



OPEN

A comparative genomic analysis of lichen-forming fungi reveals new insights into fungal lifestyles

Hyeunjeong Song¹, Ki-Tae Kim², Sook-Young Park², Gir-Won Lee^{3,9}, Jaeyoung Choi^{4,10}, Jongbum Jeon¹, Kyeongchae Cheong⁴, Gobong Choi¹, Jae-Seoun Hur⁵ & Yong-Hwan Lee^{1,4,6,7,8}✉

Lichen-forming fungi are mutualistic symbionts of green algae or cyanobacteria. We report the comparative analysis of six genomes of lichen-forming fungi in classes Eurotiomycetes and Lecanoromycetes to identify genomic information related to their symbiotic lifestyle. The lichen-forming fungi exhibited genome reduction via the loss of dispensable genes encoding plant-cell-wall-degrading enzymes, sugar transporters, and transcription factors. The loss of these genes reflects the symbiotic biology of lichens, such as the absence of pectin in the algal cell wall and obtaining specific sugars from photosynthetic partners. The lichens also gained many lineage- and species-specific genes, including those encoding small secreted proteins. These genes are primarily induced during the early stage of lichen symbiosis, indicating their significant roles in the establishment of lichen symbiosis. Our findings provide comprehensive genomic information for six lichen-forming fungi and novel insights into lichen biology and the evolution of symbiosis.

Lichens exist in symbiosis, in which at least one fungus (mycobiont) lives in a mutually beneficial relationship with photosynthetic algae and/or cyanobacteria (photobiont)^{1,2}. Since this dual nature was discovered by Schwendener in 1867³, numerous studies have demonstrated that basidiomycetes yeast⁴⁻⁷, as well as diverse microbiomes⁸, may cohabitate within lichen thalli. In lichen association, dominant fungal partners which produce basic morphological structure of lichens are determine the classification of lichens. The thallus structure composed of the fungal component retained the water for drought tolerance in extreme conditions^{9,10} as well as has a role as a shelter; protecting the photobionts from the external environment². Moreover, the algal partner synthesizes carbohydrate products by photosynthesis and transfers this carbon source to the fungal partner to maintain the lichen association¹¹.

Lichenization is common among fungi, with approximately 21% of fungal species forming lichens². Lichenized symbiosis is not derived from a single phylogenetic clade¹², but found in the Ascomycota classes Lecanoromycetes, Eurotiomycetes, Leotiomycetes, Dothideomycetes, and Arthoniomycetes, as well as in a few Basidiomycota classes^{1,2}. Previous phylogenetic studies have suggested that lichenization evolved independently at least five times in distantly related lineages¹². Such studies have also demonstrated that lichenization has been continuously maintained from the common ancestor of Lecanoromycetes, but was lost during the evolution of Lecanoromycetes. Due to this complex evolutionary history, many hypotheses have been proposed to account for the evolutionary time required for lichenization and its loss and re-evolution¹³.

To date, many studies have been conducted to elucidate the symbiotic nature of lichens. The successful re-association of lichen symbionts under laboratory conditions has facilitated microscopic observations of the fungal-algal interface during lichen establishment^{14,15}. Thus, the early stages of lichenization, which ranges from 'pro-contact' to 'growth together', have been well investigated¹⁵⁻¹⁷; however, its late stages remain poorly understood. Although the aim of many studies is to identify symbiosis-related genes, until recently, we lacked the

¹Interdisciplinary Program in Agricultural Genomics, Seoul National University, Seoul 08826, Korea. ²Department of Agricultural Life Science, Suncheon National University, Suncheon 57922, Korea. ³Research Institute for Agriculture and Life Sciences, Seoul National University, Seoul 08826, Korea. ⁴Department of Agricultural Biotechnology, Seoul National University, Seoul 08826, Korea. ⁵Korean Lichen Research Institute, Suncheon National University, Suncheon 57922, Korea. ⁶Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul 08826, Korea. ⁷Center for Fungal Genetic Resources, Seoul National University, Seoul 08826, Korea. ⁸Plant Immunity Research Center, Seoul National University, Seoul 08826, Korea. ⁹Present address: SML Genetree Co. Ltd, Seoul 08826, Korea. ¹⁰Present address: Smart Farm Research Center, Korea Institute of Science and Technology, Gangneung 25451, Korea. ✉email: yonglee@snu.ac.kr

Lichen species	Genome size (Mb)	Number of scaffolds	Number of genes	GC content (%)	References
<i>Endocarpon pusillum</i> R61883	37.13	59	9252	48.54%	Park et al., 2014
<i>Endocarpon pusillum</i> Z07020	37.33	308	9238	45.56%	Wang et al., 2014
<i>Gyalolechia flavorubescens</i>	34.47	36	9695	41.79%	Park et al., 2013
<i>Umbilicaria muehlenbergii</i>	34.81	7	8294	46.82%	Park et al., 2014
<i>Cladonia metacorrallifera</i>	36.68	30	9030	43.81%	Park et al., 2014
<i>Cladonia macilenta</i>	37.12	240	8773	42.85%	Park et al., 2013

Table 1. Genome statistics of the lichen-forming fungi.

genetic transformation tools required to perform gene manipulation in lichen biology^{18,19}. Thus, recent molecular studies have applied genetic transformation systems to elucidate lichen symbiosis.

However, the slow growth of several lichens and the difficulty of their culture in the laboratory have further required the development of genomic-level studies to gain an evolutionary understanding of lichen symbiosis. Genomics have advanced greatly since the sequencing of *Xanthoria parietina*^{20–26}. Numerous studies have approached lichen symbiosis from a genomic perspective to identify evolutionary process of lichenization and symbiosis-related genes. *Endocarpon pusillum* was the first lichen to have been subjected to genomic analysis; early studies reported its symbiosis-related genes involved in nitrogen/sugar transport and metabolism with their expression during the re-synthesis stages²⁶. Although continuous genomic studies investigating the key factors of lichen symbiosis^{20,24,27}, recent descriptions of several additional genome sequences²⁸, and the application of systems biology approach to lichen associations²⁹ improve the knowledge of lichen symbiotic systems but determining how a symbiotic lifestyle evolved remains challenging. Mycorrhizal fungi, which are mutualistic symbionts associated with > 90% of land plants, have been studied extensively to identify their symbiotic nature. Large-scale genomic sequencing of mycorrhizal fungi has revealed that convergent evolution occurred via the loss of plant cell wall-degrading enzymes (PCWDEs) and the enrichment of transposable elements (TEs) and mycorrhiza-induced small secreted proteins (MiSSPs)^{30–32}. Several molecular studies have also reported that secreted proteins play a crucial role in mycorrhizal symbiotic associations^{33–35}.

In this study, we aim to conduct a comparative genomic analysis of four Lecanoromycetes species (*Gyalolechia flavorubescens*³⁶, *Cladonia macilenta*³⁷, *Cladonia metacorrallifera*³⁸, and *Umbilicaria muehlenbergii*³⁹) and two Eurotiomycetes species (*Endocarpon pusillum* Z07020²⁶ and R61883⁴⁰) of lichen-forming fungi, which are evolutionary distant species. An additional 50 genomes from fungi with diverse lifestyles, including mycorrhizal fungi and close relatives of lichen-forming fungi, were also used to support the identification of the lichen-specific genomic features. We were going to perform a gene family gain/loss analyses in comparison with non-lichenized fungi to identify specific gene families of lichen-forming fungi. Finally, we re-synthesized *G. flavorubescens* and its algal partner *Trebouxia gelatinosa* and performed a time-series transcriptomic analysis of this re-synthesized lichen through RNA sequencing to reveal unique features of lichen symbiosis.

Results

Phylogenomic relationships and genomic similarity among lichen-forming fungi. The Lecanoromycetes and Eurotiomycetes lichen-forming fungi have similar genome sizes, ranging from 34.5 to 37.3 Mb, and similar numbers of genes (8294–9695) (Table 1). We conducted phylogenomic analysis of the six lichen-forming fungi including 50 fungal genome sequences (Supplementary Dataset S1). The phylogenomic tree showed that the four Lecanoromycetes species (*G. flavorubescens*, *C. macilenta*, *C. metacorrallifera*, and *U. muehlenbergii*) and the two *E. pusillum* isolates were distantly related (Fig. 1A). This finding is consistent with previous reports that lichenization events evolved independently in multiple lineages^{1,12,41}. The time-calibrated phylogeny suggests that Lecanoromycetes diverged approximately 258 million years ago from an ancestral fungus that may have been lichen-forming, and that divergence between *E. pusillum* isolates and the plant pathogen *Phaeoemoniella chlamydospora* occurred approximately 52 million years ago.

We analyzed the synteny of lichen-forming fungal genomes using *C. macilenta* as a reference. Dot plots revealed that both *Cladonia* species had a robust syntenic relationship, with several inverted blocks (Fig. 1B); however, as the evolutionary distance of species from *C. macilenta* increased, the syntenic relationship weakened. Because *C. macilenta* and *E. pusillum* belong to different classes, their syntenic relationship is nearly random, despite their both being lichen-forming fungi. The syntenic region of the two *Cladonia* species had 65.7–66.5% similarity, whereas the syntenic similarities between *C. macilenta* and other lichen-forming *G. flavorubescens* and *U. muehlenbergii* in Lecanoromycetes were 6.1–6.6% and 6.8–7.3%, respectively (Supplementary Table S1), and the two *Endocarpon* species in Eurotiomycetes had 3.4% and 3.5% similarity compared with *C. macilenta*.

