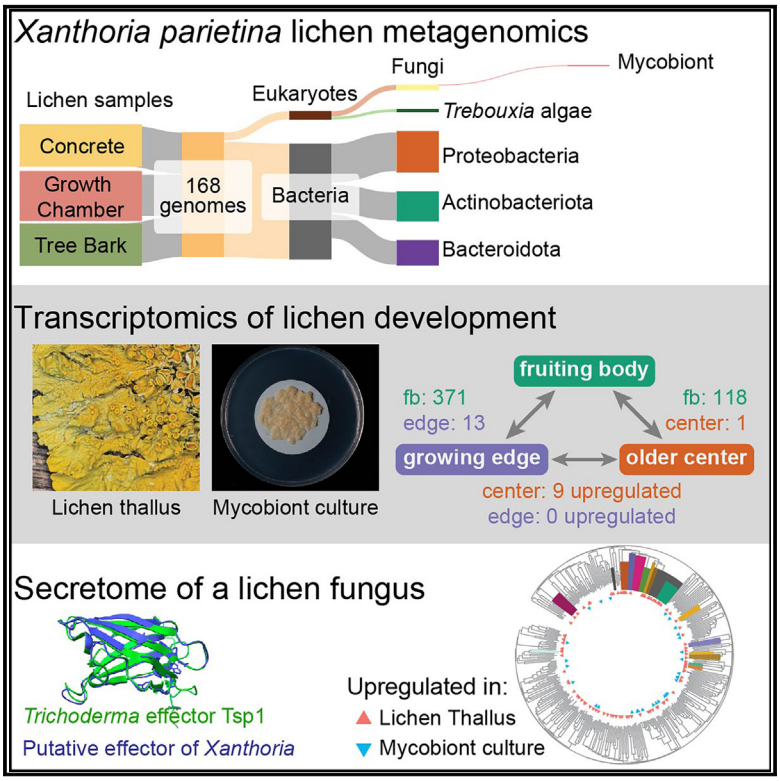
Article

Current Biology

Complexity of the lichen symbiosis revealed by metagenome and transcriptome analysis of *Xanthoria parietina*

Graphical abstract



Highlights

- Metagenomes of *Xanthoria parietina* lichens contain over 150 genomes
- Transcriptomics identifies gene functions associated with the symbiotic state
- The gene expression of a lichen fungus differs between developmental stages
- Potential effectors present in the secretome of a lichen fungus

Authors

Gulnara Tagirdzhanova, Klara Scharnagl, Neha Sahu, ..., Hanna Johannesson, Dan MacLean, Nicholas J. Talbot

Correspondence

nick.talbot@tsl.ac.uk

In brief

Lichens are symbiotic associations of fungi, algae, and bacteria that result in large, self-assembling organisms. Tagirdzhanova et al. use metagenomics to profile taxonomic diversity within the sunburst lichen and apply metatranscriptomics and protein structural modeling to identify functions critical for the symbiotic development of lichen fungi.

Tagirdzhanova et al., 2025, Current Biology *35*, 799–817
February 24, 2025
https://doi.org/10.1016/j.cub.2024.12.041

CellPress

Current Biology

CellPress
OPEN ACCESS

Article

Complexity of the lichen symbiosis revealed by metagenome and transcriptome analysis of *Xanthoria parietina*

Gulnara Tagirdzhanova,[1,8,10] Klara Scharnagl,[1,2] Neha Sahu,[1] Xia Yan,[1] Angus Bucknell,[1] Adam R. Bentham,[1,7] Clara Jégousse,[1] Sandra Lorena Ament-Velásquez,[3] Ioana Onuț-Brännström,[4] Hanna Johannesson,[5,6] Dan MacLean,[1] and Nicholas J. Talbot[1,9,11,12,*]

[1]The Sainsbury Laboratory, University of East Anglia, Norwich Research Park, Colney Lane, Norwich NR47UH, UK
[2]University & Jepson Herbaria, University of California, Berkeley, Valley Life Sciences Building, Berkeley, CA 94720, USA
[3]Department of Zoology, Stockholm University, Stockholm 106 91, Sweden
[4]Department of Ecology and Genetics, Uppsala University, Norbyv. 18D, Uppsala 752 36, Sweden
[5]Department of Ecology, Environmental and Plant Sciences, Stockholm University, Stockholm 106 91, Sweden
[6]The Royal Swedish Academy of Sciences, Lilla Frescativägen 4A, Stockholm 114 18, Sweden
[7]Present address: Centre for Programmable Biological Matter, Department of Biosciences, Durham University, Durham DH1 3LE, United Kingdom
[8]X (formerly Twitter): @metalichen
[9]X (formerly Twitter): @talbotlabTSL
[10]Bluesky: metalichen.bsky.social
[11]Bluesky: talbotlabtsl.bsky.social
[12]Lead contact
*Correspondence: nick.talbot@tsl.ac.uk
https://doi.org/10.1016/j.cub.2024.12.041

SUMMARY

Lichens are composite, symbiotic associations of fungi, algae, and bacteria that result in large, anatomically complex organisms adapted to many of the world's most challenging environments. How such intricate, self-replicating lichen architectures develop from simple microbial components remains unknown because of their recalcitrance to experimental manipulation. Here, we report a metagenomic and metatranscriptomic analysis of the lichen *Xanthoria parietina* at different developmental stages. We identified 168 genomes of symbionts and lichen-associated microbes across the sampled thalli, including representatives of green algae, three different classes of fungi, and 14 bacterial phyla. By analyzing the occurrence of individual species across lichen thalli from diverse environments, we defined both substrate-specific and core microbial components of the lichen. Metatranscriptomic analysis of the principal fungal symbiont from three different developmental stages of a lichen, compared with axenically grown fungus, revealed differential gene expression profiles indicative of lichen-specific transporter functions, specific cell signaling, transcriptional regulation, and secondary metabolic capacity. Putative immunity-related proteins and lichen-specific structurally conserved secreted proteins resembling fungal pathogen effectors were also identified, consistent with a role for immunity modulation in lichen morphogenesis.

INTRODUCTION

Symbiosis is one of the most widespread and successful lifestyle strategies for biological organisms. The term "symbiosis" was first coined to describe lichens: long thought to be a single organism, lichens were revealed instead to be the result of a stable relationship between a fungus and one or multiple photosynthetic microorganisms.[1] A unique and defining feature of the lichen symbiosis is a new body plan that arises only from the interaction. Stable and self-replicating over generations, the lichen phenotype does not resemble that of any of the symbionts grown in isolation. Lichen symbionts interact to create a single body (a thallus), which is often structurally complex and organized into multiple tissue-like layers. A major role in lichen development is believed to belong to the mycobiont—the fungal symbiont that contributes the vast majority of lichen biomass. Interwoven and glued together with extracellular matrix, mycobiont hyphae create the tough outer layers of the lichen thallus, with photosynthetic symbionts (photobionts) typically inhabiting the layer beneath, where they can take advantage of sunlight.[2] In addition to the mycobiont and photobiont, many lichens contain additional microorganisms, chiefly bacteria and yeasts, at least some of which are stably associated with lichens.[3,4]

The molecular mechanisms required for lichen development and growth remain unknown. Although we can hypothesize that some may be like those involved in the development of complex fungal structures, such as mushrooms, this hypothesis needs to be tested, and we also need to explain the remarkable

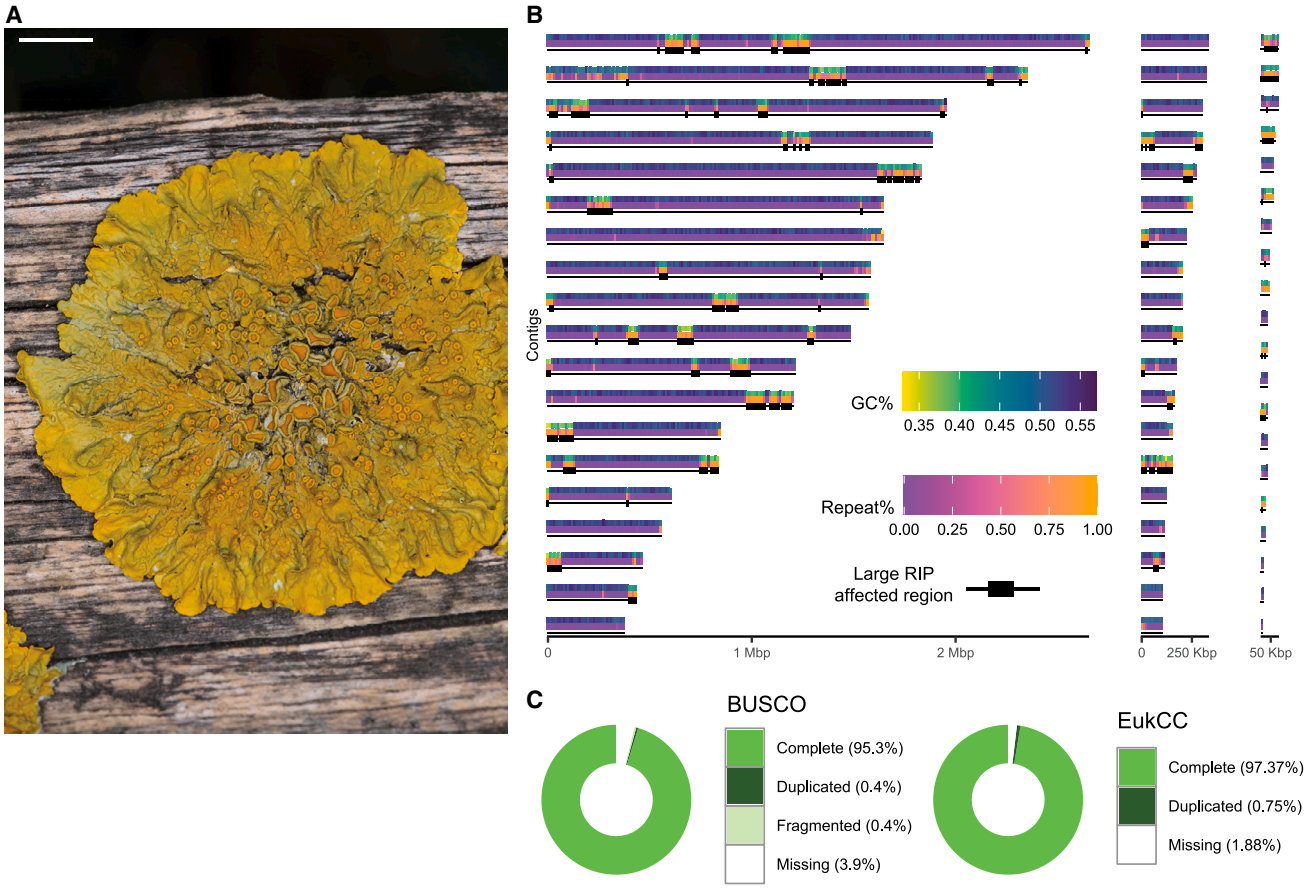


Figure 1. The genome of *Xanthoria parietina* mycobiont

(A) *X. parietina* thallus on wood. Scale bar, 5 mm.

(B) *X. parietina* mycobiont nuclear genome. Each contig is represented by three annotation tracks: GC content, repeat content, and presence of large RIP-affected regions (LRARs). The x axis corresponds to contig length. Red asterisks show telomeric repeats.

(C) Genome completeness scores estimated by BUSCO5 (ascomycota_odb10 database) and EukCC2 (NCBI 78060 Parmeliaceae database).

See also Figure S1 and Data S1.

coordination of growth between symbionts. The reason behind such limited knowledge of lichen symbiotic development lies in the recalcitrance of lichens toward laboratory experimentation. Individual symbionts often grow extremely slowly[5] and, with one exception, have never been genetically modified. The only exception, the mycobiont of *Umbilicaria* lichens, is highly unusual in its dimorphic growth habit,[6] which makes it easier to manipulate but also raises questions in respect of whether its study is applicable to other lichens. Lichen phenotypes cannot be recreated from axenic cultures in the lab, leaving us with no mechanistic insight into lichen development.

In this report, we use metagenomics and metatranscriptomics to characterize a lichen symbiosis and identify processes involved in symbiosis maintenance and development. As a model, we used *Xanthoria parietina*—a widespread lichen that has served as a model system in studies of lichen anatomy and population genetics.[7,8] *X. parietina* is believed to have no vertical co-transmission of symbionts, which disperse on their own. Hence, germinating sexual spores of the mycobiont must establish connection to a *Trebouxia* photobiont and, potentially, other members of the lichen microbiota every time a new lichen forms. We establish *X. parietina* as our model system by analyzing the genome of its mycobiont and by characterizing the diversity of microorganisms present in lichen samples. We compare mycobiont gene expression between intact lichen thalli and lab cultures and, for the first time, use different parts of lichen thalli as a proxy for developmental stages, which we compare to identify genes and molecular processes involved in lichen morphogenesis. Finally, we perform the first in-depth analysis of a lichen mycobiont secretome and identify potential symbiosis-associated lichen effector proteins.

RESULTS

Organization of the *Xanthoria* mycobiont genome

We first generated a reference genome of the *X. parietina* mycobiont. Long-read metagenomic data from an *X. parietina* thallus collected at the Norwich Research Park (Figure 1A) yielded a high-quality genome assembly of the mycobiont. Data were assembled and binned to remove sequences from any organism

other than the mycobiont. The final mycobiont genome assembly consisted of 58 contigs for a total of 29.96 Mbp, and the N50 of the assembly equaled 1.59 Mbp (Figure 1B). The assembly had completeness scores of 96.1%, according to BUSCO5 (benchmarking universal single-copy orthologs), and of 98.1%, according to EukCC2, where completeness score is defined as 100% minus the percentage of missing markers (Figure 1C). *De novo* annotation of the genome resulted in 10,727 gene models and 11,185 transcripts. The genome size and completeness and annotation statistics are consistent with other high-quality genomes from the class Lecanoromycetes published to date.[9–12] In the genome, we identified 59 biosynthetic gene clusters (Data S1A).

The repeat content of the genome was 12.7%, with long terminal repeat elements accounting for nearly half the repeated component of the genome (Data S1B). Repeats were not evenly spread across contigs, instead forming regional clusters that corresponded to genome regions with lower GC content (Figure 1B). By screening the genome for signatures of repeat-induced point mutation (RIP),[13] we discovered that these low-GC/high-repeat regions can also be considered large RIP-affected regions (LRARs) (Figure 1B). In total, we identified 158 LRARs that account for 8.5% of the genome (Data S1C).

To predict the ploidy level of the *Xanthoria* mycobiont, we analyzed minor allele frequency (MAF) distribution in eight metagenomic samples (see below) using the newly produced mycobiont genome for variant calling. Although some samples showed haploid-like patterns, other patterns had unusually high numbers of peaks in the distribution (Figure S1), consistent with neither haploid, diploid, or triploid signals. This pattern might be explained by the presence of multiple mycobiont genotypes within one thallus, which might be caused by fusion of several genetically distinct thalli or by the presence of multiple paternal genotypes from sexual recombination in separate apothecia.

168 MAGs can be isolated from *Xanthoria* metagenomes

To characterize the organismal composition of *X. parietina* thalli and account for all species detected in shotgun sequencing experiments, we generated eight deeply sequenced metagenomes (minimum of 24.7 Gbp of raw data and 85x mycobiont genome coverage) from samples of *X. parietina* collected from Norwich Research Park from different substrates and growth conditions: concrete (n = 2), tree bark (collected fresh, n = 3), and tree bark (incubated in a growth chamber for 19–21 months, n = 3). From these metagenomes, we extracted and annotated 168 medium- and high-quality non-redundant metagenome-assembled genomes (MAGs), each corresponding to a distinct species-level lineage (Figure 2A; Data S2A; see STAR Methods for details of MAG filtering and dereplication). All eleven eukaryotic MAGs belong to either fungi or algae. The seven fungal MAGs include the *X. parietina* mycobiont (Figures 2B and S1A) and three distantly related mycobionts of other lichen symbioses from classes Lecanoromycetes and Lichinomycetes (Figure S2A). These genomes were likely obtained due to propagules of these fungi on the surface of *X. parietina* samples. In addition, three MAGs of Chaetothyriales (Eurotiomycetes), a group of black yeasts reported from various lichens as potential endophytes or parasites,[14–16] were detected in three of the eight *Xanthoria* samples. All four algal MAGs belonged to different strains of *Trebouxia* (Figure 2C), the previously reported photobiont of *X. parietina* lichen.[7] The remaining 157 MAGs are shared between 14 bacterial phyla, with 59% from just two phyla: Proteobacteria and Actinobacteriota (Data S2A; Figure 2D). The two bacterial genera with most MAGs were *Sphingomonas* (Sphingomonadaceae, Alphaproteobacteria; n = 18) and clade CAHJXG01 (Acetobacteriaceae, Alphaproteobacteria; n = 9) (Data S2A).

