

1 **Large differences in carbohydrate degradation and**  
2 **transport potential in the genomes of lichen fungal**  
3 **symbionts**

4

5 Philipp Resl<sup>1,2</sup>, Adina R. Bujold<sup>3</sup>, Gulnara Tagirdzhanova<sup>3</sup>, Peter Meidl<sup>2</sup>, Sandra Freire Rallo<sup>4</sup>,  
6 Mieko Kono<sup>5</sup>, Samantha Fernández-Brime<sup>5</sup>, Hörður Guðmundsson<sup>7</sup>, Ólafur Sigmar  
7 Andrésson<sup>7</sup>, Lucia Muggia<sup>8</sup>, Helmut Mayrhofer<sup>1</sup>, John P. McCutcheon<sup>9</sup>, Mats Wedin<sup>5</sup>, Silke  
8 Werth<sup>2</sup>, Lisa M. Willis<sup>3</sup>, Toby Spribille<sup>3\*</sup>

9

10 <sup>1</sup>University of Graz, Institute of Biology, Universitätsplatz 2, 8010 Graz, Austria

11 <sup>2</sup>Ludwig-Maximilians-University Munich, Faculty of Biology Department 1, Diversity and  
12 Evolution of Plants, Menzingerstraße 67, 80638 Munich, Germany

13 <sup>3</sup>University of Alberta, Biological Sciences CW405, Edmonton, AB T6G 2R3 Canada

14 <sup>4</sup>Rey Juan Carlos University, Departamento de Biología y Geología, Física y Química  
15 Inorgánica, Móstoles, Spain.

16 <sup>5</sup>Swedish Museum of Natural History, Botany Department, PO Box 50007, SE10405  
17 Stockholm, Sweden

18 <sup>7</sup>Faculty of Life and Environmental Sciences, University of Iceland, Sturlugata 7, 102  
19 Reykjavík, Iceland

20 <sup>8</sup>University of Trieste, Department of Life Sciences, via L. Giorgieri 10, 34127 Trieste, Italy

21 <sup>9</sup>Division of Biological Sciences, University of Montana, Missoula, Montana. Current address:  
22 Biodesign Institute and School of Life Sciences, Arizona State University, Tempe, Arizona,  
23 United States of America.

24 \*corresponding author: toby.spribille@ualberta.ca

25

## 26 **Abstract**

27 Lichen symbioses are generally thought to be stabilized by the transfer of fixed carbon  
28 compounds from a photosynthesizing unicellular symbiont to a fungus. In other fungal  
29 symbioses, carbohydrate subsidies correlate with genomic reductions in the number of genes  
30 for plant cell wall-degrading enzymes (PCWDEs), but whether this is the case with lichen  
31 fungal symbionts (LFSs) is unknown. We predicted genes encoding carbohydrate-active  
32 enzymes (CAZymes) and sugar transporters in 17 existing and 29 newly sequenced genomes  
33 from across the class *Lecanoromycetes*, the largest extant clade of LFSs. Despite possessing  
34 lower mean numbers of PCWDE genes compared to non-symbiont *Ascomycota*, all LFS  
35 genomes possessed a robust suite of predicted PCWDEs. The largest CAZyme gene numbers,  
36 on par with model species such as *Penicillium*, were retained in genomes from the subclass  
37 *Ostropomycetidae*, which are found in crust lichens with highly specific ecologies. The lowest  
38 numbers were in the subclass *Lecanoromycetidae*, which are symbionts of many generalist  
39 macrolichens. Our results suggest that association with phototroph symbionts does not in itself  
40 result in functional loss of PCWDEs and that PCWDE losses may have been driven by adaptive  
41 processes within the evolution of specific LFS lineages. The inferred capability of some LFSs  
42 to access a wide range of carbohydrates suggests that some lichen symbioses may augment  
43 fixed CO<sub>2</sub> with carbon from external sources.

## 44 **Significance**

45 Lichen symbioses are considered self-contained autotrophic systems in which the total carbon  
46 economy is the sum of phototroph-fixed CO<sub>2</sub>, supplied to a fungus as sugars. In other fungal-  
47 plant symbioses, such as mycorrhizae, plant-derived sugar subsidies are associated with loss of  
48 plant cell wall-degrading enzymes (PCWDEs). We compared PCWDE inventories in 46  
49 genomes from the largest group of lichen fungal symbionts (LFSs) with non-symbionts from  
50 across *Ascomycota*. We found that despite lower overall gene numbers, all LFSs retain  
51 PCWDEs, and some possess gene numbers and functional diversity on par with non-symbionts.  
52 Our results suggest that association with a phototroph does not necessarily result in PCWDE  
53 loss, and some lichens may obtain carbon from sources other than CO<sub>2</sub> fixation.

54

55

## 56 **Key words**

57

58 CAZymes, glycomics, *Lecanoromycetes*, metagenomes, PCWDEs, saprotroph, symbiosis

## 59 **Introduction**

60 Stable fungal associations with single-celled photosynthetic organisms, usually referred to as  
61 lichens, feature prominently in the history of the discovery and study of symbiosis. In  
62 describing the pairing of fungi with algae and/or cyanobacteria within lichens for the first time,  
63 the Swiss botanist Simon Schwendener proposed that lichen fungal symbionts derive nutrition  
64 from “assimilates” of their photosynthesizing symbionts (1). Almost a hundred years later,  
65 Smith and colleagues revealed these transferred photosynthates to be polyols and glucose in  
66 the case of algae and cyanobacteria, respectively (2). They and others traced the transfer of

67 algal fixed carbon into fungal cells, where they found it to be converted into mannitol and  
68 arabitol (3, 4). The fungal-photosymbiont relationship is widely interpreted as conferring net  
69 independence from external carbohydrates on the resulting lichen thallus. Accordingly, lichen  
70 fungal symbionts have been classified as biotrophs (5), and the symbiotic outcome, the lichen  
71 thallus, as a “photosynthetic carbon autotroph” (6) or “composite autotroph” (7).