Repetitive sequence content was also analyzed in lichen-forming fungi. The simple repeat content account for approximately 1% of all lichen-forming fungal genomes, but most DNA transposons were observed only in *E. pusillum* R61883, rather than other lichen-forming fungi (Fig. 1C). The portion of retroelements differ among lichen-forming fungi, Lecanoromycetes fungi (less than 1%), *E. pusillum* (1.5%), and *U. muehlenbergii* (4%). The total composition of repeat sequences in lichen-forming fungi was lower than that of other fungal species (Supplementary Fig. S1).

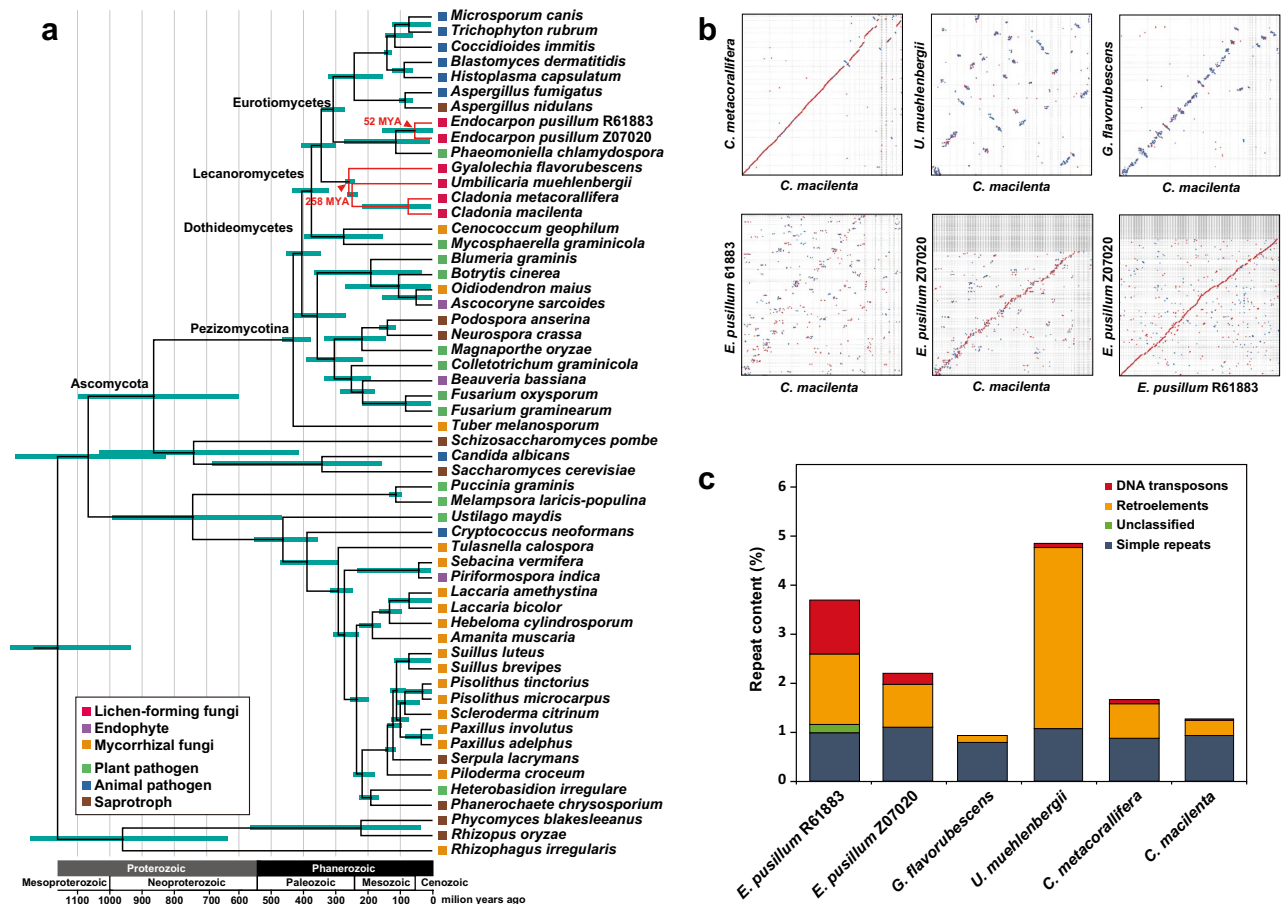


Figure 1. Phylogenomic and synteny relationships among lichen-forming fungi and their repeat contents. (A) The phylogenomic tree shows that lichen-forming fungi are derived from many ancestors. Red branches indicate lichen-forming fungal lineages, and colored squares indicate lifestyle of a fungal species. The scale of the phylogenomic tree is millions of years, as calculated using the mcmctree function in the Phylogenetic Analysis by Maximum Likelihood software package. Blue error bars at each node indicate 95% highest posterior density (HPD) for node age. (B) Synteny dot plots of lichen-forming fungi. Red (blue) dots indicate forward (reverse) matches. (C) Repetitive sequence contents of lichen-forming fungi identified using RepeatMasker software. DNA transposons, retroelements, and unclassified repeats are classes of interspersed repeats.

Gene family expansion and contraction during the evolution of lichen-forming fungi. Gene family expansion and contraction were analyzed based on orthologous genes across 56 fungal species including six lichen-forming fungi (Fig. 2). We estimated changes in gene family size when the two lichen-forming fungal clades Lecanoromycetes and *E. pusillum* diverged from different non-lichenized common ancestors. In Lecanoromycetes, 106 families expanded and 3049 contracted. Among the *E. pusillum* isolates, 238 families expanded and 886 contracted. Contractions were dominant in the lichen lineages, leading to a small total gene number in the lichen-forming fungi (Table 1; Supplementary Dataset S1). In both lichen-forming fungal clades, the cytochrome P450 (CYP) family expanded, whereas the glycoside hydrolase (GH), transcription factor (TF), and major facilitator superfamily (MFS) contracted (Table 2).

Loss of plant cell wall degrading enzymes (PCWDEs) in lichen associations. In both pathogens and symbionts, PCWDEs play essential roles in plant host cell wall remodeling for fungal colonization⁴². However, gene family expansion and contraction analysis (Fig. 3A) and the profiles of carbohydrate-active enzyme (CAZyme) genes (Supplementary Fig. S3) revealed a remarkable reduction of PCWDEs in six lichen-forming fungi. Plants and green algae have similar cell wall components, such as cellulose and hemicellulose, whereas pectin is unique to land plants and Charophycean green algae⁴³. Lichen-forming fungi have fewer CAZyme genes involved in PCWDEs compared with plant-associated fungi, and a similar number compared with animal pathogens (Supplementary Fig. S3). Almost all polysaccharide lyase (PL) family genes, which act mainly in pectin degradation, have been lost in lichen-forming fungi. Only a few genes acting on cellulose (Auxiliary activity [AA] family 9, GH5, and GH3), hemicellulose (GH5, GH27, GH31, Carbohydrate esterase [CE] family 1, GH2, GH43, and GH3), and pectin (CE1, GH2, GH43, and GH3) are conserved in lichen-forming fungi (Fig. 3A). The number of PCWDE genes is dramatically decreased among lichen-forming fungi (Fig. 3A); decreasing patterns

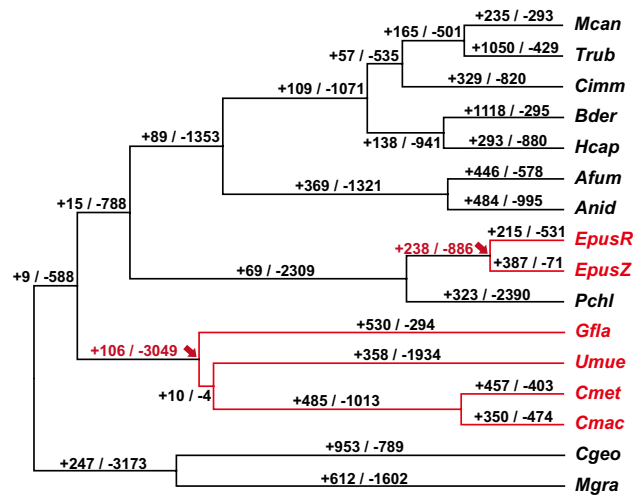


Figure 2. Gene family evolution in lichen-forming fungi and their relatives. Estimation of gene family expansion and contraction in lichen-forming fungi using the CAFE computational tool ($P > 0.01$). Red arrows indicate branch points where lichen-forming fungi diverged from non-lichenized ancestors. + and – indicate the numbers of expanded and contracted gene families, respectively. Only 16 species belonging to the Lecanoromycetes, Eurotiomycetes, and Dotidomycetes closely related to lichen-forming fungi were shown, and the analysis results for all 56 species were placed in the Supplementary Figure S2. Species abbreviations: EpusR, *Endocarpon pusillum* R61883; EpusZ, *E. pusillum* Z07020; Gfla, *Gyalolechia flavorubescens*; Umue, *Umbilicaria muehlenbergii*; Cmet, *Cladonia metacorallifera*; Cmac, *Cladonia macilenta*. Abbreviations of other species names are provided in Supplementary Dataset S1.