Next, we mapped the presence/absence of each MAG across eight *X. parietina* samples by mapping metagenomic reads onto the MAG catalog. We compared lichen samples collected from different substrates: concrete, tree bark (collected fresh), and tree bark (incubated in a growth chamber for 19–21 months). Clustering lichen samples based on the occurrence matrix revealed that samples collected from lichens growing on concrete differed from bark samples (Figure 3A). Concrete samples also had the highest number of unique MAGs (Figure 3B). The role of growth substrate in determining taxonomic composition of lichen-associated microorganisms is also confirmed by our analysis of an additional sample of a different *Xanthoria* species, *X. calcicola*, collected from concrete. This sample clustered with *X. parietina* samples from concrete and shared most lineages present in these samples (Figures S2B and S2C). At the same time, differences between substrates were lower when diversity was considered at higher taxonomic levels: the percentage of lineages present in all three substrate types, for instance, was 19% for species-level lineages and 30% for family-level lineages (Figure 3B).

To assess the impact of long-term incubation of lichen samples in the growth chamber, we compared fresh lichen samples collected from bark to the samples incubated for 19–21 months. Surprisingly, the taxonomic composition did not appear affected, as incubated samples had profiles similar to those of the fresh samples collected from the same substrate (Figures 3A and 3B). The two genus-level bacterial clades unique to growth chamber samples are an Acetobacteraceae clade LMUY01 and *Friedmanniella*, both previously reported from lichen metagenomes,[17,18] making it plausible that these originated from within the thallus. At the same time, the total number of MAGs per sample in growth chamber samples was lower than in fresh samples with comparable sequencing depth, potentially due to a decrease of bacterial diversity on the surface of lichen samples.

In addition to lineages present occasionally or in one type of substrate only, we detected generalist lineages present in all surveyed lichen thalli. As expected, every sample contained one mycobiont MAG, and at least one MAG was assigned to the *Trebouxia* photobiont (Figures 3C and 3D). The four detected photobiont lineages often co-occurred in various constellations, and photobiont identity did not appear to depend on substrate (Figure 3C). We also detected 13 bacterial MAGs universally present, of which four belonged to *Sphingomonas* (Figure 3E; Data S2B and S2C). Each metagenomic sample included at least six different *Sphingomonas* MAGs, which were not substrate dependent (Figure 3E). By contrast, Acetobacteraceae clade CAHJXG01 showed substrate dependency. Although every sample contained at least one CAHJXG01 MAG, none of the MAGs were present universally. Instead, they formed two clusters based on substrate

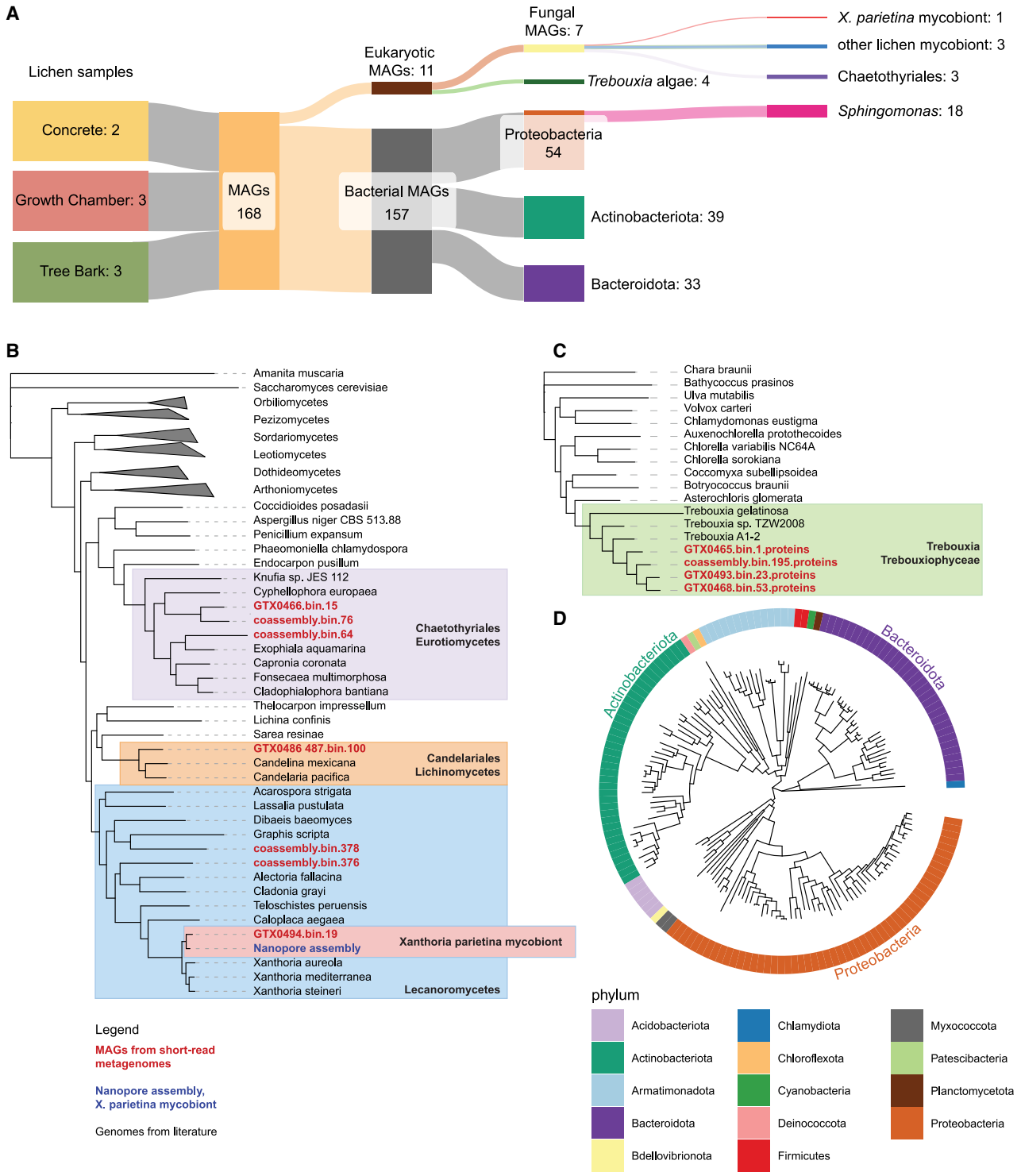


Figure 2. Maximum-likelihood trees of genomes assembled from *X. parietina* metagenomes

(A) Sankey plot showing the lichen samples used in metagenomic analysis and resulting MAGs and their taxonomic assignments.

(B) Fungal phylogenomic tree. Metagenome-assembled genomes (MAGs) assembled from the eight metagenomes shown in red, long-read genome assembly shown in blue. To clarify the taxonomic position of MAGs, we added reference genomes with known identities (shown in black; Data S2D). The genome of *Amanita muscaria* was used as an outgroup to root the tree.

(C) Algal phylogenomic tree. MAGs assembled from metagenomes are shown in red, and reference genomes are shown in black (Data S2D). The genome of *Chara braunii* was used as an outgroup.

(D) Bacterial phylogenomic tree. The color track shows taxonomic assignment.

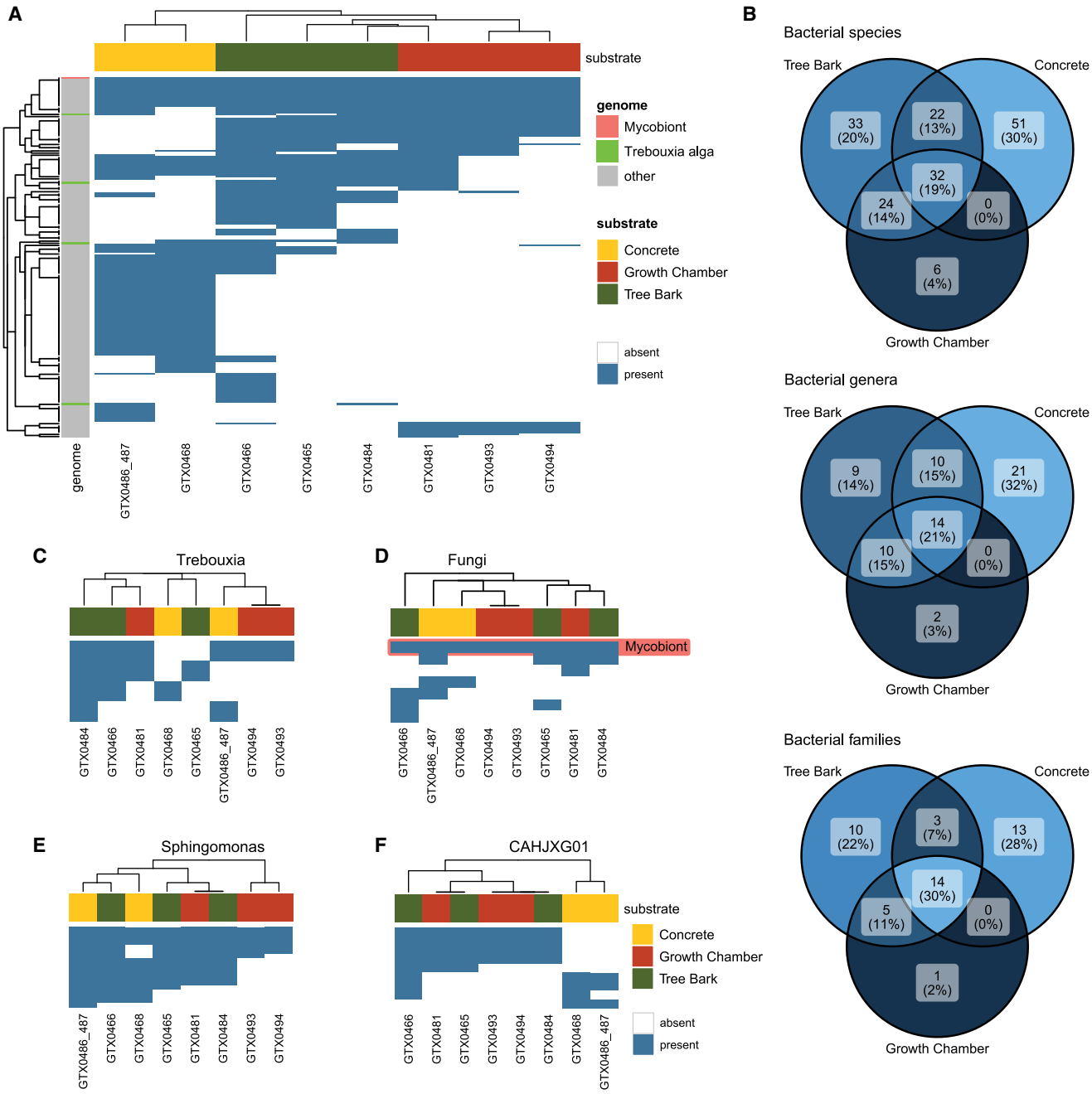


Figure 3. Species diversity detected in *X. parietina* metagenomes

(A) Presence/absence map of 168 MAGs assembled from *X. parietina* in eight metagenomes, divided by substrate.

(B) Venn diagrams showing shared and unique bacterial taxa (on the level of species = MAG, genus, and family) detected in *X. parietina* metagenomes from different substrates.

(C–F) Presence/absence map of selected lineages: (C) algae, (D) fungi, (E) *Sphingomonas*, and (F) Acetobacteraceae clade CAHJXG01.

See also Figure S2.

(Figure 3F). Both *Sphingomonas* and CAHJXG01 are frequent in lichens,[17] and other generalist bacteria have been reported from lichens too.[19–21] We conclude that lichen thalli contain a large number of associated microorganisms that can be putatively split into a substrate-dependent lichen microbial community and a core lichen community.

Differentially expressed genes disproportionately lack functional annotation and come from lichen-enriched orthogroups

To identify cellular processes involved in the lichen symbiosis, we compared gene expression of the mycobiont between intact lichen thalli at distinct developmental stages (17 samples from

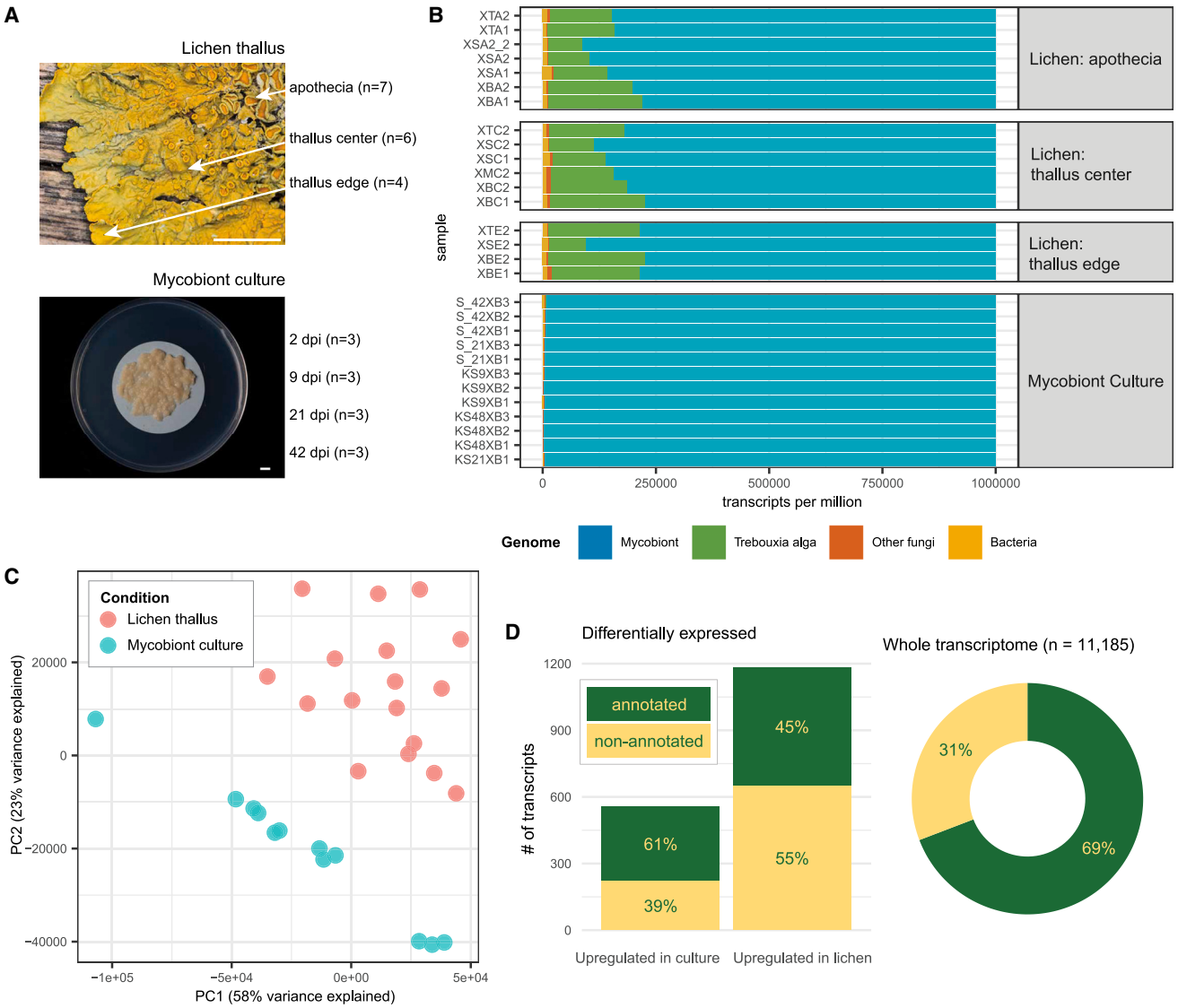


Figure 4. Transcriptomics of mycobiont culture and thalli of *X. parietina*

(A) Samples used for RNA-seq: four time points of the fungus in culture (2 dpi, 9 dpi, 21 dpi, and 42 dpi) and three developmental stages of lichen thallus (growing edge, center, and apothecia). Scale bars, 5 mm. Apothecia are fruiting bodies formed by the fungus, but apothecia of *X. parietina* contain algal cells in the margin, and, therefore, the 7%–22% share of algal transcripts is expected.