72

73 Fungi are assimilative heterotrophs and thus require a robust machinery of enzymes for  
74 scavenging and transporting extracellular nutrients, including carbohydrates. In arbuscular  
75 mycorrhizal and ectomycorrhizal fungi, the stable supply of glucose from plants is thought to  
76 have led to erosion or loss in many families of plant cell wall-degrading enzymes (PCWDEs)  
77 (8, 9), reflecting a common pattern of compensated trait loss in symbioses (10). So what  
78 happened to PCWDEs in lichen fungal symbionts? Multiple lines of indirect evidence have  
79 emerged over the last 40 years to suggest the retention of PCWDEs in lichen fungal symbionts  
80 (LFSs) or their secondary evolutionary derivatives. First, molecular phylogenetic studies have  
81 shown multiple independent origins of putative saprotroph lineages from lichen fungal  
82 symbiont ancestors, both ancient (11) and recent (12,13,14). How these newly evolved lineages  
83 acquired the carbohydrate breakdown arsenal they would presumably need for life without an  
84 alga has not been explained. Second, some fungi near the symbiont-to-saprotroph transition  
85 have been shown to switch between the two lifestyles facultatively, so-called “optional lichens”  
86 (15, 16); in these cases, the fungus appears not to be obligately dependent on the alga for  
87 nutrition. Third, many lichen symbioses exhibit anomalous “substrate specificity”, i.e., they are  
88 restricted to specific organic substrates and unable to colonize others, suggesting lack of  
89 nutritional autonomy (17, 18, 19). Fourth, lichen fungi are capable of growing axenically *in*  
90 *vitro* on a variety of sugars other than sugar alcohols, including crystalline cellulose and sucrose  
91 (reviewed by 20). Finally, enzymes involved in breakdown of lichen-exogenous carbohydrates

92 have been isolated from lichens in nature, including cellulases and lignolytic peroxidases  
93 (reviewed by 21). These phenomena are easy to treat as exceptions, but their distribution across  
94 the fungal symbiont tree hints at deeper underlying fungal capabilities, which if combined with  
95 phototroph symbiosis could lead to a kind of “hybrid” lifestyle, in which carbohydrates are  
96 obtained from multiple sources. The possibility of multiple carbon sources for lichens was even  
97 suggested by Schwendener himself in his original 1869 paper, in which he predicted that two  
98 tracks of nutrient acquisition would ultimately be proven: one for lichens that have minimal  
99 substrate contact, which he predicted to depend mostly on algal assimilates; and one for lichens  
100 that closely hug organic substrates such as tree bark or wood (1).

101

102 Even with these results and hypotheses, however, the possibility of carbon assimilation from  
103 sources other than the phototroph symbiont is still generally discounted in lichen symbiosis  
104 research. Lichen ecophysiologicalists currently calculate total lichen carbon budgets based on the  
105 rate and total amount of algal carbon fixation (6, 22). Implicit to this, though to our knowledge  
106 never explicitly stated, is that lichen fungal symbionts must lack the ability to break down  
107 lichen-exogenous carbohydrates, i.e., carbohydrates found neither in the lichen fungus nor the  
108 phototroph symbiont.

109

110 Unlike for many fungi, phenotypic carbohydrate use profiles have seldom been developed for  
111 LFSs. This is in part due to the recalcitrance of most LFSs to culturing and their extreme slow  
112 growth, if culturing is successful. Knowledge gaps around unculturable or slow-growing fungi  
113 are common, but have been offset in recent years by genome sequencing. Coupled with the  
114 development of widely available databases such as CAZy (23), it has become possible to infer  
115 carbohydrate active enzymes (CAZymes) for species for which a genome, but no experimental

116 evidence, is currently available. Comparative genomic overviews of CAZyme repertoires are  
117 now available for many symbiotic fungi (9), but no survey exists of comparative CAZyme  
118 arsenals in LFSs.

119

120 Given the common assumption that lichen symbiont complementarity confers collective  
121 autotrophy on the symbiosis, and historical assumptions that they evolved from saprotrophic  
122 ancestors, we hypothesized that LFSs would exhibit functional losses in PCWDEs coinciding  
123 with the beginning of stable association with phototroph symbionts, similar to what has been  
124 found in mycorrhizal fungi (9,24). To test this, however, we would need to map CAZymes  
125 across a well-sampled phylogeny and reconstruct the ancestral state of the last common  
126 ancestor, which is considerably older than many of the origins of e.g. ectomycorrhizal fungi  
127 (25). Here we map the occurrence of genes encoding PCWDEs at two levels: across  
128 *Ascomycota* including the origin of the *Lecanoromycetes*, the largest extant lineage of LFSs;  
129 and across representatives of major groups within *Lecanoromycetes*, representing different  
130 ecological substrate specificities (specialists and generalists) as well as major morphological  
131 outcomes of the lichens they occur in (crusts, macrolichens). Our survey of 46 lecanoromycete  
132 genomes, 29 of which we sequenced for this project, reveals a complex pattern of retention and  
133 loss that lends support to Schwendener's hypothesis of hidden saprotrophy in some lichens and  
134 is not unequivocally consistent with PCWDE erosion upon acquisition of phototroph  
135 symbionts.

