The lichen symbiosis re-viewed through the genomes of *Cladonia grayi* and its algal partner *Asterochloris glomerata*

Daniele Armaleo1, Olaf Müller1,2, François Lutzoni1, Olafur S. Andrésson3, Guillaume Blanc4, Helge B. Bode5, Frank R. Collart6, Francesco Dal Grande7, Fred Dietrich3, Igor V. Grigoriev8,9, Suzanne Joneson1,10, Alan Kuo8, Peter E. Larsen6, John M. Logsdon Jr11, David Lopez12, Francis Martin13, Susan P. May1,14, Tami R. McDonald1,15, Sabeeha S. Merchant9,16, Vivian Miao17, Emmanuelle Morin13, Ryoko Oono18, Matteo Pellegrini19, Nimrod Rubinstein20,21, Maria Virginia Sanchez-Puerta22, Elizabeth Savelkoul11, Imke Schmitt7,23, Jason C. Slot24, Darren Soanes25, Péter Szövényi26, Nicholas J. Talbot27, Claire Veneault-Fourrey13,28 and Basil B. Xavier3,29

Abstract

**Background:** Lichens, encompassing 20,000 known species, are symbioses between specialized fungi (mycobionts), mostly ascomycetes, and unicellular green algae or cyanobacteria (photobionts). Here we describe the first parallel genomic analysis of the mycobiont *Cladonia grayi* and of its green algal photobiont *Asterochloris glomerata*. We focus on genes/predicted proteins of potential symbiotic significance, sought by surveying proteins differentially activated during early stages of mycobiont and photobiont interaction in coculture, expanded or contracted protein families, and proteins with differential rates of evolution.

**Results:** A) In coculture, the fungus upregulated small secreted proteins, membrane transport proteins, signal transduction components, extracellular hydrolases and, notably, a ribitol transporter and an ammonium transporter, and the alga activated DNA metabolism, signal transduction, and expression of flagellar components. B) Expanded fungal protein families include heterokaryon incompatibility proteins, polyketide synthases, and a unique set of G-protein α subunit paralogs. Expanded algal protein families include carbohydrate active enzymes and a specific subclass of cytoplasmic carbonic anhydrases. The alga also appears to have acquired by horizontal gene transfer from prokaryotes novel archaeal ATPases and Desiccation-Related Proteins. Expanded in both symbionts are signal transduction components, ankyrin domain proteins and transcription factors involved in chromatin remodeling and stress responses. The fungal transportome is contracted, as are algal nitrate assimilation genes. C) In the mycobiont, slow-evolving proteins were enriched for components involved in protein translation, translocation and sorting.

(Continued on next page)
Background
Simon Schwendener in 1869 [1] correctly recognized lichens as intimate symbioses between specialized fungi and phototrophic unicellular green algae as the main symbionts. Cyanobacteria were later recognized also as primary phototrophs in many lichens. In addition to the main symbiotic partners, lichens harbor diverse communities of prokaryotes and fungi as cohabitants [2–6]. Recently, highly coevolved basidiomycete yeasts were discovered in the cortex of many lichens [7], sometimes causing disease [8]. The detailed interactions of the various cohabitants with the main symbionts are being investigated [9]. Typically, lichens thrive in above-ground niches with limited water in diverse environments, often withstanding extreme heat, desiccation, or cold [3, 10]. Widespread across terrestrial ecosystems, often dominant carbon and nitrogen fixers in alpine, subalpine, and high latitude habitats, the estimated 18,000 to 20,000 lichen species [11], mostly ascomycetes, represent about 20% of all known fungi [12]. There are only about 120 lichen photroph species (photobionts) [10, 13], far fewer than the 20,000 known lichen fungal species (mycobionts). Lichens are named based on their mycobiont since the fungus is the most conspicuous partner and since the same photobiont species (alga or cyanobacterium) can be found in several different lichens.

Lichens can reproduce somatically through propagules comprising both symbionts, or sexually through meiotic fungal spores that must combine with the appropriate photobiont to re-form a lichen. Sexual reproduction is not commonly seen in trebouxoid lichen algae, although evidence supporting it has been found (Heterothallism probably evolved from homothallism in Cladonia; genetic evidence for sex in Asterochloris section). Lichens are well known for their unique and abundantly produced secondary metabolites [14, 15]. The genetic, physiological, and structural integration of mycobionts and photobionts has produced a vast array of beautifully differentiated partnerships [16], with only occasional instability [17–19]. The fungal and lichen fossil record [20, 21] has placed fossils resembling extant lichen taxa in the Devonian-early Carboniferous, 415–350 million years ago, and perhaps simpler mycobiont-photobiont associations even earlier [22, 23]. Despite their intimate coexistence for hundreds of millions of years and the construction of complex interfaces between them [10, 24], lichen symbionts have not lost their genetic and cellular independence. Cell membranes are not breached in their interactions and genomes are not merged, although some have extended to lichens the concept of genome acquisition [25]. There is no evidence of horizontal gene transfer (HGT) between the symbionts, yet both have acquired genes from other sources (A low-GC region in Asterochloris is a remnant of a large virus insertion, an HGT-mediator, Inferences from differential transcription about nutritional fluxes at the symbiotic interface, and Photobiont expanded families sections). The partners of many lichens have been isolated and grown separately in axenic culture, but free-living stages of lichen fungi and their algae remain mostly cryptic in nature [18, 26–29]. However, these stages are not insignificant for the lichen life cycle [29]. The laboratory reconstitution of cultured lichen symbionts into fully developed lichens has a checkered history [30], where the reproducibility needed for molecular investigations is still elusive. Molecular studies addressing functional aspects of this mutualistic symbiosis are few [9, 31–35], but the recent publication of several lichen -omics papers and datasets [9, 36–46] heralds expansion of this field.

Here we present the first parallel genomic analysis of both primary symbionts in a lichen, the fungus (mycobiont) Cladonia grayi and the alga (photobiont) Asterochloris glomerata, and use several approaches to identify genes/proteins of potential symbiotic relevance. Our analysis is based exclusively on the symbionts’ nucleic acid sequences, and the proteins involved are predicted. Cladonia grayi (Fig. 1) belongs to a genus with
worldwide distribution, part of the class of Lecanoromycetes that includes 70% of the known lichens [47] (phylum Ascomycota, subphylum Pezizomycotina). The unicellular photobiont, *A. glomerata*, belongs to the most common order of lichen algae, the Trebouxiales [13, 48]. There are few sequenced genomes from unicellular chlorophyte algae [49–55]: some are naturally free-living and some, like *Coccomyxa subellipsiidea* [54] and *Chlorella variabilis*, are facultative symbionts [50, 54, 56]. Genomic analysis of lichens will not only increase the molecular and ecological understanding of a large and understudied portion of the fungal and algal phyla but also complement the emerging genomics of other symbioses involving mycorrhizal [57–60], endophytic [61, 62], or plant pathogenic fungi [63].

## Results and discussion

### General characteristics of the *C. grayi* and *A. glomerata* genomes

#### Genome sizes and gene organization

The mycobiont is a single-spore isolate from *C. grayi* and the photobiont is *Asterchloris glomerata* isolated from *C. grayi* soredia [64]. Table 1 includes basic features of the two nuclear and the three organelle genomes. Organelle genomes are briefly discussed in Additional file 1, and details are in Xavier et al. [65] and Xavier’s thesis [66]. Whole genome assemblies and annotations are at [67] for the mycobiont and at [68] for the photobiont. Relationships of *C. grayi* and *A. glomerata* within broad phylogenetic contexts, genome sizes, and proportions of repeated and unique sequences are shown in Fig. 2. The nuclear genomes of lichen symbionts are not reduced in size nor gene content compared to free-living relatives, in contrast to the reductions observed in many host-dependent bacteria [69]. With its 35 Mb genome and 11,400 gene models, the *C. grayi* mycobiont falls in the average size range for most Ascomycota [70]. Other lichen fungi fall in the same range, between 26 and 59 Mb [46]. Large increases in the number of transposable elements significantly affect genome size in many biotrophic fungi [71], including the ectomycorrhizal ascomycetes *T. melanosporum* [58], *E. granulatus* [72], and *C. geophilum* [73] but this was not observed in *C. grayi* (Fig. 2). Like other Chlorophyta, *A. glomerata* has more and larger introns than fungi. Its genome (56 Mb and 10,000 gene models) is significantly

### Table 1 Genome Basics

| Genome Type | C. grayi | A. glomerata |
|-------------|----------|-------------|
| Coverage    | 15x      | 24.8 x      |
| Number of scaffolds | 414 | 151 |
| Genome size (Mb) | 35 | 56 |
| Number of predicted genes | 11,388 | 10,025 |
| Number expressed in thallus | 9800 (86%) | 7700 (77%) |
| Genes per million bases | 288 | 173 |
| Average # of introns per gene | 3 | 9 |
| Average gene (mRNA) length | 1800 (1650) | 4240 (1400) |
| Intergenic DNA | ~ 45% | ~ 26% |
| Repetitive DNA | ~ 10% | ~ 5% |

### Table 1 Genome Basics

| Organelle Genomes | Basepairs | Proteins | Unknown ORFs | tRNAs |
|-------------------|-----------|----------|--------------|-------|
| Fungal mitochondrion | 50,836 | 15 | 1 | 26 |
| Algal mitochondrion | 11,093 | 32 | 18 | 25 |
| Chloroplast | 217,546 | 73 | 1 | 30 |
Fig. 2 Phylogenies, genome sizes and sequence distribution. **Left side:** Fungal (top) and algal (bottom) PhyML trees (LG + G + F + I) for *C. grayi* and *A. glomerata* involving, respectively, a random sample of 6000 and 4000 ungapped sites extracted from a concatenated alignment of 2137 and 683 orthologous protein families containing 794,828 and 159,356 ungapped sites. Bootstrap support values label internodes. Scales indicate nucleotide substitutions per site.

**Right side:** Bars are proportional to genome size, and different shadings indicate the proportions of recent and older sequence replicas or of unique sequences. Duplicated sequences in genomes were revealed by BLAST alignment of the genomic sequence against itself at the nucleotide (BLASTN) or amino acid (TBLASTX) levels. The duplicated regions include regular genes as well as repeated elements (not yet fully characterized), but microsatellites and low complexity sequences were filtered out. Sequences that matched in both BLASTN and TBLASTX searches were only counted in the BLASTN category. Only alignments with e-values $<1 \times 10^{-15}$ in both the BLASTN and TBLASTX analyses were considered.
smaller than that of *C. reinhardtii* (120 Mb) [49] but is larger than that of other Trebouxiophyceae like *C. subellipsoidae* C-169 (49 Mb) [54] and *C. variabilis* NC64A (46.2 Mb) [50] (Fig. 2). *C. subellipsoidae* C-169 is free-living, but the genus includes lichenized species [56, 74]. *Chlorella* NC64A is a facultative symbiont of ciliates and is host to large dsDNA viruses. Our analyses suggest that *A. glomerata* has also been host to large DNA viruses (A low-GC region in *Asterochloris* is a remnant of a large virus insertion, an HGT-mediator section), although a live virus has not yet been isolated from it. Over evolutionary time, chromosomal rearrangements left little synteny among the genomes of *A. glomerata*, *Coccomyxa* C-169 and *Chlorella* NC64A (Additional file 2).

Additional file 3 shows a KEGG-based categorization of the mycobiont and photobiont gene models. In this broad overview only the environmental information processing category (signal transduction) appears overrepresented in both symbiotic partners. Among the free-living *Aspergillus* species, from the closely related class Eurotiomycetes (Fig. 2), the signal transduction genes constitute between 1.4 and 1.84% of the annotated genes [75, 76], while in *C. grayi* the proportion is 6.2%. In *A. glomerata* the proportion of signal transduction genes is 7.8%, while among other Chlorophyta they represent 5–6% of the total [77]. These broad comparisons are only suggestive of an expansion of signal transduction components in the *C. grayi* partners because methodologies and annotations differ. A specific analysis of signal transduction functions (Specific survey of mycobiont and photobiont signal transduction components section) also reveals diversification in some of the *C. grayi* and *A. glomerata* components. This bilateral restructuring may underpin the multifaceted interactions between partners [10].

**A low-GC region in *Asterochloris* is a remnant of a large virus insertion, an HGT-mediator**

In 98% of the *Asterochloris* nuclear genome the GC content is between 56 and 62%. However, it is significantly lower (49%) in two large genomic regions, each located at one end of two scaffolds (~441 Kb on scaffold 80 and ~102 Kb on scaffold 120). Each low-GC region reaches the scaffold’s extremity with an array of duplicated 141-bp sequence units (totaling 1300 bp on scaffold 80 and 2053 bp on scaffold 120). These repeated sequences are found nowhere else in the *Asterochloris* genome. Figure 3a has the two scaffolds joined at the repeats, forming a single low-GC contiguous chromosomal region. Genomic contiguity has been confirmed by PCR and sequencing across the junction (Armaleo, not shown). The joined low-GC regions contain 462 predicted protein coding genes (Additional file 4), 236 of which exhibit significant matches in GenBank (BLASTP e-value <1e-5). Of these, 45% have their best match in double-stranded DNA viruses [78]. While most genes in the algal genome have many introns, only 36 of the 462 protein coding genes in this region are predicted to have them, and only 24 of 462 (5.2%) match chlorophyte genes. This differs markedly from the rest of the genome (Fig. 3b), where most genes have best matches in chlorophytes (69%). The sharp switch in nucleotide composition and phylogenetic affinity strongly suggest that the low-GC region is a remnant of a large integrated viral genome, about 540 kb long. Nucleo-cytoplasmic large DNA viruses (NCLDV) form a monophyletic class of viruses that infect a variety of eukaryotes [78, 79], including other algae and protists [80–82]. A phylogenetic analysis places the *Asterochloris* virus within the Phycodnaviridae family, sister to viruses that infect other green algae (Additional file 4). The genome of the *A. glomerata* virus may be the largest among alga-infecting NCLDV sequenced to date (ranging from 154 Kb for a *Feldmannia* sp. virus [83] to 473 Kb for a *Chrysochromulina ericina* virus [84]). Viral DNA thus is a major vehicle of HGT in *A. glomerata* and other algae [85]. The significance of this group of virally transferred genes for the symbiosis is unclear at this time. However, some of the genes in the viral region are actively transcribed, which suggests that they may eventually become functional in the photobiont. Other genes with potential symbiotic significance have been introduced into trebouxiod algae probably by HGT from bacteria [44] and archaea (Photobiont expanded families section). Trebouxiod algae ancestors may have even acquired genes from fungi before the origins of lichens [86].

