH43s, active on hemicellulose; and no CAZymes active on pectin (Figure S1). Instead, *Geoglossales* CAZyme toolkits resemble those of mycorrhizal mutualists with reduced saprotrophic machinery.[43–45] The observed CAZyme numbers are, however, consistent with suggestions that some *Geoglossales* form mycorrhizae, e.g., with bryophytes[46] or ericoid shrubs.[47–49] The suggestion of symbiosis with bryophytes, in turn, is coincidental with our recovery of *Geoglossales* as sibling to *Trizodia*, a genus originally described as a ''bryosymbiont''.[50] The remainder of lineages we recover in *Lichinomycetes* also appear to be symbionts. *Symbiotaphrina* is a yeast-like symbiont of anobiid beetles,[51] and *Xylona* is an endophyte.[8] *Sarea* fruits on oozing tree resin, but its isolation from living heartwood and sapwood[52] suggests it, too, may be an endophyte. Whether these lineages engage in mutualistic relationships deserves further study.

Perhaps the most striking feature of *Lichinomycetes* is their small genomes and reduced gene counts. Though the uniformly low diversity of CAZymes and BGCs was unexpected given the divergent morphologies and ecologies of these fungi, decreased metabolic flexibility is consistent with the lack of confirmed pathogens or saprotrophs in *Lichinomycetes*. The similarity of genome size and gene number across the annotated genomes is strong evidence that their MRCA already possessed a reduced genome. Both the limited lifestyle diversity in the group and the reconstructed age of its MRCA raise the intriguing possibility that this is one of the earliest extant lineages of symbionts in *Ascomycota*.

The persistently low metabolic diversity over 300 million years of evolution in *Lichinomycetes* stands out against the background of *Ascomycota* that diverged at similar times. Both CAZymes and BGCs exhibit high variability in most fungal groups over evolutionary timescales, likely as a consequence of the ease with which they are shared via various mechanisms, including horizontal gene transfer (HGT) mediated by transposable elements.[53,54] One possibility is that one of the mechanisms associated with acquisition of novel DNA was affected in the lichinomycete MRCA by the same event that led to the reduction in genome size, constraining the ability of its progeny to reacquire metabolic flexibility. Vertically transmitted bacterial symbionts also have small genomes with reduced coding capacity and can be structurally stable for hundreds of millions of years.[55] Any constraint on acquisition of novelty in fungi is likely to involve a different mechanism, if it exists at all. Future exploration of novel DNA acquisition in *Lichinomycetes* relative to other fungi, including the propensity of HGT and transposable elements, will help answer this question.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - DNA sampling
  - Sequencing
  - Genome assembly
  - Sample identification
  - Phylogenomics
  - Time calibration
  - Functional genomic annotations
  - Ancestral node reconstruction
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Annotation statistics
  - Principal components analyses of functional annotations
  - High fidelity pfam identification
  - Pfam statistical analyses
  - Assessment of species and metabolite diversity

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.11.014.

ACKNOWLEDGMENTS

We thank Sophie Dang (University of Alberta) for technical help in the lab. Viktor Kučera provided helpful input on the ecology of *Geoglossales*, and Jozef Pavlík and Nikolai Vladimirov kindly allowed us to use their photos. Many thanks to Emile Gluck-Thaler and John McCutcheon for their input on genome size reduction in *Lichinomycetes*. This project would not have been possible without a computing resources allocation from the Digital Research Alliance of Canada. D.D.E. acknowledges the Colombian Ministry of Science (Minciencias) Doctoral Studies abroad grant 860 of 2019 for funding. The work of K.K. and J.K. was financially supported by the grant of the Ministry of Education, Youth and Sports of the Czech Republic, INTEREXCELLENCE, INTERACTION, no. LTAUSA18188. T.S. acknowledges an NSERC Discovery Grant and the Canada Research Chairs Program for funding.

AUTHOR CONTRIBUTIONS

T.S., P.R., and D.D.E. designed the study; D.L.H. and R.L. advised on taxon sampling and nomenclature; D.D.E. and P.R. carried out all analyses; D.V., C.C.G.A., A.A., O.Č., A.H., K.K., and J.K. provided material of hard-to-find fungal species; G.T. helped with analyses; D.D.E. and P.R. curated and deposited data and scripts; D.D.E. and T.S. wrote the manuscript; and all authors contributed to editing and improving the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 11, 2022
Revised: August 30, 2022
Accepted: November 7, 2022
Published: November 23, 2022

REFERENCES

1. Kirk, P.M., Cannon, P.F., Minter, D.W., and Stalpers, J.A. (2008). Dictionary of the Fungi, 10th ed. (CAB International).
2. Spatafora, J.W., Aime, M.C., Grigoriev, I.V., Martin, F., Stajich, J.E., and Blackwell, M. (2017). The fungal tree of life: from molecular systematics to genome-scale phylogenies. Microbiol. Spectr. 5. FUNK-0053-2016.
3. James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., et al. (2006).

Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature *443*, 818–822.