136

## 137 **Results**

### 138 **Data set and phylogenomic reconstruction**

139 We assembled a data set of 83 fungal genomes from the phylum *Ascomycota*, including 46  
140 from the class *Lecanoromycetes* (Supplementary Table 1). Because the few published  
141 lecanoromycete genomes are not representative of deep evolutionary splits in the group, we  
142 generated 29 new genomes for this study, including 18 as metagenome-assembled genomes  
143 (MAGs; Supplementary Table 2). Completeness and quality metrics were comparable for  
144 genomes derived from culture and MAGs (Supplementary Figures 1 and 2). Phylogenomic  
145 analysis based on 1310 inferred universally present single-copy orthologs (Fig. 1;  
146 Supplementary Figure 3 and Supplementary Table 7) recovered major clades and sister group  
147 relationships found in recent studies, both among class-level clades of *Ascomycota* (26) as well  
148 as within the *Lecanoromycetes* (27, 28). For each genome, we performed *ab initio* gene  
149 predictions and obtained functional annotations (CAZymes, InterPro IDs, Pfams). To these we  
150 assigned activity on the common plant cell wall substrates cellulose, hemicellulose, lignin and  
151 pectin following (9).

152

### 153 **CAZymes**

154 Many CAZyme families are shared across *Ascomycota*. Glycosyl transferases (GTs), which are  
155 involved in glycosylation and the synthesis of polysaccharides, differ little across all analyzed  
156 genomes and do not exhibit any significant reduction in *Lecanoromycetes*, suggesting that a  
157 core *Ascomycota* synthetic machinery remains largely unchanged across the evolution of these  
158 groups (Fig. 1). While significant within-group variation exists, the mean number of glycoside  
159 hydrolase (GH) genes in lecanoromycete genomes is 40.7% lower compared to other  
160 *Ascomycota* genomes ( $p=0$ ; Supplementary Table 9, 10). CAZymes with Auxiliary Activities  
161 (AA), Carbohydrate Binding Modules (CBM) and Carbohydrate Esterases (CE) are also  
162 reduced significantly in *Lecanoromycetes* (Supplementary Table 10). The bulk of these  
163 differences can be attributed to a small number of CAZyme families (Supplementary Figure

164 11), most of which are plant cell wall degrading enzymes (PCWDEs). Additionally three  
165 CAZyme families widespread in *Ascomycota* were not detected in *Lecanoromycetes*. These  
166 include PL4, which contains pectin degrading rhamnogalacturonan endolyases; CBM67, which  
167 binds to L-rhamnose and frequently occurs in multi-domain protein with enzymes in GH78 and  
168 PL1; and AA13, which contains lytic polysaccharide monooxygenases (LPMOs) involved in  
169 starch breakdown. In contrast to these reductions, numerous CAZyme families are not reduced  
170 at all in *Lecanoromycetes* compared to other *Ascomycota*. Indeed some, including those  
171 involved in degradation of endogenous fungal cell wall polysaccharides, such as GH128 and  
172 AA5, are even expanded in *Lecanoromycetes* (Supplementary Figure 11).

173

174 All lecanoromycete genomes retained genes encoding enzymes predicted to act on plant cell  
175 wall carbohydrates, including cellulose and hemicellulose (e.g., GH5 and 43), lignin (AA1, 2  
176 and 5). Depending on the symbiont configuration, these carbohydrates may be produced by the  
177 phototroph partner and/or be exogenous to the lichen symbiosis (see Discussion).  
178 Phylogenetically corrected principal components analyses (PCA) of CAZyme family numbers  
179 in the sampled genomes reveals differences in the amount of variation in CAZyme composition  
180 among lecanoromycete and other ascomycete genomes (Fig. 2). For CAZymes predicted to act  
181 on cellulose, hemicellulose and pectin, the number of predicted gene families varies less among  
182 lecanoromycete genomes than among comparable classes of *Ascomycota* (Fig. 2A, 2C),  
183 reflected in tight clustering in *Lecanoromycetes* versus wide scattering in other ascomycete  
184 classes. For lignin-degrading enzymes the variation is similar (Fig. 2B). Within  
185 *Lecanoromycetes*, however, the greatest amount of variation in predicted gene sets is exhibited  
186 in the subclass *Ostropomycetidae* in cellulose/hemicellulose (Fig. 2A) and pectin degradation  
187 (Fig. 2C), resulting in scattering in the PCA ordination. Genomes from the subclass



188 *Lecanoromycetidae*, by contrast, possess similar predicted CAZyme sets, resulting in tight  
189 clusters in the PCA ordination.

190 Genomes from the *Ostropomycetidae* possess higher numbers of predicted CAZyme genes that  
191 act on cellulose and hemicellulose than those of *Lecanoromycetidae*, which are reflected in  
192 significantly higher numbers of GHs and CEs ( $p=0$  and  $p=0.0024$  respectively; Supplementary  
193 Table 10). The differences are driven by one lineage in particular, represented by five genomes  
194 from the orders *Ostropales* and *Gyalectales* and referred to here as the OG clade (Fig. 1), and  
195 are much less if the two suborders are compared without the OG clade genomes. The OG clade  
196 contains numbers of predicted CAZyme genes for cellulose and hemicellulose breakdown that  
197 are over fourfold more numerous than the lowest lecanoromycete numbers, which are in  
198 *Cetradonia* and *Peltigera*, and equal or exceed the numbers in well-studied eurotiomycete  
199 saprotrophs such as *Penicillium*. These disparities are accounted for in large part by gene  
200 assignments to two CAZyme families, GH5 and GH43. Because both of these are large,  
201 heterogeneous families that include multifunctional CAZymes, we mapped putative GH5 and  
202 GH43 orthologs from the analyzed genomes against sequences of experimentally validated  
203 enzymes and predicted which are secreted (Fig. 3). Extensive gene duplication in  
204 *Ostropomycetidae* is found in gene sequences close to characterized cellulases (GH5 subfamily  
205 5; EC 3.2.1.4) and from both lecanoromycete subclasses in sequences close to characterized  
206 1,3-beta-glucosidases (GH5 subfamily 9; EC 3.2.1.58) and endo-1,4-beta-mannosidases (GH5  
207 subfamily 7; EC 3.2.1.78). Notably, many putative lecanoromycete CAZyme genes from GH5  
208 do not closely cluster with any characterized sequences and form their own clades with  
209 sequences from other classes of *Ascomycota* (Fig. 3A).