Interpro ID	Term	Number of changed gene families in both lichen-forming fungi clades
Expanded		
IPR001128	Cytochrome P450	3
Contracted		
IPR013781	Glycoside hydrolase, catalytic domain	16
IPR007219	Transcription factor domain, fungi	12
IPR011701	Major facilitator superfamily	10
IPR001138	Zn(2)-C6 fungal-type DNA-binding domain	10
IPR005828	Major facilitator, sugar transporter-like	7

Table 2. Commonly expanded and contracted gene families in lineages of lichen-forming fungi.

of PCWDE genes have been similarly observed in ectomycorrhizal fungi, which cannot penetrate host plant cell walls during colonization³⁰.

We conducted gene tree–species tree reconciliation analysis to further infer the evolutionary relationships of PCWDE genes in lichen-forming fungi and their relatives (Fig. 3B–D). Lichen-forming fungi belonging to Lecanoromycetes lost 15 cellulose-degrading enzyme genes from their ancestral gene repertoire (Fig. 3B). *E. pusillum* underwent two steps of gene loss: Chaetothyriomycetidae lost 6 genes, and then *Endocarpon* lost 12 cellulose-degrading enzyme genes. The plant pathogen *P. chlamydospora*, which is also a member of Chaetothyriomycetidae, regained several PCWDE genes subsequent to their loss in Chaetothyriomycetidae. The propensity of gene loss pattern in hemicellulose and pectin genes is similar to that of cellulose (Fig. 3C,D). The Eurotiomycetes lineage, which comprises only animal pathogens, also underwent massive loss of PCWDE genes whereas Leotiomyces and Sordariomyces, which include many plant-associated fungi, gained repertoires of PCWDEs for host invasion. These results indicate that most genes related to the degradation of cellulose, hemicellulose, and pectin have been lost in lichen-forming fungi, but that these events occurred independently during the evolution of lichenization in different evolutionary lineages.

Loss of sugar transporters during lichenization. The MFS is the largest family of secondary transporters related to the movement of diverse solutes, especially sugar uptake⁴⁴. However, MFS-type transporters underwent extensive contraction in six lichen-forming fungi, including the sugar porter (2.A.1), anion:cation symporter (2.A.1.14), aromatic acid:H + symporter (2.A.1.15), and siderophore-iron transporter (2.A.1.16) families (Fig. 4A). Because the type of sugar alcohols in symbiosis depends on the photosynthetic partners⁴⁵, we further characterized the sugar transporters in lichen-forming fungi using dataset of *G. flavorubescens* expression

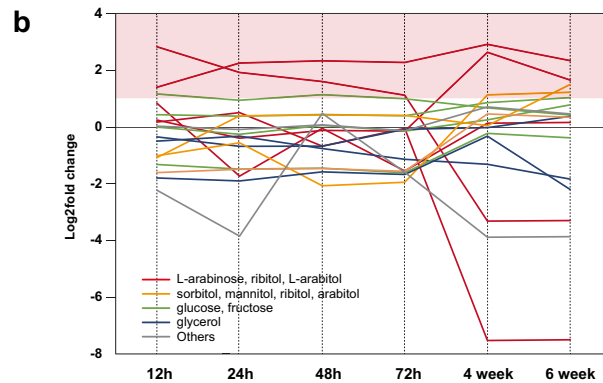
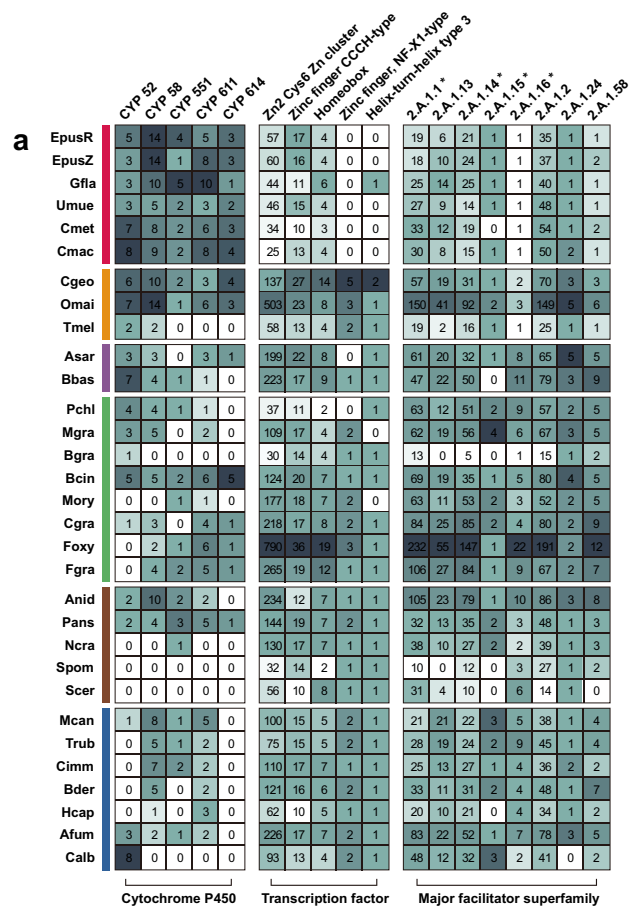


Figure 4. Distribution of expanded and contracted gene families in lichen-forming fungi. **(A)** Comparative analysis of selected families among the cytochrome P450 (CYP) and transcription factor (TF) families and major facilitator superfamily (MFS), which expanded or contracted in lichen-forming fungi. Colored boxes indicate fungal species lifestyles. Asterisk indicates rapidly contracted gene families. **(B)** Expression patterns of diverse polyol transporters in *G. flavorubescens*.

Only *E. pusillum* R61883 underwent possible duplication in specific TF families, including homeodomain-like (IPR009057) and helix-turn-helix psq (IPR007889) families (Supplementary Fig. S4 and S5A). We hypothesize that transposons and transposition of DNA-mediated genes (GO:0006313) particularly abundant in this sample may have caused these duplications (Supplementary Fig. S5A). Because DNA transposons were near (3 kb) the duplicated homeodomain-like and helix-turn-helix psq families, we hypothesize that the expansion of TF families was influenced by repeat elements (Fig. 4B).

Expanded cytochrome P450 (CYP) genes and secondary metabolites involved in lichen symbiosis. CYPs are heme-containing monooxygenases involved in a variety of metabolic processes⁴⁸. Genes in the CYP52, CYP58, CYP551, CYP611, and CYP614 families are more numerous in lichen-forming fungi than in other analyzed genomes (Fig. 4A). Expanded CYP genes in lichen-forming fungi are separated from those

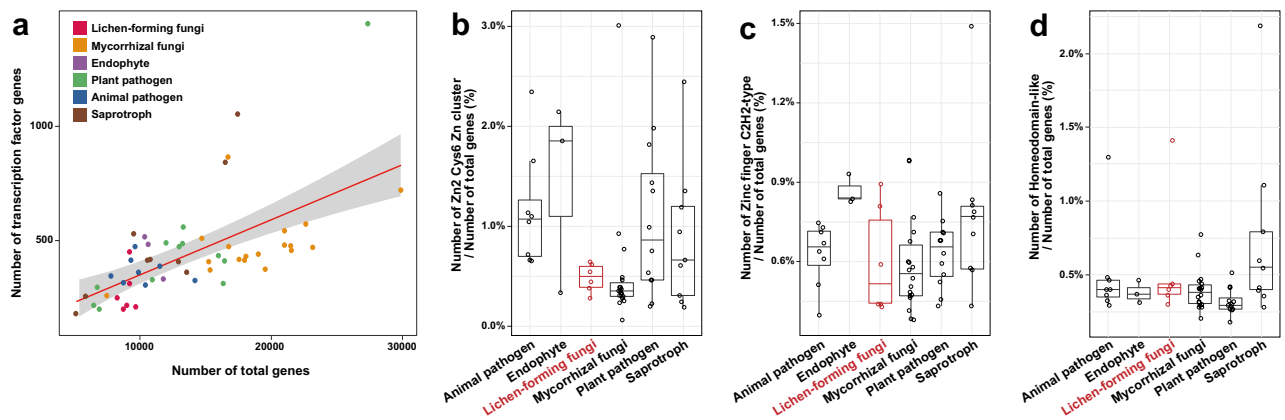


Figure 5. Major contributors of TF gene family contraction in lichen-forming fungi. **(A)** Correlation between the number of TF genes and the total number of genes. Red and grey lines are regression and error lines, respectively. Colored dots indicate fungal species lifestyles. **(B–D)** The three major TF families known to affect the total TF size. Red boxes indicate the distribution of lichen-forming fungi. **(B)** Zn2 Cys6 Zn cluster; **(C)** zinc finger C2H2-type; and **(D)** homeodomain-like.

of fungi with other lifestyles, indicating that this feature evolved uniquely from other fungi. CYP52 and CYP58 are involved in n-alkane and fatty acid assimilation and trichothecene biosynthesis⁴⁹. The CYP551, CYP611, and CYP614 families have not been characterized, but may be involved in the symbiotic lifestyle because most of these CYP genes are lichen genes (Fig. 4A; Supplementary Fig. S6).