4. Schoch, C.L., Sung, G.-H., López-Giráldez, F., Townsend, J.P., Miadlikowska, J., Hofstetter, V., Robbertse, B., Matheny, P.B., 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*, 224–239.
5. Shen, X.X., Steenwyk, J.L., LaBella, A.L., Opulente, D.A., Zhou, X., Kominek, J., Li, Y., Groenewald, M., Hittinger, C.T., and Rokas, A. (2020). Genome-scale phylogeny and contrasting modes of genome evolution in the fungal phylum *Ascomycota*. Sci. Adv. *6*, eabd0079.
6. Reeb, V., Lutzoni, F., and Roux, C. (2004). Contribution of RPB2 to multilocus phylogenetic studies of the euascomycetes (*Pezizomycotina, Fungi*) with special emphasis on the lichen-forming *Acarosporaceae* and evolution of polyspory. Mol. Phylogenet. Evol. *32*, 1036–1060.
7. Schoch, C.L., Wang, Z., Townsend, J.P., and Spatafora, J.W. (2009). *Geoglossomycetes* cl. nov., *Geoglossales* ord. nov. and taxa above class rank in the *Ascomycota* Tree of Life. Persoonia *22*, 129–138.
8. Gazis, R., Miadlikowska, J., Lutzoni, F., Arnold, A.E., and Chaverri, P. (2012). Culture-based study of endophytes associated with rubber trees in Peru reveals a new class of *Pezizomycotina: Xylonomycetes*. Mol. Phylogenet. Evol. *65*, 294–304.
9. Prieto, M., and Wedin, M. (2013). Dating the diversification of the major lineages of *Ascomycota* (*Fungi*). PLoS One *8*, e65576.
10. Voglmayr, H., Fournier, J., and Jaklitsch, W.M. (2019). Two new classes of *Ascomycota: Xylobotryomycetes* and *Candelariomycetes*. Persoonia *42*, 36–49.
11. Beimforde, C., Schmidt, A.R., Rikkinen, J., and Mitchell, J.K. (2020). *Sareomycetes* cl. nov.: A new proposal for placement of the resinicolous genus *Sarea* (*Ascomycota, Pezizomycotina*). Fungal Syst. Evol. *5*, 197–282.
12. Miadlikowska, J., Kauff, F., Hofstetter, V., Fraker, E., Grube, M., Hafellner, J., Reeb, V., Hodkinson, B.P., Kukwa, M., Lücking, R., et al. (2006). New insights into classification and evolution of the *Lecanoromycetes* (*Pezizomycotina, Ascomycota*) from phylogenetic analyses of three ribosomal RNA- and two protein-coding genes. Mycologia *98*, 1088–1103.
13. Hofstetter, V., Miadlikowska, J., Kauff, F., and Lutzoni, F. (2007). Phylogenetic comparison of protein-coding versus ribosomal RNA-coding sequence data: A case study of the *Lecanoromycetes* (*Ascomycota*). Mol. Phylogenet. Evol. *44*, 412–426.
14. Arnold, A.E., Miadlikowska, J., Higgins, K.L., Sarvate, S.D., Gugger, P., Way, A., Hofstetter, V., Kauff, F., and Lutzoni, F. (2009). A phylogenetic estimation of trophic transition networks for ascomycetous Fungi: Are lichens cradles of symbiotrophic Fungal diversification? Syst. Biol. *58*, 283–297.
15. Schmull, M., Miadlikowska, J., Pelzer, M., Stocker-Wörgötter, E., Hofstetter, V., Fraker, E., Hodkinson, B.P., Reeb, V., Kukwa, M., Lumbsch, H.T., et al. (2011). Phylogenetic affiliations of members of the heterogeneous lichen-forming fungi of the genus *Lecidea* sensu Zahlbruckner (*Lecanoromycetes, Ascomycota*). Mycologia *103*, 983–1003.
16. Beimforde, C., Feldberg, K., Nylinder, S., Rikkinen, J., Tuovila, H., Dörfelt, H., Gube, M., Jackson, D.J., Reitner, J., Seyfullah, L.J., and Schmidt, A.R. (2014). Estimating the Phanerozoic history of the Ascomycota lineages: Combining fossil and molecular data. Mol. Phylogenet. Evol. *78*, 386–398.
17. Miadlikowska, J., Kauff, F., Högnabba, F., Oliver, J.C., Molnár, K., Fraker, E., Gaya, E., Hafellner, J., Hofstetter, V., Gueidan, C., et al. (2014). A multigene phylogenetic synthesis for the class *Lecanoromycetes* (*Ascomycota*): 1307 fungi representing 1139 infrageneric taxa, 317 genera and 66 families. Mol. Phylogenet. Evol. *79*, 132–168.
18. Carbone, I., White, J.B., Miadlikowska, J., Arnold, A.E., Miller, M.A., Kauff, F., U'Ren, J.M., May, G., and Lutzoni, F. (2017). T-BAS: tree-based alignment selector toolkit for phylogenetic-based placement, alignment downloads and metadata visualization: An example with the Pezizomycotina tree of life. Bioinformatics *33*, 1160–1168.
19. Lutzoni, F., Nowak, M.D., Alfaro, M.E., Reeb, V., Miadlikowska, J., Krug, M., Arnold, A.E., Lewis, L.A., Swofford, D.L., Hibbett, D., et al. (2018). Contemporaneous radiations of fungi and plants linked to symbiosis. Nat. Commun. *9*, 5451.
20. Nelsen, M.P., Lücking, R., Boyce, C.K., Lumbsch, H.T., and Ree, R.H. (2020). No support for the emergence of lichens prior to the evolution of vascular plants. Geobiology *18*, 3–13.
21. Minh, B.Q., Hahn, M.W., and Lanfear, R. (2020). New methods to calculate concordance factors for phylogenomic datasets. Mol. Biol. Evol. *37*, 2727–2733.
22. Shen, X.-X., Hittinger, C.T., and Rokas, A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. Nat. Ecol. Evol. *1*, 0126.
23. Shimodaira, H. (2002). An approximately unbiased test of phylogenetic tree selection. Syst. Biol. *51*, 492–508.
24. Resl, P., Bujold, A.R., Tagirdzhanova, G., Meidl, P., Freire Rallo, S., Kono, M., Fernández-Brime, S., Guðmundsson, H., Andrésson, Ó.S., Muggia, L., et al. (2022). Large differences in carbohydrate degradation and transport potential among lichen fungal symbionts. Nat. Commun. *13*, 2634.
25. Hashimoto, A., Masumoto, H., Endoh, R., Degawa, Y., and Ohkuma, M. (2021). Revision of *Xylonaceae* (*Xylonales, Xylonomycetes*) to include *Sarea* and *Tromera*. Mycoscience *62*, 47–63.
26. Li, Y., Steenwyk, J.L., Chang, Y., Wang, Y., James, T.Y., Stajich, J.E., Spatafora, J.W., Groenewald, M., Dunn, C.W., Hittinger, C.T., et al. (2021). A genome-scale phylogeny of the kingdom Fungi. Curr. Biol. *31*, 1653–1665.e5. e5.
27. Schultz, M., Arendholz, W.R., and Büdel, B. (2001). Origin and evolution of the lichenized ascomycete order *Lichinales*: Monophyly and systematic relationships inferred from ascus, fruiting body and SSU rDNA evolution. Plant Biol. (Stuttg). *3*, 116–123.
28. Kauff, F., and Lutzoni, F. (2002). Phylogeny of the *Gyalectales* and *Ostropales* (*Ascomycota, Fungi*): among and within order relationships based on nuclear ribosomal RNA small and large subunits. Mol. Phylogenet. Evol. *25*, 138–156.
29. Spatafora, J.W., 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*, 1018–1028.
30. Chen, K.H., Miadlikowska, J., Molnár, K., Arnold, A.E., U'Ren, J.M., Gaya, E., Gueidan, C., and Lutzoni, F. (2015). Phylogenetic analyses of eurotiomycetous endophytes reveal their close affinities to *Chaetothyriales, Eurotiales*, and a new order – *Phaeomoniellales*. Mol. Phylogenet. Evol. *85*, 117–130.
31. Gazis, R., Kuo, A., Riley, R., LaButti, K., Lipzen, A., Lin, J., Amirebrahimi, M., Hesse, C.N., Spatafora, J.W., Henrissat, B., et al. (2016). The genome of *Xylona heveae* provides a window into fungal endophytism. Fungal Biol. *120*, 26–42.
32. Prieto, M., Schultz, M., Olariaga, I., and Wedin, M. (2019). *Lichinodium* is a new lichenized lineage in the *Leotiomycetes*. Fungal Divers. *94*, 23–39.
33. Díaz-Escandón, D., Hawksworth, D.L., Powell, M., Resl, P., and Spribille, T. (2021). The British chalk specialist *Lecidea lichenicola* auct. revealed as a new genus of *Lichinomycetes*. Fungal Biol. *125*, 495–504.
34. Mitchell, J.K., Garrido-Benavent, I., Quijada, L., and Pfister, D.H. (2021). Sareomycetes: more diverse than meets the eye. IMA Fungus *12*, 6.
35. Tschermak-Woess, E., and Poelt, J. (1976). *Vezdaea*, a peculiar lichen genus, and its phycobiont. In Lichenology: Progress and Problems, D.H. Brown, D.L. Hawksworth, and R.H. Bailey, eds. (Academic Press), pp. 89–105.
36. Leal, J.A., Prieto, A., Bernabé, M., and Hawksworth, D.L. (2010). An assessment of fungal wall heteromannans as a phylogenetically informative character in ascomycetes. FEMS Microbiol. Rev. *34*, 986–1014.