210

211 The subclass *Ostropomycetidae* also possesses a larger proportional representation of genes  
212 coding for enzymes predicted to be involved in hemicellulose, specifically xylan, breakdown.  
213 Families GH6, GH7, GH11 and GH62, GH67 and GH131 were absent in *Lecanoromycetidae*  
214 and are only present in *Ostropomycetidae* in the OG clade and the saprotroph *Agyrium*  
215 (Supplementary Figure 11). Hemicellulose breakdown also involves several CEs of which CE2  
216 is present in the majority of *Ostropomycetidae* genomes and absent in *Lecanoromycetidae*  
217 (Supplementary Figure 11). CE2 contains acetylxylan esterases (AXEs; (29)); acetylxylan is  
218 a major component of hemicellulose. An analysis of predicted GH43 gene sequences and their  
219 subcellular location compared to those of characterized enzymes shows that most characterized  
220 GH43s are in *Ostropomycetidae* (subfamilies 1, 6, 24, 26 and 36), but as with GH5s, the  
221 majority of sequences are not close to any characterized CAZymes (Fig. 3B).

222

223 CAZymes involved in lignin degradation are frequently characterized in fungi as encompassing  
224 Class II peroxidases (PODs) and selected oxidative CAZymes (17). In *Lecanoromycetes*,  
225 CAZymes associated with lignin degradation are predicted in all lecanoromycete genomes, but  
226 again the highest numbers are found in *Ostropomycetidae* (Fig. 1). The numbers of predicted  
227 genes in *Graphis* and *Varicellaria* are among the highest in the analyzed *Ascomycota* genomes.  
228 Most of these genes are members of AA1 and AA3 which contain laccases (EC 1.10.3.2; AA1;  
229 Supplementary Figure 9) and glucose-methanol-choline oxidoreductases in three subfamilies  
230 involved in lignocellulose degradation (Supplementary Figure 9). Class II PODs, by contrast,  
231 are virtually absent across *Lecanoromycetes* (Fig. 1).

232

233 The largest differences in predicted PCWDEs among lecanoromycete genomes come from  
234 CAZymes acting on pectins. All but three of the *Ostropomycetidae* genomes possess a few

235 predicted CAZymes involved in pectin degradation (Fig. 1). These include GHs (GH28, 49,  
236 53, GH79, GH108), CEs (CE8, CE12, CE15) and the only polysaccharide lyases (PL; from  
237 PL1 and PL3, containing pectate lyases, EC 4.2.2.2) predicted in *Lecanoromycetes*  
238 (Supplementary Figure 8). *Lecanoromycetidae*, by contrast, almost completely lack gene  
239 predictions for pectin degradative enzymes and completely lack PLs. Similar to CAZyme  
240 predictions for cellulose, hemicellulose and lignin degradation, the largest numbers of  
241 predicted genes involved in pectin degradation come from the OG clade of *Ostropomycetidae*.  
242 In these genomes, predicted gene numbers for CAZymes involved in pectin degradation even  
243 exceed those of known model saprotrophs in *Eurotiomycetes*, though the difference is not  
244 significant.

245

246 In addition to the PCWDEs outlined above, all sampled *Lecanoromycetes* possess a gene  
247 assigned to GH32, a family which includes invertases involved in the conversion of sucrose to  
248 glucose or fructose. Invertases are generally interpreted as an indicator of utilization of  
249 apoplastic plant sucrose. However, lecanoromycete GH32s do not cluster closely with any  
250 characterized invertases (Supplementary Figure 13).

251

## 252 **Predictions of gene family contraction and expansion**

253 In order to establish whether CAZyme patterns in *Lecanoromycetes* are due to gene gain or  
254 loss, we reconstructed ancestral gene numbers for each of the three main groups of PCWDEs  
255 (cellulose/hemicellulose, pectin, lignin). In general, the pattern reflects sequential gene loss.  
256 All three PCWDE groups are reduced already at the split between *Eurotiomycetes* and  
257 *Lecanoromycetes* (ELS node; Fig. 1), especially enzymes involved in (hemi-)cellulose and  
258 pectin breakdown. Much of the signal for this pattern comes from GH5 and GH43 which

259 contain many well-characterized cellulases and hemicellulases, respectively (Fig. 2;  
260 Supplementary Figure 7). Another apparent shedding of CAZymes, especially of those  
261 involved in pectin degradation, occurs concomitantly with the appearance of LFSs at the base  
262 of *Lecanoromycetes* (ALSS and ALSL nodes; Fig. 1). Here, too, the signal comes mainly from  
263 GHs. In several cases our ancestral state reconstruction indicates complete losses, despite the  
264 prediction of several of the relevant families in *Ostropomycetidae* genomes (Supplementary  
265 Figure 8).

266 Gene family expansion analyses of all CAZymes revealed six significantly expanded families  
267 along different branches in the phylogeny. The only GHs expanded within *Lecanoromycetes*  
268 are GH43, which are expanded at the last common ancestor of *Ostropomycetidae* (Fig. 1);  
269 another, GH18, was expanded in four model runs at the base of *Sordariomycetes*. Other  
270 significantly expanded CAZyme families in *Lecanoromycetes* include Auxiliary Activity (AA)  
271 families AA3, AA7 and AA9, which are variously expanded in *Xylographa*, the OG clade and  
272 *Cladonia*. Auxiliary Activity families typically act in concert with other CAZymes and the  
273 three expanded AA families all contain genes involved in cellulose breakdown. The only other  
274 expanded CAZyme family was CE10, now considered to be a family of esterases acting on  
275 non-carbohydrate substrates, which was expanded at the last common ancestor of *Cladonia*  
276 species and *Umbilicariomycetidae*.