(B) Proportion of RNA-seq reads mapped to different categories of genomes. Transcript per million (TPM) values are summed across four groups: the mycobiont, *Trebouxia* algae, other fungi, and bacteria.

(C) Principal component analysis plot for RNA-seq based on TPM values. The samples are colored by sample type.

(D) Proportion of differentially expressed transcripts with and without functional annotations (defined as any annotation with InterProScan or PFAM domains or any assignment to UniProt, the Carbohydrate-Active Enzyme (CAZy) database, Merops, Gene Ontology, and Kyoto Encyclopedia of Genes and Genomes). Right panel shows percentage of transcripts with and without functional annotations across the entire transcriptome.

See also Data S3.

seven thalli; see below) and axenically grown mycobiont in culture (12 samples from four time points; Figure 4A). We pseudoaligned RNA sequencing (RNA-seq) data to a reference produced by compiling predicted transcriptomes from the long-read mycobiont genome and non-mycobiont MAGs isolated from *X. parietina* metagenomes. The majority of reads in all libraries aligned to the mycobiont transcriptome (Figure 4B), and, because the mycobiont is responsible for >90% of the biomass of the lichen thallus,[22] we focused on its gene expression. Analysis of gene expression in the photobiont proved impossible at this stage due to the presence of multiple different algal strains. Principal component analysis of the mycobiont data furthermore showed that gene expression differed significantly between lichen samples and mycobiont culture (Figure 4C).

We identified 1,749 differentially expressed mycobiont genes, of which 1,185 were upregulated in lichen thallus and 564

CellPress
OPEN ACCESS

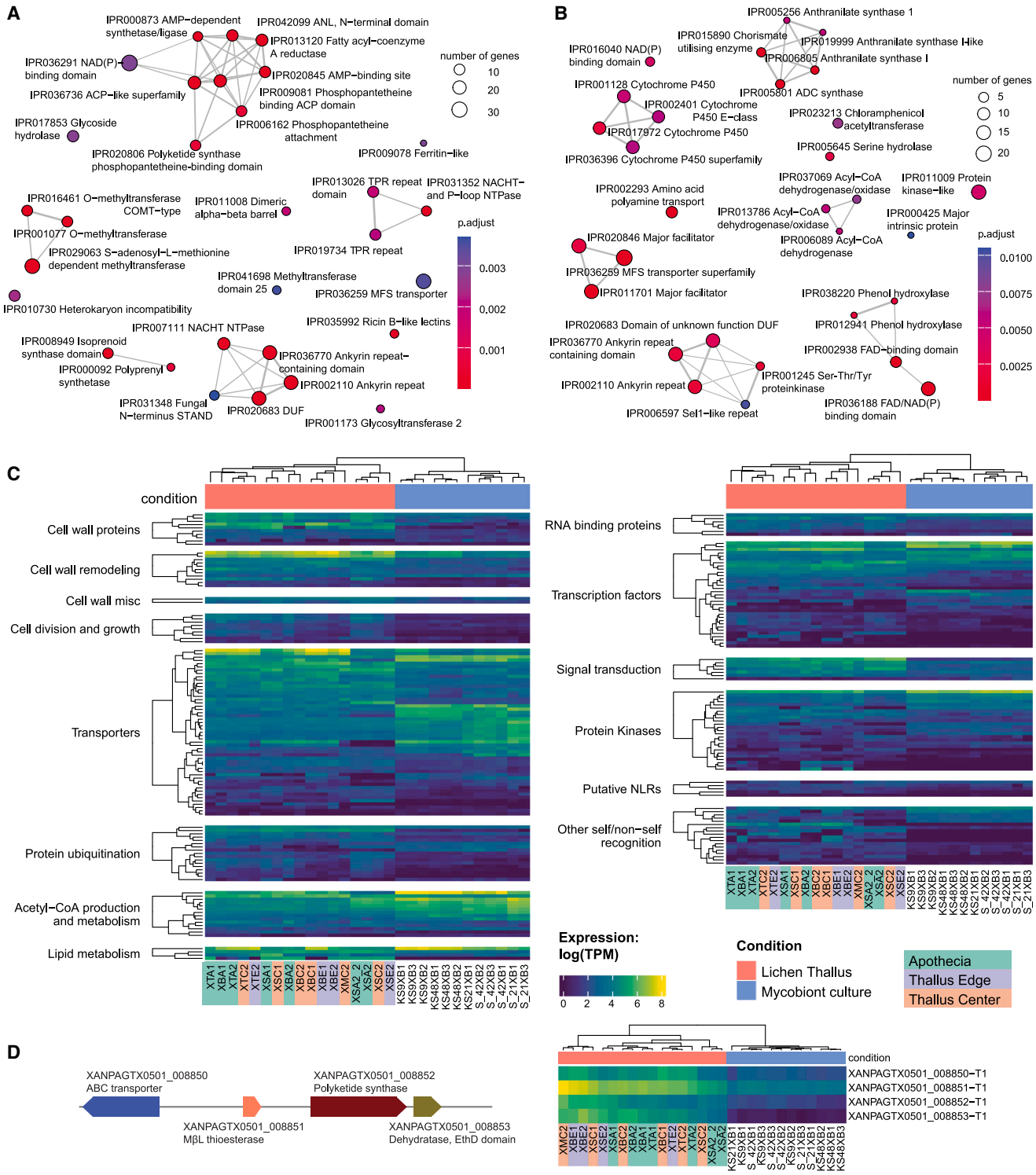


Figure 5. Differential gene expression of the mycobiont of *X. parietina*

(A and B) Enrichment plots showing InterProScan domains enriched in genes upregulated in: (A) lichen thalli and (B) mycobiont culture. The size of the node represents the number of genes annotated with a given domain in the gene set. Two nodes are connected if domains are present together within at least one gene; the width of the edge corresponds to the number of such genes.

(C) Heatmap showing gene expression (as log(TPM), where TPM stands for transcripts per million). We only show differentially expressed genes assigned to one of the gene categories potentially involved in fungal multicellularity[24] and symbiosis. Only categories with more than two genes are shown.

(legend continued on next page)

upregulated in mycobiont culture (Data S3A and B). Differentially regulated genes (DEGs) were observed to lack functional annotation more frequently than across the entire transcriptome (Figure 4D). In total, 31% of the transcriptome failed to be assigned any function, as is typical for genome annotations of lichen fungi,[10,23] but, among lichen-upregulated transcripts, this value reached 55%. Similarly, differentially expressed genes more often came from orthogroups identified as overrepresented in the genomes of lichen mycobionts (21% vs. 11% in the whole transcriptome). We conclude that gene expression associated with the symbiotic state includes untapped protein diversity and might include completely unknown gene functions.

Genes related to cell division, cell wall biogenesis, secondary metabolism, and protein ubiquitination are upregulated in the lichen symbiosis

We next investigated the identity of gene functions differentially expressed in the lichen symbiosis compared with mycobiont culture. Transporters from the Major Facilitator Superfamily were overrepresented in both lichen thallus and mycobiont-culture-upregulated genes (Figures 5A and 5B), with similar numbers of genes encoding transporters upregulated either in lichen thalli or mycobiont culture (Data S3B and S3C; Figure 5C). However, genes encoding transporters believed to play an important role in lichen symbiosis, such as putative polyol and ammonium transporters,[1] were lichen thallus upregulated. Among nine genes highly similar to known polyol transporters, one (XANPAGTX0501_001653-T1) was upregulated in lichen thalli and also assigned to a lichen-enriched orthogroup. We also identified one lichen-thallus-upregulated gene encoding an ammonium transporter (XANPAGTX0501_004972-T1). No genes encoding putative polyol and ammonium transporters were upregulated in mycobiont culture. Genes encoding proteins from other key functional groups implicated in fungal symbioses and/or fungal multicellularity,[24] such as transcription factors (TFs) and protein kinases, were also differentially upregulated in either lichen thalli or mycobiont culture (Figure 5C). At the same time, specific groups of TFs showed patterns of differential expression. For example, homeobox domain TFs and zinc-finger C2H2-type TFs were upregulated only in lichen thalli. Similarly, three of four differentially expressed zinc-finger RING-type TFs were lichen thallus upregulated. Representatives of these families of TFs have previously been linked to fruiting body development in mushroom-forming fungi,[24] consistent with a role in lichen tissue development. Conversely, the majority (six of eight) of differentially expressed Zn (II)2Cys6 zinc cluster TFs were upregulated in mycobiont culture.

In addition to 40 differentially expressed TFs, we identified other transcriptional regulators that were upregulated in lichen thalli. Five genes encoding proteins with RNA-binding domains were upregulated in lichen thalli, for example, as well as one RNA-dependent RNA polymerase (XANPAGTX0501_002123-T1). More notably, a group of genes linked to protein ubiquitination was upregulated in lichen thalli (Figure 5C). These included eight genes encoding F-box proteins and four genes encoding BTB/POZ proteins (Data S3C). Both these families are hypothesized to be involved in post-translational protein modification during formation of complex structures in mushroom development,[24] highlighting potential similarities in developmental biology of complex multicellular fungal structures.

Genes encoding proteins associated with cell division and growth—such as helicases, Rad21/Rec8-like proteins, and ribonucleases—were upregulated in lichen thalli (Figure 5C). Given the extremely slow growth of lichen fungi in culture,[26] this observation raises questions regarding whether the growth of mycobionts is conditional on the presence of other symbionts.

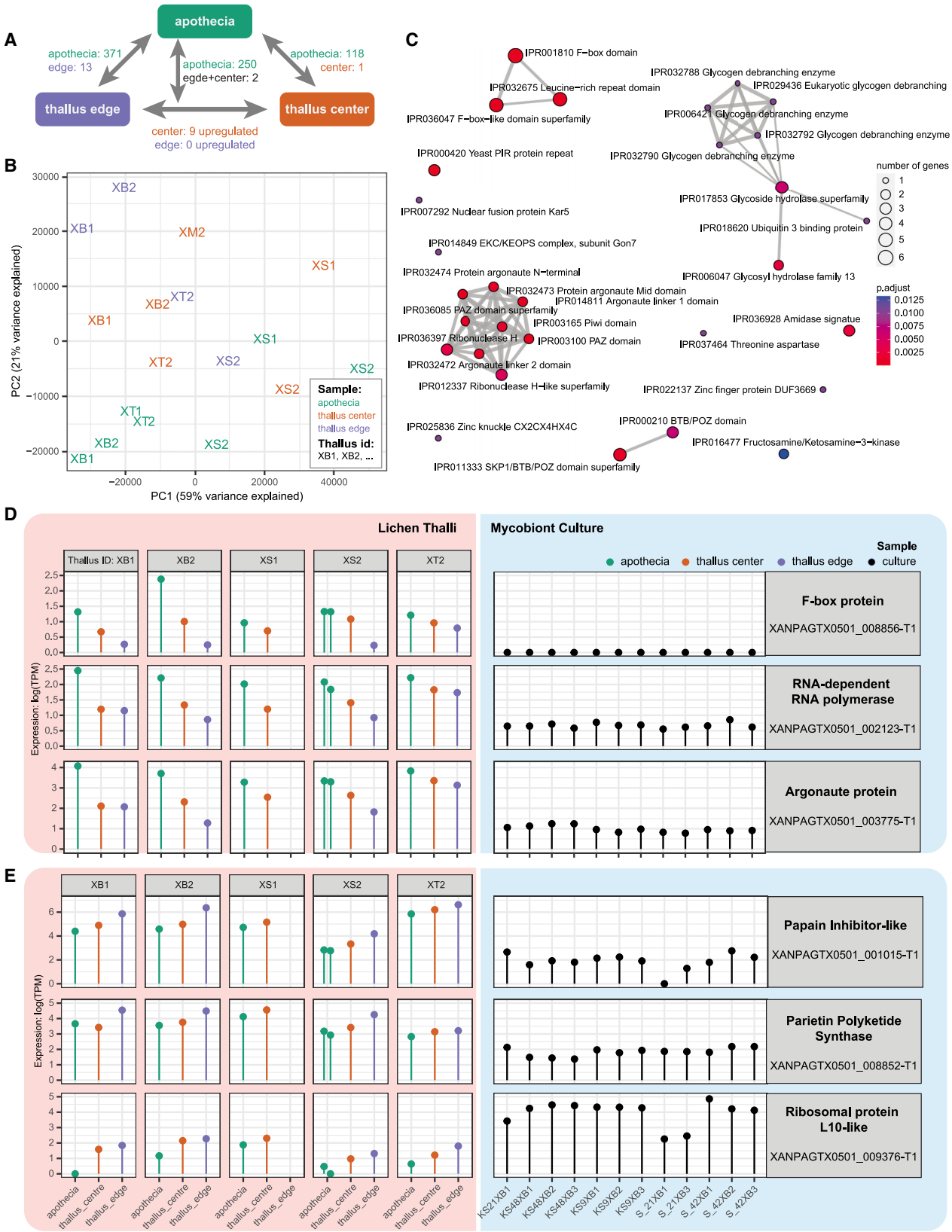
Genes associated with cell wall biosynthesis, cell wall remodeling, and cell wall proteins were often upregulated in lichen thalli (Figure 5C; Data S3C). Genes assigned to ricin-B-like lectins, for instance, are overrepresented in lichen-thallus-upregulated genes (Figure 5A). In addition, three genes with matches to concanavalin-A-like lectin/glucanase domains were lichen thallus upregulated (Data S3C). Four aspartic peptidases A1, involved in cell wall remodeling,[24] as well as seven carbohydrate-active enzymes (CAZymes) active on glucans and chitin, were upregulated in lichen thalli.

The biosynthetic gene cluster (BGC) putatively responsible for biosynthesis of the anthraquinone parietin, the pigment of *Xanthoria*[27] responsible for its yellow color, was upregulated in lichen thalli (Figure 5D). Cluster Xp_GTX0501_17_Cluster_1 is a type I polyketide BGC with similarity to the BGC of TAN-1612 (Data S1A), a compound from *Aspergillus nidulans* structurally similar to anthraquinones.[25,28] Five more BGCs were lichen thallus upregulated, including Xp_GTX0501_4_Cluster_4, which is similar to a BGC linked to alkaloid peramine (Data S1A). Such a BGC has been reported from the *X. parietina* mycobiont previously,[28] which reported a lack of A1 and R domains in the peramine synthase and deemed it non-functional. Based on our gene expression analysis, we can hypothesize that this BGC instead produces a different compound, which is induced during symbiosis. Although only one BGC was upregulated in the mycobiont, some BGCs contained a mixture of thallus- and mycobiont-upregulated genes (Figure S3). Overall, somewhat contrary to expectation, genes related to secondary metabolism do not show a pattern of being lichen thallus upregulated.

Differentially expressed genes showed clear spatial clustering, suggesting that epigenetic regulation of gene expression might play a role in lichens (Data S3B). Using a sliding window of 30 kbp, we therefore scanned the mycobiont genome to identify clusters of three or more jointly upregulated genes. We detected 92 lichen-thallus-upregulated and 49 mycobiont-culture-upregulated clusters. The majority (n = 83 and 45, respectively) had fewer than 10 genes; however, the largest cluster contained 19 lichen-thallus-upregulated genes and was 87 kbp in length.

(D) Putative parietin biosynthetic gene cluster. The left track shows the structure of the cluster: ABC transporter, metallo-beta-lactamase-type thioesterase, polyketide synthase, and a dehydratase with an EthylD domain. The structure of Xp_GTX0501_17_Cluster_1 (Figure 5D) is identical to the anthraquinone BGCs identified in lichen genomes by Llewellyn et al.[25] The heatmap shows expression levels for the four genes included in the cluster.
See also Figures S3 and S4.