37. Molnár, K., and Farkas, E. (2010). Current results on biological activities of lichen secondary metabolites: A review. Zeitschrift für. Naturforschung C *65*, 157–173.

38. Honegger, R. (2012). 15 The Symbiotic Phenotype of Lichen-Forming Ascomycetes and Their Endo- and Epibionts. In The Mycota. B, Hock., ed. (Springer Verlag GmbH), pp. 287–339.

39. Johnson, G.T. (1940). Contributions to the study of the *Trypetheliaceae*. Ann. Mo. Bot. Gard. *27*, 1.

40. Spribille, T., Resl, P., Stanton, D.E., and Tagirdzhanova, G. (2022). Evolutionary biology of lichen symbioses. New Phytol. *234*, 1566–1582.

41. Wang, Z., Binder, M., Schoch, C.L., Johnston, P.R., Spatafora, J.W., and Hibbett, D.S. (2006). Evolution of helotialean fungi (*Leotiomycetes, Pezizomycotina*): A nuclear rDNA phylogeny. Mol. Phylogenet. Evol. *41*, 295–312.

42. Naranjo-Ortiz, M.A., and Gabaldón, T. (2019). Fungal evolution: diversity, taxonomy and phylogeny of the Fungi. Biol. Rev. Camb. Philos. Soc. *94*, 2101–2137.

43. Kohler, A., Kuo, A., Nagy, L.G., Morin, E., Barry, K.W., 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*, 410–415.

44. Murat, C., Payen, T., Noel, B., Kuo, A., Morin, E., Chen, J., Kohler, A., Krizsán, K., Balestrini, R., Da Silva, C., et al. (2018). Pezizomycetes genomes reveal the molecular basis of ectomycorrhizal truffle lifestyle. Nat. Ecol. Evol. *2*, 1956–1965.

45. Miyauchi, S., Kiss, E., Kuo, A., Drula, E., Kohler, A., Sánchez-García, M., Morin, E., Andreopoulos, B., Barry, K.W., Bonito, G., et al. (2020). Large-scale genome sequencing of mycorrhizal fungi provides insights into the early evolution of symbiotic traits. Nat. Commun. *11*, 5125.

46. Ohenoja, E., Wang, Z., Townsend, J.P., Mitchel, D., and Voitk, A. (2010). Northern species of earth tongue genus *Thuemenidium* revisited, considering morphology, ecology and molecular phylogeny. Mycologia *102*, 1089–1095.

47. Nitare, J. (1982). *Geoglossum arenarium*, sandjordtunga – ekologi och utbredning i Sverige. Svensk Botanisk Tidskrift *76*, 349–357.

48. Baba, T., Janošík, L., Koukol, O., and Hirose, D. (2021). Genetic variations and *in vitro* root-colonizing ability for an ericaceous host in *Sarcoleotia globosa (Geoglossomycetes)*. Fungal Biol. *125*, 971–979.

49. Kučera, V., Slovák, M., Janošík, L., and Beenken, L. (2021). A new species of *Sabuloglossum* (*Geoglossaceae, Ascomycota*) from montane areas. Plant Biosystems, (in press).