277

## 278 **Transporters**

279 As an additional line of evidence for potential use of exogenous carbohydrates, we mapped the  
280 numbers of predicted sugar transporters across all genomes (Fig. 1). Most predicted  
281 transporters were more or less evenly represented across most genomes, but two groups of  
282 transporters exhibited a concerted pattern of absence within *Lecanoromycetes*. Cellodextrin

283 transporters, close orthologs of experimentally demonstrated cellodextrin transporters from  
284 *Aspergillus* and *Penicillium* (Supplementary Figure 5), were predicted in all *Ostropomycetidae*  
285 genomes except *Schaereria*, and were most numerous in *Lignoscripta*, *Stictis*, *Ptychographa*  
286 and *Xylographa*, all of which are LFSs of lichens with high wood or bark specificity (Fig. 1).  
287 Cellodextrin transporters are involved in transmembrane import of cellobiose and other  
288 cellodextrins, which are short beta-linked glucose fragments of cellulose. Only nine of the  
289 remaining 21 lecanoromycete genomes had any predicted cellodextrin transporters, and in  
290 these, gene numbers were well below the average observed in *Ostropomycetidae*. We detected  
291 a similar pattern for the predicted occurrence of maltose transporters (Fig. 1), involved in  
292 transmembrane import of alpha-linked starch breakdown products: 22 of 24 *Ostropomycetidae*  
293 genomes had predicted maltose transporters, but only five lecanoromycete genomes outside of  
294 *Ostropomycetidae* had any, mostly in early diverging lineages.

295

## 296 **Evidence of cellulase functionality**

297 To validate the functionality of putative cellulases found in lichens, we selected *X. bjoerkii*  
298 genes *Xylbjo000518* and *Xylbjo004565*, hereafter called *cellulase A* and *B*, for further analysis  
299 based on their similarity with Cel5A cellulases, for which there has been substantial structural  
300 and functional characterization. The cellulase domain of each gene was expressed in  
301 *Escherichia coli* as a C-terminal fusion with the maltose binding protein, which enhances  
302 protein expression and stability and provides a convenient handle for purification. The partially  
303 purified proteins were active on both cellulose ( $\beta$ -1,4-linked glucose) and barley  $\beta$ -glucan  
304 (alternating  $\beta$ -1,4/ $\beta$ -1,3-linked glucose) but not xylan ( $\beta$ -1,4-linked xylose containing side  
305 branches of  $\alpha$ -arabinofuranose and  $\alpha$ -glucuronic acids) (Fig. 4A-C). Both exhibited a pH  
306 optimum of 5, similar to other Cel5A enzymes but differ slightly in their temperature optima.

307 While both are active at 20 °C, only *cellulase A* is also active at higher temperatures. Closely  
308 related orthologs of *cellulase A* and *cellulase B* were recovered in all lecanoromycete genomes.  
309 *Lecanoromycetidae* mostly only had a single ortholog with the exception of *Lobaria*,  
310 *Pseudocyphellaria*, *Pseudevernia* and *Sticta* which each have two orthologs (Fig. 4D). In  
311 contrast, the majority of *Ostropomycetidae* genomes have two or more orthologs, and the  
312 highest numbers can be found in *Xylographa* species and *Stictis* which each have five (Fig.  
313 4D).

314

## 315 **Discussion**

316 Our analysis of CAZyme and sugar transporter genes paints a picture of a lecanoromycete  
317 PCWDE arsenal that is larger, more diverse and more consistently present than expected.  
318 Understanding when and where in the fungal mycelium the PCWDEs are deployed will be  
319 critical to determining their biological significance. Our data show a reduction in mean  
320 numbers of lecanoromycete degradative CAZyme genes relative to other *Ascomycota*, but also  
321 reveal significant differences within *Lecanoromycetes*. Genes for the breakdown of cellulose,  
322 hemicellulose and pectin are disproportionately enriched in the subclass *Ostropomycetidae*,  
323 and some genomes, notably in the OG clade, retain overall CAZyme numbers and functionality  
324 similar to those of well-known model saprotrophs such as *Aspergillus* and *Penicillium*.  
325 Furthermore, LFSs associated with the same genus of algal phototroph, *Trebouxia*, can retain  
326 multiple genes coding for pectin degradation as well as cellodextrin transporters (in *Lambiella*,  
327 *Loxospora*, *Ptychographa* and *Xylographa*), or they can largely lack these genes (in *Ramalina*).  
328 This implies that association with *Trebouxia* does not in itself result in gene loss, and the exact

329 nature of the fungal-phototroph interaction, and other aspects of symbiosis biology, may need  
330 to be considered.

331

332 A central question remains: what are the targets of LFS PCWDEs? The two most obvious  
333 candidates are the phototroph itself, on the one hand, and lichen-exogenous plant tissues, such  
334 as wood and tree bark, on the other. The first possibility — targeting of the alga — echoes  
335 suggestions in studies of ectomycorrhizal fungi, where PCWDEs have been postulated to play  
336 a role in “cell softening” of root tissues in their vascular plant symbionts upon contact initiation  
337 (30). Cellulases and other CAZymes have been postulated to be involved in haustorial  
338 penetration of the algal cell wall (31, 32) and degradation of algal cell walls in fresh lichen  
339 growth tips (33). They could also play a role in digestion of dead algal cell walls, especially if  
340 algal populations turn over during the “lifespan” of the thallus, as has long been suspected (34).  
341 Transcriptomic studies of isolates, co-cultures and natural lichens could provide evidence of  
342 this. However, the one study to date to report CAZyme differential expression (35) only  
343 detected upregulation of multifunctional GHs that could also be involved in fungal cell wall  
344 modification (GH2 and GH12), but none of the core lecanoromycete cellulases or  
345 hemicellulases we report here. Most common algal symbionts are thought to contain cellulose  
346 in their cell walls but no pectin (36); it is unclear if any lichen algal symbiont possesses pectin  
347 in its cell walls.