Lichen-forming fungi synthesize various unique secondary metabolites². We found more polyketide synthase (PKS) genes in lichen-forming fungi, mainly in Lecanoromycetes, than in 2). Reconciliation analysis revealed that the gain of these PKS genes occurred after lichen-forming fungi emerged from non-lichenized ancestors (Supplementary Fig. S7). Although the *E. pusillum* isolates are distantly related to Lecanoromycetes lichens, lichen-forming fungi shared many PKS genes in the phylogenetic analysis (Supplementary Fig. S8); we identified a lichen-specific PKS group that consisted entirely of lichen species, except *G. flavorubescens*. Sequence similarity analysis based on a blast search revealed that this is a lichen-specific PKS gene, with no ortholog in other fungal species (Supplementary Fig. S9A). Instead, *G. flavorubescens* have species-specific PKS genes (Fig. 7B). This genomic evidence is consistent with previous findings of the presence of unique secondary metabolites synthesized in lichen-forming fungi^{50–52}.

Transcriptomic data for the resynthesis of *G. flavorubescens* and *T. gelatinosa* were used to investigate the relationships among two expanded gene families (PKS and CYP) and lichen symbiosis. Several PKS and CYP genes were highly expressed only during the early stage (at 12, 24, 48, and 72 h), whereas other genes were induced only during the late stage (4 and 6 weeks) (Supplementary Fig. S10). All lichen-specific PKS genes were induced only during the early stage of symbiosis (Supplementary Fig. S10B). The expanded gene families appear to be involved in producing various compounds and secondary metabolites, as previously described^{2,50–52}. However, the expression patterns of lichen-specific PKS genes indicate that the lichen-specific PKS products are induced during the early stage of symbiosis.

Lichen-specific genes of six lichen-forming fungi. In addition to the loss of unnecessary genes in lichen-forming fungi, we attempted to identify newly gained genes that may contribute to their unique symbiotic lifestyle. Ortholog clustering analysis of the six lichen-forming fungi identified 3051 core groups, whereas clustering with only four Lecanoromycetes identified 3,468 core groups (Supplementary Fig. S11A and B; Supplementary Table S3). Thus, the number of core gene clusters in lichen-forming fungi remained consistent regardless of the lichen-forming fungi in different classes.

We identified 5498 lichen-specific orthogroups, including species-specific genes, after clustering with an additional 50 fungal genomes. The number of core groups was substantially reduced among lichen-forming fungi, leaving only one lichen-specific core group (Supplementary Fig. S11C and D). This finding suggests that no universal lichen-forming fungal gene sets are involved in their symbiosis and that, rather than core genes, they have many genus- or species-specific genes (Supplementary Table S3), which are also important in mycorrhizal symbiosis³⁰.

Lichen-specific genes were functionally annotated through gene ontology (GO) analysis using biological process terms. GO terms revealed no association with approximately 90% of the genes in *E. pusillum* R61883, 97% in *E. pusillum* Z07020, 99% in *G. flavorubescens*, 99% in *U. muehlenbergii*, 88% in *C. metacorrallifera*, and 89% in *C. macilenta* (Supplementary Fig. S12A); therefore, these genes were likely newly gained during the evolution of lichen symbiosis. Common function of the functionally annotated genes in the six lichen-forming fungi included oxidation–reduction processes, protein phosphorylation, transmembrane transport, carbohydrate metabolic processes, and transcription regulation (Supplementary Fig. S12B). However, genes involved in DNA-mediated transposition were found only in *Endocarpon* species, primarily in *E. pusillum* R61883. This difference appears to be related to the abundance of DNA transposons and the expansion of specific TF families in *E. pusillum* R61883, as mentioned above (Supplementary Fig. S5).

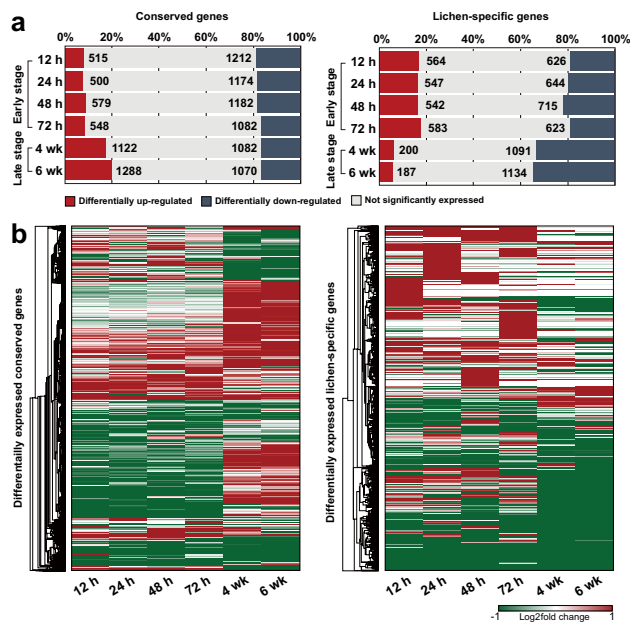


Figure 6. Symbiosis-induced genes in *G. flavorubescens*. Gene expression profiles of conserved and lichen-specific genes in *G. flavorubescens*. (A) Differentially up- (\log_2 fold change > 1) and downregulated (\log_2 fold change ≤ -1) genes at each time stage. Gray bars indicate genes that were not significantly expressed. (B) Expression patterns of lichen-specific and conserved genes.

Symbiosis-induced genes in *G. flavorubescens*. Genome-wide expression profiling was performed using *G. flavorubescens* and its algal partner *T. gelatinosa* to determine the possible roles of lichen-specific genes and conserved genes during lichenization. Of the conserved genes with orthologs in non-lichenized fungi, 17–20% were upregulated 4–6 weeks PCI, whereas 8–9% were upregulated at 12–72 h PCI (\log_2 fold-change > 1) (Fig. 6A). Most genes that were upregulated during the late-stage were not affected during the early stage (Fig. 6B), indicating that different conserved genes are associated with the early and late stages. Lichen-specific genes exhibited a different pattern from conserved genes, with more genes upregulated during the early stage (17–18%) than the late stage (6%) (Fig. 6A). Similar to conserved genes, different genes were upregulated during the early and late stages, suggesting that lichen-specific genes that are highly induced during the early stage are no longer necessary during the late stage. Different lichen-specific upregulated genes were involved at each time point, even within each stage (Fig. 6B).

GO enrichment analysis was performed on functionally annotated conserved genes that were differentially expressed during the early and late stages (Supplementary Table S4). Genes that were up- or downregulated only during the early or late stage were defined as differentially expressed. Genes that were differentially upregulated during the early stage were significantly enriched in terms of epigenetic mechanisms, including chromosome organization (GO:0051276), DNA repair (GO:0006281), peptidyl-amino acid modification (GO:0018193), protein acylation (GO:0043543), and histone modification (GO:0016570). In contrast, terms related to glucose (GO:0006,096 and GO:0009070) and lipid (GO:0042157, GO:0042158, and GO:1903509) metabolism were significantly enriched during the late stage, suggesting that the early and late stages play different roles in lichen symbiosis.

Glucose metabolism is important in lichen-forming fungi, which absorb photosynthetic products from their algal partners and convert them into glucose or fructose for fungal metabolism²⁶. Glycolysis/gluconeogenesis pathway analysis of the differentially expressed genes mapped only genes induced during the late stage to the pathways (Supplementary Fig. S13), indicating that conversion of the obtained monosaccharides into energy sources occurs actively during the late stage.

Small secreted proteins (SSPs) in lichen-forming fungi are involved in establishment and maintain the symbiosis.

Although small secreted proteins (SSPs) are virulence factors in pathogenic fungi and are important for symbiosis in mycorrhizal fungi⁵³, their roles in lichen-forming fungi remain unclear. We found that lichen-forming fungi had 286–482 secreted proteins and 107–207 SSPs (Supplementary Fig. S14); these numbers were smaller than those for other fungi with different lifestyles, especially plant-associated symbionts (Fig. 7A). Most SSPs found in lichen-forming fungi were genus- or species-specific according to the blast results (Fig. 7B; Supplementary Fig. S15). Sequence identity analysis of SSPs in *C. macilenta* revealed that 32% of the SSPs were *Cladonia*-specific and 39% were species-specific, whereas 43% of SSPs in *E. pusillum* R61883 were *Endocarpon*-specific and 31% were species-specific (Fig. 7B). In contrast, 76% of SSPs in *G. flavorubescens* and 75% in *U. muehlenbergii* were species-specific, as both species lacked closely related species, in our dataset (Supplementary Fig. S15). Although both of these lichen-forming fungal species belong to Lecanoromycetes, and