50. Stenroos, S., Laukka, T., Huhtinen, S., Döbbeler, P., Myllys, L., Syrjänen, K., and Hyvönen, J. (2010). Multiple origins of symbioses between ascomycetes and bryophytes suggested by a five-gene phylogeny. Cladistics. *26*, 281–300.

51. Noda, H., and Kodama, K. (1996). Phylogenetic position of yeastlike endosymbionts of Anobiid beetles. Appl. Environ. Microbiol. *62*, 162–167.

52. Robinson-Jeffrey, R.C., and Loman, A.A. (1963). Fungi isolated in culture from red heartwood stain and advanced decay of lodgepole pine in Alberta. Can. J. Bot. *41*, 1371–1375.

53. McDonald, M.C., Taranto, A.P., Hill, E., Schwessinger, B., Liu, Z., Simpfendorfer, S., Milgate, A., and Solomon, P.S. (2019). Transposon-mediated horizontal transfer of the host-specific virulence protein ToxA between three fungal wheat pathogens. MBio *10*, e01515-19–e01519.

54. Qiu, H., Cai, G., Luo, J., Bhattacharya, D., and Zhang, N. (2016). Extensive horizontal gene transfers between plant pathogenic fungi. BMC Biol. *14*, 41.

55. McCutcheon, J.P., Boyd, B.M., and Dale, C. (2019). The life of an insect endosymbiont from the cradle to the grave. Curr. Biol. *29*, R485–R495.

56. Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data.

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

58. Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017). MetaSPAdes: A new versatile metagenomic assembler. Genome Res. *27*, 824–834.

59. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

60. Alneberg, J., Bjarnason, B.S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson, A.F., Quince, C., and Quince, C. (2014). Binning metagenomic contigs by coverage and composition. Nat. Methods *11*, 1144–1146.

61. Laetsch, D.R., and Blaxter, M.L. (2017). BlobTools: Interrogation of genome assemblies. F1000Res. *6*, 1287.

62. Saary, P., Mitchell, A.L., and Finn, R.D. (2020). Estimating the quality of eukaryotic genomes recovered from metagenomic analysis with EukCC. Genome Biol. *21*, 244.

63. Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics *31*, 3210–3212.

64. Waterhouse, R.M., Seppey, M., Simão, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E.V., and Zdobnov, E.M. (2018). BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol. Biol. Evol. *35*, 543–548.

65. Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A., and Zdobnov, E.M. (2021). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Mol. Biol. Evol. *38*, 4647–4654.

66. Uritskiy, G.V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome *6*, 158.

67. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods *14*, 417–419.

68. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map Format and SAMtools. Bioinformatics *25*, 2078–2079.

69. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009). BLAST+: architecture and applications. BMC Bioinformatics *10*, 421.

70. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 1530–1534.

71. Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol. Biol. Evol. *30*, 772–780.

72. Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973.

73. Zhang, C., Rabiee, M., Sayyari, E., and Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics *19*, 153.

74. Lemoine, F., and Gascuel, O. (2021). Gotree/Goalign: toolkit and Go API to facilitate the development of phylogenetic workflows. NAR genom. bioinform. *3*, lqab075.

75. Palmer, J.M., and Stajich, J. (2020). Funannotate v1.8.1: Eukaryotic genome annotation (Zenodo). https://doi.org/10.5281/zenodo.4054262.

76. Smit, A.F.A., Hubley, R., and Green, P. (2013–2015). RepeatMasker Open-4.0. http://www.repeatmasker.org.

77. Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y.O., and Borodovsky, M. (2008). Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res. *18*, 1979–1990.

78. Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics *24*, 637–644.

79. Majoros, W.H., Pertea, M., and Salzberg, S.L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. Bioinformatics *20*, 2878–2879.

80. Korf, I. (2004). Gene finding in novel genomes. BMC Bioinformatics *5*, 59.

81. Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. Genome Biol. *9*, R7.

82. Lowe, T.M., and Eddy, S.R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. *25*, 955–964.

83. 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*, 1236–1240.

84. Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., and Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. Mol. Biol. Evol. *34*, 2115–2122.

85. Käll, L., Krogh, A., and Sonnhammer, E.L.L. (2007). Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. Nucleic Acids Res. *35*, W429–W432.

86. Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S.Y., Medema, M.H., and Weber, T. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Res. *47*, W81–W87.

87. Blin, K., Shaw, S., Kloosterman, A.M., Charlop-Powers, Z., van Wezel, G.P., Medema, M.H., and Weber, T. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res. *49*, W29–W35.

88. Nielsen, H. (2017). In Predicting Secretory Proteins with SignalP BT - Protein Function Prediction: Methods and Protocols, D. Kihara, ed. (Humana Press), pp. 59–73.

89. Revell, L.J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. *3*, 217–223.

90. Buchfink, B., Xie, C., and Huson, D.H. (2014). Fast and sensitive protein alignment using DIAMOND. Nat. Methods *12*, 59–60.

91. Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., Rattei, T., Mende, D.R., Sunagawa, S., Kuhn, M., et al. (2016). EGGNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. *44*, D286–D293.

92. Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., et al. (2021). Pfam: The protein families database in 2021. Nucleic Acids Res. *49*, D412–D419.

93. Yin, Y., Mao, X., Yang, J., Chen, X., Mao, F., and Xu, Y. (2012). dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. *40*, W445–W451.

94. Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., and Finn, R.D. (2018). The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res. *46*, D624–D632.

95. UniProt Consortium (2021). UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. *49*, D480–D489.

96. Tagirdzhanova, G., Saary, P., Tingley, J.P., Díaz-Escandón, D., Abbott, D.W., Finn, R.D., and Spribille, T. (2021). Predicted input of uncultured fungal symbionts to a lichen symbiosis from metagenome-assembled genomes. Genome Biol. Evol. *13*, evab129.

97. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A., and Jermiin, L.S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods *14*, 587–589.

98. Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., and Vinh, L.S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. *35*, 518–522.