CellPress
OPEN ACCESS



(legend on next page)

Although for about a quarter of the identified clusters (n = 24 and 10, respectively), the pattern can be explained by the differential expression of BGCs, the majority of spatial clusters do not overlap with the identified BGCs, suggesting that the explanation concerns epigenetic regulation. The spatial clusters did not appear linked to repeat elements, as only a small fraction (n = 8 and 2, respectively) overlapped with genomic regions with repeat content exceeding 50% (Figure S4). Altogether, genes assigned to spatial clusters accounted for 47% of all differentially expressed genes. Interestingly, a similar pattern has been previously reported during fruiting body development in a non-lichenized ascomycete.[29]

NLR-like genes are differentially expressed by *Xanthoria* mycobiont

We detected 23 genes in the mycobiont, potentially encoding nucleotide oligomerization domain (NOD)-like receptors or NLRs—a group of proteins involved in self/non-self recognition and immunity in plants, animals, and fungi[30] (Figure S4). These putative NLRs share structural features with known fungal NLRs[31]: a nucleotide-binding domain (either NB-ARC or NACHT), a repeat domain (either ankyrin, WD40, or tetratricopeptide repeats), and a variable effector domain. In half of putative NLRs, we identified effector domains from three functional groups: enzymatic domains (alpha/beta hydrolase and nucleoside phosphorylase domains), cell-death-inducing domains (HET and HeLo), and domains from InterProScan families IPR031359 and IPR031352 that lack a described function (Data S1D). Unexpectedly, in the C terminus of one putative NLR, we found a papain-like protease domain that matched ubiquitin-specific proteases. The remaining 11 NLRs contained no conserved effector domain recognizable by InterProScan or PFAM—as is typical for fungal NLRs, whose effector domains are underrepresented in existing databases.[31]

Four putative NLRs are upregulated in lichen thalli, including one with a pore-forming HeLo domain (Figures 5C and S4). Even though one NLR was identified as mycobiont culture upregulated, its expression levels were low in only some of the lichen samples, whereas in others, they were equivalent to expression levels in mycobiont culture samples. This inconsistency between lichen samples was more typical for genes potentially involved in self/non-self recognition (meaning NLRs and other genes with HET or HaLo domains) compared with other analyzed functional groups (Figure 5C). At the same time, of the 73 genes with HET or HaLo domains, 14 were consistently lichen thallus upregulated. We conclude that a subset of NLRs may be associated with lichen development.

Tissue-specific gene expression in lichen architectures

To investigate lichen morphogenesis, we carried out RNA-seq analysis of three distinct stages of lichen thallus development—the growing edge, representing actively growing thallus; the center, representing more mature thallus tissue; and the apothecia, which represent sites of sexual reproduction and ascospore formation (Figure 4A). We observed the largest number of upregulated genes in apothecia (Figure 6A; Data S3D). The highest number of DEGs was identified in a comparison of apothecia and the growing thallus edge, whereas central thallus tissue was less differentiated from either tissue type. In the edge/center comparison, all but one center-upregulated gene was also among apothecia-upregulated genes (when we compared apothecia to the combined set of edge and center thallus tissue samples) (Figure S5). This pattern might be explained by apothecium primordia being present in the thallus center—in *X. parietina*, apothecia are clustered in the central part of the thallus[8]—thereby affecting the expression profile, making it more similar to that of the fruiting body yet too small to be detected and excluded during sample preparation. Our ability to detect tissue-specific patterns was also complicated by variation between lichen thalli, as expression profiles appeared to depend both on developmental stage and on each individual thallus preparation (Figure 6B). To identify DEGs between different developmental stages, we therefore controlled for thallus identity, yet some tissue-specific genes may still have evaded detection.

Functional domains involved in protein ubiquitination and RNA interference were enriched among apothecium-upregulated genes (Figure 6C). The majority of lichen-thallus-upregulated genes associated with these functions were also upregulated in apothecia compared with other developmental stages (Data S3D; Figure S5). Most notably, XANPAGTX0501_008856-T1 encodes an F-box protein expressed in all lichen samples and none of the mycobiont culture samples. This gene had higher levels of expression in apothecia (b-value = 2.2, controlling for the thallus identity; Figure 6D). A similar pattern was observed in eight of 14 lichen-thallus-upregulated genes encoding ubiquitination proteins, including four additional F-box proteins and all lichen-thallus-upregulated genes associated with RNA interference (Figure 6D; Data S3D). By contrast, the 27 lichen-thallus-upregulated genes encoding transporters and 23 lichen-thallus-upregulated genes encoding TFs only included one and five apothecia-upregulated genes, respectively. Among the latter, we identified a homeobox protein gene (Figure S5), which is known to govern fruiting body formation in fungi.[32] Genes involved in karyogamy (*KAR5*) and conidiation (*CON6*) were also among genes

Figure 6. Differential gene expression of the mycobiont symbiont of *X. parietina* across three different developmental stages

(A) Overview of differential gene expression between the stages.

(B) PCA plot for lichen-derived RNA-seq based on TPM values. The samples are colored by the developmental stage.

(C) Enrichment plot showing InterProScan domains enriched in the apothecia-upregulated genes.

(D) Expression of three apothecia-upregulated genes involved in protein ubiquitination (F-box protein) and RNA interference (RNA-dependent RNA polymerase and Argonaute protein) across studied samples. The lichen samples are grouped based on the lichen thallus they derived from and colored based on developmental stage. The three shown genes were all upregulated in lichen samples compared with mycobiont culture and in apothecia compared with other developmental stages.

(E) Expression of three thallus-edge-upregulated genes. In addition to being upregulated in the thallus edge compared with apothecia, two of these genes are upregulated in lichen thalli compared with the mycobiont culture (putative polyketide synthase from the parietin gene cluster and a putative papain inhibitor), and one is upregulated in the mycobiont culture (ribosomal protein L10-like).

See also Figures S5 and S6.

upregulated in apothecia, as well as a gene encoding glycogen-debranching enzyme (Figure S5), consistent with gene expression profiles of mushroom development.[24] The gene model corresponding to the mating-type locus *MAT1-2-1* was also upregulated in apothecia, although the question remains whether *MAT* is actually functional because it contains a premature stop codon (Figure S6). Although 71% (n = 177 out of 250) of apothecia-upregulated genes were also upregulated in lichen thalli compared with mycobiont culture, 17 apothecia-upregulated genes were even more strongly expressed in culture. These genes included XANPAGTX0501_001643-T1, yet another apothecia-upregulated F-box gene. Similar to the lichen thallus/mycobiont culture comparison, the functions of most DEGs remain unknown. Of the 250 genes upregulated in apothecia compared with the other developmental stages, 171 (68%) have no functional annotations.

Contrary to our expectations, few genes were upregulated in the thallus edge compared with the center of the thallus. Because growth in *X. parietina* happens primarily at its narrow marginal rim, we expected numerous upregulated genes associated with active growth. However, no edge-upregulated gene was detected that could be linked to growth. One possible exception is XANPAGTX0501_009376-T1, which contains a ribosomal protein L10-like domain (*RPL10*; Figure 6E). Although ribosomal proteins in general are associated with growth,[24] the profile of this gene does not match growth-associated genes discussed earlier. Instead of being lichen thallus upregulated, it was more highly expressed in the mycobiont compared with any lichen sample. Alternatively, *RLP10* could be induced by stress, as is known for plant *RLP10*,[33] which is specifically expressed under UV light. Notably, another gene upregulated in the edge encodes the polyketide synthase linked to the biosynthesis of parietin—the key photoprotective pigment in *Xanthoria* (Figure 6E). Nearly half (n = 6 out of 13) of edge-upregulated genes are predicted to encode secreted proteins, including a putative papain inhibitor (Figure 6E).

The mycobiont secretome contains putative effector proteins

We next investigated whether mycobionts possess potential secreted effector proteins that potentially modulate cellular functions or impair immunity within symbiotic partners. Effectors are well known in pathogenic and mutualistic fungi.[34–37] We identified 608 putative secreted proteins in the predicted proteome of the *X. parietina* mycobiont, of which 154 were lichen thallus upregulated and 40 mycobiont culture upregulated (Figures S7A and S7B). Genes encoding putative secreted proteins were spread unevenly in the genome and formed several spatial clusters, which were apparently unrelated to clusters of repeat elements (only eight such genes out of 608 were in genomic regions with repeat content exceeding 50%; Figure S7C). As effectors are often sequence-unrelated, we carried out structural predictions using AlphaFold2 to identify structurally related proteins within the predicted secretome. Structures with a quality score pTM $\geq$ 0.5 (n = 393) were used to construct a structural phylogenetic tree using FoldTree.[38] We divided the tree into 84 structural clusters, which together included 311 proteins (Figures 7A and 7B; Data S4), and the remaining 82 proteins were considered singletons. In addition to structural analysis, we also screened the secretome using two effector-predicting tools: EffectorP and deepredeff (Figure S7D), although these provided inconsistent results.

The predicted secretome included proteins similar to known effectors. A large group of proteins (clusters 18–24a), for example, showed similarity to killer toxins Kp4 and a newly described effector from the plant pathogen *Zymoseptoria tritici* (Figures 7A and 7C). Collectively, these clusters accounted for 8% of the secretome and included 47 proteins, of which 18 were upregulated in lichen thalli compared with the mycobiont (Data S4). Protein XANPAGTX0501_009887-T1 was also lichen thallus upregulated and highly similar to Tsp1 (Figure 7D), an effector from *Trichoderma virens* that suppresses plant immunity by stimulating the salicylic acid pathway.[39] Other clusters of potential effectors include proteins with folds similar to known fungal effectors: CFEM proteins,[40] ribonucleases,[41] and NTF2/SnoaL proteins[42] (Figure 7A). In the list of putative effectors, we also included thaumatin and gamma-crystallin-like proteins, as these families were identified as probable effectors.[43,44] Except for the ribonuclease cluster, these clusters had at least one lichen-thallus-upregulated protein, and none contained any mycobiont-culture-upregulated proteins (Figure 7E). Similarly, of five proteins identified as ricin-B-like lectins, four were lichen thallus upregulated, consistent with their proposed role in symbiont recognition.[45]

Secreted enzymes also account for over a third of the secretome (n = 207) and are primarily represented by CAZymes and proteases. The most numerous enzyme cluster was formed by AA7 (Figure 7B)—oligosaccharide oxidase family expanded in lecanoromycete fungi[46] and is active on many substrates. Other major groups included GH16, a multifunctional family of glycoside hydrolases, and families active on beta-glucans (GH128, GH72, and GH12), which might target the mycobiont's own cell wall. Metallopeptidases M35 were also numerous, and, curiously, we identified several putative protease inhibitors (cluster 83), two of which were upregulated in lichen thalli. Unlike putative effectors, secreted enzymes were often upregulated in the mycobiont culture (n = 19, out of 53 differentially expressed).

Combining sequence-based and structure-based annotation allowed us to assign putative functions to the majority of the secretome, although some assignments, especially based on the hits to the AlphaFold database, require significant further validation. However, the remaining 205 proteins contained no identified InterProScan or Pfam domain and yielded no significant match to a characterized protein when searched against structural databases (Data S4). These proteins might play a role in symbiosis, as the percentage of lichen-thallus-upregulated proteins in the "novel" set was even higher than in the secretome (33%, compared with 25% in the whole secretome and 11% across the whole transcriptome). The majority (n = 165) of novel proteins failed to produce structural models with quality scores above the set threshold (pTM $\geq$ 0.5) and were consequently excluded from clustering. Others, however, were included and formed six clusters (cl04, cl14, cl15, cl48, cl52, and cl82) composed entirely of proteins lacking annotation. Notably, cl04 consisted of eight proteins, two of which were differentially expressed and lichen thallus upregulated (Figure 7A). Proteins from this cluster were classified as effectors by deepredeff but not EffetorP. Although all of them

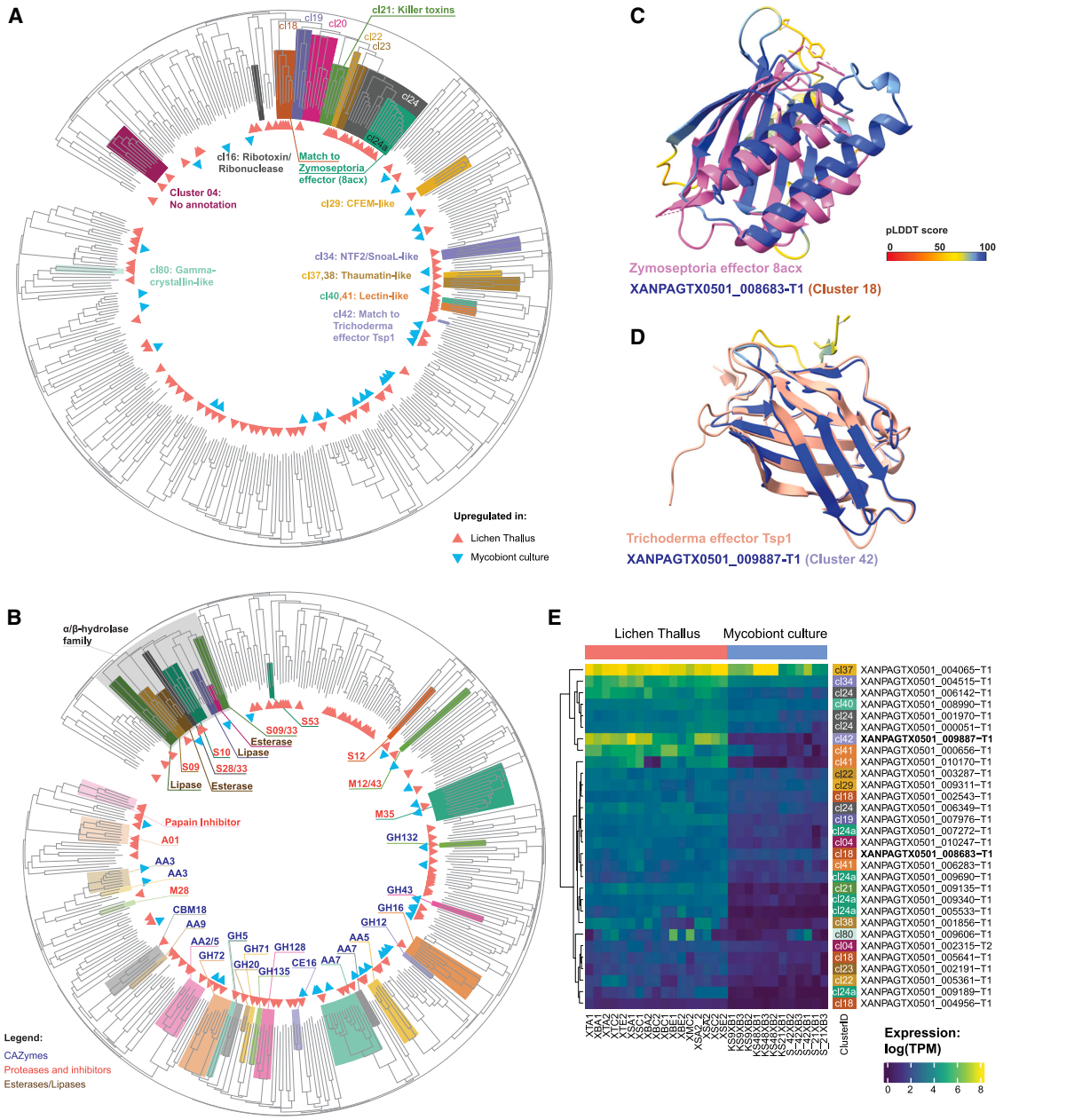


Figure 7. Structural clustering of the predicted secretome of the *X. parietina* mycobiont

(A and B) The structural phylogenetic tree produced from structural models of predicted secreted proteins. Structural models with high confidence (pTM ≥ 0.5) were analyzed using FoldTree, and the resulting tree split into 84 structural clusters (Data S4). The DEGs are indicated with triangles. (A) Putative effector clusters and other clusters of interest are highlighted. (B) Clusters formed by various hydrolases: carbohydrate-active enzymes (CAZymes), proteases, lipases, and esterases (as indicated by color). For the CAZymes and proteases, we give their family assignments.

(C) Predicted structure of XANPAGTX0501_008683-T1 (cluster 18) superimposed onto a *Zymoseptoria* effector (PDB: 8acx).