**Heterothallism probably evolved from homothallism in Cladonia; genetic evidence for sex in Asterochloris**

Typically in ascomycetes, two kinds of mating genes, *MAT1–1* and *MAT1–2*, cooperate in mating [87]. They are referred to as idiomorphs because, while they share the same locus, *MAT1*, their encoded proteins are different transcription factors: *MAT1–1* is characterized by an α1-domain [88] and *MAT1–2* by a MAT A_− HMG domain [88]. These may be linked also to other idiomorph-specific genes. When both *MAT1–1* and *MAT1–2* are in the same genome, the fungus is self-fertile (homothallic), but when a genome contains only *MAT1–1* or only *MAT1–2*, the fungus is self-sterile (heterothallic) and mating occurs only between different mating types; both heterothallic and homothallic species of lichen fungi have been found [89–93]. *C. grayi* produces typical ascomycetous fruiting bodies with apothecia (Fig. 1); the *MAT* locus in the single-spore isolate Cgr/DA2myc/ss, like in many Pezizomycota [87, 94], is located between *Apn2* and *Sla2*, genes for a nuclease and a cytoskeleton assembly protein, respectively (Fig. 4a). The locus in Cgr/DA2myc/ss contains the core
gene \textit{MAT1--1} and an associated gene [87], \textit{MAT1--1-7}, but no \textit{MAT1--2} (Additional file 5 and Fig. 4a), as do the sequenced mating-type loci of single-spore isolates from two other \textit{Cladonia} species (Additional file 5). This expands earlier RAPD-PCR and RFLP data on three other \textit{Cladonia} species, also heterothallic [91]. PCR amplification and sequencing of the region between \textit{Sla2} and \textit{Apn2} from DNA of a natural \textit{C. grayi} thallus revealed a single \textit{MAT1--2} gene (Additional file 5 and Fig. 4a; sequence accession MH795990), further supporting heterothallism. The sequences of the two loci revealed the presence of vestigial sequences suggesting that, in \textit{Cladonia}, heterothallism might have evolved from homothallism (Additional file 5 and Fig. 4b). This appears to diverge from the trends in other ascomycete genera where homothallism is thought to have evolved from heterothallism [94--96], although not in all cases [97].

Sexual reproduction is generally assumed not to occur in trebouxoid algae like \textit{Asterochloris} [98]. They reproduce vegetatively, both in the lichen thallus or in culture, through non-motile autospores and aplanospores or through flagellated zoospores [99]. Law and Lewis proposed that sexual reproduction in the photobiont, the inhabitant in the mutualism, should be selected against [100]. However, the occurrence of sex in these algae is indicated by isolated microscopic observation of presumptive gametes in the 1920s [99], 1960s (figure 28 [101]; page 135 and figure 28 [102]) and in 2015 [99], and molecular evidence of genetic recombination was also uncovered through phylogenetic analysis of a \textit{Trebouxia} population in \textit{Letharia} lichens.
The identification of many meiotic genes in the *A. glomerata* genome and their expression in coculture (Additional file 5) add further evidence for the occurrence of meiosis in trebouxoid algae, probably in their free-living stages [29]. The gametes observed in *Asterochloris* [99] are flagellated, as are its vegetative zoospores (9 + 2 type [104]). Not surprisingly, the motility proteins present in *Asterochloris* match those of other flagellated chlorophytes but are mostly absent from non-motile chlorophytes (Fig. 5 and Additional file 5). The critical implications of fungal and algal sexual reproduction for the lichen symbiosis and for symbiont autonomy are discussed in Conclusions.

**Search for symbiosis-specific genes I: differential expression in coculture vs. monoculture**

Table 2 lists the genes discussed in the next three sections. The system to survey the gene sets differentially transcribed when fungus and alga first contact each other is based on three parallel cultures on filters placed on low nutrient agar medium [33]: the aposymbiotic fungus, the aposymbiotic alga, and a coculture of the two (Fig. 6 and Additional file 6). RNA was extracted from the individual cultures after 21 days of growth on the filters. Differential expression data (Methods) were analyzed in two ways. In the first, we identified GO annotations and metabolic pathways significantly enriched...
among genes induced or repressed more than 1.5x (0.6 > \log_2 > -0.6) in coculture relative to monoculture.

In the second, a more extensive approach, we focused only on induced genes and added information from PFAM domains and literature searches to define each gene’s putative function in more detail than was possible through GO terms alone.

**GO term-centered analysis of transcription in coculture**

The limited results are listed in Additional file 6. In coculture, the alga shows only differential activation of ubiquitin-dependent catabolic processes. The fungus induces redox-active enzymes and proteases, a finding also reflected in the data from the more extensive approach. Surprisingly, however, in presence of the alga the fungus also down-regulates several genes involved in respiratory ATP generation; the biological significance of this remains to be determined, pending experimental confirmation.

**Extended analysis of transcription induced in coculture**

In this analysis we used day-21 ratios of coculture (Co) vs. monoculture (Mo) RPKM directly rather than the

---

**Fig. 5** Flagellar proteins. **a** Number of candidate flagella proteins in chlorophytes. Reference *C. reinhardtii* proteins of the CiliaCut protein set (blue bars) and flagella proteome (green bars) were searched for putative orthologs in sequenced motile and non-motile chlorophytes using the reciprocal best BLASTP hit criterion. **b** The 314 candidate *A. glomerata* flagella proteins identified from multiple sources of evidence (see Methods). **c** Distribution of flagella proteins across Chlorophytes. The left cladogram shows the likely evolutionary relationships of sequenced Chlorophytes. The f mark indicates organisms known to build motile flagella. Crei: *Chlamydomonas reinhardtii*; Volvox: *Volvox carteri*; NC64A: *Chlorella variabilis* NC64A; Otau: *Ostreococcus tauri*; Oluc: *Ostreococcus lucimarinus*; M. CCMP: *Micromonas pusilla* CCMP1545; M. RCC: *Micromonas* sp. RC299. Presence (dot) or absence (circle) of putative orthologs identified by reciprocal best BLASTP hit of *C. reinhardtii* outer dynein proteins, inner dynein proteins, radial spoke proteins, central pair proteins and intraflagellar transport proteins.
log₂ values. Ratios are abbreviated here as Co/Mo. In each symbiont, a few hundred genes changed expression in coculture relative to monoculture, while most genes remained unaffected (Co/Mo ≅ 1) (Fig. 7). The extended transcription analysis was limited to the genes induced in coculture. Due to differences in Co/Mo ranges for the symbionts (Additional file 6), we defined two different Co/Mo induction thresholds: ≥ 2 for the fungus, ≥1.3 for the alga (Fig. 7). This yielded 795 up-regulated genes out of 11,388 in C. grayi (7%) and 471 out of 10,024 in A. glomerata (4.7%) (Additional file 6). Induced gene products were inferred by BLAST [105] and, based on GO terms, PFAM domains and literature searches, were grouped into three categories summarized in Fig. 8: unknown and unique to each symbiont, insufficiently defined, and better defined (Additional file 6).

Relative to their overall genomic frequency of 7.2% (821/11388), mycobiont secreted proteins are disproportionately enriched to 18.5% (147/795) in coculture (p = 2.8E-28). The genomic average protein length is 477 AA, while the induced and secreted proteins are smaller, averaging 341 AA (Fig. 9 and Additional file 6), with

Table 2 Compilation of genes/proteins of potential symbiotic significance discussed in Search for symbiosis-specific genes I, II, and III sections

| Fungus: Cladonia grayi |  |
|------------------------|--|
| Gene groups induced in coculture | Small secreted proteins / Transcription / Cell wall turnover / Protein turnover / Metabolism / Membrane transport / Defense / Extracellular hydrolases |
| Selected examples | Polyol transporter / Ammonium transporter / Calcium channel inhibitor / lectins / DNA methyltransferase / Ga, RGS protein, dual specificity phosphatase |
| Expanded protein families | HET domain / Ank domain / Mem permeases / Unknown transmembrane proteins / Fructosamine kinases / Polyketide synthases / Signal transduction components / Stress-related TFs |
| Contracted protein families | Transportome: Carbohydrate transporters / Major Facilitator Superfamily (MFS) / ATP Binding Cassette (ABC) superfamily / Aminoacid-Polyamine organo Cation (APC) family / Oligopeptide Transporter (OPT) family / Proton-dependent Oligopeptide Transporter (POT) family |
| Slow evolvers | Proteostasis maintenance / Aldehyde dehydrogenases / Major Facilitator Superfamily (MFS) |
| Selected examples | Mechanosensitive calcium channel / Sugar transporters |
| Fast evolvers | Signal transduction / Membrane trafficking / Stress protection |
| Selected examples | Superoxide dismutase / Trehalose synthase |

| Alga: Asterochloris glomerata |  |
|-----------------------------|--|
| Gene groups induced in co-culture | Secreted proteins / Transcription / Cell wall turnover / Protein turnover - ubiquitin / DNA processes / Signal transduction / Protein trafficking / Flagellum synthesis |
| Selected examples | Thioredoxin / Kinesin / Fasciclin domain proteins / Mechanosensitive calcium channel |
| Expanded/new protein families | Fam_16: DNA binding-recombination proteins / Kinases / Carbohydrate active enzymes (CAZ) / Ank domain proteins / Archaeal ATPases / Desiccation-Related-Proteins / Magnesium transporters / Signal transduction components / Stress-related TFs |
| Contracted protein families | Nitrate assimilation |
| Slow evolvers | Two kinases and one clathrin vesicle adaptor |
| Fast evolvers | Seven diverse proteins |

Fig. 6 Cultures of C. grayi and A. glomerata reconstituted on filters in Petri dishes
Fig. 7 Differential fungal (C. grayi) and algal (A. glomerata) gene expression in coculture vs. monoculture. RPKM expression ratios are sorted from high to low. Genes considered unaffected in coculture are labeled gray (Co/Mo ~ 1). Those labeled black above or below the gray range are considered induced or repressed, respectively. Induction and repression thresholds correspond respectively to 2 and 0.5 for the fungus and 1.3 and 0.77 for the alga. Notice the smaller range of differential expression induction displayed by the alga under our experimental conditions (Additional file 6).

Fig. 8 Classes of genes differentially induced during early fungus-alga interactions in coculture. The pie charts divide the induced genes for each symbiont into three broad classes (numbers of genes in parentheses). The “better defined” genes are subdivided in groups roughly comparable between the symbionts (gray and white boxes). The area of each box is proportional to the percent of genes it contains relative to all better defined genes (265 for the fungus and 243 for the alga). The number behind each group’s name indicates its enrichment factor relative to the whole genome (see Methods). The hatched areas represent groups with less than 10 genes each. The p values for the enrichment of the indicated groups within the induced genes are all < 0.05, and most are << 10\(^{-3}\).
most unknown and unique to \textit{C. grayi}. The smallest proteins tend to be also among the most strongly induced. \textit{Cladonia} shares biotrophic overexpression of small secreted proteins with the mycorrhizal basidiomycete \textit{Laccaria bicolor} \cite{57}, the mycorrhizal ascomycete \textit{Tuber melanosporum} \cite{58}, and with pathogenic fungi \cite{106}. The \textit{Cladonia} secretome (821 proteins) is comparable in size to the secretomes of ectomycorrhizal and many other fungi \cite{107}. Also in the alga, relative to their overall genomic frequency of 3.6\% (365/10025), secreted proteins appear significantly enriched ($p=0.00003$) among the 471 induced proteins (32/471 = 6.8\%), but they are not significantly smaller than average (Fig. 9 and Additional file 6). It remains to be seen whether the algal secretome in early symbiosis shares any features with the secretion responses of higher plants to fungal infection \cite{108}.