99. Rokas, A., and Carroll, S.B. (2006). Bushes in the tree of life. PLoS Biol. *4*, 1899–1904.

100. Kishino, H., Miyata, T., and Hasegawa, M. (1990). Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. J. Mol. Evol. *31*, 151–160.

101. Kishino, H., and Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. J. Mol. Evol. *29*, 170–179.

102. Shimodaira, H., and Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. *16*, 1114–1116.

103. To, T.H., Jung, M., Lycett, S., and Gascuel, O. (2016). Fast dating using least-squares criteria and algorithms. Syst. Biol. *65*, 82–97.

104. Schluter, D., Price, T., Mooers, A.Ø., and Ludwig, D. (1997). Likelihood of ancestor states in adaptive radiation. Evolution *51*, 1699–1711.

105. Revell, L.J. (2009). Size-correction and principal components for interspecific comparative studies. Evolution *63*, 3258–3268.

106. De Cáceres, M., and Legendre, P. (2009). Associations between species and groups of sites: indices and statistical inference. Ecology *90*, 3566–3574.

107. Lücking, R., Hodkinson, B.P., and Leavitt, S.D. (2016). The 2016 classification of lichenized fungi in the Ascomycota and Basidiomycota – Approaching one thousand genera. Bryologist *119*, 361–416.

108. Elix, J.A., Kalb, K., Rupprecht, J., Schobert, R., and Rambold, G. (2012). liaslight.lias.net/Identification/Navikey/Metabolites/. In LIAS metabolites – A Database for the Rapid Identification of Secondary Metabolites of Lichens, G. Rambold, ed. last visited: Saturday, 09-Apr-2022 09:25:44 CEST.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Herbarium specimen of *Watsoniomyces obsoletus* used for metagenome sequencing	Herbarium ALTA	https://www.ualberta.ca/museums/museum-collections/cryptogamic-herbarium.html
Herbarium specimens of *Bathelium mastoideum* (2) and *Bogoriella megaspora* (1) used for metagenome sequencing	Herbarium ABL	http://sweetgum.nybg.org/science/ih/herbarium-details/?irn=123895
Herbarium specimen of *Candelina mexicana* used for metagenome sequencing	Herbarium CUVC	http://sweetgum.nybg.org/science/ih/herbarium-details/?irn=126294
All remaining 29 herbarium specimens used for generating new metagenomes	Herbarium UPS	http://www.evolutionsmuseet.uu.se/databaser/databaseseng.html
Fungal culture of *Cirrosporium novae-zelandiae* used for generating genome	Westerdijk Fungal Biodiversity Institute (CBS)	https://wi.knaw.nl/page/fungal_table
Critical commercial assays		
QIAamp DNA Investigator Kit	QIAGEN	Cat#56504
DNeasy Plant Mini Kit	QIAGEN	Cat#69104
Illumina DNA Prep, (M) Tagmentation	Illumina	Cat#20018704
Nextera DNA CD Indexes	Illumina	Cat#20018707
TruSeq DNA PCR-Free Low Throughput Library Prep Kit	Illumina	Cat#20015962
Deposited data		
Raw Sequencing Reads	This paper	NCBI: SRR19097865–SRR19097892 (https://www.ncbi.nlm.nih.gov)
Assembled and annotated genomes	This paper	NCBI: JAPETH000000000– JAPEUK000000000 (https://www.ncbi.nlm.nih.gov)
Software and algorithms		
FastQC v0.11.9	Andrews[56]	https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
Trimmomatic v0.39	Bolger et al.[57]	https://github.com/usadellab/Trimmomatic
SPAdes v3.13.1, v3.14.1, v3.15.1	Nurk et al.[58]	https://github.com/ablab/spades
Bowtie2 v2.3.5.1	Langmead and Salzberg[59]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
CONCOCT v1.1.0	Alneberg et al.[60]	https://github.com/BinPro/CONCOCT
BlobTools	Laetsch and Blaxter[61]	https://github.com/DRL/blobtools
EukCC v0.1.5.1	Saary et al.[62]	https://github.com/Finn-Lab/EukCC
BUSCO v2.0.0, v3.0.2, v5.2.1	Simão et al.,[63] Waterhouse et al.,[64] Manni et al.[65]	http://busco.ezlab.org
MetaWrap(Quant_bins Module) v1.3.2	Uritskiy et al.[66]	https://github.com/bxlab/metaWRAP
Salmon v0.13.1 (within MetaWrap)	Patro et al.[67]	https://github.com/COMBINE-lab/salmon
Samtools v1.9	Li et al.[68]	http://www.htslib.org
BLAST+ v2.2.31, v2.6.0, v2.9.0, v2.11+	Camacho et al.[69]	https://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/
IQTree v2.0.7, v2.1.3	Minh et al.[70]	http://www.iqtree.org
Phylociraptor v0.9.3	Phylociraptor	https://github.com/reslp/phylociraptor
MAFFT v7.464	Katoh and Standley[71]	https://mafft.cbrc.jp/alignment/software/
trimAL v1.4.1	Capella-Gutiérrez et al.[72]	http://trimal.cgenomics.org
ASTRALv5.7.1	Zhang et al.[73]	https://github.com/smirarab/ASTRAL