(D) Predicted structure of XANPAGTX0501_009887-T1 (cluster 42) superimposed onto a *Trichoderma* effector Tsp1 (PDB: 7cwj).

(E) Heatmap showing expression (as log(TPM), where TPM stands for transcripts per million) of DEGs corresponding to clusters of interest (shown in A). The right annotation track shows number of the cluster. The two proteins from (C) and (D) are highlighted in bold.

See also Figure S7 and Data S4.

were assigned to lichen-enriched orthogroups, some of them showed similarity to uncharacterized proteins from various phytopathogenic and endophytic fungi, including *Alternaria alternata* and *Mollisia scopiformis* (Data S4), raising questions about their potential as novel lichen effectors.

DISCUSSION

In this study, we set out to explore how the intricate self-replicating architectures of lichens are formed from the symbiotic association of morphologically simple microorganisms. Our aim

was to define the constituent microbiome of a single lichen species and identify mechanisms that orchestrate lichen growth and development. To do this, we carried out a metagenomic analysis of a very common lichen *X. parietina*, investigated its development by transcriptional profiling, and used a combination of informatic and structural modeling to define potential determinants of lichen morphogenesis.

Our analysis of metagenomic data underlines the intrinsic complexity of lichen symbioses. Although it has previously been assumed that one mycobiont individual corresponds to one thallus, our data suggest that multiple genotypes can be present within a single thallus. Non-standard MAF distributions in several of our samples (already reported from lichens before[47]) can potentially result from either the genotypes of male parents present in zygotes and ascospores within the fungal apothecia or by several different mycobiont "individuals" that live so closely together as to appear as one. Evidence for such "chimeric" thalli has been reported previously,[48,49] and an experiment showed that a thallus fragment of *X. parietina* can fuse seamlessly back into its parent thallus.[50] Many of our samples also contained multiple lineages of the photobiont *Trebouxia*, as known for many lichens,[51,52] including *X. parietina*.[53] Multiple algal strains are hypothesized to offer benefit to the symbiosis by providing more plasticity but may also reflect opportunistic acquisition of photobiont partners during lichen development.

Some lichen symbioses are now known to associate with certain bacteria and non-mycobiont fungi.[3,4] However, the organismal composition of *X. parietina* thalli was undescribed until now. Our metagenomic analysis reveals considerable diversity of microorganisms in addition to the mycobiont and the photobiont, including 157 distinct bacterial lineages. Although basidiomycete yeasts, known from several lichen groups,[3,54] were not detected, ascomycete black yeasts were present occasionally. Of over a hundred bacterial species, the majority were also present sporadically or in samples collected from a specific substrate, while a smaller subset was present across all studied samples. A similar pattern was observed in the lung lichen *Lobaria*, where the bacterial microbiome can be split into a variable portion, influenced by local environment, and a stable "core" portion.[55] In *X. parietina*, the core microbiota includes multiple lineages from the genus *Sphingomonas*, one of the bacteria most frequently detected in lichens.[17] Evidence from other lichen symbioses suggests that *Sphingomonas* is tightly associated with *Trebouxia* photobionts and can use polyols produced by *Trebouxia*.[56,57] Whether *Sphingomonas* and other bacteria are commensals or mutualistic symbionts in *X. parietina* remains to be tested; however, their presence needs to be taken into account. The presence of bacteria might affect the behavior of eukaryotic symbionts, both in the natural symbiotic state and in laboratory experiments. In the analysis of our own transcriptomic data from the mycobiont culture, we detected little presence of bacteria (Figure 4B), which might have affected the mycobiont gene expression. In fact, as one of the standard methods of isolating lichen symbionts in culture relies on homogenizing lichen thalli,[5] the presence of lichen-associated bacteria cannot be ruled out for many previous studies of lichen symbionts in culture. In addition, the potential presence of bacteria needs to be considered while analyzing lichen-derived omics data since otherwise bacterial data might be misassigned to one of the eukaryotic partners, creating artifacts of the analysis.

Using transcriptional profiling, we next identified biological processes that differ between the mycobiont in its natural state—as a member of the complex lichen symbiosis—and in aposymbiotic culture. As in previous reports,[58–60] polyol and ammonium transporters were upregulated in lichen thalli, consistent with the hypothesis that mycobionts use alga-produced polyols in exchange for ammonium. Cell wall proteins, including lectins, were also upregulated in symbiosis, matching a predicted role in hyphal adhesion and symbiont recognition.[45] Another class of lichen-thallus-upregulated cell wall protein, fungal hydrophobins, have also been shown to be differentially expressed in *X. parietina* and are hypothesized to play a role in creating lichen architectures.[61] Secondary metabolism, cell wall synthesis and remodeling, and cell division functions are also associated with lichen thalli.

We were particularly interested in defining how lichen gene expression is regulated during development. It appears that changes of gene expression during lichenization may be partially driven by epigenetic regulation because differentially expressed genes tend to form clusters instead of being randomly dispersed across the fungal genome. Among identified differentially expressed TFs, we identified different classes of TFs being upregulated in different states, consistent with orchestration of a specific developmental program, while post-transcriptional and post-translational regulation also probably play an important role, as we identified lichen-thallus-upregulated genes potentially involved in RNA interference and targeted protein ubiquitination. Among the latter, F-box and BTB/POZ proteins, which have been linked to complex multicellularity in mushroom-forming fungi and non-lichenized ascomycetes,[24,62] were especially prevalent. Surprisingly, however, we did not identify many differentially expressed genes among key signaling pathways such as MAPK-dependent signaling and TOR signaling.

Studies of complex multicellularity in fungi primarily focus on fruiting bodies,[24] which have anatomy similar to lichens (as both are essentially formed by tightly packed hyphae) but that also differ from lichens in two important ways. First, lichen thalli are formed through concerted growth of multiple symbionts. Second, in mushroom-forming fungi, complex architectures emerge only to serve a specific function—usually sexual reproduction, although not exclusively, as in, for instance, sclerotia—and the fungus exists primarily as simple mycelium. For mycobionts in lichens, however, complex lichen thallus architectures represent the *only* known natural mode of existence, and lichen thalli are not primarily linked to reproduction. In many lichens, for example, none of the symbionts are known to undergo sexual reproduction. Thalli of *X. parietina* contain apothecia of its mycobiont, yet most of the body is formed by vegetative hyphae and photobiont cells. This prompted us to try to separate the mechanisms behind complex structures from those specific to fruiting body development and sexual reproduction. We identified genes upregulated in apothecia compared with vegetative parts of the thallus. In addition to expected genes involved in sexual reproduction, we identified numerous genes linked to RNA interference and targeted protein ubiquitination, consistent with a role in fruiting body development. The majority of apothecium-upregulated genes were also upregulated in lichen thalli

compared with cultures (as expected, given that mycobionts never reproduce sexually in culture[2]). Several exceptions—genes upregulated in fruiting bodies *and* the mycobiont culture—are of special interest with regard to a role in fungal development. Also of interest are genes upregulated in lichen thalli compared with culture but *not* in apothecia, as these might represent core machinery required for forming complex lichen structures. These include the majority of lichen-thallus-upregulated cell wall proteins and TFs, as well as genes linked to self/non-self recognition (discussed below). When considered together, our study provides evidence that a similar toolbox is used for complex multicellularity by lichen mycobionts and non-lichenized ascomycete and basidiomycete fungi, consistent with a recent hypothesis,[63] and suggesting conservation in higher-order fungal developmental biology.

The longevity of lichen thalli is another feature separating them from complex non-lichen fungal fruiting bodies. *X. parietina* and similar lichens grow primarily in the thin outer rim of the thallus, meaning that lichen tissue is younger at the margins of a thallus. We therefore used different parts of thalli as a proxy for lichen developmental stages. We aimed to identify genes involved in active growth of a lichen thallus and tissue differentiation. However, our analysis yielded only a few thallus-edge-upregulated genes. The lack of growth-associated gene expression in the edge might be explained by both biological and technical reasons. Although growth in *X. parietina* happens primarily at its narrow marginal rim, central parts of thalli are also capable of regenerative growth.[8] In addition, lichen growth does not occur continuously and instead switches on and off depending on microclimate.[64] Both factors might therefore complicate detection of growth-related genes in a transcriptomic study. Genes upregulated in the thallus edge include several secreted proteins and a gene cluster linked to pigment biosynthesis. This can be seen as evidence that lichen tissue in this developmental stage experiences stress and secretes proteins in order to affect other microorganisms or modify its substrate. Our study is the first to compare gene expression profiles of different parts of a lichen architecture and has allowed broad classification of lichen-associated gene functions by developmental stage.

Our final aim was to investigate factors that mediate the interaction between a mycobiont and other symbionts. The lichen symbiosis likely involves bidirectional recognition between symbionts and potentially the recruitment of appropriate strains. Although much is known about fungal self/non-self recognition systems, research mostly focuses on the mechanisms for different strains within one species to recognize each other, and our understanding of how fungi recognize other organisms remains poor.[65] NLRs and HET domain proteins are hypothesized to play a role in fungal immunity and fungal symbioses, but experimental validation is still lacking.[65] We identified putative NLR-encoding genes in the *X. parietina* mycobiont, several of which were upregulated in lichen thalli, suggesting a role in lichen development or maintenance. We should note, however, that our data derived from established lichen thalli, and, therefore, we were unable to capture gene expression changes during initial establishment of the symbiosis. Future research will need to test the role of NLRs in recognition between lichen symbionts.

In addition to recognizing each other, symbionts may possess machinery to influence one another during lichen formation. Secreted effector proteins are often used by fungi, both pathogenic and mutualistic, to establish contact and suppress the immune response of their plant host.[34–37] Because green algae share some features of the plant immune system,[66] we recently hypothesized[67] that mycobiont effectors might play a role in lichen symbioses. By predicting and analyzing the secretome of the mycobiont, we identified putative effectors, many of which were upregulated in lichen thalli. Effectors often evolve so rapidly that they lose sequence similarity,[68] so we predicted protein structures for all 608 secreted proteins and used structures to group similar secreted proteins together. In this way, we were able to identify a large group of proteins with similarity to killer toxin Kp4 and several fungal effectors.[69,70] We conclude that effector-like proteins are encoded by the mycobiont, consistent with manipulation of other symbionts during lichen development. In addition, our analysis revealed completely novel proteins that show no similarity to any characterized protein that might represent novel lichen-specific families of effectors.

Conclusion

In summary, metagenomic and metatranscriptomic analysis of *X. parietina* has identified biological processes involved in lichen development. Our results show that lichen morphogenesis shares features with the development of multicellular structures by non-lichenized fungi, such as sclerotia and mushrooms, but it is also clear that lichen formation involves a large amount of unknown biology. The majority of genes upregulated in the symbiotic state cannot, for example, be assigned any function based on similarity to databases. We attempted to push this boundary by making structural predictions for proteins secreted by the mycobiont, which allowed identification of structurally related putative effectors but also highlighted the large number of completely novel proteins present in lichens. Our study will therefore provide a resource for future research on developmental biology of this elusive group of organisms.

RESOURCE AVAILABILITY

Lead contact

Nicholas J. Talbot (nick.talbot@tsl.ac.uk) is the lead contact for this study.

Materials availability

X. parietina lichen samples collected for this study have been deposited to The University and Jepson Herbaria (Data S5A).

Data and code availability

- Raw metagenomic and metatranscriptomic sequencing data, as well as assembled and annotated genomes, have been deposited at ENA (ENA: PRJEB78723 and ENA: PRJEB38537).
- All original code generated for data analysis is available on GitHub (https://github.com/metalichen/2024-Multipartite-complexity-omics-Xanthoria).
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by grants from the Leverhulme Trust (RPG-2018-139), the Gatsby Charitable Foundation, the Halpin Family, and the Biotechnology and Biological Sciences Research Council (BBS/E/J/000PR9798) to N.J.T. We thank Jesper Svedberg and Markus Hiltunen Thorén for suggestions on data analysis, Paul Dyer for providing the mycobiont culture, Phil Robinson

for providing photographs, and Alison MacFadyen for helping with depositing data.

AUTHOR CONTRIBUTIONS

N.J.T., K.S., and G.T. conceived the project. K.S., X.Y., and G.T. performed lab work. G.T., N.S., C.J., D.M., A.B., A.R.B., S.L.A.-V., I.O.-B., and H.J. designed and performed bioinformatic analysis. G.T. prepared figures and tables. G.T. and N.J.T. drafted the manuscript, and all authors contributed to the editing.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

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

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Lichen thalli
  - *Xanthoria parietina* mycobiont culture
- METHOD DETAILS
  - Mycobiont genome sequencing and assembly
  - RNA extraction and sequencing
  - Mycobiont genome annotation
  - Identification of lichen-enriched orthogroups
  - Metagenome sequencing and analysis
  - Confirming sample identities
  - Ploidy analysis
  - Transcriptomic analysis
  - Protein structure prediction and analysis
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cub.2024.12.041.

Received: August 20, 2024
Revised: November 15, 2024
Accepted: December 17, 2024
Published: January 30, 2025

REFERENCES

1. Spribille, T., Resl, P., Stanton, D.E., and Tagirdzhanova, G. (2022). Evolutionary biology of lichen symbioses. New Phytol. *234*, 1566–1582.
2. Honegger, R. (1993). Developmental biology of lichens. New Phytol. *125*, 659–677.
3. Spribille, T., Tuovinen, V., Resl, P., Vanderpool, D., Wolinski, H., Aime, M.C., Schneider, K., Stabentheiner, E., Toome-Heller, M., Thor, G., et al. (2016). Basidiomycete yeasts in the cortex of ascomycete macrolichens. Science *353*, 488–492.
4. Grube, M., Cernava, T., Soh, J., Fuchs, S., Aschenbrenner, I., Lassek, C., Wegner, U., Becher, D., Riedel, K., Sensen, C.W., et al. (2015). Exploring functional contexts of symbiotic sustain within lichen-associated bacteria by comparative omics. ISME J. *9*, 412–424.
5. Yoshimura, I., Yamamoto, Y., Nakano, T., and Finnie, J. (2002). Isolation and Culture of Lichen Photobionts and Mycobionts. In Protocols in Lichenology, I.C. Kranner, R.P. Beckett, and A.K. Varma, eds. (Springer Berlin Heidelberg), pp. 3–33.
6. Park, S.-Y., Jeong, M.-H., Wang, H.-Y., Kim, J.A., Yu, N.-H., Kim, S., Cheong, Y.H., Kang, S., Lee, Y.-H., and Hur, J.-S. (2013). *Agrobacterium tumefaciens*-mediated transformation of the lichen fungus, *Umbilicaria muehlenbergii*. PLoS One *8*, e83896.
7. Wyczanska, M., Wacker, K., Dyer, P.S., and Werth, S. (2023). Local-scale panmixia in the lichenized fungus *Xanthoria parietina* contrasts with substantial genetic structure in its *Trebouxia* photobionts. Lichenologist *55*, 69–79.
8. Honegger, R. (1996). Experimental studies of growth and regenerative capacity in the foliose lichen *Xanthoria parietina*. New Phytol. *133*, 573–581.
9. Llewellyn, T., Mian, S., Hill, R., Leitch, I.J., and Gaya, E. (2023). First Whole Genome Sequence and Flow Cytometry Genome Size Data for the Lichen-Forming Fungus *Ramalina farinacea* (Ascomycota). Genome Biol. Evol. *15*, evad074. https://doi.org/10.1093/gbe/evad074.
10. McKenzie, S.K., Walston, R.F., and Allen, J.L. (2020). Complete, high-quality genomes from long-read metagenomic sequencing of two wolf lichen thalli reveals enigmatic genome architecture. Genomics *112*, 3150–3156.
11. Gerasimova, J.V., Beck, A., Werth, S., and Resl, P. (2022). High Diversity of Type I Polyketide Genes in as Revealed by the Comparative Analysis of 23 Lichen Genomes. J. Fungi (Basel) *8*, 449. https://doi.org/10.3390/jof8050449.
12. Allen, J.L., Jones, S.J.M., and McMullin, R.T. (2021). Draft Genome Sequence of the Lichenized Fungus *Bacidia gigantensis*. Microbiol. Resour. Announc. *10*, e0068621.
13. Gladyshev, E. (2017). Repeat-Induced Point Mutation and Other Genome Defense Mechanisms in Fungi. Microbiol. Spectr. *5*. https://doi.org/10.1128/microbiolspec.FUNK-0042-2017.
14. Cometto, A., Leavitt, S.D., Grube, M., De Hoog, S., and Muggia, L. (2023). Tackling fungal diversity in lichen symbioses: molecular and morphological data recognize new lineages in Chaetothyriales (Eurotiomycetes, Ascomycota). Mycol. Prog. *22*. https://doi.org/10.1007/s11557-023-01901-9.
15. Muggia, L., Gueidan, C., Knudsen, K., Perlmutter, G., and Grube, M. (2013). The lichen connections of black fungi. Mycopathologia *175*, 523–535.
16. Quan, Y., Muggia, L., Moreno, L.F., Wang, M., Al-Hatmi, A.M.S., da Silva Menezes, N., Shi, D., Deng, S., Ahmed, S., Hyde, K.D., et al. (2020). A re-evaluation of the Chaetothyriales using criteria of comparative biology. Fungal Divers. *103*, 47–85.
17. Tagirdzhanova, G., Saary, P., Cameron, E.S., Allen, C.C.G., Garber, A.I., Escandón, D.D., Cook, A.T., Goyette, S., Nogerius, V.T., Passo, A., et al. (2024). Microbial occurrence and symbiont detection in a global sample of lichen metagenomes. PLoS Biol. *22*, e3002862.
18. Leiva, D., Fernández-Mendoza, F., Acevedo, J., Carú, M., Grube, M., and Orlando, J. (2021). The Bacterial Community of the Foliose Macro-lichen Peltigera frigida Is More than a Mere Extension of the Microbiota of the Subjacent Substrate. Microb. Ecol. *81*, 965–976.
19. Park, Y., Noh, H.-J., Hwang, C.Y., Shin, S.C., Hong, S.G., Jin, Y.K., Lee, H., and Lee, Y.M. (2022). *Hymenobacter siberiensis* sp. nov., isolated from a marine sediment of the East Siberian Sea and *Hymenobacter psoromatis* sp. nov., isolated from an Antarctic lichen. Int. J. Syst. Evol. Microbiol. *72*. https://doi.org/10.1099/ijsem.0.005290.
20. Phongsopitanun, W., Matsumoto, A., Inahashi, Y., Kudo, T., Mori, M., Shiomi, K., Takahashi, Y., and Tanasupawat, S. (2016). *Actinoplanes lichenis* sp. nov., isolated from lichen. Int. J. Syst. Evol. Microbiol. *66*, 468–473.
21. Jiang, L., An, D., Wang, X., Zhang, K., Li, G., Lang, L., Wang, L., Jiang, C., and Jiang, Y. (2020). *Methylobacterium planium* sp. nov., isolated from a lichen sample. Arch. Microbiol. *202*, 1709–1715.
22. Beck, A., and Mayr, C. (2012). Nitrogen and carbon isotope variability in the green-algal lichen *Xanthoria parietina* and their implications on mycobiont-photobiont interactions. Ecol. Evol. *2*, 3132–3144.
23. 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*, evab047. https://doi.org/10.1093/gbe/evab047.