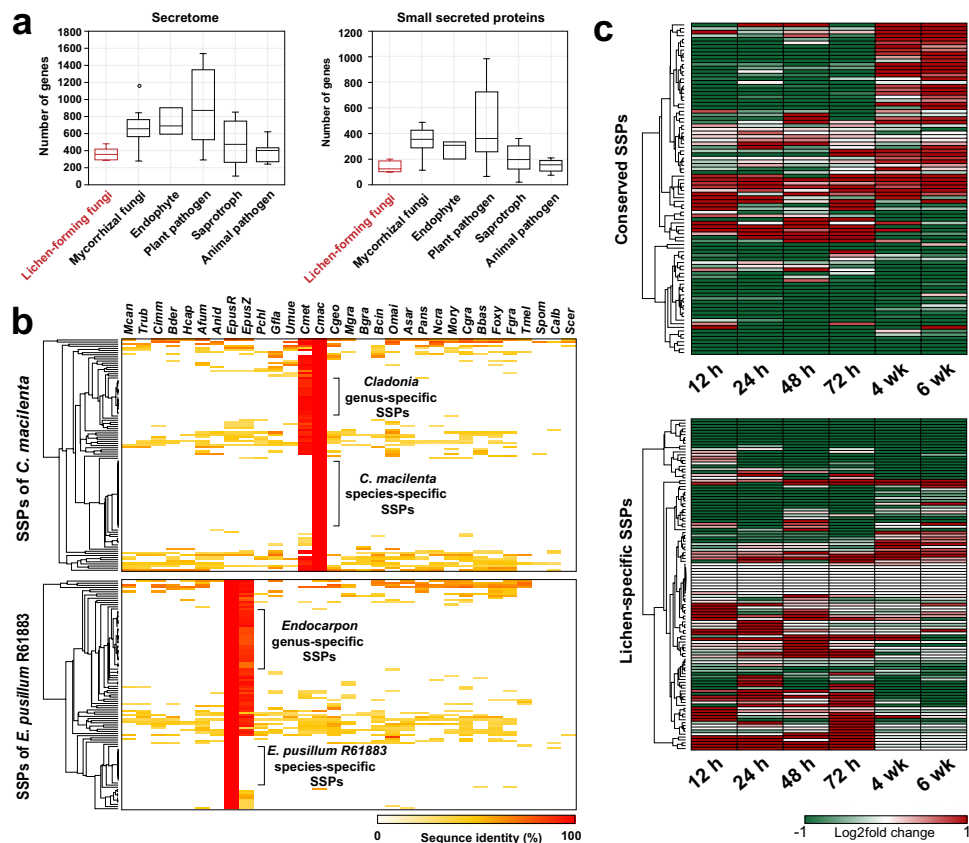


Figure 7. Lichen-specific small secreted proteins (SSPs) from *G. flavorubescens* induced in early lichen symbiosis. **(A)** Box plot of the number of secretomes and SSP distributions in lichen-forming fungi compared with other lifestyles. **(B)** Lichen-specific SSPs in *C. macilenta* and *E. pusillum* R61883. Ortholog SSPs of 56 fungal species were identified using blast and SSPs of *C. macilenta* and *E. pusillum* R61883 as references ($E = 1 \times 10^{-5}$). Abbreviations for fungal species are provided in Supplementary Dataset S1. **(C)** SSP expression in *G. flavorubescens* was classified as conserved (top) or lichen-specific (bottom).

both have *Trebouxia* spp. as algal partners, their SSPs differed markedly based on sequence similarity, suggesting that SSPs of lichen-forming fungi may not just dependent on their photobiont, but have been independently gained during speciation.

Like all genes, SSPs in *G. flavorubescens* had different expression profiles for conserved or specific genes. Most conserved SSPs were highly upregulated during the late stage of resynthesis, although some were also constitutively expressed during the early stages (Fig. 7C). However, most genus- and species-specific lichen SSPs were upregulated at different time points during the early stage.

Discussion

The main goal of symbiosis research is to determine how the beneficial associations evolved and to identify genes involved in the establishment and functioning of symbioses. However, to date, these features have been only partially investigated in lichen-forming fungi. All that is known about the evolution of lichen symbionts independently from non-lichenized ancestors was acquired from studies of small subunit and large subunit rDNA markers^{12,41}. We used whole-genome sequences to determine phylogenetic relationships more accurately. Our comparative analysis revealed that the lichen-forming fungi experienced massive reductions in unnecessary genes during symbiosis with their algal partners. Newly acquired lineage- and species-specific genes are involved in establishing lichen symbiosis, whereas conserved genes maintain the relationship.

PCWDEs are involved in host cell wall remodeling in mycorrhizal symbiosis⁵⁴ and confer virulence to fungal plant pathogens⁵⁵. However, we found that lichen-forming fungi experienced large contractions in PCWDE genes compared with their non-symbiotic ancestors. Pectin-degrading enzymes may no longer be necessary for algal host association in lichen symbiosis, because this cell wall component is unique to Charophyceae algae and land plants, but is not present in Chlorophyte green algae, which can form lichen⁴³. Consequently, most pectin-degrading enzyme genes have been lost in lichen-forming fungi derived from non-lichenized fungi. Nevertheless, cellulose and hemicellulose enzyme genes underwent contractions similar to pectin, although they are common cell wall components in both plants and green algae⁴³. Lichen-forming fungi have simple wall-to-wall apposition or develop highly differentiated non-breaking symbiotic structures called intraparietal haustoria, as well as intracellular haustoria, which penetrate algal cell walls⁵⁶. Although the haustoria of the lichen species

used in this study were not observed directly, we suggest that they do not penetrate the host algal cell walls for colonization, as proposed in previous studies^{56–58}. These symbiotic forms may lead to the non-functionalization of PCWDEs, which leads to gene loss⁵⁹. However, a recent study using a vast number of lichen-forming fungal genomes revealed that not all lichen-forming fungi lost large numbers of PCWDE genes²⁸. Taking this findings into consideration, the overall trend of our data suggests the loss of PCWDE genes in lichen, with several exceptional cases. Ectomycorrhizal fungi have similar symbiotic relationships with their hosts and have lost many PCWDE genes, unlike endomycorrhizal fungi and plant pathogens^{30,32,60}. Their symbiotic structure also does not penetrate the plant cell walls^{54,61}, which suggests that the loss of the ability to degrade cell walls in lichen-forming and mycorrhizal fungi is a consequence of their symbiotic fungus–host interface.

Many MFS-type transporters were also lost, although carbohydrate movement from the algal partner to the lichen-forming fungus is important in lichen symbiosis⁶². Although previous lichen genomic studies have also reported reductions in sugar transporters^{20,26}, we found that these losses were common in lichen lineages, not only in each species. Lichen-forming fungi likely use specific transporters, as they receive different mobile carbohydrates, such as ribitol, sorbitol, and glucose, from their algal partners⁴⁵. We propose that the extensive loss of sugar transporter genes is a result of the dispensability of common sugar transporters. The upregulation of ribitol transporter genes during the late stage of *G. flavorubescens* symbiosis supports this hypothesis, whereas other sugar transporters exhibited no significant changes. Algal partners export carbohydrates only during symbiosis⁶², such that the completion of lichen symbiosis and the initiation of nutrient exchange occur between 72 h and 4 weeks PCI.

As a result of contractions in diverse gene families (e.g., PCWDEs, sugar transporters, and TFs), the genome size and total number of genes of lichen-forming fungi are lower than those of other fungal species, especially plant-associated fungi. This is unsurprising because gene losses are widespread among all organisms⁵⁹, and genome reduction is a dominant evolutionary process resulting in the loss of non-functionalized genes⁶³. Symbionts have significantly reduced genomes due to their dependence on photosynthetic partners^{64,65}. The loss of energy-production genes in the mitochondrial genomes of lichen-forming fungi is an example of this reductive evolution⁶⁵. Although to date genome streamlining has been evaluated only in bacterial and mitochondrial genomes, we suggest that it can also occur in the nuclear genome. The loss of TF genes is another consequence of this evolutionary mechanism, and the loss of many genes and dependency on the host may influence the size of TF families.

Both massive gene losses and independent gene gains have occurred in lichen-forming fungi. Each species has many unique genes. In this study, we attempted to identify lichen-specific core genes, but found only one orthogroup. Similarly, mycorrhizal fungi lack universal symbiosis genes⁶⁰; instead, newly gained lichen-specific genes, including lineage- and species-specific genes, appear to be more related to their lifestyles, with high expression during early symbiosis, when they influence their partners¹⁴. Most SSPs were also genus- or species-specific and had similar expression patterns. The transient expression of each set of specific genes suggests that they are activated differently during each period of the early stage. Based on the sequential expression of effector proteins in the plant pathogen *Colletotrichum higginsianum*, different SSPs may be involved during each stage of host–pathogen interaction⁶⁶. Because the functions of most lichen-specific genes are unknown, and the SSPs of symbionts play essential roles in maintaining mycorrhizal symbiotic relationships^{33,34}, lineage- and species-specific genes may play significant roles in the establishment of lichen symbiosis. The functionally annotated conserved genes that may be involved in maintaining their symbiotic relationships were induced mainly during the late stage of lichen symbiosis, when recognition and contact with the partner are completed, growth is continued¹⁵, and metabolic processes such as nutrient exchange are activated. For example, genes involved in glycolysis are expressed differentially during the late stage of lichen symbiosis, allowing the use of sugars obtained from photosynthetic partners. These findings indicate that lichen-specific genes and conserved genes play roles in different stages of lichen resynthesis.