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Gotree v0.4.2	Lemoine and Gascuel[74]	https://github.com/evolbioinfo/gotree
Funannotate v1.7.4, v1.8.1, v1.8.3, v1.8.5, v1.8.6, v1.8.8	Palmer and Stajich[75]	https://github.com/nextgenusfs/funannotate
RepeatMasker v4.1.0	Smit et al.[76]	http://repeatmasker.org
Gene-Mark-ES v4.59_lic	Ter-Hovhannisyan et al.[77]	http://exon.gatech.edu/GeneMark/gmes_instructions.html
Augustus v3.3.3 (within Funannotate)	Stanke et al.[78]	http://bioinf.uni-greifswald.de/augustus/
GlimmerHMM v3.0.4 (within Funannotate)	Majoros et al.[79]	https://ccb.jhu.edu/software/glimmerhmm/
SNAP v2013_11_29 (within Funannotate)	Korf[80]	https://github.com/KorfLab/SNAP
EVidence Modeler v1.1.1 (within Funannotate)	Haas et al.[81]	https://evidencemodeler.github.io
tRNAscan-SE v1.4.0 (within Funannotate)	Lowe et al.[82]	http://trna.ucsc.edu/tRNAscan-SE/
InterProScan v5.48-83.0	Jones et al.[83]	https://github.com/ebi-pf-team/interproscan
eggnog-mapper v1.0.3	Huerta-Cepas et al.[84]	https://github.com/eggnogdb/eggnog-mapper
HMMER v3.1b2 (within eggnog-mapper)	HMMER	http://hmmer.org
phobius v1.0.1	Käll et al.[85]	https://phobius.sbc.su.se
antiSMASH v5.0.2, v6.0.0	Blin et al.[86,87]	http://antismash.secondarymetabolites.org
Signalp v4.1	Nielsen[88]	https://services.healthtech.dtu.dk/service.php?SignalP-4.1
Phytools	Revell[89]	http://www.phytools.org
Diamond v2.0.7.145 (within eggnog-mapper)	Buchfink et al.[90]	https://github.com/bbuchfink/diamond
Databases		
NCBI Genome (for 467 genomes; Data S1C)	NCBI Genome	https://www.ncbi.nlm.nih.gov/genome/
NCBI Short Read Archive (for *Dibaeis baeomyces* and *Graphis scripta* raw reads)	NCBI SRA	https://www.ncbi.nlm.nih.gov/sra
eggNOG db v4.5.4	Huerta-Cepas et al.[91]	http://eggnog45.embl.de/#/app/home
PFAM db v33.1 (within Funannotate)	Mistry et al.[92]	https://pfam.xfam.org
dbCAN v9.0 (within Funannotate)	Yin et al.[93]	https://bcb.unl.edu/dbCAN2/
MEROPS db v12.0 (within Funannotate)	Rawlings et al.[94]	https://www.ebi.ac.uk/merops/
UniProtKB db v.2020_06 (within Funannotate)	Uniprot Consortium[95]	https://www.uniprot.org/help/uniprotkb
Mycobank	Mycobank	https://www.mycobank.org
Other		
Scripts and code for analysis	Figshare	https://doi.org/10.6084/m9.figshare.19558762

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to lead contact, Toby Spribille (toby.spribille@ualberta.ca).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- All newly generated genomic data is available in NCBI allocated under the Bioproject accession: PRJNA833438. Raw reads are available under SRA accessions: SRR19097865–SRR19097892. Genome assemblies are available in NCBI with accession numbers: JAPETH000000000– JAPEUK000000000.
- All scripts and data used in this study are available at 10.6084/m9.figshare.19558762.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

A total of 30 specimens for a total of 29 species were used to generate new genomic information, including samples from Austria (1), Brazil (3), Canada (15), Colombia (1), the Czech Republic (1), Mexico (1), the United Kingdom (2) and the U.S.A. (5) and one fungal culture (CBS125236) originally derived from a specimen from New Zealand. Individual details of each sample are provided in Table S1.

METHOD DETAILS

DNA sampling

Based on phylogenies published over the past 17 years, we targeted our sampling in the *Lichinomycetes* (4 samples), *Geoglossomycetes* (3), *Coniocybomycetes* (2), *Candelariomycetes* (4), *Xylobotryomycetes* (1), and *Sareomycetes* (1), as well as *incertae sedis* species from *Vezdaeales* (3, including a specimen that later was identified as a *Trizodia*), *Thelocarpales* (3), *Caeruleum* (1) and *Piccolia* (1). In addition, we sampled early diverging lineages and poorly sampled groups in *Lecanoromycetes* (1), *Arthoniomycetes* (2), *Leotiomycetes* (1) and *Dothideomycetes* (3). We used old and fresh material to collect samples from ascocarps or thallus pieces that were dried at room temperature (Table S1); for the sole representative of *Xylobotryomycetes*, *Cirrosporium novae-zelandiae*, we used mycelium from a living culture. Each sample was frozen to –80°C and ground using a TissueLyser II (Qiagen, Germany). The DNA extraction was performed using the QIAamp DNA Investigator Kit (Qiagen) following the manufacturer's instructions, with an overnight lysis and a double incubation during the elution step (2 x 10 min, in ATE).

Sequencing

We constructed libraries using between 10 to 300 ng of DNA per sample with the Nextera DNA Flex Library Prep kit (Illumina), with an average insert size ranging from 500 to 700 bp. Libraries were normalized to 4 nM and sequenced in Canada's Michael Smith Genome Sciences Centre B.C. Cancer with PE 150 Illumina Sequencing with an average coverage of 230X (Table S1).

Genome assembly

The resulting paired end raw reads were quality checked using FastQC.[56] Adapters and low-quality areas were trimmed using Trimmomatic v0.39[57] (Option: ILLUMINACLIP:/Trimmomatic-0.39/adapters/TruSeq3-PE.fa:2:30:10:2 LEADING:30 TRAILING:30 SLIDINGWINDOW:4:15 MINLEN:80) and later inspected with FastQC. The filtered paired-end reads were assembled using metaSPAdes[58] with kmer sizes 21,31,51,71,81,101,127 (-k) and the parameter '–meta' to improve metagenome recovery. We used Bowtie2 v2.3.5.1[59] to align and index reads to the resulting contigs. Binning was carried out using CONCOCT v1.1.0[60] and BlobTools.[61] Taxonomy, completeness and contamination of each bin was checked using EukCC v0.1.5.1,[62] BUSCO v5.2.1[65] and relative coverage, which was estimated using the metaWRAP v1.3.2[66] module QUANT_Bin based on Salmon v0.13.1.[67]

We inspected each assembly using a custom R script based on Tagirdzhanova et al.[96] plotting reads, relative coverage and GC content, grouped by the binning results to classify for potential mismatches and collapsing bins into MAGs based on GC%, coverage and non-redundant BUSCOs when needed (Table S1). Assembly and bin statistics were calculated using Samtools v 1.9.[68]

Sample identification

In cases where we needed to resolve ambiguity, we validated sample identifications with our assembled MAGs by recovering the ribosomal DNA from raw assemblies using BLAST+ v2.9,[69] and filtering sequences based on relative coverage. We aligned the obtained sequences against sequences retrieved from NCBI nt, and constructed individual trees using IQTree v2.0.7.[70] Five samples were assessed to likely represent undescribed species within their genera and are referred to by isolate numbers in the analyses.