24. Nagy, L.G., Vonk, P.J., Künzler, M., Földi, C., Virágh, M., Ohm, R.A., Hennicke, F., Bálint, B., Csernetics, Á., Hegedüs, B., et al. (2023). Lessons on fruiting body morphogenesis from genomes and transcriptomes of Agaricomycetes. Stud. Mycol. *104*, 1–85.
25. Llewellyn, T., Nowell, R.W., Aptroot, A., Temina, M., Prescott, T.A.K., Barraclough, T.G., and Gaya, E. (2023). Metagenomics Shines Light on the Evolution of "Sunscreen" Pigment Metabolism in the Teloschistales (Lichen-Forming Ascomycota). Genome Biol. Evol. *15*, evad002. https://doi.org/10.1093/gbe/evad002.
26. Zakeri, Z., Junne, S., Jäger, F., Dostert, M., Otte, V., and Neubauer, P. (2022). Lichen cell factories: methods for the isolation of photobiont and mycobiont partners for defined pure and co-cultivation. Microb. Cell Factories *21*, 80.
27. Solhaug, K.A., and Gauslaa, Y. (1996). Parietin, a photoprotective secondary product of the lichen *Xanthoria parietina*. Oecologia *108*, 412–418.
28. Berry, D., Mace, W., Rehner, S.A., Grage, K., Dijkwel, P.P., Young, C.A., and Scott, B. (2019). Orthologous peramine and pyrrolopyrazine-producing biosynthetic gene clusters in *Metarhizium rileyi*, *Metarhizium majus* and *Cladonia grayi*. Environ. Microbiol. *21*, 928–939.
29. Rodenburg, S.Y.A., Terhem, R.B., Veloso, J., Stassen, J.H.M., and van Kan, J.A.L. (2018). Functional Analysis of Mating Type Genes and Transcriptome Analysis during Fruiting Body Development of Botrytis cinerea. mBio *9*, e01939-17. https://doi.org/10.1128/mBio.01939-17.
30. Uehling, J., Deveau, A., and Paoletti, M. (2017). Do fungi have an innate immune response? An NLR-based comparison to plant and animal immune systems. PLoS Pathog. *13*, e1006578.
31. Wojciechowski, J.W., Tekoglu, E., Gąsior-Głogowska, M., Coustou, V., Szulc, N., Szefczyk, M., Kopaczyńska, M., Saupe, S.J., and Dyrka, W. (2022). Exploring a diverse world of effector domains and amyloid signaling motifs in fungal NLR proteins. PLoS Comput. Biol. *18*, e1010787.
32. Coppin, E., Berteaux-Lecellier, V., Bidard, F., Brun, S., Ruprich-Robert, G., Espagne, E., Aït-Benkhali, J., Goarin, A., Nesseir, A., Planamente, S., et al. (2012). Systematic deletion of homeobox genes in *Podospora anserina* uncovers their roles in shaping the fruiting body. PLoS One *7*, e37488.
33. Falcone Ferreyra, M.L., Pezza, A., Biarc, J., Burlingame, A.L., and Casati, P. (2010). Plant L10 ribosomal proteins have different roles during development and translation under ultraviolet-B stress. Plant Physiol. *153*, 1878–1894.
34. Mentlak, T.A., Kombrink, A., Shinya, T., Ryder, L.S., Otomo, I., Saitoh, H., Terauchi, R., Nishizawa, Y., Shibuya, N., Thomma, B.P.H.J., et al. (2012). Effector-mediated suppression of chitin-triggered immunity by *Magnaporthe oryzae* is necessary for rice blast disease. Plant Cell *24*, 322–335.
35. Zeng, T., Rodriguez-Moreno, L., Mansurkhodzaev, A., Wang, P., van den Berg, W., Gasciolli, V., Cottaz, S., Fort, S., Thomma, B.P.H.J., Bono, J.-J., et al. (2020). A lysin motif effector subverts chitin-triggered immunity to facilitate arbuscular mycorrhizal symbiosis. New Phytol. *225*, 448–460.
36. Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., and Kahmann, R. (2015). Fungal Effectors and Plant Susceptibility. Annu. Rev. Plant Biol. *66*, 513–545.
37. Snelders, N.C., Rovenich, H., and Thomma, B.P.H.J. (2022). Microbiota manipulation through the secretion of effector proteins is fundamental to the wealth of lifestyles in the fungal kingdom. FEMS Microbiol. Rev. *46*, fuac022. https://doi.org/10.1093/femsre/fuac022.
38. Moi, D., Bernard, C., Steinegger, M., Nevers, Y., Langleib, M., and Dessimoz, C. (2023). Structural phylogenetics unravels the evolutionary diversification of communication systems in gram-positive bacteria and their viruses. Preprint at bioRxiv. https://doi.org/10.1101/2023.09.19.558401.
39. Gupta, G.D., Bansal, R., Mistry, H., Pandey, B., and Mukherjee, P.K. (2021). Structure-function analysis reveals *Trichoderma virens* Tsp1 to be a novel fungal effector protein modulating plant defence. Int. J. Biol. Macromol. *191*, 267–276.
40. Zuo, N., Bai, W.-Z., Wei, W.-Q., Yuan, T.-L., Zhang, D., Wang, Y.-Z., and Tang, W.-H. (2022). Fungal CFEM effectors negatively regulate a maize wall-associated kinase by interacting with its alternatively spliced variant to dampen resistance. Cell Rep. *41*, 111877.
41. Pennington, H.G., Jones, R., Kwon, S., Bonciani, G., Thieron, H., Chandler, T., Luong, P., Morgan, S.N., Przydacz, M., Bozkurt, T., et al. (2019). The fungal ribonuclease-like effector protein CSEP0064/BEC1054 represses plant immunity and interferes with degradation of host ribosomal RNA. PLoS Pathog. *15*, e1007620.
42. de Guillen, K., Lorrain, C., Tsan, P., Barthe, P., Petre, B., Saveleva, N., Rouhier, N., Duplessis, S., Padilla, A., and Hecker, A. (2019). Structural genomics applied to the rust fungus *Melampsora larici-populina* reveals two candidate effector proteins adopting cystine knot and NTF2-like protein folds. Sci. Rep. *9*, 18084.
43. Seong, K., and Krasileva, K.V. (2021). Computational Structural Genomics Unravels Common Folds and Novel Families in the Secretome of Fungal Phytopathogen. Mol. Plant Microbe Interact. *34*, 1267–1280.
44. Saunders, D.G.O., Win, J., Cano, L.M., Szabo, L.J., Kamoun, S., and Raffaele, S. (2012). Using hierarchical clustering of secreted protein families to classify and rank candidate effectors of rust fungi. PLoS One *7*, e29847.
45. Singh, R.S., and Walia, A.K. (2014). Characteristics of lichen lectins and their role in symbiosis. Symbiosis *62*, 123–134.
46. 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.
47. Tripp, E.A., Zhuang, Y., and Lendemer, J.C. (2017). A review of existing whole genome data suggests lichen mycelia may be haploid or diploid. Bryologist *120*, 302–310.
48. Mansournia, M.R., Wu, B., Matsushita, N., and Hogetsu, T. (2012). Genotypic analysis of the foliose lichen *Parmotrema tinctorum* using microsatellite markers: association of mycobiont and photobiont, and their reproductive modes. Lichenologist *44*, 419–440.
49. Murtagh, G.J., Dyer, P.S., and Crittenden, P.D. (2000). Sex and the single lichen. Nature *404*, 564.
50. Moxham, T.H. (1981). Fusion of a detached lobe onto the parent thallus in the lichen *Xanthoria parietina*. Bryologist *84*, 363.
51. Muggia, L., Leavitt, S., and Barreno, E. (2018). The hidden diversity of lichenised Trebouxiophyceae (Chlorophyta). Phycologia *57*, 503–524.
52. Onuț-Brännström, I., Benjamin, M., Scofield, D.G., Heiðmarsson, S., Andersson, M.G.I., Lindström, E.S., and Johannesson, H. (2018). Sharing of photobionts in sympatric populations of *Thamnolia* and *Cetraria* lichens: evidence from high-throughput sequencing. Sci. Rep. *8*, 4406.
53. Nyati, S., Werth, S., and Honegger, R. (2013). Genetic diversity of sterile cultured *Trebouxia* photobionts associated with the lichen-forming fungus *Xanthoria parietina* visualized with RAPD-PCR fingerprinting techniques. Lichenologist *45*, 825–840.
54. Černajová, I., and Škaloud, P. (2019). The first survey of Cystobasidiomycete yeasts in the lichen genus *Cladonia*; with the description of *Lichenozyma pisutiana* gen. nov., sp. nov. Fungal Biol. *123*, 625–637.
55. Grimm, M., Grube, M., Schiefelbein, U., Zühlke, D., Bernhardt, J., and Riedel, K. (2021). The Lichens' Microbiota, Still a Mystery? Front. Microbiol. *12*, 623839.

56. Tran, K.N., Pham, N., Jang, S.-H., and Lee, C. (2020). Purification and characterization of a novel medium-chain ribitol dehydrogenase from a lichen-associated bacterium *Sphingomonas* sp. PLoS One *15*, e0235718.

57. Kono, M., Tanabe, H., Ohmura, Y., Satta, Y., and Terai, Y. (2017). Physical contact and carbon transfer between a lichen-forming *Trebouxia* alga and a novel *Alphaproteobacterium*. Microbiology (Reading) *163*, 678–691.

58. Armaleo, D., Müller, O., Lutzoni, F., Andrésson, Ó.S., Blanc, G., Bode, H.B., Collart, F.R., Dal Grande, F., Dietrich, F., Grigoriev, I.V., et al. (2019). The lichen symbiosis re-viewed through the genomes of *Cladonia grayi* and its algal partner *Asterochloris glomerata*. BMC Genomics *20*, 605.

59. Wang, Y.-Y., Liu, B., Zhang, X.-Y., Zhou, Q.-M., Zhang, T., Li, H., Yu, Y.-F., Zhang, X.-L., Hao, X.-Y., Wang, M., et al. (2014). Genome characteristics reveal the impact of lichenization on lichen-forming fungus *Endocarpon pusillum* Hedwig (Verrucariales, Ascomycota). BMC Genomics *15*, 34.

60. Kono, M., Kon, Y., Ohmura, Y., Satta, Y., and Terai, Y. (2020). In vitro re-synthesis of lichenization reveals the genetic background of symbiosis-specific fungal-algal interaction in *Usnea hakonensis*. BMC Genomics *21*, 671.

61. Scherrer, S., Haisch, A., and Honegger, R. (2002). Characterization and expression of *XPH1*, the hydrophobin gene of the lichen-forming ascomycete *Xanthoria parietina*. New Phytol. *154*, 175–184.

62. Riquelme, M., Aguirre, J., Bartnicki-García, S., Braus, G.H., Feldbrügge, M., Fleig, U., Hansberg, W., Herrera-Estrella, A., Kämper, J., Kück, U., et al. (2018). Fungal Morphogenesis, from the Polarized Growth of Hyphae to Complex Reproduction and Infection Structures. Microbiol. Mol. Biol. Rev. *82*, e00068-17. https://doi.org/10.1128/MMBR.00068-17.

63. Naranjo-Ortiz, M.A., and Gabaldón, T. (2020). Fungal evolution: cellular, genomic and metabolic complexity. Biol. Rev. Camb. Philos. Soc. *95*, 1198–1232.

64. Palmqvist, K., Dahlman, L., Jonsson, A., and Nash, T.H. (2008). The carbon economy of lichens. In Lichen Biology, T.H. Nash, ed. (Cambridge University Press), pp. 182–215.

65. Daskalov, A. (2023). Emergence of the fungal immune system. iScience *26*, 106793.

66. Han, G.-Z. (2019). Origin and evolution of the plant immune system. New Phytol. *222*, 70–83.

67. Scharnagl, K., Tagirdzhanova, G., and Talbot, N.J. (2023). The coming golden age for lichen biology. Curr. Biol. *33*, R512–R518.

68. Seong, K., and Krasileva, K.V. (2023). Prediction of effector protein structures from fungal phytopathogens enables evolutionary analyses. Nat. Microbiol. *8*, 174–187.

69. Lu, S., and Faris, J.D. (2019). *Fusarium graminearum* KP4-like proteins possess root growth-inhibiting activity against wheat and potentially contribute to fungal virulence in seedling rot. Fungal Genet. Biol. *123*, 1–13.

70. Allen, A., Snyder, A.K., Preuss, M., Nielsen, E.E., Shah, D.M., and Smith, T.J. (2008). Plant defensins and virally encoded fungal toxin KP4 inhibit plant root growth. Planta *227*, 331–339.

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

72. Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., and Wang, Z. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ *7*, e7359.

73. 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.

74. Seppey, M., Manni, M., and Zdobnov, E.M. (2019). BUSCO: Assessing Genome Assembly and Annotation Completeness. Methods Mol. Biol. *1962*, 227–245.

75. Flynn, J.M., Hubley, R., Goubert, C., Rosen, J., Clark, A.G., Feschotte, C., and Smit, A.F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. Proc. Natl. Acad. Sci. USA *117*, 9451–9457.

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

77. van Wyk, S., Harrison, C.H., Wingfield, B.D., De Vos, L., van der Merwe, N.A., and Steenkamp, E.T. (2019). The RIPper, a web-based tool for genome-wide quantification of Repeat-Induced Point (RIP) mutations. PeerJ *7*, e7447.