The better-defined gene products induced in coculture, 265 in \textit{C. grayi} and 243 in \textit{A. glomerata}, were subdivided for enrichment analysis in broad functional subgroups of at least 10 genes each (Fig. 8, Additional file 6). Except for the limited enrichment (1.4-fold) of signal transduction and defense genes in the fungus, most subgroups in both symbionts were enriched between 2- and 9-fold and with high significance in the set of induced genes relative to the whole genome (Fig. 8). At these early stages, both partners respond to the other’s presence by activating transcription, cell wall metabolism and protein turnover. The mycobiont’s specific responses center on upregulating membrane transporters, secreted hydrolases, and small proteins, broadly resembling the symbiotic responses of the EM fungus \textit{L. bicolor} \cite{112}. The photobiont’s specific responses center on growth and motility through the predominant activation of DNA and signal transduction processes, as well as protein trafficking and zoospore formation (flagella). This correlates with the alga’s rapid initial growth visible in coculture compared to monoculture (Additional file 6), possibly reflecting the growth observed in nature in free-living trebouxoid algae near suitable fungal hyphae \cite{29} and increasing the chances of successful formation of new lichen initials. This issue is taken up again in Conclusions, where the importance for lichen evolution of re-lichenization of free-living algae released from natural thalli is discussed. It needs underlining that this coculture system and similar ones \cite{113} do not proceed beyond formation of poorly differentiated lichenoids (Additional file 6), and thus they do not capture the complete interaction network extended in space and time needed for proper lichenization in nature. Individual genes of potential symbiotic significance are discussed in Inferences from differential transcription about nutritional fluxes at the symbiotic interface section and in Additional file 6.

\textbf{Inferences from differential transcription about nutritional fluxes at the symbiotic interface}

Whereas hexose sugars are the carbohydrates transferred from plant to fungus in mycorrhizae \cite{114}, polyols are the
means by which trebouxoid algae transfer carbon to the fungus in lichens [115], and the first putative polyol transporter gene in a lichen fungus was identified recently [116]. We identified a family of five putative ribitol transporters in *C. grayi* (Additional file 6 and Fig. 10); in coculture, only one of the five was induced (Table 3 and Fig. 10), suggesting its involvement in importing ribitol to the fungus. The alga has dozens of putative sugar/MFS transporters whose expression remains mostly unchanged in coculture, except for seven that are weakly induced (Co/Mo 1.13–1.27). Their involvement, if any, in ribitol export to the fungus remains to be determined.

It has been known since the 1970s that nitrogen is transferred from lichen mycobionts to eukaryotic photobionts [121, 122], and more recently it was suggested that the fungus acts as a nitrogen gateway for the alga [123]. This hypothesis is consistent with the results of a study of the lichen *Cladonia portentosa*, which found that the response of the algal proteome to nitrogen excess was more muted than that of the fungal proteome [43]. As is thought to be the case with mycorrhizal fungi [114], a primary form of nitrogen released by the lichen mycobiont to the alga is likely to be ammonium, although in lichens amino acids might be involved as well [124]. Ammonium transporters generally import NH$_4^+$ into the cell, but they can also export it [125–128]. Of particular interest are two ancient likely horizontal gene transfers of ammonium transporters from prokaryotes to fungi [129]. Transporters from the earliest transfer are found in all fungi, whereas transporters from the later transfer were derived from hyperthermoacidophilic Archaea and were retained primarily by lichenizing fungi [46, 129], suggesting a specific role in the symbiosis. Their selective retention might have been driven by the lichens' frequent exposure to high temperatures and by the very low pH of their apoplast due to the high content of acidic secondary metabolites. Of the four MEthylammonium Permease-type (MEP) ammonium transporters in the *C. grayi* mycobiont, two belong to the general fungal class (CLAGR_004047-RA and CLAGR_009781-RA) and two to the class primarily retained by lichenizing fungi (CLAGR_005848-RA and CLAGR_003366-RA). The *Cladonia* mycobiont also has four transporters belonging to the Gpr1/FUN34/YaaH family, whose functions are debated [130] but include the postulated ability to export NH$_4^+$ in yeast [128], where they are named ATO (Ammonia Transport Outward). The differential transcription of these eight transporter genes is interesting: relative to

---

**Fig. 10** A predicted ribitol transporter in *C. grayi*. **a** Differential transcription in coculture vs. monoculture of five putative sugar transporters in *C. grayi*. They were the top five BLAST hits obtained by querying the genome with the sequences of two functionally validated fungal D-sorbitol/D-mannitol/ribitol transporters [117]. Only CLAGR_004844-RA is induced in coculture (Co/Mo on Y axis; 1 means no induction). **b** CLAGR_004844-RA amino acid sequence. The 12 transmembrane domains [118] are indicated in bold purple. Consensus amino acids for sugar transporters [119, 120] are highlighted in cyan. **c** Protein phylogeny (PhyML, 100 bootstraps) of the five *C. grayi* transporters. The green branches correspond to nodes with bootstrap support ≥70%. The *C. grayi* proteins are labeled brown. The other transporters are identified by GI number and by fungal species. Taxa are also labeled as ascomycetes (Asco) or basidiomycetes (Basidio), the latter used as outgroup. The CLAGR_004844-RA clade is highlighted yellow. Bar indicates amino acid substitutions/site.
fungal monoculture, in coculture with the alga five are repressed, two are induced (one strongly), and one is unchanged (Table 3). The repression of a majority of these fungal transporters in coculture echoes the selective repression by the ectomycorrhizal fungus *A. muscaria* [131] of a fungal NH$_4^+$ importer at the symbiotic interface, thought to prevent re-absorption by the fungus of exported NH$_4^+$. Awaiting experimental verification, we speculate that the ATO transporter CLAGR_005116-RA (Co/Mo 5.6) and perhaps the Mep1b-like ammonium transporter CLAGR_003366-RA (Co/Mo 1.5) are candidates for mediating ammonium export to the alga, while the others, probably importers, are repressed to reduce re-absorption by the fungus. In the cocultured alga relative to monoculture, expression of four of its five ammonium transporters does not change and that of the fifth is undetectably low (Table 3). Overall, however, the transcriptional shifts in coculture of the fungal ammonium transporters are consistent with NH$_4^+$ being a major mediator of nitrogen transfer from mycobiont to photobiont. Phosphate is an important nutrient known to be transferred from mycorrhizal fungi to their plants [114], but there is no suggestive transport polarity or coculture induction pattern among the many putative phosphate transporters in *A. glomerata* and *C. grayi* that could shed light on this system. (For other individual genes induced in coculture, see Table 3, Additional file 6).

### Table 3 Selected fungal (CLAGR) and algal (Aster) genes of potential symbiotic significance differentially expressed in coculture

| Gene name          | Hypothetical function      | Co/Mo ratio | Gene name          | Hypothetical function      | Co/Mo ratio |
|--------------------|----------------------------|-------------|--------------------|----------------------------|-------------|
| CLAGR_004844-RA    | Ribitol transporter         | 2.6         | Aster-02360        | Ammonium transporter       | 1.0         |
| CLAGR_005116-RA    | ATO (Amm. exporter)         | 5.6         | Aster-02123        | Ammonium transporter       | 0.9         |
| CLAGR_006601-RA    | ATO (Amm. exporter)         | 0.9         | Aster-04935        | Ammonium transporter       | 1.0         |
| CLAGR_003233-RA    | ATO (Amm. exporter)         | 0.3         | Aster-08051        | Ammonium transporter       | 1.0         |
| CLAGR_006538-RA    | ATO (Amm. exporter)         | 0.4         | Aster-08052        | Ammonium transporter       | n.d.        |
| CLAGR_000407-RA    | Mep2-type Amm. transp.      | 0.9         |                   |                            |             |
| CLAGR_009781-RA    | Mep3-type Amm. transp.      | 0.6         |                   |                            |             |
| CLAGR_005848-RA    | Mep1a-type Amm. transp.     | 0.6         |                   |                            |             |
| CLAGR_003366-RA    | Mep1b-type Amm. transp.     | 1.5         |                   |                            |             |

**B. Other genes discussed in Additional file 6**

| Gene name          | Hypothetical function      | Co/Mo ratio | Gene name          | Hypothetical function      | Co/Mo ratio |
|--------------------|----------------------------|-------------|--------------------|----------------------------|-------------|
| CLAGR_010764-RA    | KP4 killer toxin-like       | 150         | Aster-04252        | Thioredoxin                | 73          |
| CLAGR_008646-RA    | Lectin                     | 10          | Aster-01625        | Kinesin motor              | 2.3         |
| CLAGR_010932-RA    | Lectin                     | 4           | Aster-01936        | Fasciclin domain protein   | 1.7         |
| CLAGR_011186-RA    | Ga signaling subunit       | 2           | Aster-06761        | Fasciclin domain protein   | 1.5         |
| CLAGR_002910-RA    | Regulator of Ga signaling  | 4           | Aster-03695        | Calcium channel            | 1.3         |
| CLAGR_002710-RA    | Dual specificity phosphatase| 2           |                   |                            |             |
| CLAGR_007359-RA    | DNA methyltransferase      | 3           |                   |                            |             |
| CLAGR_000113-RA    | Calcium transporter         | 2           |                   |                            |             |

**Search for symbiosis-specific genes II: evolution of gene family size**

Lineage-specific expansion or contraction of multigene families is often associated with lineage-specific functional shifts in eukaryotes [57, 58, 132–134]. We undertook one broad and one circumscribed analysis of multigene families. For the broad survey, multigene families in *C. grayi* and *A. glomerata* were identified using MCL [135] and analyzed for changes in family size using CAFE [136], with the taxa listed in Methods. In *C. grayi*, 390 families were expanded, 3369 showed no change, and 769 families had undergone contraction by comparison to a putative common ancestor. In *Asterochloris*, 648 families were expanded, 2729 showed no change, and no families were contracted. The circumscribed analysis was limited to the *C. grayi* transportome and predicted 458 *C. grayi* membrane transporters using the TransportTP online tool [137] with the Transporter Classification Database [138]. We discuss only families for which we can suggest symbiotic roles and highlight whenever a significant fraction of a gene family is also induced in coculture. We assume that the overlap of induction with expansion/contraction (Additional file 1)
increases the likelihood that the genes involved play symbiotic roles. We did not find significantly contracted families in A. glomerata, although a separate search focused on nitrogen assimilation revealed a reduced set of nitrate assimilation genes in the alga (Specific survey of photobiont proteins involved in nitrate and CO2 assimilation section).

**Mycobiont expanded families**

In the C. grayi fungus, the most notable expansions involve 204 HET incompatibility proteins, 13 of which are also induced in coculture; 156 Ankyrin domain proteins; an 80-member family (Fam_5) of multi-transmembrane-domain proteins of unknown function with 10 coculture-induced members; Fam_668, a fructosamine kinase family with 7 members of which 3 are coculture induced; the family of polyketide synthases (PKSs) and Non-Ribosomal Peptide Synthetases with at least 29 members. HET proteins might prevent incompatible C. grayi hyphae from fusing with a C. grayi mycobiont already lichenized with its photobiont. Ankyrin domains could be involved in inter-protein contacts at the symbiotic interface. Fructosamine kinases are known to reverse aging-associated protein damage produced by glycation [139, 140]. A few of the PKSs [34, 141] have been shown to be involved in the synthesis of the primary and most abundant secondary metabolites well known in lichens, but the large number of PKSs in C. grayi (Fig. 11 and Additional file 7) points to a vast and still uncharacterized metabolic potential. Further description of these expanded families and elaboration of the hypotheses on their possible symbiotic significance are found in Additional file 7. A separate analysis identified expanded families of signal transduction components (Specific survey of mycobiont and photobiont signal transduction components section and Additional file 8), some also differentially expressed in coculture (Additional files 6 and 8).

**Mycobiont contracted families**

Carbohydrate transporter Fam_1, MFS (Major Facilitator Superfamily) Fam_2 and Fam_7, and amino acid permease Fam_22 are dramatically reduced, and some members of these families are also induced in coculture (Additional file 7) or score as slow-evolvers as discussed in Slow-evolving proteins and anti-stress strategies in the mycobiont section. The contraction of these families is part of the overall contraction observed in the specific analysis of the mycobiont’s transportome (Additional file 7). A reduced transportome has been observed also in the ectomycorrhizal ascomycete T. melanosporum [58]. This is consistent with the loss of generalist nutrient uptake and the reliance of the fungus on a few specialized carbon sources from the photobiont, which are polyolys in green algal lichens like C. grayi [115] (see Inferences from differential transcription about nutritional fluxes at the symbiotic interface section).

**Photobiont expanded families**

Additional file 7: The most notable expansions in A. glomerata involve the 100-member Fam_16 of unknown proteins unique to this alga, with 20% of the members induced in coculture; Kinase families totaling 52 members; Carbohydrate Active Enzymes (CAZ) with a total of 40 members; a 29-member Ankyrin domain protein family; a probably HGT-derived family of 26 archaeal ATPase-like proteins (Additional file 7 and Fig. 12). A survey using Phyre2, a protein structure prediction program [142, 143], suggests DNA-related functions for some Fam_16 members. The expansion of kinases matches the expanded signal transduction capability mentioned in Genome sizes and gene organization section. The expanded CAZ families might enhance the structural and biochemical versatility of the photobiont’s cell surface. The parallel expansion of Ank-domain proteins in C. grayi and A. glomerata might mediate reciprocal boundary interactions. Finally, the putative ATPases of archaeal origin and the Desiccation-Related-Proteins of bacterial origin might be involved in the lichen alga’s resistance to desiccation and elevated temperatures. Further description of these expanded families and elaboration of the hypotheses on their possible symbiotic significance can be found in Additional file 7.