Phylogenomics

We selected the most complete genome (Reference to Representative status in NCBI Ref-Seq when available) per genus from all *Ascomycota* genomes available at the time of this study in the NCBI database, for a total of 469 genomes (Data S1C). We used the Phylociraptor v0.9.3 pipeline to combine the external dataset and our 30 newly generated genomes, for a total of 499 genomes, into a phylogeny. We set Phylociraptor to use the Dikarya_odb9 (1312 genes) data set as reference to extract BUSCO genes per genome using BUSCO v3.0.2.[64] Then, we excluded all genomes with less than 50% completeness (based on single copy genes alone) and, for the selection of loci to be included, excluded BUSCO genes occurring in less than three genomes, from posterior analyses (Data S1D). A final set of 481 genomes (including all the newly generated genomes in this study) was used for the aligning step using MAFFT v7.464[71] (–auto) per BUSCO gene, and trimmed *a posteriori* using trimAl v1.4.1[72] (–gappyout, to trim based on gap distribution). All gene alignments with less than 10 parsimony-informative sites were excluded from the analysis for a final set of 1292 genes. Then, we calculated evolutionary models per gene specifying nuclear models (-msub nuclear) and including FreeRate heterogeneity calculations (-m MFP).[97] We constructed individual gene trees per alignment using IQTree v2.0.7 with a 1000 ultrafast bootstrap[98] (–bb 1000) approach and the calculated models per gene. We built a species tree using ASTRALv5.7.1[73] to calculate a coalescent model based on the maximum shared quartets between the 1292 gene trees and estimating a local posterior probability per branch. To build a concatenated tree, we combined the calculated evolutionary models and the alignments to construct a maximum-likelihood tree using IQTree v2.0.7 with a 1000 ultrafast bootstrapping (–bb 1000) approach.

We examined our resulting phylogenomic trees in two different ways: one qualitative, comparing topology between both trees, and one quantitative, calculating concordance factors with the alignments and all individual gene trees. We used Gotree v0.4.2[74] with the 'compare' option for trees and edges, to check for discrepancies between concatenated and species tree topology and branch lengths (Data S1H). Since informative sites resampling does not scale symmetrically in large genomic datasets, bootstrap values become less informative.[99] We accordingly calculated concordance factors as a support metric in IQTree v2.0.7 using both gene trees for gene Concordance Factor (gCF) and gene alignments to calculate site Concordance Factor (sCF).[21]

To follow up the secondary support, we ran the concatenated analysis constraining to three topological hypotheses testing the three hypothetical relationships between *Lichinomycetes* and the sibling clade *Lecanoromycetes+Eurotiomycetes* (LE) (H1: Others (*Lichinomycetes*+LE); H2: *Lichinomycetes* (Others+(LE)); H3: LE (Others+*Lichinomycetes*)). Then, we combined the resulting most parsimonious topologies, with the initial concatenated run, together with the concatenated and coalescence topology and performed an AU test[23] with 10000 resamples by bootstrapping[100] (RELL method, -zb 10000) and including the weighted KH[101] and weighted SH[102] tests (-zw). Then we combine the topologies into two sets; one for methods (Concatenated+Coalescence) and one for hypothesis (H1+H2+H3) and calculate log-likelihoods of each site in the concatenated alignment per topology (-wsl). Afterwards, we calculated the log-likelihood per gene to test which topology maximizes the likelihood using Phylogenetic Signal Parser v1.1.[22] In addition, we estimated the relative signal (low and high) that maximizes the likelihood per site in each topology. We tested for statistical differences using a G-test based on a 1:1 for methods and a 1:1:1 ratio for topological hypotheses. All the analyses were performed with and without branch-lengths.

Time calibration

We estimated divergence times for our concatenated tree using the distance–based calibration method LSD2[103] (least square dating), included in IQTree v2.1.3 (–date), using six fossil records from *Ascomycota* matching our sampled groups proposed in Nelsen et al.[20] and Lutzoni et al.,[19] limiting the selection to fossil dated nodes. We set the calibration with a fixed outgroup based on *Basidiomycota*, set tips to the present time (–date-tip 0) and calculated a confidence interval based on 100 iterations (–date-ci 100). The calibration was based on a *Pezizomycotina* node dated between 589 and 405 Mya, a *Diaporthales* node dated between 193.8 and 133.3 Mya, *Hypocreales* node dated between 145.6 and 100.1 Mya, a *Glomerellales* node with a constrained lower bound of 60.5 Mya, a *Coniocybales* node with a constrained lower bound of 40.7 Mya and a fixed *Aspergillaceae* node at 40.7 Mya (41.2 ± 0.5).