78. Palmer, J.M., and Stajich, J. (2020). Funannotate v1.8.1: Eukaryotic genome annotation (Zenodo). 10.5281/ZENODO.4054262.

79. Lomsadze, A., Ter-Hovhannisyan, V., Chernoff, Y.O., and Borodovsky, M. (2005). Gene identification in novel eukaryotic genomes by self-training algorithm. Nucleic Acids Res. *33*, 6494–6506.

80. Stanke, M., and Waack, S. (2003). Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics *19* (suppl 2), ii215–ii225.

81. Testa, A.C., Hane, J.K., Ellwood, S.R., and Oliver, R.P. (2015). CodingQuarry: highly accurate hidden Markov model gene prediction in fungal genomes using RNA-seq transcripts. BMC Genomics *16*, 170.

82. 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.

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

84. 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.

85. Chan, P.P., and Lowe, T.M. (2019). tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences. Methods Mol. Biol. *1962*, 1–14.

86. Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R., and Finn, R.D. (2018). HMMER web server: 2018 update. Nucleic Acids Res. *46*, W200–W204.

87. Buchfink, B., Reuter, K., and Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat. Methods *18*, 366–368.

88. 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.

89. UniProt Consortium (2023). UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. *51*, D523–D531.

90. Rawlings, N.D., Waller, M., Barrett, A.J., and Bateman, A. (2014). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. *42*, D503–D509.

91. 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.

92. Cantalapiedra, C.P., Hernández-Plaza, A., Letunic, I., Bork, P., and Huerta-Cepas, J. (2021). eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. Mol. Biol. Evol. *38*, 5825–5829.

93. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res. *47*, D309–D314.

94. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, G.A., Bileschi, M.L., Bork, P., Bridge, A., Colwell, L., et al. (2023). InterPro in 2022. Nucleic Acids Res. *51*, D418–D427.

95. 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.

96. Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., and Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Res. *35*, W182–W185.

97. Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007). WoLF PSORT: protein localization predictor. Nucleic Acids Res. *35*, W585–W587.

98. Hallgren, J., Tsirigos, K.D., Pedersen, M.D., Almagro Armenteros, J.J., Marcatili, P., Nielsen, H., Krogh, A., and Winther, O. (2022). DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. Preprint at bioRxiv. https://doi.org/10.1101/2022.04.08.487609.

99. Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., and Nielsen, H. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. *37*, 420–423.

100. Sperschneider, J., and Dodds, P.N. (2022). EffectorP 3.0: Prediction of Apoplastic and Cytoplasmic Effectors in Fungi and Oomycetes. Mol. Plant Microbe Interact. *35*, 146–156.

101. Kristianingsih, R., and MacLean, D. (2021). Accurate plant pathogen effector protein classification ab initio with deepredeff: an ensemble of convolutional neural networks. BMC Bioinformatics *22*, 372.

102. Kourelis, J., Sakai, T., Adachi, H., and Kamoun, S. (2021). RefPlantNLR is a comprehensive collection of experimentally validated plant disease resistance proteins from the NLR family. PLoS Biol. *19*, e3001124.

103. Allio, R., Schomaker-Bastos, A., Romiguier, J., Prosdocimi, F., Nabholz, B., and Delsuc, F. (2020). MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. Mol. Ecol. Resour. *20*, 892–905.

104. Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol. *20*, 238.

105. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 10.

106. Li, D., Liu, C.-M., Luo, R., Sadakane, K., and Lam, T.-W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics *31*, 1674–1676.

107. Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. *25*, 1043–1055.

108. Olm, M.R., Brown, C.T., Brooks, B., and Banfield, J.F. (2017). dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J. *11*, 2864–2868.

109. Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics *34*, 3094–3100.

110. Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics *36*, 1925–1927.

111. Parks, D.H., Chuvochina, M., Rinke, C., Mussig, A.J., Chaumeil, P.-A., and Hugenholtz, P. (2022). GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nucleic Acids Res. *50*, D785–D794.

112. Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. *32*, 268–274.

113. Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589–595.

114. 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.

115. Bao, W., Kojima, K.K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. Mobile DNA *6*, 11.

116. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. Bioinformatics *30*, 2068–2069.

117. 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.

118. 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.

119. Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res. *22*, 568–576.

120. Knaus, B.J., and Grünwald, N.J. (2017). vcfr: a package to manipulate and visualize variant call format data in R. Mol. Ecol. Resour. *17*, 44–53.

121. Kopylova, E., Noé, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics *28*, 3211–3217.

122. Henderson, G., Yilmaz, P., Kumar, S., Forster, R.J., Kelly, W.J., Leahy, S.C., Guan, L.L., and Janssen, P.H. (2019). Improved taxonomic assignment of rumen bacterial 16S rRNA sequences using a revised SILVA taxonomic framework. PeerJ *7*, e6496.

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

124. Pimentel, H., Bray, N.L., Puente, S., Melsted, P., and Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. Nat. Methods *14*, 687–690.

125. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. Omics *16*, 284–287.

126. Pignatelli, M., Serras, F., Moya, A., Guigó, R., and Corominas, M. (2009). CROC: finding chromosomal clusters in eukaryotic genomes. Bioinformatics *25*, 1552–1553.

127. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nat. Methods *19*, 679–682.

128. van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M., Söding, J., and Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. Nat. Biotechnol. *42*, 243–246.

129. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

130. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., et al. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. *50*, D439–D444.

131. Letunic, I., and Bork, P. (2024). Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res. *52*, W78–W82. https://doi.org/10.1093/nar/gkae268.

132. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., and Ferrin, T.E. (2023). UCSF ChimeraX: Tools for structure building and analysis. Protein Sci. *32*, e4792.

133. Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., Lieber, M., et al. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nat. Protoc. *8*, 1494–1512.

134. Gunawardhana, P.L.T. (2017). Genetic and kinetic basis of inorganic nutrient uptake in lichensUniversity of Nottingham, PhD thesis. https://eprints.nottingham.ac.uk/41845/.

135. Joneson, S., Armaleo, D., and Lutzoni, F. (2011). Fungal and algal gene expression in early developmental stages of lichen-symbiosis. Mycologia *103*, 291–306.

136. Hiltunen, M., Ament-Velásquez, S.L., and Johannesson, H. (2021). The Assembled and Annotated Genome of the Fairy-Ring Fungus

Marasmius oreades. Genome Biol. Evol. *13*, evab126. https://doi.org/10.1093/gbe/evab126.

137. Xavier, B.B., Miao, V.P.W., Jónsson, Z.O., and Andrésson, Ó.S. (2012). Mitochondrial genomes from the lichenized fungi *Peltigera membranacea* and *Peltigera malacea*: features and phylogeny. Fungal Biol. *116*, 802–814.

138. Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., Schulz, F., Jarett, J., Rivers, A.R., Eloe-Fadrosh, E.A., et al. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat. Biotechnol. *35*, 725–731.

139. Ament-Velásquez, S.L., Tuovinen, V., Bergström, L., Spribille, T., Vanderpool, D., Nascimbene, J., Yamamoto, Y., Thor, G., and Johannesson, H. (2021). The Plot Thickens: Haploid and Triploid-Like Thalli, Hybridization, and Biased Mating Type Ratios in Letharia. Front. Fungal Biol. *2*, 656386.

140. Kent, W.J. (2002). BLAT–the BLAST-like alignment tool. Genome Res. *12*, 656–664.

141. Jain, C., Rodriguez-R, L.M., Phillippy, A.M., Konstantinidis, K.T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat. Commun. *9*, 5114.

142. R Core Team (2020). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Nine samples of Xanthoria parietina lichen used for metagenomic sequencing	This paper	GT0240, GT0242-GT0244, GT0246, GT0249-GT0252
Seven samples of Xanthoria parietina lichen used for metatranscriptomic sequencing	This paper	XB1, XB2, XM2, XS1, XS2, XT1, XT2
Critical commercial assays		
RNeasy Plant Mini Kit	QIAGEN	Cat# 74904
DNeasy Plant Mini Kit	QIAGEN	Cat# 69104
NucleoBond High Molecular Weight DNA Kit	Macherey–Nagel	Cat# 740160.20
Circulomics Short Read Eliminator Kit	PacBio	Cat# 102-208-300
DNA ligation V14 kit	Oxford Nanopore	Cat# SQK-LSK114
Deposited data		
Raw metagenomic data, metagenomic assemblies, metagenome-assembled genomes, annotated genome of Xanthoria parietina mycobiont	This paper	ENA: PRJEB78723
Transcriptomic and metatrascriptomic data	This paper	ENA: PRJEB38537
Experimental models: Organisms/strains		
Xanthoria parietina mycobiont culture		N/A
Software and algorithms		
bowtie2	Langmead and Salzberg[71]	https://github.com/BenLangmead/bowtie2
MetaBAT2	Kang et al.[72]	https://bitbucket.org/berkeleylab/metabat/
EukCC v2.1.2	Saary et al.[73]	https://eukcc.readthedocs.io/
BUSCO5	Seppey et al.[74]	https://busco.ezlab.org/
RepeatModeler v2.0.3	Flynn et al.[75]	https://www.repeatmasker.org/RepeatModeler/
RepeatMasker v4.1.2	Smit et al.[76]	https://www.repeatmasker.org/
RIPper	van Wyk et al.[77]	https://theripper.hawk.rocks
Funannotate v1.8.15	Palmer and Stajich[78]	https://funannotate.readthedocs.io/
Genemark-ES v4.62	Lomsadze et al.[79]	https://genemark.bme.gatech.edu/
Augustus v3.3.2	Stanke and Waack[80]	https://bioinf.uni-greifswald.de/augustus/
CodingQuarry v2.0	Testa et al.[81]	https://sourceforge.net/projects/codingquarry/
GlimmerHMM v3.0.4	Majoros et al.[82]	https://ccb.jhu.edu/software/glimmerhmm/
SNAP 2006-07-28	Korf[83]	https://hpc.nih.gov/apps/snap.html
EVidence Modeler v1.1.1	Haas et al.[84]	https://github.com/EVidenceModeler/EVidenceModeler/wiki
tRNAscan-SE v2.0.9	Chan and Lowe[85]	https://lowelab.ucsc.edu/tRNAscan-SE/
HMMER v3.3.2	Potter et al.[86]	http://hmmer.org/
diamond v2.1.6	Buchfink et al.[87]	https://github.com/bbuchfink/diamond/wiki
PFAM v35.0	Mistry et al.[88]	http://pfam.xfam.org/
UniProtDB v2023_01	UniProt Consortium[89]	https://www.uniprot.org/
MEROPS v12.0	Rawlings et al.[90]	https://www.ebi.ac.uk/merops/
dbCAN v11.0	Yin et al.[91]	https://bcb.unl.edu/dbCAN2/index.php
Emapper v2.1.12	Cantalapiedra et al.[92]	https://github.com/eggnogdb/eggnog-mapper
Eggnog v5.0	Huerta-Cepas et al.[93]	http://eggnog5.embl.de/
InterProScan v5.42-78.0	Paysan-Lafosse et al.[94]	https://www.ebi.ac.uk/interpro/
antiSMASH v7.0	Blin et al.[95]	https://fungismash.secondarymetabolites.org/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
KAAS	Moriya et al.[96]	https://www.genome.jp/kegg/kaas/
WolfPSORT	Horton et al.[97]	https://wolfpsort.hgc.jp/
deepTMHMM	Hallgren et al.[98]	https://dtu.biolib.com/DeepTMHMM
SignalP v5	Almagro Armenteros et al.[99]	https://services.healthtech.dtu.dk/services/SignalP-5.0/
EffectorP v3.0	Sperschneider and Dodds[100]	https://effectorp.csiro.au/
deepredeff v.01.1	Kristianingsih and MacLean[101]	https://cran.r-project.org/web/packages/deepredeff/
RefPlantNLR	Kourelis et al.[102]	https://github.com/JKourelis/refplantnlR
MitoFinder v1.4.1	Allio et al.[103]	https://github.com/RemiAllio/MitoFinder
Orthofinder v2.5.4	Emms and Kelly[104]	https://github.com/davidemms/OrthoFinder
cutadapt v1.17	Martin[105]	https://cutadapt.readthedocs.io/
MEGAHIT v1.2.6	Li et al.[106]	https://github.com/voutcn/megahit
CheckM v1.2.0	Parks et al.[107]	https://ecogenomics.github.io/CheckM/
dRep v2.5.0	Olm et al.[108]	https://drep.readthedocs.io/en/latest/
Minimap2 v2.24-41122	Li[109]	https://github.com/lh3/minimap2
GTDB-Tk v1.7.0	Chaumeil et al.[110]	https://github.com/Ecogenomics/GTDBTk
GTDB v202	Parks et al.[111]	https://gtdb.ecogenomic.org/
IQ-TREE v2.2.2.2	Nguyen et al.[112]	http://www.iqtree.org/
metaMap	https://github.com/alexmsalmeida/metamap	https://github.com/alexmsalmeida/metamap
BWA v0.7.17-r1188	Li and Durbin[113]	https://github.com/lh3/bwa
Samtools v1.1	Li et al.[114]	https://www.htslib.org/
RepBase v18.08	Bao et al.[115]	https://www.girinst.org/repbase/
Prokka v1.14.6	Seemann[116]	https://github.com/tseemann/prokka
MAFFT v7.271	Katoh and Standley[117]	https://mafft.cbrc.jp/alignment/software/
trimAL v1.2	Capella-Gutiérrez et al.[118]	http://trimal.cgenomics.org/
Picard v2.21.2	https://broadinstitute.github.io/picard/	https://broadinstitute.github.io/picard/
Varscan v2.3.9	Koboldt et al.[119]	https://sourceforge.net/projects/varscan/
vcfR v1.15.0	Knaus and Grünwald[120]	https://cran.r-project.org/web/packages/vcfR
SortMeRNA v3.0.3	Kopylova et al.[121]	https://github.com/sortmerna/sortmerna
Silva v132	Henderson et al.[122]	https://www.arb-silva.de/
kallisto v0.46.2	Bray et al.[123]	https://pachterlab.github.io/kallisto/
sleuth v0.30.1	Pimentel et al.[124]	https://pachterlab.github.io/sleuth/
ClusterProfiler v4.2.2	Yu et al.[125]	https://doi.org/10.18129/B9.bioc.clusterProfiler
CROC	Pignatelli et al.[126]	https://github.com/emepyc/croc
ColabFold v1.5.0	Mirdita et al.[127]	https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb
FoldSeek v8.ef4e960	van Kempen et al.[128]	https://search.foldseek.com/
PDB	Berman et al.[129]	https://www.rcsb.org/
AlphaFold DB	Varadi et al.[130]	https://alphafold.ebi.ac.uk/
FoldTree	Moi et al.[38]	https://github.com/DessimozLab/fold_tree
iTOL v6	Letunic and Bork[131]	https://itol.embl.de/
ChimeraX v1.6.1	Meng et al.[132]	https://www.rbvi.ucsf.edu/chimerax/
Trinity v2.14.0	Haas et al.[133]	https://github.com/trinityrnaseq/trinityrnaseq/wiki
Other		
Scripts and code for analysis	This paper	https://github.com/metalichen/2024-Multipartite-complexity-omics-Xanthoria

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Lichen thalli

Lichen thalli were collected in Norwich Research Park (Norwich, UK; 52.623133°N, 1.221621°E). For the metagenomes, eight thalli of *X. parietina* were collected from tree bark and concrete (Data S5A). Three were incubated in a growth chamber for 19-21 months under a 12-h night/day light cycle. These thalli were sprayed weekly alternating between deionized water and liquid Bold's Mineral Medium (BMM). The rest of the thalli were sourced from the field and immediately used for DNA extraction following air drying. One additional thallus of *X. calcicola* was collected from concrete substrate. For the metatranscriptomes, seven thalli were collected from tree bark, tree twigs, concrete, and metal substrates (Data S5B).