**Specific survey of mycobiont and photobiont signal transduction components**

This analysis confirms the suggestion from the data in Additional file 3 that signal transduction protein families are expanded in C. grayi and A. glomerata (Genome sizes and gene organization section). For C. grayi, these include PTH11-type receptors [147] and components of MAPK pathways described in Additional file 8, together with coculture expression data collected to identify signal transduction candidates potentially involved in early fungus-alga interactions (Additional files 6 and 8). The analysis also revealed an expanded set of five divergent Go subunit paralogs in C. grayi (Fig. 13). The Asterochloris expansions (Additional file 8) involve a more diverse set of families than Cladonia’s.

**Specific survey of mycobiont and photobiont transcription factor (TF) families**

Overall results are displayed in Fig. 14, and specific families are listed in Additional file 9. With regard to overall TF family gain or loss, neither Cladonia nor Asterochloris stand out within their respective phylogenies. With regard to family expansion and contraction, Asterochloris also conforms to the general algal phylogeny trend of
expansions outnumbering contractions. In contrast, Cladonia appears to have experienced more expansions of TF families than not lichenized fungi. One Cladonia expansion involves a group of proteins characterized by PCI domains [148] (Additional file 9). PCI-domain proteins are involved in many diverse processes, and we could not find specific roles for these TFs. With regard to families with known functions, both Cladonia and Asterochloris have expanded TFs involved in chromatin remodeling (SWI/SNF [149] in Cladonia, SWI/SNF_Baf60 [150] and Sir2 [151] in Asterochloris) and stress responses (SGT [152] and CSD [153] in Cladonia, WRKY [154] in Asterochloris). Asterochloris has also contracted the SET family and the CCHC zinc finger family. The contracted SET family comprises protein lysine methyltransferases, involved in methylation of histones and other proteins [155]. We hypothesize that selection on both symbionts’ ability to shuttle reversibly between free-living and symbiotic states [31] and interact in the thallus in different ways across time and space [10] produced the needed genomic plasticity through the parallel expansion of chromatin-remodeling functions. In contrast, the significant SET family contraction seems odd, given the multifaceted effects of protein lysine methylation [156]. We suggest that the stresses placed on lichens by repeated, rapid and large oscillations in their exposure to light, temperature and hydration might have led to the expansion of stress-related TFs in both symbionts.

Fig. 11 Polyketide Synthase (PKS) and Non-Ribosomal Peptide Synthetase (NRPS) genes in C. grayi. The three protein categories in a, b, c are named on the right. CLAGR_009784 is a PKS-NRPS hybrid. The only PKS whose likely downstream product (grayanic acid) is known is CLAGR_002732 [34]. The length of the horizontal lines is proportional to gene length (vertical gray lines delimit 3-Kb segments). Graphic symbols for protein domains are indicated at top. Genes with a suz_1 prefix were reannotated manually. See Additional file 7 for further details.
These environmental stresses appear to affect the evolution of many lichen adaptations [44, 157–160] as discussed in Slow-evolving proteins and anti-stress strategies in the mycobiont section for the fungus (where several converge on maintaining the resilience of the proteome) and in Additional file 7 for the alga.

Specific survey of photobiont proteins involved in nitrate and CO₂ assimilation

We compared A. glomerata to related algae for possible adaptations involving nitrogen and carbon assimilation proteins that did not surface in the protein family screens. The top of Fig. 15 displays the types and organization of nitrate assimilation genes linked to a phylogenetic tree for Chlorophytes. A. glomerata has a reduced set of nitrate assimilation genes compared to other green algal genomes in this figure. The nitrate assimilation cluster (HANT-AC) [161] present in most microalgae [162] is reduced to nitrite reductase (NIR) and the NRT3 transporter in A. glomerata, and the total number of NNP transporters is reduced in Coccomyxa subellipsoida C-169 and A. glomerata. A. glomerata has

Fig. 12 Protein family tree of archaeal ATPases. The phylogram was constructed using FastTree [144] on a MAFFT [145] amino acid alignment of 91 putative archaeal ATPases from prokaryotes and eukaryotes. Branches with bootstrap values ≥ 0.77 are thickened. Bar indicates amino acid substitutions/site. To the ATPases present in the published Galdieria phylogeny [146], we added all the proteins above Methanocaldococcus J1_MJO532 in this figure. The eukaryotic taxa shown represent most of those currently known to harbor putative archaea-derived ATPases. ATPases from Galdieria (Gs) are marked red, from green algae and plants green, from fungi brown. All branches in black are prokaryotic, except for the amoeba Dictyostelium at the base. Branch labels include the taxon name or symbol and a protein identifier. The Asterochloris (Aster) proteins are indicated by their gene names in the JGI database [68]. The phylogeny suggests several independent HGT events, but it cannot exclude a very ancient HGT from Archaea to a common eukaryotic ancestor followed by losses in most eukaryotes. Asterochloris, Galdieria, and Selaginella have the largest families of archaeal ATPases (with 26, 12, and 7 members, respectively). See also Additional file 7.
lost the clustered nitrate reductase (NAR), but it retains a nitrate reductase paralog (NAR-P) also found in \textit{C. subellipsoidea}. Nitrate cannot be taken up directly by the lichenized alga enclosed in fungal tissue but has to be first metabolized into other compounds by the mycobiont [123]. These compounds may include amino acids, excreted by the fungus and presumably taken up by the alga [124], and ammonium (see Inferences from differential transcription about nutritional fluxes at the symbiotic interface section). Congruent with these findings, \textit{A. glomerata} isolated from \textit{C. grayi} can grow on nitrate only after a period of adaptation to axenic culture (Armaleo, unpublished). Day 21 gene expression in monoculture and coculture is graphed on the bottom of Fig. 15. Interestingly, nitrate assimilation genes in \textit{A. glomerata} are turned down about 50% in coculture relative to monoculture, which may contribute to the regulation of growth experienced by the alga when it lichenizes with the fungus [163]. We hypothesize that, while \textit{A. glomerata} is capable of autonomous nitrate assimilation utilizing nitrate transporters and NAR-P when free-living, suppression of these mechanisms during symbiosis has relaxed selection on their strict maintenance, resulting in the loss of a full HANT-AC and a reduction in the number of nitrate transporters. The extreme reduction of the nitrate assimilation toolkit in \textit{A. glomerata} relative to other non-symbiotic chlorophytes could be viewed as a parallel to the transportome contraction in \textit{C. grayi} (Mycobiont contracted families section and Additional file 7), each contraction resulting from the almost exclusive nutritional reliance of each partner on the other.

**Fig. 13** A unique set of \( G \alpha \) subunits is present in \textit{Cladonia}. The protein phylogeny (PhyML, 100 bootstraps) of the eight \textit{C. grayi} \( G \alpha \) subunits clusters into three major MAG A, MAG B, and MAG C clades (highlighted). The unique MAG C paralogs are shown on the bottom. The green branches correspond to nodes with bootstrap support \( \geq 67\% \). The \textit{C. grayi} proteins are labeled brown. Bar indicates amino acid substitutions/site. See also Additional file 8.
As algal photosynthesis is central to the lichen symbiosis, we surveyed the carbonic anhydrases (CAs) of *Asterochloris*, enzymes that catalyze the interconversion between CO$_2$ and HCO$_3^-$ . This adds to previous extensive work published on CAs in lichen algae [165, 166]. Delivery of CO$_2$ to the enclosed lichen photobionts is extremely variable, depending on the specific anatomy and physiology of its mycobiont-photobiont combination [167] and on rapid changes in temperature and in the supply of nutrients, water, and light [168]. When lichen thalli become fully saturated with water, the diffusion of CO$_2$ to the algae is further limited by the swollen hyphal tissue surrounding them. This reduces photosynthetic rates [169, 170] even if the hydrophobicity of the lichen interior maintains air-filled spaces for rapid CO$_2$ diffusion [10]. Our data, detailed in Additional file 10, suggest that *Asterochloris* parallels *Chlamydomonas* and *Chlorella* in having one $\alpha$ CA functioning in the pyrenoid but differs in having expanded a specific subclass of cytoplasmic $\beta$ CAs, whose potential relationship to the symbiosis needs exploring.

**Search for symbiosis-specific genes III: proteins with anomalous rates of evolution**

We hypothesize that, after a burst of adaptive selection at the origin of Lecanoromycetes 300–350 million years ago [171–174], proteins that acquired fundamental symbiotic roles would have since stabilized under purifying selection and evolved at slower rates, as suggested for a polyketide synthase specific for a secondary metabolite unique to *C. grayi* [34]. Conversely, faster evolving proteins are likely to represent more recent changes. We compared amino acid substitution rates of *C. grayi* and *A. glomerata* proteins to those of eleven free-living ascomycetes and six unicellular chlorophytes respectively using three methods (Methods). We focus on the proteins whose rates were “slow” or “fast” relative to nonlichen species by at least two methods to
reduce false positives. This yielded 38 slow- and 11 fast-evolving candidates in the fungus and 3 slow- and 7 fast-evolving candidates in the alga (Slow-evolving proteins and anti-stress strategies in the mycobiont and Fast-evolving proteins sections).

**Slow-evolving proteins and anti-stress strategies in the mycobiont**

Additional file 11 (Excel sheet “Fungus slow-evolvers”) lists a total of 72 proteins that score in multiple ways (highlighted in orange) for potential symbiotic significance. The first 38 were identified as slow-evolving by at least two of the rate methods. Of the 38, 22 are universal eukaryotic components of the protein translation, translocation and sorting machinery (Additional file 11). This machinery maintains proteostasis, the dynamic equilibrium of the proteome, through a complex network guiding protein folding and functionality from synthesis to modification, to sorting, and to degradation [175–177]. Studies with several eukaryotic systems, including yeast, show that downturns in protein translation and translocation are essential to protect cells from dehydration, heat, and hyperosmotic stresses [178–184]. In yeast for instance, many temperature-sensitive mutations impairing a variety of ribosome assembly steps also dramatically increase desiccation tolerance [185], highlighting how slowdowns in ribosomal assembly may reduce protein misfolding and aggregation during desiccation. However, lichens must withstand daily cycles of dehydration, rehydration, thermal and UV radiation stress so extreme that they would kill most other organisms [157, 186]. Therefore, we hypothesize that such exceptional circumstances in early lichen evolution selected for “upgraded” fungal components of the
ribosomal biogenesis, mRNA processing, and protein trafficking networks, which were already part of the normal Environmental Stress Response (ESR) [184]. The upgrades were then stabilized under purifying selection and integrated with other defenses [157] involving, for instance, antioxidants [159, 187] and synthesis of photoprotective anthraquinones [188] that lichens share with non-extremophiles. There is some correlative support in the literature for the centrality of ribosomal function in the lichen response to stress. In the lichen Lobaria pulmonaria [37] as much as 35% of the expressed fungal proteome is involved in ribosomal and protein turnover functions. In the lichen Cladonia rangiferina [39] global transcriptional responses to desiccation and rehydration also suggest involvement of the protein translation machinery. In the cultured lichen fungus Endocarpon pusillum, ribosomal protein genes are highly induced during PEG-induced dehydration stress [159]. We speculate that the unusually high number of nuclear rDNA introns (Additional file 12) present in lichen fungi [189, 190] could be another possible adaptation to control the rate of ribosomal assembly for stress (e.g. desiccation) protection. A major consequence may be the well-known slow growth phenotype of lichens and lichen mycobionts. Other double-scoring slow evolving proteins of interest are listed and discussed in Additional file 11. These include a putative osmosensing calcium channel and two homogeneous groups of proteins: 18 membrane transporters and 7 aldehyde dehydrogenases (ALDHs).

Fast-evolving proteins
Additional file 11 includes a description of 11 putative fast-evolving mycobiont proteins involved in signal transduction, membrane trafficking and stress protection, as well as slow- and fast-evolving proteins in the photobiont.

Conclusions
Symbiotically relevant genes affect stress resistance, nutritional, signaling, and structural interactions
Our analysis identified several proteins expected to influence the symbionts’ environmental stress resistance. They include slow evolving mycobiont proteins (Slow-evolving proteins and anti-stress strategies in the mycobiont section, Additional file 11) enriched for universal eukaryotic components of the protein translation, translocation and sorting machinery which manages the proteome under normal as well as stress conditions like dehydration and heat [178–182]. Also, some of the faster evolving C. grayi proteins, probably under adaptive selection, appear to be involved in protection from stress (Additional file 11). In A. glomerata, the likely horizontally transferred archaean ATPases and Desiccation-Related Proteins may contribute to its heat and desiccation resistance (Photobiont expanded families section, Additional file 7). Transcription factor family expansion includes proteins involved in chromatin remodeling and stress responses in both symbionts (Specific survey of mycobiont and photobiont transcription factor (TF) families section).

We also identified proteins governing several symbiotically relevant nutrient interactions. The fungus in coculture induced two transporters potentially central in the carbon and nitrogen exchange at the symbiotic interface: one an importer for ribitol, the carbon source provided by trebouxoid algae to their fungal partners [115] (Inferences from differential transcription about nutritional fluxes at the symbiotic interface section and Additional file 6), the other a possible ammonium exporter (Inferences from differential transcription about nutritional fluxes at the symbiotic interface section), pointing at NH3- as a major nitrogen source provided by the mycobiont to the photobiont. Reliance of each partner on the other as a restricted nutrient source is also reflected by the contraction of the sugar transportome in C. grayi (Mycobiont contracted families section) and by the reduced nitrate assimilation potential in A. glomerata (Specific survey of photobiont proteins involved in nitrate and CO2 assimilation section). Un-known is the significance for symbiotic carbon fixation of the expansion of a specific A. glomerata subclass of cytoplasmic carbonic anhydrases, enzymes that catalyze the interconversion between CO2 and HCO3− (Specific survey of photobiont proteins involved in nitrate and CO2 assimilation section and Additional file 10).