The evolutionary pattern of gene loss of lichen-forming fungi is similar to that of ectomycorrhizal fungi. Both symbionts lost the ability to degrade cell walls and gained lineage-specific genes that may be involved in symbiosis; this evolutionary process is well known in mycorrhizal fungi^{30,32}. However, ectomycorrhizal fungi still retain PCWDEs, including GH28, GH88, CE8, and GH30, which are induced in mycorrhizal symbiosis for host cell wall modification³⁰, whereas lichen-forming fungi have lost most of these genes. The number of effector proteins remains unknown⁶⁷. However, because lichen-forming fungi have fewer SSPs than mycorrhizal fungi or other plant-associated fungi, we suggest that the number of SSPs depends on the complexity of their host. Because green algae are the ancestors of land plants and have evolved to become more complex in terms of cellular organization⁶⁸, lichen-forming fungi may not require many SSPs to interact with the defense mechanism of their living host. Most gene family expansion occurred during the speciation of each mycorrhizal fungus (Supplementary Fig. S2), which suggests that lichen-forming fungi and mycorrhizal fungi underwent different unknown evolutionary processes to develop their lifestyles.

This study is the first comparative analysis of diverse lichen-forming fungi using whole genomes to clarify elements of lichen symbiosis. We found that the loss of non-essential genes, such as specific families of PCWDEs, sugar transporters, and TFs, streamlined the genomes of lichen-forming fungi, providing new insights on lichen symbiosis. Lineage- and species-specific genes, including SSPs, play a role during the early stage of lichen symbiosis, and may be involved in recognition between lichen-forming fungi and their partners. These findings advance understating of the evolution of symbiotic lifestyles and the determinants contributing to lichen symbiosis. Genomic resources may contribute to future molecular functional studies of the unrevealed biological functions of significant factors in lichen symbiosis.

Materials and methods

Genome resources of fungal species and ortholog clustering. We sequenced genome of five lichen-forming fungi including *G. flavorubescens* KoLRI002931 (accession no. AUPK01000000)³⁶, *C. macilenta* KoLRI003786 (AUPP01000000)³⁷, *C. metacoralifera* KoLRI002260 (AXCT02000000)³⁸, *U. muehlenbergii* KoLRI000956 (JFDN01000000)³⁹ and *E. pusillum* R61883 (JFDM01000000)⁴⁰. *E. pusillum* Z07020 (APWS00000000)²⁶ in previous researches and the other 50 fungal genomes used for comparative analysis were downloaded from Broad Institute (<http://www.broadinstitute.org/>), JGI fungal genome portal MycoCosm (<http://jgi.doe.gov/fungi>) and NCBI-GenBank database. The predicted protein sequences from 56 fungal genomes were clustered by OrthoFinder v2.2.7 with the default program settings⁶⁹. The conserved and lichen-specific genes were annotated by GO term annotation using InterProScan version 60⁷⁰.

Phylogenetic analysis and divergence time estimation. Total proteins of 56 fungal genomes were used to construct a whole genome-based phylogenomic tree using CVtree3 with k-tuple 7⁷¹. Initial curation of the divergence time for the major fungal taxa was achieved by Timetree⁷², and the divergence times were estimated by MCMCTree in PAML package version 4.8⁷³ using molecular markers, including actin (ACT1), translation elongation factor EF1- α (TEF1), RNA polymerase II large subunits (RPB1 and RPB2) and β -tubulins (TUB1 and TUB2). The final phylogenomic tree with divergence times was visualized by MEGA version 7.0.26⁷⁴.

Repetitive sequence and whole genome synteny analysis. The repeat contents were analyzed using TRF and rmBlastN, parts of RepeatMasker v4.0.5 package with RepBase 21.05 fungi library⁷⁵. For pairwise genomic comparisons, MUMmer v3.23⁷⁶ was used for aligning and comparing the whole genomes between lichen-forming fungi and the other fungal genomes.

Gene family evolution analysis and gene family annotation. CAFE (Computational analysis of gene family evolution) v2 was used to find out the gene families with significant changes in size ($P < 0.01$)⁷⁷. The time-calibrated phylogenetic tree and gene families identified by ortholog clustering were used for this analysis. Functional annotations of expanded and contracted gene families were identified by domain-based InterProScan v60⁷⁰. The cytochrome P450 genes were firstly identified with Fungal Cytochrome P450 Database (FCPD)⁴⁸, and then BLAST analysis against the P450 database in David Nelson cytochrome P450 web site⁷⁸ for nomenclature. The secondary metabolite biosynthesis gene clusters, including PKS, NRPS, and DMATs were identified by SMURF⁷⁹. Candidate MFS transporters were obtained using the Transporter Classification Database (TCDB)⁸⁰. Diverse polyol and monosaccharide transporters of *G. flavorubescens* were predicted by BLAST search (identity > 30, query coverage > 60) with functionally characterized transporter genes listed in⁸¹. The transcription factors were predicted by InterProScan v60 using the previously annotated DNA-binding domains⁴⁷, and PCWDE encoding CAZyme families were annotated by HMMER search against dbCAN CAZyme domain HMM database⁸². The secretomes and SSPs were predicted by the method described previously⁵³. The maximum-likelihood phylogenetic trees of CYP, PKS, PCWDEs, homeodomain-like and helix-turn-helix psq TF genes were constructed using RAxML version 8.2.9 with a bootstrap value of 1000⁸³. Aligning protein sequences using ClustalW 2.1⁸⁴ and remove poorly aligned regions by trimAl v1.2⁸⁵ were preceded before phylogenetic analysis. We reconciled the gene trees of PCWDEs resulting from this analysis with the species tree using NOTUNG 2.9⁸⁶.

RNA extraction and expression analysis. Actively growing *G. flavorubescens* KoLRI002931 mycelia were collected and macerated the mycelia into 10 mL of sterilized distilled water using a homogenizer (Ika, T10 basic, German). The macerated mycelia were dropped on malt extract agar medium (Difco), incubated at 15 °C for 4–6 weeks, and then covered 50 μ L of 2 weeks old *Trebouxia gelatinosa* cell suspension (1×10^8 /mL) which is partner alga of *G. flavorubescens*. The plates were incubated at 15 °C without light. For harvesting samples at different time points during re-synthesis between *G. flavorubescens* and *T. gelatinosa*, all samples were collected 0 h, 12 h, 24 h, 48 h, 72 h, 4 weeks, and 6 weeks after re-synthesis, immediately frozen using liquid nitrogen, and stored at -80 °C until processed. The whole samples on the medium were collected from three replicates of three biological repeats except 4 and 6 weeks. Total RNA was extracted using an Easy-Spin Total RNA Extraction Kit (iNtRON Biotechnology, Seoul, Korea). RNA sequencing was performed at Macrogen Inc. (Seoul, Korea) using Illumina HiSeq platform.

The NGS QC Toolkit ver. 2.3.3⁸⁷ was used to remove adaptors, low-quality sequences and sequences containing more than 5% N to obtain clean reads. Since the sequences of *G. flavorubescens* and *T. gelatinosa*, the partner alga, were mixed in the clean reads, the algal reads were eliminated using BWA (0.7.9a-r786)⁸⁸. Then the paired-end clean reads were aligned to *G. flavorubescens* genome using TopHat v2.0.12⁸⁹ and the gene expression levels were calculated as FPKM (Fragment Per Kilobase of transcript per Million mapped reads) using cufflink v2.2.1⁹⁰ and cuffdiff v2.21⁹¹. The FPKM value of 0 h, 12 h, 24 h, 48 h, and 72 h PCI were calculated with three biological repeats, and 4 weeks and 6 weeks PCI were calculated without repeats. Fold changes were calculated simply using a modified function, $\log_2\left(\frac{[\text{FPKM}_{\text{SYMBIOSIS}} + 1]}{[\text{FPKM}_{\text{MYCELLA}} + 1]}\right)$. The hierarchical clustering of protein expressions in the heatmaps was performed using the Euclidean clustering distance by Morpheus run by Broad Institute (<https://software.broadinstitute.org/morpheus>). Significantly enriched GO terms in symbiosis induced genes were identified using R package topGO version 2.38.1 with threshold $p < 0.05$ ⁹². KEGG pathway mapping analysis using differentially expressed genes was performed by KEGG Automatic Annotation Server (KAAS) web sites⁹³.

Data availability

The transcriptome sequence data of *G. flavorubescens* have been deposited in the NCBI Sequenced Read Archive (SRA) under accession number PRJNA210248.