Functional genomic annotations

From the 481 genomes included in the phylogenomic analyses, we selected 115 genomes for functional annotations, of which 114 belong to *Pezizomycotina*, including 29 of the 30 newly generated genomes, and one representative genome from *Saccharomycotina* (*Yarrowia lipolytica*). We annotated the selected genomes using the Funannotate[75] pipeline. Cleaning, sorting and masking steps were run within the default settings, with the -exhaustive argument to improve cleaning, and masking was set to tantan (default) and RepeatMasker v4.1.0.[76] Initial predictions were obtained using Gene-Mark-ES v4.59_lic under self-training settings with a branch point model specified for gene organization of fungal genomes (–ES -fungus).[77] The functional prediction includes an initial training step using Augustus v3.3.3,[78] followed by BUSCO v2.0.0[63] (Dikaria_odb9), Gene-Mark-ES, GlimmerHMM v3.0.4[79] and SNAP v2013_11_29,[80] followed by EVidence Modeler v1.1.1[81] to create a weighted consensus from all predictions. The tRNAs were identified using tRNAscan-SE v1.4.0.[82] With the consensus predictions, the pipeline annotated protein families and domains using InterProScan v5.48-83.0.[83] External annotations based on orthology was performed using eggnog-mapper v1.0.3[84] with eggNOG db v4.5.4,[91] followed by the annotation for transmembrane signal peptides and secondary metabolism biosynthetic gene clusters (BGCs) using phobius v1.0.1[85] and antiSMASH,[86,87] respectively. All functional predictions and external annotations were parsed together using the Funannotate annotation step including: a final step of annotations with an HMMsearch for Pfams using PFAM db v33.1[92] and CAZymes using dbCAN v9.0;[93] a blastp search against MEROPS db v12.0[94] and UniProtKB db v.2020_06[95] using Diamond v2.0.7.145[90]; prediction of secreted proteins using SignalP v4.1[88]; and a Gene Ontology term (GO term) search from the eggNOG annotations.

To compare all the annotated genomes, we used the Funannotate compare tool, and manually parsed secondary metabolite BGCs from antiSMASH results (Data S1P).

Ancestral node reconstruction

To analyze evolutionary trends, we estimated ancestral node states based on present day CAZymes numbers using the R package Phytools[89] with the function anc.ML under an Ornstein-Uhlenbeck character evolution modeling.[104] We reconstructed ancestral nodes for CAZymes families, and putative CAZymes degrading pectin, cellulose and lignin, using our time-calibrated concatenated tree pruned with the annotated taxa set. All estimations were kept for posterior plotting, irrespective of the convergence, and the characters that did not converge were highlighted. Ancestral reconstructed nodes were later included in the PCA sets.

QUANTIFICATION AND STATISTICAL ANALYSIS

All calculations and plots were run and tested using R Studio version 4.1.2 (arm64 built and x86_64 build).

Annotation statistics

To compare the annotations, we used custom R scripts to generate individual heat maps for each of the metrics recovered during the annotation. We used the clades obtained from the phylogeny as an *a priori* clustering, to calculate basic statistics per clade. All descriptive comparisons were later tested using ANOVAs and Kruskal-Wallis tests depending on the data distribution, with their own *post hoc* Tukey HSD and Dunn tests, respectively, against *Lichinomycetes* (Table S2; Data S1N).

Principal components analyses of functional annotations

To explore the multidimensionality or aggregation of the annotation sets, we constructed Principal Component Analyses (PCA) for all Pfams, all CAZyme classes and all BGC types, using both summarized— as displayed in Figure 4 (CAZyme classes, Pfam annotations, using only Pfams that have a standard deviation > 1 over all annotated genomes; BGCs: NRPS, NRPS-like, T1PKS, Others)— and entire sets using the function 'prcomp' from base package using a prior Log transformation in all cases but CAZymes. Then, we use the function 'phyl.pca' from R package Phytools to include a pruned version of the time-calibrated tree as a covariable to reduce any variability induced by the sampling size and non-independence, phylogenetically correcting the PCA (pPCA) under a maximum likelihood-optimized Brownian Motion character evolution model.[105] The differences between pPCA and PCA in dimensionality, clustering and loadings were used as a qualitative proxy for phylogenetic signal.

High fidelity pfam identification

To calculate statistically high fidelity Pfams per class, we used the R package 'indicspecies' with the function 'multipatt' for a multilevel pattern analysis, using Pfams per taxa in place of communities and clustering, based on pre-defined phylogenetic clades. The function constructs a group-equalized biserial point correlation matrix, consisting of Pearson's correlations with an abundance quantitative vector and a binary grouping vector equalizing for group sizes,[106] to estimate probability of group association. Then, based on 9999 permutations and an alpha of 0.05, we collected all the significant combinations shared from one to all classes. We ran the analysis both with and without *Saccharomycetes*, to remove outgroup dominance of rare Pfam occurrences. The collected Pfam combinations—up to two classes—were later plotted and visually explored.

Pfam statistical analyses

All Pfams were tested for differences in distribution between *Lichinomycetes* and non-*Lichinomycetes*. We used two approaches: first, a Wilcoxon signed-rank test, testing for significant differences per Pfam between *Lichinomycetes* and non-*Lichinomycetes* (p-value < 0.01), and second, a Fisher's exact test to evaluate independence (Fisher p-value < 0.01) between occurrences and total recovered sequences per Pfam between *Lichinomycetes* (L) and non-*Lichinomycetes* (n-L). To better understand the distribution we set two ratios: one based on average (mean L / mean n-L), and the other one based on logical distinctions of Pfam occurrence ratios corrected by sample size (Data S1R). The sample size correction consisted of highlighting Pfams with occurrences in more than half of *Lichinomycetes* and over-represented compared against non-*Lichinomycetes* occurrences corrected by the total number of classes (10) to avoid single occurrences. To evaluate for potential missed annotations or low hit numbers due to the heterogeneity of our sampling set compared to the Pfam reference database, we extracted the number of predicted proteins containing annotated Pfams and divided it by the total number of predicted proteins to estimate Pfam annotation success ratio. Then, we evaluated significant differences between *Lichinomycetes* and non-*Lichinomycetes* in annotation rate quality using a Kruskal-Wallis test.

Assessment of species and metabolite diversity

Species diversity of lineages within *Lichinomycetes* was assessed using Mycobank (see key resources table) and Lücking et al. (Data S1U).[107] Secondary metabolite diversity within these lineages was assessed based on Elix et al. (Data S1T).[108]