Gene families whose characteristics suggest involvement in other symbiotically relevant interactions comprise: signal transduction proteins, expanded in both symbionts, which include a unique new set of five MAG C paralogs in the mycobiont (Specific survey of mycobiont and photobiont signal transduction components section and Additional file 8); ankyrin domain protein families (Mycobiont expanded families and Photobiont expanded families sections), also expanded in both partners, perhaps involved in increased protein-protein interactions at the boundaries between them; algal glycosyl transferase families, whose expansion could be necessary to adapt extracellular surfaces to the varied contacts in which the photobiont engages (Photobiont expanded families sections, and Additional file 7); the expanded set of mycobiont polyketide synthases producing compounds for a mostly undiscovered array of biochemical functions (Mycobiont expanded families section); HET incompatibility protein families in C. grayi, some possibly involved in the competition among fungal genotypes to secure the appropriate alga (Mycobiont expanded families section). Also the induction in coculture of secreted proteins in the fungus and less prominently in the alga (Extended analysis of transcription induced in coculture section) suggests their involvement in a variety of unexplored symbiotic functions.
The evolution of lichenization involved changes in many conserved genes scattered throughout the symbionts’ genomes

Most of the symbiotically relevant genes suggested here have homologs in non-lichen fungi and algae, and we assume that they are variants modified by symbiosis. This indicates that lichens evolved mainly through the accumulation of scattered regulatory and structural changes in available genes rather than through sudden key innovations. This in turn suggests that, to establish a basic and reversible nutritional dependency at very early evolutionary stages, the free-living ancestors of myco- and photobionts might have required at first only a few or even no changes. A model of such basic interactions was developed experimentally using the fungus S. cerevisiae and the alga C. reinhardtii [191]. Selection towards increasing stability and environmental adaptability would have then transformed such precarious mutualistic/antagonistic and reversible states into lichens over evolutionary time. This scenario suggests that there were multiple pathways for fungi, algae and cyanobacteria to evolve into lichens, which is consistent with the emerging consensus that ascocarps could have had a few independent evolutionary origins [171, 192]. It is also compatible with the fact that lichens display a wide array of structures with different levels of complexity, from leprose and crustose to fruticose and foliose, and with the overall staggering variety of interactions throughout the biosphere between fungi and photosynthetic organisms [193, 194].

Sexual reproduction, symbiont autonomy and equivalence

While the advantages accrued by lichens through sexual reproduction of their mycobionts are fairly clear [91, 92, 97, 195–197] and consistent with the behavior of a mutualistic exhibitant [100], the advantages of sex for the algal inhabitant were previously expected to be limited [100], mostly based on the assumption that the lichen alga is asexual and that it dies when its lichen thallus dies [198]. As indicated in Heterothallism probably evolved from homothallism in Cladonia; genetic evidence for sex in Asterochloris section, however, there is strong evidence that sex in trebouxoid algae does occur and that free photobiont populations exist on the substrates near lichen thalli [27, 29, 199–201], so that sexually produced variation and algal adaptations to lichenization are likely to be incorporated and selected for in lichenized populations. W.B. Sanders proposed in 2005 [29] that the transient free-living state is necessary for completion of the algal sexual cycle, which is consistent with the fact that the rare direct observations of algal sexual stages have always involved aposymbiotic cells (Heterothallism probably evolved from homothallism in Cladonia; genetic evidence for sex in Asterochloris section). Encounters of germinating mycobiont spores and free-living photobionts [199] could produce lichens with new combinations of both genomes, expanding and fine-tuning a lichen’s adaptation to different ecological niches. Photobiont contributions to lichen adaptability are in fact highlighted by several studies. For a given lichen, the correlation of algal genotypic variation with habitat appears stronger than that of the fungal genotype [103, 202–205]. Other analyses indicate that both fungal and algal genotypes substantially influence a lichen’s ecological adaptability [206, 207], and that multiple algal genotypes can coexist within single thalli [208–210] and move horizontally among fungal genotypes [207, 208, 211–214] even when the predominant means of photobiont transmission is vertical through vegetative propagules [201, 208]. We propose therefore that, over hundreds of millions of years of tight coexistence, genomic and functional autonomy in each partner were maintained by the benefits of periodically detaching each partner from the symbiosis for sex and for partner switching, which increased the overall adaptability of the lichenized symbionts. Beyond sex, the list of phycobiont investments in the symbiosis is long. It includes the adaptations suggested in Photobiont expanded families, Specific survey of mycobiont and photobiont signal transduction components, Specific survey of mycobiont and photobiont transcription factor (TF) families, Specific survey of photobiont proteins involved in nitrate and CO2 assimilation sections, the intrinsic resilience of the free-living lichen alga to desiccation [44, 160], the increased resistance to PSII photoinhibition in the symbiotic vs. free-living alga [215], and correlated structural and physiological adaptations: algal morphology changes significantly between lichenized and free-living states ([216] and references therein) and, when the lichen is hydrated, about 50% of the fixed carbon can be converted to lichen biomass [217] with an energy conversion efficiency comparable to that with which chloroplast photosynthesis translates into plant biomass [218]. Ribosomal DNA introns occur in lichenized trebouxoid algae [219–221], including the Group IB intron in the LSU gene of A. glomerata [221]. We speculate, as for the mycobiont rDNA introns (Slow-evolving proteins and anti-stress strategies in the mycobiont section), that these photobiont rDNA introns may be involved in mediating desiccation tolerance. The evolutionary introduction of A. glomerata or Trebouxiales to lichenization is not yet known but, based on the only published broad Chlorophyta chronogram [222] (which does not include Trebouxiales), the timing could be compatible with that estimated for lichen fungi [173, 174]. It is therefore possible that the Asterochloris/Trebouxia lineage has been adapted to symbiosis for hundreds of millions of years. Metaphorically, the lichen alga is not the “second sex” [223]: it deserves a full seat at the symbiotic table.
Methods

Where appropriate, methodological information is included in figure and table legends.

Biological isolates

The sequenced mycobiont is the *C. grayi* single-spore isolate Cgr/DA2myc/ss [34] (Culture Bank accession # CBS 132746; GenBank ITS accession # KC592272) [224]. The sequenced photobiont is the *Asterochloris* strain Cgr/DA1pho, isolated from *C. grayi* soredia as described [64]. The species was identified by ITS1 sequencing and 100% identity to *Asterochloris glomerata* [48]. Natural *C. grayi* podetia used for EST generation were harvested, and chemotype was confirmed by Thin Layer Chromatography as described [34].

Genomic DNA extraction

The *C. grayi* mycobiont was grown for ~ 9 weeks in 50-ml cultures of HMY (MY according to Hamada [225]: 10 g/L malt extract; 4 g/L yeast extract; 4 g/L sucrose) in 125-ml flasks shaking at room temperature, with periodic medium changes and grinding to improve growth. Typically, after 3 weeks of growth, a 50-ml mycelial culture was first ground and then expanded as follows. A Polytron homogenizer (Kinematica GmbH, model PT 10/35) fitted with an autoclaved standard 10-mm diameter probe (Brinkmann generator PTA 10S) was set up in a laminar flow hood with the probe at a 45° angle. The 50-ml culture was then steriley ground in the tilted flask for 5 s at half-maximal setting. Using wide bore pipets, four 10-ml aliquots from the ground culture were transferred to four 125-ml flasks, each with 40 ml of fresh HMY. After every grinding, absence of contamination in each flask was tested by spreading 1 ml of the ground and diluted culture onto an HMY test plate and periodically monitoring growth by microscopy. Mycelia were harvested from the liquid cultures by filtration through nylon mesh, rinsed on the filter with sterile TE, lyophilized, weighed and stored at -80° C. DNA was extracted from 850 mg mycelia (dry weight) and periodically monitored growth by microscopy. Mycelia were harvested from the liquid cultures by filtration through nylon mesh, rinsed on the filter with sterile TE, lyophilized, weighed, and stored at -80° C. DNA was extracted from 850 mg mycelia (dry weight). The sample was thoroughly ground in a mortar with liquid nitrogen; the powder was transferred to a 50-ml polypropylene tube and resuspended in 18 ml (~ 20/1, v/w) of lysis buffer (40 mM TrisHCl pH 8; 20 mM Na Acetate; 1 mM EDTA; 1% w/v SDS) plus freshly added RNase to 0.1 mg/ml. The suspension was repeatedly mixed with a pipet while being kept at 65° C for 5 min. A 5 M NaCl solution (6 ml) was mixed in, and the suspension was centrifuged at 3 K rpm for 5 min on a tabletop centrifuge to pellet cell debris, polysaccharides and some proteins. The supernatant was extracted with an equal volume of a 1:1 phenol:chloroform mix saturated with 100-mM Tris HCl at pH 8. The aqueous phase was brought to 65° C, and 0.11 volumes of CTAB solution (10% CTAB; 0.7 M NaCl) were mixed in. The mixture was extracted with an equal volume of chloroform. An equal volume of room-temperature isopropanol was mixed into the aqueous phase, and the solution was aliquoted into six 12-ml polypropylene tubes and centrifuged at 5° C for 15 min at 6 K rpm. After completely removing the supernatants, the procedure was repeated from the lysis buffer step in a concentrated and simplified fashion. Each pellet was first left undisturbed at 65° C for 3–5 min in 150 μl of lysis buffer plus RNase to facilitate resuspension. The six pellets were then thoroughly resuspended by pipetting, suspensions were pooled into one of the six tubes, and 1/3 volume (300 μl) of 5 M NaCl was mixed into the pool. The solution was aliquoted into two 1.5-ml microfuge tubes, undissolved residue was removed by benchtop centrifugation for 5 min at 15 K rpm, and the supernatants were extracted once with chloroform. Two volumes of ice-cold 100% ethanol were added to the aqueous phases, and the precipitated DNA was pelleted by benchtop centrifugation at room temperature. The pellets were rinsed with 70% ethanol, air-dried in a laminar flow hood for 30 min and left undisturbed to rehydrate on ice in TE buffer overnight. They were then resuspended by gentle pipetting, and the solutions were pooled. The total yield was 127 μg of high molecular weight DNA, verified by Qubit quantitation and gel electrophoresis. *Asterochloris glomerata* was incubated on AMY plates (MY according to Ahmadjian [226]: 20 g/L malt extract; 2 g/L yeast extract) for 4 weeks at room temperature under constant light (photon flux = 40 μmol m⁻² s⁻¹). Algae were harvested from one plate, transferred to sterile deionized water in a microfuge tube, pelleted, lyophilized, weighed and stored at -80° C. DNA was extracted from 83 mg algae (dry weight) as described [64], with one modification: after the final resuspension of the DNA in TE, the volume of TE was increased to 1 ml, the sample was transferred to a 12-ml Falcon tube, and the DNA was re-precipitated by adding 330 μl of 5 M NaCl followed by 2.7 ml of cold 100% EtOH. The solution was mixed by inversion and centrifuged at 6 K rpm for 10 min at 5° C. The pellet was rinsed with 70% EtOH, air-dried in a laminar flow hood for 30 min and left undisturbed to rehydrate on ice in TE buffer overnight. It was then resuspended by gentle pipetting. The total yield was 110 μg of high molecular weight DNA, verified by Qubit quantitation and gel electrophoresis.

Genomic DNA sequencing, assembly, annotation and data storage

For the mycobiont, 50,883,772 Illumina single-end reads (36 b, 52 x coverage) were pre-assembled with Velvet
podetia) and the suspension was vortexed until all air pockets were removed, indicating an even permeation of the solvent into the tissues. Podetia were kept in RNAlater at room temperature for about 1 h with occasional mixing. Excess RNAlater was removed by centrifugation, followed by gentle blotting of the podetia with filter paper. Podetia were lyophilized overnight and stored at -80°C. Podetia were thoroughly ground in liquid nitrogen with a mortar, and RNA was extracted with the RNAqueous Total RNA Isolation Kit (Life Technologies) according to the manufacturer’s protocol with two modifications. For the first, the initial buffer volume in which the ground powder was resuspended was increased to 5 ml Lysis Buffer mixed with 680 μl of Plant Isolation Aid for 200 mg of air-dried podetia. For the second, a DNase-treatment step was inserted as follows. After loading the lysate onto the cartridge and centrifuging to remove residual Lysis Buffer-Plant Isolation Aid solution, 500 μl of Wash 2/3 solution were centrifuged through the cartridge. The cartridge was then centrifuged again without added liquid to remove all traces of Wash 2/3. After moving the cartridge to a new collection tube, the matrix was wetted by adding 60 μl of DNase solution (52 μl RNase-free water, 6 μl 10xDNase buffer Promega M198A, 2 μl RQ1 RNase-free DNase Promega M610A). Under these conditions, nucleic acids detach from but remain within the matrix. After incubating the cartridge with DNase at 37°C for 15 min, 300 μl of a 1:1 mix of Lysis Buffer: 64% ethanol were added to allow the RNA to bind again to the matrix. The cartridge was kept at room temperature for 5 min, and the liquid was removed by centrifugation. The manufacturer’s protocol was resumed by adding 700 μl of Wash 1. The final yield was 27 μg of purified RNA. Agilent microchip electrophoresis showed a RIN = 9. The RNA was used to prepare a 454 cDNA library of 740,680 reads averaging around 500 bases in length. Of the reads, 76.2% mapped onto the mycobiont genome and 12.1% onto the photobiont genome. The residual unmapped reads are from other eukaryotic or prokaryotic inhabitants of natural C. grayi thalli. Reads were assembled into ESTs using Newbler (Roche), Velvet [227] and Trinity [235], and mapped to the assembled symbiont genomes to aid in the annotation (see preceding section).