Xanthoria parietina mycobiont culture

A pure culture of *X. parietina* mycobiont was kindly provided by Prof. Paul Dyer, University of Nottingham, UK. The culture was originally obtained from a single thallus collected in Wetton (Peak District, UK) in 2012. About 20 apothecia were excised from the thallus and discharged ascospores were collected onto a Petri dish with Malt Extract Yeast Extract (MEYE) media and grown at 18°C. The culture was maintained on solid MEYE medium with addition of streptomycin and ampicillin. The detailed procedure for the isolation of mycobiont can be found in Gunawardhana.[134]

METHOD DETAILS

Mycobiont genome sequencing and assembly

A fragment of a *X. parietina* thallus was cleared from all visible contaminants and all apothecia removed with a razor blade. Lichen material was homogenized with a Geno/Grinder homogenizer (SPEX SamplePrep, Metuchen NJ, USA) at 1300 rpm for 1 min. DNA was extracted with the NucleoBond High Molecular Weight DNA Kit (Macherey–Nagel, Düren, Germany). Short fragments were removed with Circulomics Short Read Eliminator Kit (Pacific Biosciences, Menlo Park, CA, USA) with 25 kbp cut-off. The resulting 0.6 μg of high-molecular weight DNA were used for long-read sequencing. The library was prepared using a DNA ligation V14 kit (Oxford Nanopore Technologies, Oxford, UK) and sequenced using a PromethION Flow Cell FLO-PRO114M (Oxford Nanopore Technologies, Oxford, UK). We used Dorado v0.2.4 (Oxford Nanopore Technologies, Oxford, UK) for base-calling. Contigs were assembled with Flye v2.9-b1780 with 'overlap 10 K, error rate 0.01, no-alt-contigs, meta' flags. Long-read sequencing and assembly were performed by Future Genomics (Leiden, Netherlands).

RNA extraction and sequencing

For transcriptomes of the mycobiont culture, the culture was ground under sterile conditions and bulked up in liquid MEYE medium grown under a 12-h night/day light cycle at 18°C. We plated the culture on nitrocellulose filters and incubated on 2% agar plates with BMM:MEYE 99:1 medium (following Joneson et al.[135]). The cultures were harvested at 2, 9, 21, and 42 days post inoculation; each time point had three replicates. The cultures were snap-frozen in liquid nitrogen and RNA was extracted using an RNeasy Plant Mini Kit (QIAGENE, Hilden, Germany). Total RNA was sent to Novogene UK (Cambridge, UK) and sequenced on an Illumina HiSeq2500 platform to PE150 data.

For lichen metatranscriptomes, we collected fresh partially hydrated samples and placed them in RNAlater (Thermo Fisher Scientific, Waltham, MA, USA) for two days. Next, each sample was separated into three developmental stages (central part, thallus edge, and apothecia) manually with a razor blade. Between 5-8 apothecia, depending on size, were used for the apothecia RNA extractions, which included the entire apothecia with the thalline margin. RNA was extracted and sequenced as described above. From seven thalli, we produced 17 metatranscriptomes, of which six were derived from the central part, four for the thallus edge, and seven from apothecia (Data S5B).

Mycobiont genome annotation

First, we removed non-mycobiont sequences from the assembly, by using a metagenomic binning approach. We used short-read data produced from the same lichen sample as the long-read assembly (Data S5A), and this was aligned to the long-read assembly with bowtie2.[71] Using the resulting alignment, we binned the assembly with MetaBAT2.[72] To identify the bins corresponding to the mycobiont genome, we used EukCC v2.1.2.[73] We also ran a BLASTx search against the NCBI-nr database using each contig as a query. The final MAG contained 58 contigs and was created by combining two bins and three unbinned contigs with hits to Lecanoromycete fungi. The quality of the MAG was assessed with EukCC and BUSCO5[74] using the ascomycota_odb10 database. The mitochondrial genome was detected using the same BLASTx search. To identify telomeric repeats, we used a script from Hiltunen et al.[136] with TAA[C]+ as a query.

Prior to gene annotation, we masked repeat elements in the genome. We created a custom repeat library using RepeatModeler v2.0.3[75] with the -LTRStruct flag. Using the repeat library, we masked repeats in the genome using stand-alone RepeatMasker v4.1.2 (https://www.repeatmasker.org/). To annotate repeat-induced point mutations we used RIPper.[77] Gene prediction and functional annotation was done with the Funannotate pipeline v1.8.15.[78] Gene prediction parameters were generated using the funannotate *train* module with the transcriptomic data from the mycobiont culture as an input. For gene prediction, we used the *predict* module, which performed *ab initio* prediction with Genemark-ES v4.62,[79] Augustus v3.3.2,[80] CodingQuarry v2.0,[81] GlimmerHMM

v3.0.4,[82] and SNAP 2006-07-28.[83] We created consensus models with EVidence Modeler v1.1.1[84] and annotated tRNA with tRNAscan-SE v2.0.9.[85] To create functional annotations, we used the *annotate* module, which runs HMMER v3.3.2 and diamond v2.1.6[87] searches against the following databases: PFAM v35.0,[88] UniProtDB v2023_01,[89] MEROPS v12.0,[90] dbCAN v11.0,[91] and BUSCO ascomycota_odb10.[74] In addition, we annotated the predicted proteins using Emapper v2.1.12[92] and Eggnog v5.0 database, InterProScan v5.42-78.0,[94] antiSMASH v7.0 web server,[95] and KAAS web server.[96] To further improve the gene annotation prediction, we employed a sequence homology-based approach. We used the orthogroup clustering method (see Identify lichen-enriched orthogroups below), focusing only on *X. parietina* genes. From these, we leveraged the previously identified functional annotations from Funannotate and assigned gene functions to orthogroups. If at least 40% of the genes within an orthogroup were annotated in *X. parietina*, we assigned the remaining *Xanthoria* genes the same functional annotation.

To predict the secretome, we used three tools: WolfPSORT,[97] deepTMHMM,[98] and SignalP v5.[99] We defined a protein as being putatively secreted using three criteria: signal peptide identified by SignalP, no transmembrane domains identified by deepTMHMM, and the probability of being secreted of $\geq$0.6 according to WolfPSORT. To identify spatial clusters of genes encoding secreted proteins, we used CROC[126] with default parameters (window size of 30,000 bp, sliding window offset of 10,000 bp, p-value <0.05, minimum three genes in a cluster, Benjamini & Hochberg method of multiple correction). All secreted proteins were analyzed with two effector-predicting tools: EffectorP v3.0[100] and deepredeff v.01.1.[101] To identify NLR-like proteins, we used a custom script filtering proteins based on their InterProScan domains (Data S1D); the list of domains typical for fungal NLRs we took from Uehling et al.[30]; the visualization script was partially based on RefPlantNLR.[102] To annotate the *MAT* locus, we ran a BLASTp search against the predicted proteome. As a query, we used *MAT* genes from *X. polycarpa* (GenBank IDs: CAI59767.1, CAI59768.1, CAI59769.1, CAI59770.1, CAI59771.1, CAI59772.1). We used the same queries to screen metagenomic assemblies and raw reads (see below). To identify putative polyol transporters, we used a BLASTp search with four known transporters (GenBank IDs: AAX98668.1, CAR65543.1, CAG86001.1, NP_010036.1) as queries; hits with the e-value<1e-100 were considered.

We annotated the mitochondrial genome using MitoFinder v1.4.1.[103] As a reference, we used the mitochondrial genome of *Peltigera malacea* from Xavier et al.[137] We added the missing *rrnL* annotation manually based on the BLASTn search results.

Identification of lichen-enriched orthogroups

To identify orthogroups enriched in lichen-forming fungi, we analyzed a dataset of 44 fungal species, including the newly acquired *X. parietina* genome and reference genomes from 17 lichen-forming and 26 non-lichen-forming fungi (Data S2D). We employed Orthofinder v2.5.4[104] to classify proteins from these species into orthogroups. The copy number matrix from these orthogroups, was then subjected to the fisher.test function in R to identify orthogroups that have an overrepresentation of genes present predominantly in lichen-forming fungi when compared to fungi that do not form lichens. This function uses an ABCD matrix to calculate the enrichment, where A represents the total number of genes in a specific orthogroup in lichen-forming fungi, B represents the total number of genes in the same orthogroup among non-lichen forming fungi, C represents the total number of genes in remaining orthogroups in lichen-forming fungi and D represents the total number of genes in remaining orthogroups among non-lichen-forming fungi. The orthogroups significantly enriched with lichen genes were ones with a Benjamin-Hochberg corrected p-value $\leq$0.05.

Metagenome sequencing and analysis

Eight samples of *X. parietina* and one of *X. calcicola* were collected from either bark or concrete (Data S5A). Of them, three *X. parietina* samples collected from bark were subjected to incubation in a growth chamber for 19-21 months (see above). Prior to DNA extraction, the samples were air-dried and homogenized as described above. DNA was extracted with a DNeasy Plant Mini Kit (QIAGENE, Hilden, Germany) and sequenced on an Illumina NovaSeq 6000 platform by Novogene UK (Cambridge, UK).

Metagenomic data from *X. parietina* samples were cleared from human contamination by aligning to the reference human genome with bowtie2.[71] We removed adapters using cutadapt v1.17.[105] The filtered data were assembled using MEGAHIT v1.2.6.[106] We ran both individual assemblies for each sample, and co-assembly of all *X. parietina* samples. Next, all assemblies were binned with MetaBAT v2.15.[72] To identify eukaryotic MAGs and assign them preliminary taxonomic assignments, we screened all bins with EukCC2.[73] For prokaryotic MAGs, we used CheckM v1.2.0.[107] Next, we selected all high and medium quality MAGs (completeness $\geq$50%, contamination <10%)[138] and dereplicated them using dRep v2.5.0[108] at 95% ANI (average nucleotide identity) and 40% AF (alignment fraction) thresholds in order to obtain species-level representatives.

To produce taxonomic assignments for the eukaryotic MAGs, we combined them with reference genomes (Data S2D) and built phylogenomic trees (Data S2E-F). The MAGs were split into two groups – fungal and algal – based on the annotations from EukCC. To the fungal tree we also added the long-read assembly of the mycobiont. For the two reference algal genomes that lacked annotations (Data S2D), we ran BUSCO5 with the chlrophyta_odb10 database and used the predicted proteins for the analysis. The species tree was generated using OrthoFinder v2.5.4. The MAG of the mycobiont was identified based on its position in the phylogenomic tree. To confirm this, we aligned it against the long-read genome assembly of *X. parietina* mycobiont using Minimap2 v2.24-41122.[109] We further identified the three non-mycobiont fungal MAGs from classes Lecanoromycetes and Lichinomycetes by computing a maximum-likelihood phylogenomic tree of RPB2 marker gene (Figure S2A). We used BLAST to extract RPB2 sequences with GenBank sequences LC322052.1 and AYN25505.1 as queries. We combined these sequences with 91 reference sequences (Data S2G) and aligned them using MAFFT v7.271121 with the –maxiterate 1000 flag. The alignment was clipped using trimAL v1.2122 to remove positions present in <70% of sequences. The phylogeny was calculated using IQ-TREE116.

To assign taxonomy to the bacterial MAGs, we used GTDB-Tk v1.7.0[110] with the GTDB database v202.[111] From the alignment of 120 marker genes produced by GTDB-Tk, we generated a maximum-likelihood phylogeny using IQ-TREE v2.2.2.2[112] (Data S2H).

To map the presence/absences of species-level lineages across the metagenomic samples, we used the metaMap pipeline (https://github.com/alexmsalmeida/metamap). Reads from all metagenomes, including the additional *X. calcicola* sample, were aligned against the entire MAG catalog with BWA v0.7.17-r1188.[113] Secondary alignments were removed using Samtools v1.10.[114] All MAGs covered $\geq$50% in a given metagenome were counted as present. To calculate the depth of coverage, we multiplied the number of reads aligned to the MAG by the read length and divided by the total length of the MAG.

For metatranscriptomic analysis, all MAGs except for the mycobiont MAG, were annotated. First, we filtered each MAG using Funannotate modules *clean* and *sort* to remove contigs shorter than 500 bp and showing >95% overlap with other contigs. We masked repeats in the eukaryotic MAGs using RepeatMasker and the RepBase database v18.08.[115] For fungal MAGs we used fngrep.ref, which contains repeats from across fungi; for algal MAGs we used chlrep.ref, which contains annotated repeats from *Chlamydomonas*. Next, we ran the *predict* module as described above. For training Augustus, we used the BUSCO dikarya_odb9 database for fungi and chlorophyta_odb10 for algae. Bacterial genomes were annotated using Prokka v1.14.6.[116]

Confirming sample identities

To confirm the identity of mycobionts from metagenomic and genomic samples, we ran a BLASTn search to extract the ITS region (ITS1, 5.8S ribosomal RNA gene, ITS2) using JF831902.1 *X. parietina* as the query. We combined the extracted ITS sequences with 338 reference sequences from various Teloschistaceae (Data S5C) and computed a maximum-likelihood phylogeny using IQ-TREE as described above (Figure S2D).

Ploidy analysis

To calculate the minor allele frequency distributions of the mycobiont genome, we adapted the pipeline from Ament-Velásquez et al.[139] We aligned the metagenomic short-read data to the long-read mycobiont genome assembly using BWA v0.7.17-r1188[113] with PCR duplicated marked with Picard v2.21.2 (https://broadinstitute.github.io/picard/). We called variants using Varscan v2.3.9[119] with the flags –p-value 0.1 –min-var-freq 0.005. We removed contigs shorter than 100 kpb and filtered out variants overlapping with repeat elements. The resulting vcf file was processed using the vcfR library v1.15.0.[120]

Transcriptomic analysis

We trimmed the data to remove adaptors and poly-A tails with cutadapt.[105] To remove rRNA contamination, we used SortMeRNA v3.0.3[121] using the Silva database v132.[122] Next, we created a reference index by combining predicted coding sequences from the annotated MAGs and the long-read mycobiont genome. We pseudoaligned the transcriptomic data to the index using kallisto v0.46.2.[123] For differential gene expression analysis of the mycobiont, we used sleuth v0.30.1.[124] We only included data mapped to the transcripts from the predicted mycobiont transcriptome, and the transcript per million values were normalized across all samples. Genes were identified as differentially expressed if they had |b-value| >1 and P-adjust < 0.05. To compare samples from different developmental stages, we controlled for the thallus identity following https://pachterlab.github.io/sleuth_walkthroughs/pval_agg/analysis.html. For enrichment analysis, we used ClusterProfiler v4.2.2.[125] To identify clusters of differentially-expressed genes, we used CROC[126] with default parameters (see above).

To estimate genetic differences between the lichen sample used for culturing and the thalli used for metatranscriptomics, we ran the following analysis. We selected two RNA-seq libraries, one representing the culture (KS48XB1) and one representing metatranscriptomic samples (XBE1). We assembled each library independently using Trinity v2.14.0.[133] To select transcripts originating from the mycobiont, we aligned the assemblies to the predicted transcriptome (obtained by *de novo* annotation of the reference genome) using BLAT v37x1.[140] Next, we used FastANI v1.3[141] to calculate ANI. The predicted transcriptome showed 99.22% similarity to the XBE1 transcriptome and 99.28% similarity to the KS48XB1 sample, suggesting that both samples are equally close to the sample used for the reference genome.

Protein structure prediction and analysis

We predicted structures of the proteins from the predicted secretome using ColabFold v1.5.0.[127] We used FoldSeek v8.ef4e960[128] to search the structures against two databases: PDB[129] (downloaded on 2023.12.11) and AlphaFold[130] (downloaded on 2024.04.18). We only retained the hits with e-value <0.001. All protein structures with pTM (template modeling score) $\geq$0.5 were subjected to structural clustering. We used the 0.5 threshold following Seong and Krasileva.[68] For clustering, we first removed the signal peptide (as identified by SignalP, see above) and disordered regions, defined as residues with pLDDT (predicted local-distance difference score) $\leq$0.55. Next, we constructed a structural phylogenetic tree using FoldTree.[38] Based on the LDDT tree produced by FoldTree, we manually curated a set of clusters with similar protein structures. One cluster (cl42) included only one protein. It was designated cluster status due to its similarity to a known effector. To visualize the structural tree, we used iTOL v6.[131] The protein models were visualized using ChimeraX v1.6.1.[132]

QUANTIFICATION AND STATISTICAL ANALYSIS

We performed statistical analysis in R v4.1.2.[142] To identify lichen-enriched orthogroups, we used Fisher test (function fisher.test). In the list, we included the orthgroups with a Benjamin-Hochberg corrected p-value $\leq$0.05.