348

349 The second possibility, that LFS PCWDEs target lichen-exogenous plant polysaccharides, is  
350 supported by both experimental evidence and inference. Cellulases as well as  
351 polygalacturonases, active on pectins, have been detected in incubated whole lichens both in  
352 the presence (37, 38) and absence (39) of cellulose-containing algal photosymbionts. Cellulase

353 production furthermore has been reported to vary depending on the species of tree it is growing  
354 on (40). Activity has also been demonstrated for laccases and different types of peroxidases  
355 (21, 41, 42). Several secondary origins of putative fungal saprotrophs occur within the  
356 *Lecanoromycetes*, some of which are closely related to fungi involved in “optional lichens”, in  
357 which the fungus can occur either as an LFS or saprotroph (43). Our sampling shows cellulase  
358 gene family expansions within the latter group, especially in LFSs of wood-obligate lichens  
359 such as *Ptychographa* and *Xylographa*, and our experimental evidence for cellulase activity  
360 derives from one of these genomes (*Xylographa bjoerkii*). A further indication of  
361 lecanoromycete use of exogenous carbohydrates could be the presence of invertases. The loss  
362 of invertases in ectomycorrhizal fungi is thought to limit their ability to access plant sucrose  
363 and reinforce their dependence on a plant partner (9). Though the orthologs we found did not  
364 exhibit high similarity to characterized sequences, the prediction of a GH32 is consistent with  
365 the ability to culture LFSs on sucrose (44) and experimental evidence of invertases in lichen  
366 fungi (20, 45).

367

368 PCWDE targeting of algal cell walls or lichen-exogenous plant tissues are not mutually  
369 exclusive possibilities. However, determining where and when these genes are expressed in  
370 nature will require consideration of the life cycle and spatial extent of LFS mycelia, both of  
371 which are still poorly understood. The mycelia of sexually reproducing LFSs go through an  
372 aposymbiotic stage of unknown length, during which it has been suggested they may be  
373 saprotrophic (46). Lichens at high latitudes or under snowpack may go through seasonal  
374 fluctuations in fixed carbon input, which could be augmented by other sources (21). Even after  
375 symbiosis is established, parts of the mycelium can be free of phototroph cells, and exhibit  
376 deviations in carbohydrates that suggest other metabolic processes are in effect than in the  
377 phototroph-associated mycelium (47). Many lichens include a phototroph-free “prothallus”, in



378 which fungal hyphae radiate beyond the zone of phototroph cells (48). Others possess a  
379 “hypothallus”, a phototroph-free cushion of mycelium in direct contact with the substrate (47).  
380 Macrolichens are often anchored onto their substrates by phototroph-free “holdfasts”,  
381 “rhizines” or mycelial pegs which have traditionally been interpreted as having an exclusively  
382 structural, stabilizing function, but can extend as mycelial networks into xylem (49) or living  
383 moss mats (39). Parts of the mycelium with phototroph cells may also differ based on position  
384 relative to growth tips, which are thought to include larger proportions of living cells (50), and  
385 entire strata can be phototroph-free.

386

387 Did stable phototroph association coincide with the loss of PCWDEs in LFSs? Looking only  
388 at the reductions in mean CAZyme gene number, the answer would appear to be positive, but  
389 the occurrence of LFSs with CAZyme arsenals as large or larger than those of many  
390 saprotrophs in the OG clade shows that LFSs do not necessarily lack CAZyme arsenals. Despite  
391 their minority representation in our data set, CAZyme-rich LFSs may in fact be numerous: the  
392 *Ostropales* and *Gyalectales* which make up the OG clade include over 3200 named species,  
393 almost 17% of named LFSs (51). For phototroph acquisition to have no evolutionary  
394 consequences for the CAZyme arsenal, symbiont-derived photosynthates would likely have to  
395 provide a function other than as a substrate for growth and respiration. In fact this has already  
396 long been postulated, in the form of an osmoregulatory role for polyols in managing cell  
397 desiccation (52). Some have even suggested this may be the main role for transferred  
398 photosynthates (53, 54). A strong inference about whether LFS CAZyme gene reduction  
399 coincided with the onset of stable phototroph association requires greater certainty about the  
400 ancestral states along the lecanoromycete phylogenetic backbone. Our ancestral state  
401 reconstruction suggests a gradual loss of GHs and PLs since the last common ancestor of  
402 Lecanoromycetes and Eurotiomycetes, and only pectin degradation genes are inferred to have

403 been abruptly reduced. This reconstruction is sensitive to taxon undersampling and may be a  
404 conservative estimate. The large CAZyme arsenal of the OG clade, with dozens of unlinked  
405 genes and numerous PLs, is unlikely to have been acquired by lateral gene transfer. If the PL-  
406 rich OG clade CAZyme arsenal is in fact ancestral, this would imply that CAZyme loss is not  
407 an automatic consequence of stable association with phototroph symbionts, but rather of  
408 subsequent events or adaptations. If so, this implies that the OG clade CAZyme arsenal was  
409 lost no fewer than seven times in our tree. Support for a scenario of frequent mass gene  
410 reduction comes from the fact that one such loss appears to have happened within the OG clade  
411 itself: *Gomphillus* exhibits marked gene number reductions compared to its closest sampled  
412 relatives. Our data do not currently allow the hypothesis that the OG clade CAZyme arsenal is  
413 ancestral to be rejected, and if it is not, it becomes more likely that CAZyme loss in LFSs is  
414 driven by additional processes.