RNA extraction and sequencing for analysis of differential expression in coculture
To analyze differential gene expression during symbiont coculture in vitro, RNA was isolated from mycobiont and photobiont grown on filters in agar plates. The culturing details are as described [33]. Briefly, mycobiont and photobiont were routinely propagated as described for the DNA extractions. To set up filter cultures,
separate fungal and algal cultures were harvested and filtered first through gauze and then through 74 μm nylon to obtain uniform suspensions either of small mycelial fragments or of algal cells. Cell density of the separated suspensions was measured by OD, and fixed amounts (dry weight equivalents: 2.9 mg mycobiont and/or 1.3 mg photobiont) were uniformly delivered by suction onto 0.22 μm pore-size nitrocellulose filters (Millipore, GSWP04700) to obtain cultures of either fungus alone, alga alone, or the two together. The seeded filters were placed on 1.5% agarose (Sigma-Aldrich A9539) plates (one filter per plate) with “99:1” medium [33] and incubated for 21 days at room temperature under a constant photon flux of 40 μmol m−2 s−1 (Fig. 6). The main methodological difference relative to Joneson et al. (2011) [33] was the addition of RNAlater at harvest. Each filter was peeled off the agarose surface and placed over three stacked Whatman 9-cm filter paper disks in the lid of a 10-cm plastic Petri dish. To exchange residual growth medium for RNAlater, 2 × 1 ml of RNAlater were evenly pipetted over the entire surface of the nitrocellulose filter culture. After the RNAlater was absorbed by the Whatman filters, the nitrocellulose filter was placed (culture facing up) into the empty bottom half of the Petri dish, and 1 ml RNAlater was uniformly pipetted over it. While submerged in the RNAlater layer, the culture was scraped off the nitrocellulose with a sterile blade, pooled to the edge in RNAlater by tilting the dish, and the suspension was pipetted into a microfuge tube. Addition of 1 ml RNAlater, scraping and pooling were repeated once more, and the two suspensions were kept in two tubes for 1 h at room temperature. To obtain a single pellet from the two tubes corresponding to one filter, the two suspensions were centrifuged sequentially in the same tube, the RNAlater was discarded, and the pellet was lyophilized and stored at -80°C. Each frozen pellet was ground with a fitting pestle in its microcentrifuge tube cooled in liquid nitrogen [31]. The ground powder derived from one filter culture was resuspended in 500 μl of Lysis Buffer-Plant Isolation Aid combination (RNAquous Total RNA Isolation Kit, Life Technologies), and RNA was extracted according to the manufacturer’s instructions except for the insertion of a DNAse treatment step as described in the preceding section. For each growth condition, the extracts from two replicate cultures were combined. RNA yields per combined cultures were 3 μg, 7 μg and 16 μg respectively from fungus alone, alga alone, and coculture. Agilent microchip electrophoresis showed RINs near 9. The RNA was used to produce cDNA libraries totaling 156 million RNAseq reads. Reads were aligned using Bowtie [236]. Gene expression was calculated using BowStrap [237] with the annotated gene models from the mycobiont and photobiont genomes at JGI. Gene expression is defined as re-sampled Reads Per Kilobase of gene sequence per Million aligned sequences (RPKM). For the GO term-centered analysis, differential expression was represented as the ratio between the base 2 logs of each gene’s coculture and monoculture RPKMs on day 21. Genes with log fold changes > 0.6 (for upregulation) and < 0.6 (for downregulation) were analyzed. Statistical significance (p < 0.05) was based on hypergeometric distribution enrichment of genes or pathways relative to their representation in the whole genome. For the more extensive analysis (Extended analysis of transcription induced in coculture section), the ratio between each gene’s coculture and monoculture RPKMs on day 21 (referred to here as Co/Mo) was taken directly to represent differential expression, and only ratios > 2 for the mycobiont or > 1.3 for the photobiont were considered induced. To combine the induced genes/proteins (795 in the mycobiont and 471 in the photobiont) into functionally related protein subgroups, we collected information from GO terms, PFAM domains, and literature searches (Additional file 6). Statistical significance (p < 0.05) for the coculture induction of each resulting subgroup (Fig. 8) was based on that subgroup’s hypergeometric distribution enrichment within all induced proteins relative to the representation, in the whole genome, of genes encoding proteins similar to those in that subgroup. Such proteins were identified (e-value <1e −5) by using the subgroup as query in a BLASTP genome search.

**Predictions of multigene families, PKSs, and secreted proteins**

Protein sequences for multigene family analysis were retrieved from the Joint Genome Institute Portal [164] and the Broad Institute Portal [238]. As references for Cladonia grayi, the following fungal taxa were used: Stagonospora nodorum (Phaeosphaeria nodorum) (16,597 models), Botrytis cinerea (Botryotinia fuckeliana) (16,389 models), Sclerotinia sclerotiorum (15,522 models), Neurospera crassa (9907 models), Aspergillus nidulans (10,753 models), Aspergillus fumigatus (9887 models), Aspergillus oryzae (12,017 models), Tuber melanosporum (7496 models), Fusarium graminearum (13,826 models), Pyrenophora triticirepentis (12,169 models) and Chaetomium globosum (11,103 models). As references for Asterochloris glomerata, the following algal taxa were used: Coccomyxa sp. C169 (9629 models), Chlorella sp. NC64A (9791 models), Chlamydomonas reinhardtii (15,113 models), Ostreococcus tauri (17,725 models), Volvox carteri (14,542 models), and Micromonas sp. CCMP1545 (10,545 models). Orthogroups of mutual best BLAST hits present in the symbiont genomes and all reference species (2455 in C. grayi, 1454 in A. glomerata) were further analyzed. Protein groups were aligned with MAFFT [145] using Tuber melanosporum or Volvox carteri as outgroups, and alignments were subsequently
trimmed with Gblocks [239, 240] (A. glomerata) and Guidance [241] (C. grayi). Prototyes were queried in an all-against-all BLASTP analysis, and protein families were subsequently predicted with MCL [135] with inflation parameters set to 3.0. In the fungal data set, 5051 protein families were identified that were phylogenetically relevant (containing at least two species and 10 sequences), of which 4528 were found in C. grayi. In the algal data set, 3545 phylogenetically relevant gene families were identified. C. grayi PKSs and NRPSs were predicted using anti-SMASH v3.0 [242], followed by manual curation. Secreted proteins in C. grayi and A. glomerata were predicted as indicated in the legend of Fig. 9.

Analysis of gene family expansions and contractions

For the broad analysis encompassing all C. grayi and A. glomerata genes, multigene families predicted as described in the previous section were analyzed for evolutionary changes in protein family size using CAFE [136]. The program uses a random birth and death model of gene gain and loss across a user-specified tree structure. The distribution of family sizes generated under the random model provides a basis for assessing the significance of the observed family size differences among taxa (p-value < 0.001). CAFE estimates for each branch in the tree whether a protein family has not changed, has expanded, or has contracted. For the analysis of C. grayi family expansion and contraction centered on the transportome, predicted proteins containing at least three transmembrane domains were extracted from the same fungal genomes used in the broad analysis and listed in the previous section. These predicted proteins were classified according to the Transport Classification database [138] using the TransportTP on-line tool [137, 243]. For several families, classifications were checked manually with predicted fungal transporter proteins retrieved from TCDB, TransportDB [244], and the Saccharomyces Genome Database [245] used as queries in BLASTP and TBLASTX searches against the 12 fungal genomes. Lineage-specific gene gains and losses were estimated using CAFE [136].

For the analysis centered on signal transduction components, predicted proteins were identified based on PFAM domains among the taxa indicated above. To identify family expansions and contractions, protocols similar to those indicated above were used. For the analysis centered on putative DNA-binding transcription factors (TFs), the proteomes of the fungal and algal taxa indicated in Additional file 9 were searched using the hmmmscan module of HMMER version 3.0 [246] against a comprehensive DNA-binding domain database assembled from three compilations [247–249] to which corresponding PFAM domain designations were assigned. GA (gathering bitscore) thresholds for each domain were retrieved from PFAM 25.0, and each hit on a protein exceeding the appropriate GA threshold was recorded. Matches where alignments covered less than 80% of the corresponding PFAM profile HMM were discarded. Putative transcription factors were classified into families based on domain structure as described by Lang et al. [249] and rules were extended by incorporating fungal-specific terms. A random birth and death model [250] was used to investigate expansions, contractions, gains, losses and ancestral states of TF families across algal and fungal phylogenies [4, 251] (Fig. 14). The best model was selected from a series of models with increasing complexity by using likelihood ratio tests to drop parameters that did not further improve the likelihood of a model. After rate parameter optimization, posterior probability of ancestral gene content of each node and the evolutionary change for each family along the branches of the tree were estimated.

Analysis of nitrate assimilation genes in chlorophytes

Homologs of nitrate assimilation-related (HANT) genes (C. reinhardtii: nitrate reductase, AAF17595.1; nitrite reductase, XP_001696787.1; NRT2, XP_001696788.1; formate/nitrite transporter, XP_001696698.1; M. psilla: nitrite reductase, XP_003057941.1; molybdopterin synthase, XP_003057940.1) were determined by BLASTP of algal genome best protein models (Asterochloris glomerata Cgr/DA1pho v1.0, Chlamydomonas reinhardtii v4.0, Chlorella sp. NC64A v1.0, Coccomyxa sp. C-169 v2.0, Micromonas pusilla CCM1545 v 2.0, Micromonas com- moda NOUM17 (RCC 299) v2.0, Volvox carteri nagar- iensis v 1.0, Ostreococcus sp. RCC809 v 2.0, Ostreococcus lucimarinus v 2.0, Ostreococcus tauri RCC4221 v3.0). Retained hits had at least 45% similarity with and covered between 50 and 150% of the length of the query. Homologs of the query genes were considered part of a nitrate assimilation cluster (HANT-AC) if separated by no more than 6 intervening gene models from another gene in the query [252]. Intervening and flanking sequence up to 10 kb of HANT-ACs were examined for unannotated nitrate assimilation-related genes by BLASTX against the NCBI non-redundant protein database; NRT3 was determined by the presence of NAR2 PFAM domain. Orthology and paralogy of nitrate assimilation genes were determined by examination of phylogenetic trees generated as follows: protein sequences were aligned with MAFFT v.7.2.2.1 [145] using default parameters, and poorly aligned regions were removed with trimAl v.1.4 using the “automated1” algorithm [253]. The best model of protein evolution (JTT, WAG, LG) was inferred by AICc using ProtTest v.3.4 [254], and maximum likelihood analysis was performed with RAxML v.8.1.20 [255]. Nitrite reductase phylogeny was estimated using the combined sequences from both
BLAST searches and was reduced to models containing a PLN02431 protein domain. Expected but missing gene homologs were considered absent by TBLASTN of the respective genome assemblies. Algal species phylogeny was estimated by analysis of RPB2 using the methods above.

A. glomerata meiosis gene identification and phylogenetic analysis; flagellar proteins

We sought A. glomerata orthologs of multiple meiosis-specific genes proposed as a “meiosis detection toolkit” for identifying genomic evidence for sexual reproduction ([256, 257] and references therein) and select paralogs: Msh4; Msh5; Asy1/Hop1; Ahp2/Hop2; Mnd1; Dmc1 (paralog Rad51—DNA repair in meiosis and mitosis); Syn1/Rec8 (paralogs Syn2, Syn3, Syn4—cohesins not specific to meiosis [258]); Spo11–1, Spo11–2 [259] (paralog Spo11–3—no reported meiotic phenotype; affects overall plant size, hormone responses, and endoreduplication [260]). We also sought orthologs of two additional meiosis-specific genes: Pch2 [261] and Mer3 [262, 263]. Arabidopsis thaliana orthologs of these 16 genes were used as TBLASTN [105] queries against the version 2 Arabidopsis thaliana orthologs of these 16 genes were publicly available green algal genomes (A. glomerata CC-169; JGI, 2013). If initial BLAST searches and was reduced to models containing a PLN02431 protein domain. Expected but missing gene homologs were considered absent by TBLASTN of the respective genome assemblies. Algal species phylogeny was estimated by analysis of RPB2 using the methods above.