Received: 23 March 2022; Accepted: 6 June 2022

Published online: 24 June 2022

References

- Grube, M. & Wedin, M. Lichenized fungi and the evolution of symbiotic organization. *Microbiol. Spectr.* <https://doi.org/10.1128/microbiolspec.FUNK-0011-2016> (2016).
- Nash, T. H. *Lichen Biology*. 2nd edn, (Cambridge University Press, 2008).
- Honegger, R. Simon Schwendener (1829–1919) and the dual hypothesis of lichens. *Bryologist* **103**, 307–313 (2000).
- Spribile, T. *et al.* Basidiomycete yeasts in the cortex of ascomycete macrolichens. *Science* **353**, 488–492. <https://doi.org/10.1126/science.aaf8287> (2016).
- Millanes, A. M., Diederich, P. & Wedin, M. *Cyphobasidium* gen. nov., a new lichen-inhabiting lineage in the Cystobasidiomycetes (Pucciniomycotina, Basidiomycota, Fungi). *Fungal Biol.* **120**, 1468–1477 (2016).
- Černajová, I. & Škaloud, P. The first survey of Cystobasidiomycete yeasts in the lichen genus *Cladonia*; with the description of *Lichenzyma pisutiana* gen. nov., sp. nov. *Fungal Biol.* **123**, 625–637 (2019).
- Tuovinen, V. *et al.* Two basidiomycete fungi in the cortex of wolf lichens. *Curr. Biol.* **29**, 476–483.e475 (2019).
- Hawksworth, D. L. & Grube, M. Lichens redefined as complex ecosystems. *New Phytol.* **227**, 1281 (2020).
- Honegger, R. Water relations in lichens. In *Fungi in the Environment*, 185–200 (2006).
- Kranner, I., Beckett, R., Hochman, A. & Nash, T. H. Desiccation-tolerance in lichens: A review. *Bryologist* **111**, 576–593. <https://doi.org/10.1639/0007-2745-111.4.576> (2008).
- Sanders, W. B. & Masumoto, H. Lichen algae: The photosynthetic partners in lichen symbioses. *Lichenologist* **53**, 347–393 (2021).
- Gargas, A., DePriest, P. T., Grube, M. & Tehler, A. Multiple origins of lichen symbioses in fungi suggested by SSU rDNA phylogeny. *Science* **268**, 1492–1495 (1995).
- Nelsen, M. P., Lucking, R., Boyce, C. K., Lumbsch, H. T. & Ree, R. H. The macroevolutionary dynamics of symbiotic and phenotypic diversification in lichens. *Proc. Natl. Acad. Sci. USA*. **117**, 21495–21503. <https://doi.org/10.1073/pnas.2001913117> (2020).
- Ahmadjian, V., Jacobs, J. B. & Russell, L. A. Scanning electron microscope study of early lichen synthesis. *Science* **200**, 1062–1064. <https://doi.org/10.1126/science.200.4345.1062> (1978).
- Athukorala, S. N., Huebner, E. & Piercey-Normore, M. D. Identification and comparison of the 3 early stages of resynthesis for the lichen *Cladonia rangiferina*. *Can. J. Microbiol.* **60**, 41–52. <https://doi.org/10.1139/cjm-2013-0313> (2014).
- Trembley, M. L., Ringli, C. & Honegger, R. Morphological and molecular analysis of early stages in the resynthesis of the lichen *Baeomyces rufus*. *Mycol. Res.* **106**, 768–776 (2002).
- Joneson, S., Armaleo, D. & Lutzoni, F. Fungal and algal gene expression in early developmental stages of lichen-symbiosis. *Mycologia* **103**, 291–306 (2011).
- Park, S.-Y. *et al.* Agrobacterium tumefaciens-mediated transformation of the lichen fungus *Umbilicaria muehlenbergii*. *PLoS One* **8**, e83896 (2013).
- Liu, R., Kim, W., Paguirigan, J. A. A., Jeong, M.-H. & Hur, J.-S. Establishment of *Agrobacterium tumefaciens*-mediated transformation of *Cladonia macilenta*, a model lichen-forming fungus. *J. Fungi* **7**, 252 (2021).
- Armaleo, D. *et al.* The lichen symbiosis re-viewed through the genomes of *Cladonia grayi* and its algal partner *Asterochloris glomerata*. *BMC Genom.* **20**, 1–33 (2019).
- Armstrong, E. E. *et al.* Draft genome sequence and annotation of the lichen-forming fungus *Arthonia radiata*. *Genome Announc.* **6**, e00281–e218 (2018).
- Dal Grande, F. *et al.* The draft genome of the lichen-forming fungus *Lasallia hispanica* (Frey) Sancho & A Crespo. *The Lichenologist* **50**, 329–340 (2018).
- Mead, O. L. & Gueidan, C. Complete genome sequence of an Australian strain of the lichen-forming fungus *Endocarpon pusillum* (Hedwig). *Microbiol. Resour. Announc.* **9**, e01079–e1020 (2020).
- Pizarro, D. *et al.* Whole-genome sequence data uncover widespread heterothallism in the largest group of lichen-forming fungi. *Genome Biol. Evol.* **11**, 721–730 (2019).
- Wang, Y., Yuan, X., Chen, L., Wang, X. & Li, C. Draft genome sequence of the lichen-forming fungus *Ramalina intermedia* strain YAF0013. *Genome Announc.* **6**, e00478–e418 (2018).
- Wang, Y. Y. *et al.* Genome characteristics reveal the impact of lichenization on lichen-forming fungus *Endocarpon pusillum* Hedwig (Verrucariales, Ascomycota). *BMC Genom.* **15**, 34. <https://doi.org/10.1186/1471-2164-15-34> (2014).
- Pizarro, D. *et al.* Genome-wide analysis of biosynthetic gene cluster reveals correlated gene loss with absence of usnic acid in lichen-forming fungi. *Genome Biol. Evol.* **12**, 1858–1868 (2020).
- Resl, P. *et al.* Large differences in carbohydrate degradation and transport potential in the genomes of lichen fungal symbionts. *bioRxiv* **44**, 17 (2021).
- Nazem-Bokae, H., Hom, E. F., Warden, A. C., Mathews, S. & Gueidan, C. Towards a systems biology approach to understanding the lichen symbiosis: opportunities and challenges of implementing network modelling. *Front. Microbiol.* **12**, 1028 (2021).
- Kohler, A. *et al.* Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nat. Genet.* **47**, 410–415. <https://doi.org/10.1038/ng.3223> (2015).
- Pellegrin, C., Morin, E., Martin, F. M. & Veneault-Fourrey, C. Comparative analysis of secretomes from ectomycorrhizal fungi with an emphasis on small-secreted proteins. *Front. Microbiol.* **6**, 1278 (2015).
- Miyachi, S. *et al.* Large-scale genome sequencing of mycorrhizal fungi provides insights into the early evolution of symbiotic traits. *Nat. Commun.* **11**, 5125. <https://doi.org/10.1038/s41467-020-18795-w> (2020).
- Kloppholz, S., Kuhn, H. & Requena, N. A secreted fungal effector of *Glomus intraradices* promotes symbiotic biotrophy. *Curr. Biol.* **21**, 1204–1209. <https://doi.org/10.1016/j.cub.2011.06.044> (2011).
- Plett, J. M. *et al.* A secreted effector protein of *Laccaria bicolor* is required for symbiosis development. *Curr. Biol.* **21**, 1197–1203. <https://doi.org/10.1016/j.cub.2011.05.033> (2011).
- Tsuzuki, S., Handa, Y., Takeda, N. & Kawaguchi, M. Strigolactone-induced putative secreted protein 1 is required for the establishment of symbiosis by the arbuscular mycorrhizal fungus *Rhizophagus irregularis*. *Mol. Plant Microbe Interact.* **29**, 277–286. <https://doi.org/10.1094/MPMI-10-15-0234-R> (2016).
- Park, S. Y. *et al.* Draft genome sequence of lichen-forming fungus *Caloplaca flavorubescens* strain KoLRI002931. *Genome Announc.* <https://doi.org/10.1128/genomeA.00678-13> (2013).
- Park, S. Y. *et al.* Draft genome sequence of *Cladonia macilenta* KoLRI003786, a lichen-forming fungus producing biruloquinone. *Genome Announc.* <https://doi.org/10.1128/genomeA.00695-13> (2013).
- Park, S. Y. *et al.* Draft genome sequence of lichen-forming fungus *Cladonia metacorallifera* strain KoLRI002260. *Genome Announc.* <https://doi.org/10.1128/genomeA.01065-13> (2014).