415

416 What additional adaptational processes could drive CAZyme loss? The most striking apparent  
417 functional losses, both in terms of CAZyme genes as well as in cellodextrin and cellobiose  
418 transporters, occurred in the *Lecanoromycetes* outside of the *Ostropomycetidae*, especially in  
419 the subclass *Lecanoromycetidae*, which largely lacks these two types of transporters. LFSs in  
420 *Lecanoromycetidae*, like those in *Ostropomycetidae*, are considered obligate symbionts of  
421 algae or cyanobacteria. They share many of the same algal symbionts, especially *Trebouxia*,  
422 and have no known lifecycle differences. However, they differ in general thallus architecture.  
423 *Ostropomycetidae* almost exclusively form crustose thalli in which thallus-to-substrate surface  
424 area contact is maximized. *Lecanoromycetidae*, and to some extent *Acarosporomycetidae* and  
425 *Umbilicariomycetidae*, include LFSs involved both in crustose thalli as well as so-called  
426 macrolichens, in which the thallus often becomes greatly expanded into leaf-, hair- or shrub-  
427 like thalli, in which surface contact is minimized. *Lecanoromycetidae* macrolichens include

428 some of the largest lichens by biomass, and in theory these would require more carbon, not  
429 less. Lichens involving *Lecanoromycetidae* differ from those involving *Ostropomycetidae* in  
430 at least two traits that could be implicated in sugar uptake. In the first, the type of fungal-algal  
431 contact differs, from intracellular haustoria in *Ostropomycetidae* to so-called intraparietal  
432 haustoria, which do not breach the cell wall, in *Lecanoromycetidae* (55). The exact  
433 consequence of this is unknown, but intracellular haustoria are characteristic of pathogenic  
434 fungi and may require a greater variety of CAZymes to penetrate the algal cell wall. In the  
435 second, macrolichens involving *Lecanoromycetidae* owe their architecture to a well-developed  
436 exopolysaccharide gel, termed cortex, which forms a rigid structural exoskeleton considered a  
437 prerequisite to macrolichen formation (56). The polysaccharide composition of this layer varies  
438 widely across lichen symbioses (57) and is poorly characterized and mapped across the  
439 phylogenetic tree. Additional evolution in cortex layering happened in many lichens involving  
440 *Lecanoromycetidae* (58), though in how many remains unclear. The cortex, which is  
441 hydrophilic, also mediates water retention (57) and has been shown to operate like a sponge  
442 for passive uptake of dissolved nutrients (59) and glucose (60). One of the largest epiphytic  
443 macrolichens has been experimentally demonstrated to take up tree-derived glucose (61).  
444 Whether this is specifically facilitated by the cortex remains to be tested, but as all  
445 environmental molecules enter macrolichens through the cortex, it seems likely. If capture of  
446 simple sugars is found to be a general function of the cortex in different environments, it could  
447 be expected to have significant evolutionary consequences for maintenance of CAZyme genes  
448 used for more costly carbohydrates.

449

450 The existence of a robust carbohydrate breakdown machinery across a large swathe of LFS  
451 evolution calls into question the assumption, underlying decades of ecophysiological work on  
452 lichens, that the lichen carbon economy is exclusively the sum of algal CO<sub>2</sub> fixation. The

453 finding that some LFSs have CAZyme arsenals on par with saprotrophs lends support to  
454 Schwendener's hypothesis that different types of lichens exist: those that depend mostly on  
455 their phototroph, and those that augment their carbon assimilation from multiple sources (1).  
456 The longstanding assumption that LFSs solely utilize fixed carbon must now be weighed  
457 against competing hypotheses, including A) that some or many LFSs build their mycelia from  
458 non-algal carbon sources, including absorbed monosaccharides and complex polysaccharides,  
459 before or during symbiosis; and B) that many different models of carbon acquisition — fixed,  
460 seasonal, facultative, scavenged and/or absorbed — may exist under the umbrella of what we  
461 currently call lichen symbiosis.

462

## 463 **Materials and Methods**

464 Comprehensive Materials and Methods as well as supplementary results are available as  
465 supplementary material.

### 466 **Used genomes**

467 We built a data set consisting of 83 fungal genomes from the phylum Ascomycota, including  
468 46 genomes from the Lecanoromycetes, twenty-nine of which were newly generated for this  
469 study, and 37 genomes from the related classes *Eurotiomycetes*, *Dothideomycetes*,  
470 *Arthoniomycetes* and *Sordariomycetes*. Within *Lecanoromycetes*, the acquisition of new  
471 genomes was targeted to include a representation of lineages including 1) LFSs of macrolichens  
472 (*Cladonia*, *Evernia*, *Lobaria*, *Peltigera*, *Pseudevernia*, *Pseudocyphellaria*, *Ramalina*, *Sticta*,  
473 *Umbilicaria*); 2) LFS lineages involved in crust-forming lichens for which carbon acquisition

474 from the substrate has been postulated, including wood specialists (*Lecidea scabridula*,  
475 *Lignoscripta*, *Ptychographa*, *Puttea*, *Xylographa* (18)), bark specialists (*Graphis*, *Loxospora*,  
476 *Schaereria*, *Stictis*, *Thelotrema*, *Varicellaria*), specialists of decaying plant matter  
477 (*Gomphillus*, *Icmadophila*), mineral soil (*Dibaeis*), rock (*Acarospora*, *Trapelia*) and a lichen-  
478 on-lichen “parasite” (*Lambiella*); 3) LFS lineages from crust-forming lichens that behave as  
479 ecological generalists (*Gyalolechia*, *Mycoblastus*, *Toensbergia*); and 4) a lecanoromycete  
480 saprotroph (*Agyrium*) and endophyte (*Cyanodermella* (62)). Our sampling simultaneously  
481 represents a cross-section of lecanoromycete evolution, with 24 genomes from the subclass  
482 *Ostropomycetidae*, 17 from the subclass *Lecanoromycetidae*, and four and one each from the  
483 species-poor subclasses *Acarosporomycetidae* and *Umbilicariomycetidae*, respectively. Ten  
484 genomes were derived from cultured samples and 18 are metagenome-assembled genomes  
485 (MAGs) assembled and binned according to (63). Culture-derived genomes differed little from  
486 MAGs in estimated completeness and gene numbers (Supplementary Figures 1 and 2).