A. glomerata meiosis gene identification and phylogenetic analysis; flagellar proteins

We sought A. glomerata orthologs of multiple meiosis-specific genes proposed as a “meiosis detection toolkit” for identifying genomic evidence for sexual reproduction ([256, 257] and references therein) and select paralogs: Msh4; Msh5; Asy1/Hop1; Ahp2/Hop2; Mnd1; Dmc1 (paralog Rad51—DNA repair in meiosis and mitosis); Syn1/Rec8 (paralogs Syn2, Syn3, Syn4—cohesins not specific to meiosis [258]); Spo11–1, Spo11–2 [259] (paralog Spo11–3—no reported meiotic phenotype; affects overall plant size, hormone responses, and endoreduplication [260]). We also sought orthologs of two additional meiosis-specific genes: Pch2 [261] and Mer3 [262, 263]. Arabidopsis thaliana orthologs of these 16 genes were used as TBLASTN [105] queries against the version 2 assembly of the A. glomerata genome and additional publicly available green algal genomes (Ostreococcus sp. RCC809, Ostreococcus lucimarinus CCE9901, Micromonas pusilla CCMP1545, Chlamydomonas reinhardtii CC-503 cw92 mt+, Volvox carteri f. nagariensis Eve (HK10 derivative), Chlorella variabilis NC64A, Coccomyxa subellipsoidea C-169; JGI, 2013). If initial A. thaliana queries failed, subsequently identified green algal orthologs were used for additional queries. All genome hits were compared to the NCBI non-redundant protein database by BLASTX and, for putative orthologs of the 16 genes listed above, manually annotated with a predicted protein sequence based on A. thaliana and revisions following amino acid multiple sequence alignment comparisons (MUSCLE v3.6 [264], the MEGA5.05 interface of MUSCLE [265]). Phylogenetic analyses (not shown) were done to test for orthology (e.g., comparing meiotic Spo11–1 or Spo11–2 from non-meiotic Spo11–3) and green algal origin (i.e., confirming that the putative A. glomerata orthologs fall into a clade with plants and green algae rather than fungal orthologs) using maximum likelihood (green algae and plants) or NJ (green algae, plants, select additional fungi) methods. This analysis identified 12 A. glomerata genes specific for meiosis and/or mitosis and one, Spo11–3, for endoreduplication (Additional file 5).

The 314 candidate flagella proteins in A. glomerata (Additional file 5) were identified by searching for putative orthologs among sequenced motile and non-motile chlorophytes using the reciprocal best BLASTP hit criterion (e-value<1e–5). The queried Chlorophytes included C. reinhardtii; Volvox carteri; Coccomyxa subellipsoidea C-169; Chlorella variabilis NC64A; O. tauri; O. lucimarinus; Micromonas pusilla CCMP1545; Micromonas sp. RC299. The queried databases included the Chlamydomonas CiliaCut protein set [49], the Chlamydomonas flagella proteome [266, 267] and protein families ubiquitous in motile chlorophytes but absent in non-motile chlorophytes (Fig. 5). The latter proteins were referred to as Chlorophyte CiliaCut in reference to work by Merchant et al. [49], and contain candidate flagella proteins. A majority (69/95 = 68%) of the 95 Asterochloris proteins in the Chlorophyte CiliaCut are orthologs to proteins of the Chlamydomonas CiliaCut or Chlamydomonas flagella proteome.

Tests of relative evolutionary rates

Three methods were used to assess evolutionary rates in C. grayi and A. glomerata relative to the 11 fungal and 6 algal taxa and alignments described in the multigene family prediction methods. The first protocol used HyPhy [268] to determine substitution rates for each individual conserved orthogroup from aligned mutual best BLAST hit orthologs present in all taxa (2455 in the ascomycetes and 1454 in the chlorophytes). Rates of C. grayi and A. glomerata genes were considered reduced or accelerated if predictions matched in at least 90% of all analyzed triplets and rates were below the significance threshold (p-value <= 0.05), while only nonsignificant values were scored in remaining triplets. In the second protocol, HyPhy was applied to the multigene families predicted through MCL [135]. To detect anomalous evolutionary rates, the third protocol used likelihood ratios applied to the same orthogroups used in the other protocols. Since it is applied here for the first time, the third protocol is detailed in Additional file 13.

Additional files

- Additional file 1: Organelle genomes and maps. (DOCX 5930 kb)
- Additional file 2: Synteny between Asterochloris vs. Coccomyxa or Chlorella. (DOCX 583 kb)
- Additional file 3: KEGG-based functional gene categories in the C. grayi symbionts. (DOCX 798 kb)
- Additional file 4: A 500-Kb DNA virus insertion left a clear footprint in the A. glomerata genome. (DOCX 100 kb)
- Additional file 5: Sexual reproduction and candidate genes. (ZIP 182 kb)
- Additional file 6: In vitro culture and gene expression. (ZIP 1530 kb)
- Additional file 7: Expanded and contracted families. (ZIP 220 kb)
- Additional file 8: Signal transduction diversification. (ZIP 3420 kb)
- Additional file 9: Transcription factor family expansions and contractions. (XLSX 30 kb)
- Additional file 10: Carbonic anhydrases in A. glomerata. (DOCX 217 kb)
- Additional file 11: Slow- and fast-evolving proteins. (ZIP 104 kb)
Abbreviations

AC: Corrected Akaike’s information criterion; ALDH: Aldehyde dehydrogenase; AM: Arbuscular mycorrhizae; ANK: Ankyrin domain; ATO: Ammonia Transport Outward; BLASTP: Basic Local Alignment Search Tool (Proteins); BLASTX: Protein databases search using a translated nucleotide query; CA: Carbonic anhydrase; CAFE: Computational Analysis of Evolutionary Trees; CCHC: Cys 2 His Cys zinc finger domain; CEG: Core eukaryotic gene; CEGMA: Core Eukaryotic Genes Mapping Approach; CoMo: Coculture to Monotone in evolution; CTAB: Cetyltrimethyl ammonium bromide; DRP: Desiccation-related Protein; dsDNA: Double-stranded DNA; EM: Ectomycorrhiza; ESR: Environmental Stress Response; EST: Expressed Sequence Tags; GO: Gene Ontology; GPCR: G Protein-Coupled Receptor; HANT-AC: High Affinity Nitrite Transporter-Assimilation Chloroplast; HET: Heterokaryon incompatibility; HGT: Horizontal gene transfer; IGS: Internal Transcribed Spacers; JTT: Jones-Taylor-Thornton model of protein evolution; KEGG: Kyoto Encyclopedia of Genes and Genomes; LAT: L-type amino acid transporter; LG: Le-Gascuel model of protein evolution; LSU: Large ribosomal subunit; MAG: Magnaporthe grisea-like Go protein; MAPK: Mitogen Activated Protein Kinase; Mb: Megabases; MCL: Markov Cluster Algorithm; MIP: Methylammonium Permease type transporters; MFS: Major Facilitator Superfamily; ML: Maximum likelihood; MY: Malt Yeast extract; NAR: Nitrate reductase; NAR-P: Nitrate reductase paralog; NCLDV: Nucleo-cytoplasmic large DNA viruses; NIR: Nitrite reductase; NJ: Neighbor Joining; NNP: Nitrate/ Nitrite Porter; NRPS: Non-Ribosomal Peptide Synthases; NRT: Nitrate transport related; ORF: Open Reading Frame; PFAM: Protein Family database; PhyML: Maximum Likelihood Phylogeny estimation; PKS: Polyketide synthase; PS II: Photosystem II; RGS: Regulators of G-protein Signaling; RIN: RNA Integrity Number; RPKM: Reads Per Kilobase of gene sequence per Million aligned sequences; SET domain: Su(var)3-9, Enhancer-of-zeste and Trithorax; SSU: Small ribosomal subunit gene; TBLASTX: Translated nucleotide databases search using a protein query; TBLASTX: Translated nucleotide databases search using a translated nucleotide query; TCDB: Transporter Classification Database; TE: Tris EDTA buffer; TF: Transcription factor; TM: Transmembrane; TransportTP: Membrane transporter prediction and characterization program; WAG: Whelan-Goldman model of protein evolution

Acknowledgements

We thank Huntington Willard (Geisinger National Precision Health), Greg Wire (Duke University) and Lisa Bukovin (BioMASON) for their support as members of the Duke Institute of Genomic Sciences and Policy; Scott Baker and Jon Magnuson (Pacific Northwest National Laboratory) for including our subcontract in the PNNL proposal; Ursula Goodenough (Washington University in St. Louis) for insightful discussions and for connecting us with Pavel Skaloud; Pavel Skaloud (Charles University in Prague) for specifying our Astenochloris strain as glomerata; Steven Kapowicz (Eastern New Mexico University) for early work on organelle genome assembly and annotation; Fay-Wei Li (Boycott Thompson Institute) and Jolanta Mjadklowska (Duke University) for help with phylogenetic analysis; Thomas Barlow (Duke University) for the photograph of C. grayi (Fig. 1).

Funding

This work was supported by an intramural seed grant (GSP 2008001) to DA, FL, and FD from the Duke Institute of Genomic Sciences and Policy (for DNA sequencing and assembly); by a three-year subcontract to DA, FL, and FD (for bioinformatics) from a DOE grant No 112442 to Pacific Northwest National Laboratory; and by a DOE-Biological Environmental Research grant No FWP 61327 to FRC and PL (for RNAseq). The work at the Joint Genome Institute is supported by the U.S. DOE Office of Science under Contract No. DE-AC02-05CH11231. The work of MP, DL and SSM was supported by the U.S. DOE Office of Science, Office of Biological and Environmental Research program under Award No DE-FC02-02ER63421. The meiosis gene identification was performed in the context of work supported by the National Science Foundation under Grant No. 1011101 to E.S. and J.M.L. The funding bodies were not involved in the design of the study, the collection, analysis, and interpretation of data or in writing the manuscript.

Availability of data and materials

The assembled and annotated genomes, as indicated in Methods, are available at the DOE Joint Genome Institute Portal (http://genome-jgi.doe.gov/). Most other relevant data are contained, or their sources indicated, in Additional files. Other datasets used and/or analysed during the current study are available from the author responsible on reasonable request. For the sequenced mycobiont Gpt/Du/myc/ss, the Culture Bank accession # is CBS 132746 the GenBank ITS accession # is KCS592272. The sequence accession # for the MAT–2 gene is NHPT95990.

Authors’ contributions

DA conceived and directed the study, wrote grant proposals, devised the experiments and general analyses, collected the natural specimens, isolated and maintained the strains, developed the coulture methods and the protocols for DNA and RNA extraction, curated and analyzed the bioinformatics data, prepared most figures and tables, wrote and revised the manuscript. OM generated most of the bioinformatics data; assemblies and annotations, gene family analyses and, with NR, evolutionary rate analyses; NR developed the likelihood-based heterotachy detection method. AK and IVG were instrumental in implementing genomic reassembly and in setting up the JGI portals. FL and FRC sequenced coculture RNA and contributed to its analysis. DA, FL and FD were co-PIs on the grant proposals; FD conducted early bioinformatic analyses setting the stage for subsequent work, and FL co-supervised with DA the work of OM. GB analyzed algal synteny, the flagellar proteome and the viral insertion in the alga. MP and DL provided general KEGG-based gene categorizations and, with SSM, specific web portals for functional gene analysis and sequence retrieval. ES and JML identified the meiosis-specific genes in the alga. SPM assisted DA in the laboratory procedures. RO contributed to the coculture setups. FM, EM and CVF analyzed gene families and the fungal transportome. JS analyzed nitrate gene clusters. SJ, FDG, IS, and HBB were involved in the PKN-NRPS analysis and SJ reannotated manually several genes. BBX, OSA, VM and MWP were involved in the assembly and analysis of the organelle genomes. VM and TRM contributed to the sequencing and analysis of the MAT loci. DS and NIT analyzed signal transduction components. PS analyzed transcription factor families. FL, VM, RO, JS, SSM, ES, and PS also provided extensive feedback on the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher’s Note

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

Author details

1Department of Biology, Duke University, Durham, USA. 2Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, USA. 3Faculty of Life and Environmental Sciences, University of Iceland, Reyjavik, Iceland. 4Aix Marseille University, Université de Toulon, CNRS, 13005, MO UIM 110, 13288 Marseille, France. 5Molekulares Biotechnologie, Fachbereich Biowissenschaften & Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, Frankfurt am Main, Germany. 6Argonne National Laboratory, Biosciences Division, Argonne, & Department of Bioengineering, University of Illinois at Chicago, Chicago, USA. 7Senckenberg Biodiversity and Climate Research Center (SBiK-F), Frankfurt am Main, Germany. 8US Department of Energy Joint Genome Institute, Walnut Creek, USA. 9Department of Plant and Microbial Biology, University of California – Berkeley, Berkeley, USA. 10College of General Studies, University of Wisconsin - Milwaukee at Waukesha, Waukesha, USA. 11Department of Biology, University of Iowa, Iowa City, USA. 12Gilead Sciences, Inc, Foster City, USA. 13INRA, Université de Lorraine, Interactions Arabes-Microorganismes,
INIRA-Nancy, Champenois, France. 14Department of Population Health and Pathobiology, College of Veterinary Medicine, North Carolina State University, Raleigh, USA. 15Department of Biology, St. Catherne University, St. Paul, USA. 16Department of Molecular and Cell Biology, University of California – Berkeley, Berkeley, USA. 17Department of Microbiology and Immunology, University of British Columbia, Vancouver, Canada. 18Department of Ecology, Evolution, and Marine Biology, University of California - Santa Barbara, Santa Barbara, USA. 19Department of Molecular, Cell, and Developmental Biology, and DOE Institute for Genomics and Proteomics, University of California, Los Angeles, USA. 20National Evolutionary Synthesis Center, Durham, USA. 21Calico Life Sciences LLC, South San Francisco, USA. 22IBAM, Facadtit de Ciencias Agrarias, CONICET, Universidad Nacional de Cuyo, Chacras de Coria, Argentina. 23Institute of Ecology, Evolution and Diversity, Fachbereich, Biowissenschaften, Goethe University Frankfurt, Frankfurt am Main, Germany. 24College of Food, Agricultural, and Environmental Sciences, Department of Plant Pathology, The Ohio State University, Columbus, USA. 25College of Life & Environmental Sciences, University of Exeter, Exeter, UK. 26Department of Systematic and Evolutionary Botany, University of Zurich, Zurich, Switzerland. 27The Sainsbury Laboratory, Norwich Research Park, Norwich, UK. 28Université de Lorraine, INRA, Interactions Arbres-Microorganismes, Faculté des Sciences et Technologies, Vandoeuvre les Nancy Cedex, France. 29Laboratory of Medical Microbiology, Vaccine & Infectious Disease Institute, University of Antwerp, Antwerp, Belgium.