39. Park, S. Y. *et al.* Draft genome sequence of *Umbilicaria muehlenbergii* KoLRILF000956, a Lichen-forming fungus amenable to genetic manipulation. *Genome Announc.* <https://doi.org/10.1128/genomeA.00357-14> (2014).
40. Park, S. Y. *et al.* Draft genome sequence of *Endocarpon pusillum* strain KoLRILF000583. *Genome Announc.* <https://doi.org/10.1128/genomeA.00452-14> (2014).
41. Lutzoni, F., Pagel, M. & Reeb, V. Major fungal lineages are derived from lichen symbiotic ancestors. *Nature* **411**, 937–940. <https://doi.org/10.1038/35082053> (2001).
42. Lionetti, V. & Metraux, J. P. Plant cell wall in pathogenesis, parasitism and symbiosis. *Front. Plant Sci.* **5**, 612. <https://doi.org/10.3389/fpls.2014.00612> (2014).
43. Popper, Z. A. *et al.* Evolution and diversity of plant cell walls: From algae to flowering plants. *Annu. Rev. Plant Biol.* **62**, 567–590. <https://doi.org/10.1146/annurev-arplant-042110-103809> (2011).
44. Pao, S. S., Paulsen, I. T. & Saier, M. H. Jr. Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **62**, 1–34 (1998).
45. Richardson, D. H., Hill, D. J. & Smith, D. C. Lichen physiology, 11: Role of alga in determining pattern of carbohydrate movement between lichen symbionts. *New Phytol.* **67**, 469. <https://doi.org/10.1111/j.1469-8137.1968.tb05476.x> (1968).
46. Hill, D. J. & Ahmadjian, V. Relationship between carbohydrate movement and symbiosis in lichens with green-algae. *Planta* **103**, 267. <https://doi.org/10.1007/Bf00386850> (1972).
47. Shelest, E. Transcription factors in fungi: TFome dynamics, three major families, and dual-specificity TFs. *Front. Genet.* **8**, 53. <https://doi.org/10.3389/fgene.2017.00053> (2017).
48. Park, J. *et al.* Fungal cytochrome P450 database. *BMC Genom.* **9**, 402. <https://doi.org/10.1186/1471-2164-9-402> (2008).
49. Shin, J., Kim, J. E., Lee, Y. W. & Son, H. Fungal cytochrome P450s and the P450 complement (CYPome) of *Fusarium graminearum*. *Toxins*. <https://doi.org/10.3390/toxins10030112> (2018).
50. Boustie, J. & Grube, M. Lichens—A promising source of bioactive secondary metabolites. *Plant Genet. Res.* **3**, 273–287 (2005).
51. Calcott, M. J., Ackerley, D. F., Knight, A., Keyzers, R. A. & Owen, J. G. Secondary metabolism in the lichen symbiosis. *Chem. Soc. Rev.* **47**, 1730–1760 (2018).
52. Ranković, B. *Lichen Secondary Metabolites* (Springer International Publishing, 2015).
53. Kim, K. T. *et al.* Kingdom-wide analysis of fungal small secreted proteins (SSPs) reveals their potential role in host association. *Front. Plant Sci.* **7**, 186. <https://doi.org/10.3389/fpls.2016.00186> (2016).
54. Balestrini, R. & Bonfante, P. Cell wall remodeling in mycorrhizal symbiosis: A way towards biotrophism. *Front. Plant Sci.* **5**, 237. <https://doi.org/10.3389/fpls.2014.00237> (2014).
55. Kubicek, C. P., Starr, T. L. & Glass, N. L. Plant cell wall-degrading enzymes and their secretion in plant-pathogenic fungi. *Annu. Rev. Phytopathol.* **52**, 427–451. <https://doi.org/10.1146/annurev-phyto-102313-045831> (2014).
56. Honegger, R. Ultrastructural studies in lichens. *New Phytol.* **103**, 797–808 (1986).
57. Honegger, R. Cytological aspects of the mycobiont–phycobiont relationship in lichens: Haustorial types, phycobiont cell wall types, and the ultrastructure of the cell surface layers in some cultured and symbiotic myco- and phycobionts. *Lichenologist* **16**, 111–127 (1984).
58. Valladares, F., Wierzchos, J. & Ascaso, C. Porosimetric study of the lichen family Umbilicariaceae: Anatomical interpretation and implications for water storage capacity of the thallus. *Am. J. Bot.* **80**, 263–272 (1993).
59. Albalat, R. & Canestro, C. Evolution by gene loss. *Nat. Rev. Genet.* **17**, 379–391. <https://doi.org/10.1038/nrg.2016.39> (2016).
60. Kuo, A., Kohler, A., Martin, F. M. & Grigoriev, I. V. Expanding genomics of mycorrhizal symbiosis. *Front. Microbiol.* **5**, 582. <https://doi.org/10.3389/fmicb.2014.00582> (2014).
61. Bonfante, P. & Genre, A. Mechanisms underlying beneficial plant-fungus interactions in mycorrhizal symbiosis. *Nat. Commun.* **1**, 48. <https://doi.org/10.1038/ncomms1046> (2010).
62. Smith, D. The movement of carbohydrate from alga to fungus in lichens. *Bull. Soc. Bot. Fr.* **115**, 129–133 (1968).
63. Wolf, Y. I. & Koonin, E. V. Genome reduction as the dominant mode of evolution. *BioEssays* **35**, 829–837. <https://doi.org/10.1002/bies.201300037> (2013).
64. McCutcheon, J. P. & Moran, N. A. Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* **10**, 13–26. <https://doi.org/10.1038/nrmicro2670> (2011).
65. Pogoda, C. S., Keepers, K. G., Lendemer, J. C., Kane, N. C. & Tripp, E. A. Reductions in complexity of mitochondrial genomes in lichen-forming fungi shed light on genome architecture of obligate symbioses. *Mol. Ecol.* **27**, 1155–1169. <https://doi.org/10.1111/mec.14519> (2018).
66. Kleemann, J. *et al.* Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen *Colletotrichum higginsianum*. *PLoS Pathog.* **8**, e1002643. <https://doi.org/10.1371/journal.ppat.1002643> (2012).
67. Harris, J. M. *et al.* What are the top 10 unanswered questions in molecular plant-microbe interactions?. *Mol. Plant Microbe Interact.* **33**, 1354 (2020).
68. Turmel, M., Pombert, J. F., Charlebois, P., Otis, C. & Lemieux, C. The green algal ancestry of land plants as revealed by the chloroplast genome. *Int. J. Plant Sci.* **168**, 679–689. <https://doi.org/10.1086/513470> (2007).
69. Emms, D. M. & Kelly, S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157. <https://doi.org/10.1186/s13059-015-0721-2> (2015).
70. Jones, P. *et al.* InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031> (2014).
71. Zuo, G. & Hao, B. CVTree3 web server for whole-genome-based and alignment-free prokaryotic phylogeny and taxonomy. *Genomics Proteomics Bioinform.* **13**, 321–331. <https://doi.org/10.1016/j.gpb.2015.08.004> (2015).
72. Hedges, S. B., Marin, J., Suleski, M., Paymer, M. & Kumar, S. Tree of life reveals clock-like speciation and diversification. *Mol. Biol. Evol.* **32**, 835–845. <https://doi.org/10.1093/molbev/msv037> (2015).
73. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591. <https://doi.org/10.1093/molbev/msm088> (2007).
74. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874. <https://doi.org/10.1093/molbev/msw054> (2016).
75. Smit, A., Hubley, R., & Green, P. 2013–2015. RepeatMasker Open-4.0 (2015).
76. Kurtz, S. *et al.* Versatile and open software for comparing large genomes. *Genome Biol.* **5**, R12. <https://doi.org/10.1186/gb-2004-5-2-r12> (2004).
77. Han, M. V., Thomas, G. W., Lugo-Martinez, J. & Hahn, M. W. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol. Biol. Evol.* **30**, 1987–1997. <https://doi.org/10.1093/molbev/mst100> (2013).
78. Nelson, D. R. The cytochrome p450 homepage. *Hum. Genomics* **4**, 59–65. <https://doi.org/10.1186/1479-7364-4-1-59> (2009).
79. Fedorova, N. D., Moktali, V. & Medema, M. H. Bioinformatics approaches and software for detection of secondary metabolic gene clusters. *Methods Mol. Biol.* **944**, 23–45. https://doi.org/10.1007/978-1-62703-122-6_2 (2012).
80. Saier, M. H. Jr. *et al.* The Transporter classification database (TCDB): Recent advances. *Nucleic Acids Res.* **44**, D372–379. <https://doi.org/10.1093/nar/gkv1103> (2016).
81. Yoshino, K. *et al.* The conservation of polyol transporter proteins and their involvement in lichenized Ascomycota. *Fungal Biol.* **123**, 318–329. <https://doi.org/10.1016/j.funbio.2019.01.006> (2019).
82. Yin, Y. *et al.* dbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **40**, W445–451. <https://doi.org/10.1093/nar/gks479> (2012).

83. Stamatakis, A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033> (2014).
84. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948. <https://doi.org/10.1093/bioinformatics/btm404> (2007).
85. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973. <https://doi.org/10.1093/bioinformatics/btp348> (2009).
86. Darby, C. A., Stolzer, M., Ropp, P. J., Barker, D. & Durand, D. Xenolog classification. *Bioinformatics* **33**, 640–649. <https://doi.org/10.1093/bioinformatics/btw686> (2017).
87. Patel, R. K. & Jain, M. NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. *PLoS ONE* **7**, e30619. <https://doi.org/10.1371/journal.pone.0030619> (2012).
88. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324> (2009).
89. Kim, D. *et al.* TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36. <https://doi.org/10.1186/gb-2013-14-4-r36> (2013).
90. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515. <https://doi.org/10.1038/nbt.1621> (2010).
91. Trapnell, C. *et al.* Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* **31**, 46–53. <https://doi.org/10.1038/nbt.2450> (2013).
92. Alexa, A. & Rahnenfuhrer, J. topGO: Enrichment analysis for Gene Ontology. R package version 2.28.0. *BioConductor*. Published online (2016).
93. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C. & Kanehisa, M. KAAAS: An automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* **35**, W182–185. <https://doi.org/10.1093/nar/gkm321> (2007).

Acknowledgements

We thank Bongsu Choi at Suncheon National University for the maintenance of the server containing the transcriptome data of lichen forming-fungi.

Author contributions

H.S. and Y-H.L. designed the research; H.S., K-T.K., G-W.L., J.C., J.J., K.C., and G.C. performed bioinformatics and genomic analysis; H.S. and G-W.L. performed transcriptome analysis; S-Y.P. and J-S.H. provided the data; H.S., K-T.K. and Y-H.L. wrote the manuscript.

Funding

This work was supported by the National Research Foundation of Korea (NRF) grants funded by Ministry of Science and ICT (MSIT) (2020R1A2B5B03096402, 2018R1A5A1023599, 2021M3H9A1096935, and 2017M3A9B8069471).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-14340-5>.

Correspondence and requests for materials should be addressed to Y.-H.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022