Received: 29 August 2018 Accepted: 20 March 2019

Published online: 23 July 2019

References
1. Honegger R. Great Discoveries in bryology and lichenology - Simon Schwendener (1829-1919) and the Dual Hypothesis of Lichens. Bryologist. 2000;103(2):107–13.

2. U'Ren JM, Lutzoni F, Miadlikowska J, Laetsch AD, Arnold AE. Host and geographic structure of endophytic and endolithic fungi at a continental scale. Am J Bot. 2012;99(5):898–914.

3. Grube M, Wedin M. Lichenized Fungi and the Evolution of Symbiotic Organization. Microbiol Spectr. 2016;4(6):1–7.

4. Arnold AE, Miadlikowska J, Higgins KL, Sarvate SD, Gugger P, Way A, Hofstetter V, Kauff F, Lutzoni F. A Phylogenetic Estimation of Trophic Organization. Microbiol Spectr. 2016;4(6):1–5.

5. Honegger R, Edwards D, Axe L. The earliest records of internally stratified bacteria in the lichen microbiome. Environ Microbiol. 2012;14(1):147–61.

6. Spribille T, Tuovinen V, Kubartova A, Ottosson E, Stenlid J, Thor G, Dahlberg A. No support for occurrence of free-living Cladonia mycobionts in dead wood. Fungal Ecol. 2015;14:130–2.

7. Sanders WB. Observing morphoscopic phases of lichen life cycles on transparent substrata placed in situ. Lichenologist. 2005;37:57–82.

8. Stocker-Worgotter E. New frontiers in bryology and lichenology - Experimental lichenology and microbiology of lichens: Culture experiments, secondary chemistry of cultured mycobionts, resynthesis, and thallus morphogenesis. Bryologist. 2001;104(4):576–81.

9. Armaleo D, Miao V. Symbiosis. In: Honegger R, editor. The Lichen Alga Trebouxia – Do they live Free-Living? Plant Syst Evol. 1988;158-2:243–7.

10. Bubrick P, Galun M, Frensdorff A. Observations on free-living Trebouxia De Puymal and Pseudotrebouxia Archibald, and evidence that both symbionts from Xanthoria petulans (L.) THFR. can be found free-living in nature. The New Phytol. 1984;97:465–7.

11. Tuovinen V, Svensson M, Kubartova A, Ottosson E, Stenlid J, Thor G, Dahlberg A. No support for occurrence of free-living Cladonia mycobionts in dead wood. Fungal Ecol. 2015;14:130–2.

12. Joneson S, Armaleo D, Lutzoni F. Fungal and algal gene expression in early developmental stages of lichen-symbiosis. Mycologia. 2011;103(2):291–306.

13. Armaleo D, Sun XM, Culberson C. Insights from the first putative biosynthetic gene cluster for a lichen depside and depsidone. Mycologia. 2011;103(4):741–54.

14. Defoort PT. Early molecular investigations of lichen-forming symbionts: 1986–2001. Annu Rev Microbiol. 2004;58:273–301.

15. Grube M, Berg G, Andrésson ÖS, Wilhelmsdottir O, Yder P, VPW M. Lichen Genomics: Prospects and Progress. In: Martin F, editor. The Ecological Genomics of Fungi. Hoboken: Wiley; 2014. p. 191–212.

16. Schneider T, Schmid E, de Castro JV Jr, Cardinale M, Eberl L, Grube M, Berg G, Riedel K. Structure and function of the symbiosis partners of the lung lichen (Lobaria pulmonaria L. Hoffm.) analyzed by metaproteomics. Proteomics. 2011;11(13):2752–6.

17. Junttila S, Armaleo D. Characterization of a transcriptome from a non-model organism, Cladonia rangiferina, the grey reindeer lichen. BMC Genomics. 2012;13:575.

18. Junttila S, Laito A, Gyenesi A, Sudd PR. Whole transcriptome characterization of the effects of dehydration and rehydration on Cladonia rangiferina, the grey reindeer lichen. Bmc Genomics. 2013;14:870.

19. Park SY, Choi J, Kim JA, Jeong MH, Kim S, Lee YH, Hur JS. Draft Genome Sequence of Cladonia macilenta KoLRI003786, a Lichen-Forming Fungus Producing Biruloquinone. Genome Announc. 2013;1(5)
185. Welch AZ, Gibney PA, Bostein D, Koshland DE. TOR and RAS pathways regulate desiccation tolerance in Saccharomyces cerevisiae. Mol Biol Cell. 2013;24(2):115–28.

186. Meessen J, Sanchez FJ, Brandt A, Balzer EM, de la Torre R, Sancho LG, de Vera JP, Ott S. Extremotolerance and resistance of lichens: comparative studies on five species used in astrobiological research. I. Morphological and anatomical characteristics. Orig Life Evol Biosph. 2013;43(3):288–303.

187. Kanner I, Cram WJ, Zorn M, Wornik S, Yoshimura I, Stabentheiner E, Pfeiffer HW. Antioxidants and photoprotection in a lichen as compared with its isolated symbiotic partners. P Natl Acad Sci U S A. 2005;102(8):3141–6.

188. Gaye E, Fernández-Brieme S, Vargas R, Lachlan RF, Guadian C, Ramírez-Mejía M, Lutzoni F. The adaptive radiation of lichen-forming Teloschistaceae is associated with sun-screening pigments and a bark-to-rock substrate shift. Proc Natl Acad Sci. 2015.

189. Bhattacharya D, Lutzoni F, Reeve B, Simon D, Nason J, Fernandez F. Widespread occurrence of spicoseoilom introns in the rDNA genes of ascomycetes. Mol Biol Evol. 2000;17(2):1971–84.

190. Gargas A, DePriest PT, Taylor JW. Positions of multiple insertions in SSU rDNA of lichen-forming fungi. Mol Biol Evol. 1995;12(2):208–18.

191. Hom EFY, Murray AW. Niche engineering demonstrates a latent capacity for delimit species in the lichen photobiont Trebouxia. Mol Phylogenet Evol. 2014;76:202–10.

192. Singh G, Dal Grande F, Divakar PK, Otte J, Crespo A, Schmitt I. Macroclimatic and coevolutionary forces influence fungal-algal association patterns in Protapomaria. In: The 8th IAL Symposium. Helsinki: University of Helsinki; 2016.

193. Sadowska-Des AD, Dal Grande F, Lumbsch HT. Unravelling the phylogenetic relationships of lichenised fungi in Switzerland and implications for its conservation. Mol Ecol. 2014;23(13):3159–72.

194. Weiss M, Waller F, Zuccaro A, Selosse MA. Unicellular symbionts - one thousand and one interactions with land plants. New Phytol. 2016;206(1):20–40.

195. Seymour FA, Crittenden PD, Dyer PS. Sex in the extremes: lichen-forming fungi. Mycologist. 2005;19(2):51–8.

196. Zoller S, Lutzoni F, Scheidegger C. Genetic variation among unrelated symbiotic green algae. Mol Ecol. 1999;8(12):2049–55.

197. Singh G, Dal Grande F, Werth S, Schiederger C. Long-term consequences of disturbances on reproductive strategies of the rare epiphytic lichen Lobaria pulmonaria: a case study of environmental preferences in symbiotic green algae. Mol Biol Evol. 1996;13(7):797–88.

198. Friedl T, Besendahl A, Pfeiffer P, Bhattacharya D. The distribution of group I introns in lichen algae suggests that lichenization facilitates intron lateral transfer. Mol Phylogenet Evol. 2000;14(3):342–52.

199. Nelsen MP, Duque-Thus R, Grube M, Friedl T, et al. Revisiting photobiont association patterns in lichen symbiosis linked to macroclimate. New Phytol. 2016;209(3):1317–29.

200. Piercey-Normore MD, DePriest PT. Algal switching among lichen symbioses. American J Bot. 2001;88(8):1490–8.

201. Kosugi M, Arita M, Shizuma R, Kashino Y, Kohle H, Sato K. Responses to Desiccation Stress in Lichens are Different from Those in Their Photobionts. Plant Cell Physiol. 2009;50(4):879–88.

202. Peksa O, Skaloud P. Changes in chloroplast structure in lichenized algae. Symbiosis. 2008;46(3):153–60.

203. Palmvist K, Dahlman, L., Jonsson, A., and Nash, T.H. The carbon economy of lichens. In: Lichen Biology. Edited ger. Cambridge: Cambridge University Press; 2004.

204. Palmvist K, Sundberg B. Light use efficiency of dry matter gain in five macro-lichens: relative impact of microclimate conditions and species-specific traits. Plant Cell Environ. 2000;23(1):1–14.

205. Bhattacharya D, Friedl T, Damberger S. Nuclear-encoded rDNA group I introns: Origin and phylogenetic relationships of insertion site lineages in the green alga. Mol Biol Evol. 1996;13(7):797–89.

206. Friedl T, Besendahl A, Pfeiffer P, Bhattacharya D. The distribution of group I introns in lichen algae suggests that lichenization facilitates intron lateral transfer. Mol Phylogenet Evol. 2001;14(3):342–52.

207. Nyati S, Bhattacharya D, Werth S, Homeger R. Phylogenetic Analysis of LSU and SSU RRNA Group I Introns of Lichen Fungi. Mycologist. 2005;19(2):51–8.

208. Korf I. Gene finding in novel genomes. BMC Bioinformatics. 2004;5:59.

209. Singh G, Dal Grande F, Divakar PK, Otte J, Crespo A, Schmitt I. Fungal-algal association patterns in lichen symbiosis linked to macroclimate. New Phytol. 2017;214(1):309–19.

210. Aoki H, Putnam N. Phylogenetic and Ecological Significance of the Problem of Coexistence versus Competition? Environ Microbiol. 2011;13(3):806–18.

211. Vähäkallio V, Väisänen S. Functional diversity in Antarctic Lichens. J Eukaryot Microbiol. 2015;62(2):196–205.

212. Watson P, Barlow M, Newell F. The Nature of Lichens. Bot J Linn Soc. 1988;96(1):3–26.

213. Singh G, Dal Grande F, Divakar PK, Otte J, Crespo A, Schmitt I. Fungal-algal association patterns in lichen symbiosis linked to macroclimate. New Phytol. 2017;214(1):309–19.

214. Piercey-Normore MD, DePriest PT. Algal switching among lichen symbioses. American J Bot. 2001;88(8):1490–8.
235. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009;10(3).

236. Larsen PE, Collart FR. BowStrap v1.0: Assigning statistical significance to their use in phylogenetic analysis. Mol Biol Evol. 2000;17(4):540–52.

237. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient software-and-tools. Accessed 5 Feb 2016.

238. Casteleiro J, Casteleiro J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol. 2007;56(4):564–77.

239. Penn O, Privman E, Landan G, Graur D, Pupko T. An Alignment Confidence Score Capturing Robustness to Guide Tree Uncertainty. Mol Biof. 2010; 28(1):1759–67.

240. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Muller R, Wohlleben W, et al. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. 2015;43(W1):W237–43.

241. TransportTP: A Transporter Prediction Server. http://bioinfo3.noble.org/transporter/.

242. Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. Nucleic Acids Res. 2017;45(D1):D320–4.

243. Wissinger B. Saccharomyces Genome Database. http://www.yeastgenome.org/. Accessed 2 July 2017.

244. Schwikowski B, Uetz P, Fields S. A Network of protein-protein interactions in yeast. Nature. 2000;405(6783):640–4.

245. Middeldorp J, Cate JH, Rabaey K, Groen AK, Schaaps D, Bork P. The Yeast Proteome: A Systematic Survey of Protein Interactions in Yeast. Mol Biol Cell. 2004;15(12):5655–77.

246. Wiwattanapatapee R, Prueksaritanont T, Kangkwan P, Araya-Panich L, Kanchanapoom S, Suwannarangsri I, Khwong S, Juangwong S, Supaphol P, et al. AntiSMASH 3.0: an updated and enhanced software-and-tools. Accessed 5 Feb 2016.

247. Park J, Park J, Jang S, Kim S, Kong S, Choi J, Ahn K, Kim J, Lee S, Kim S, et al. S.R.E. HMMER: biosequence analysis using profile hidden Markov models. http://hmmer.org/. Accessed 22 July 2017.

248. Schubert V, Weissleder A, Ali H, Fuchs J, Lermontova I, Meister A, Schubert I. The Chromosome: One is dependent on MER3, whereas the other is not. Curr Biol. 2005;15(18):692–701.

249. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32(5):1792–7.

250. Tamura K, Peterson D, Peterson N, Stecher G, Neitz M, Kumar S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mol Biof Evol. 2011;28(10):2711–9.

251. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

252. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

253. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

254. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

255. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

256. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

257. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

258. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

259. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

260. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

261. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.

262. Yang PF, Dienner DR, Yang C, Kohno T, Pauker GJ, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas reinhardtii. J. Cell Biol. 2005;170(1):103–13.