487

#### 488 **Genome assembly and filtering**

489 Raw sequence data was inspected with FastQC 0.11.7 and trimmed with trimmomatic 0.29 to  
490 remove adapter remnants and low-quality reads. Genomes were assembled using SPAdes  
491 3.12.0 or Abyss 2.0.2. To extract LFS contigs, draft assemblies were filtered using blobtools  
492 1.1.1 or CONCOCT 1.2 and assembly completeness assessed with Quast 4.6.3 and BUSCO  
493 3.0.2. We additionally extracted mitochondrial contigs by blasting mitochondrial genes  
494 downloaded from NCBI against each *de novo* assembly. We then discarded contigs with blast  
495 hits of  $e < 1e-03$  and alignment length  $> 500$  bp.

496

#### 497 **Gene calling and functional annotation**

498 We used funannotate 1.8.7 to perform gene-calling and functional annotation for all used  
499 genomes in the same way. This reduces potential biases introduced by different gene-calling  
500 and annotation methods. *De novo* sequenced genomes were repeat-masked using  
501 RepeatModeller and RepeatMasker. We then used Augustus 3.3, snap 2013\_11\_29,  
502 GeneMark-ES 4.62 and GlimmerHMM 3.0.4 and tRNA-Scan 2.0.5. Functional annotations for  
503 all predicted protein sequences were inferred using Interproscan-5.48-83.0, HMMer3 searches  
504 against dbCAN (v9; CAZymes) and Pfam (33.1) as well as eggnog-scanner searches against  
505 EggNOG (4.5.1; various annotations) databases.

506

### 507 **Phylogenomics**

508 Phylogenomic trees were calculated using the phylociraptor pipeline  
509 (<https://github.com/reslp/phylociraptor>). We ran BUSCO 3.0.2 for each genome to identify  
510 single copy-orthologous, combined amino-acid sequences of each BUSCO gene from all  
511 genomes. Only genes which were present in >50% of genomes were aligned using mafft 7.464  
512 and trimmed using trimal 1.4.1. We calculated single-gene trees using iqtree 2.0.7 and a species  
513 tree using ASTRAL 5.7.1. We created a concatenated alignment from all alignments, estimated  
514 the best substitution model for each and calculated a tree based on a partitioned analysis of the  
515 concatenated alignment using iqtree 2.0.7. We used the concatenated phylogeny to generate an  
516 ultrametric tree using r8s 1.81. This tree was used for subsequent analyses. We used custom  
517 python and R scripts to plot phylogenomic trees.

518

### 519 **Selection of CAZyme groups**

520 We selected sets of CAZyme families involved in (hemi-)cellulose, pectin and lignin  
521 degradation based on previous studies (9, 24). For the lignin set we additionally identified class  
522 II peroxidases based on similarity of all Ascomycota class II peroxidases downloaded from

523 RedOxiBase (accessed Jul. 14, 2021; <http://peroxibase.toulouse.inra.fr/>) using Orthofinder  
524 2.5.2.

525

### 526 **Ancestral state reconstruction of CAZyme families and similarity of CAZyme sets**

527 We reconstructed the ancestral size of each CAZyme family using CAZyme counts from our  
528 genome annotations and our ultrametric phylogenomic tree in R. We used anc.ML from  
529 phytools under an Ornstein-Uhlenbeck model of trait evolution. We then used custom R scripts  
530 to visualize ancestral states.

531 We calculated phylogenetically corrected PCAs in R as implemented in the phytools function  
532 *phyl.pca* using maximum-likelihood optimization. We used log-transformed gene-counts of  
533 CAZyme families and our ultrametric phylogeny as input for *phyl.pca*. The PCAs were  
534 visualized using custom R scripts.

535

### 536 **Gene family expansion analysis**

537 We analyzed gene family expansions of CAZyme families, based on CAZyme family count  
538 data from funannotate and our ultrametric phylogenomic tree using CAFE 5 (git commit  
539 08d27a1). We ran CAFE twenty times while accounting for different gene-birth rates and error-  
540 models and summarized the results using custom R and python scripts.

541

### 542 **Additional characterization of CAZymes**

543 CAZyme families known to contain important degradative enzymes (e.g., cellulases,  
544 hemicellulases) were further characterized using Saccharis v1 (git commit 9a748be). For each  
545 CAZyme family, we downloaded all characterized genes from [cazy.org](http://cazy.org) including additional  
546 information such as taxonomy, accession numbers and CAZyme (sub-)family assignments.  
547 Using Saccharis, we created MUSCLE 3.8.31 alignments for each gene family of all (from

548 genomes used in this study plus characterized genes from cazy.org) genes and created  
549 maximum-likelihood trees using Fasttree 2.1.10. For all genes included in our phylogenetic  
550 reconstructions, we also predicted subcellular locations using DeepLoc 1.0. We only  
551 considered predictions for subcellular locations if the predicted probability was >70%. We used  
552 custom R and python scripts to visualize the trees.

553

#### 554 **Heterologous expression of putative cellulases**

555 Sequences assigned to GH5 subfamily 5, with confirmed cellulolytic activity, from obligately  
556 wood inhabiting LFS *Xylographa* species were aligned to the characterized and crystalized  
557 cellulase domain of *Trichoderma reesei* (PDB: 3QR3; <https://www.rcsb.org/structure/3QR3>).  
558 Based on sequence similarity, we selected two candidate cellulase sequences from *Xylographa*  
559 *bjoerkii* for testing cellulolytic activity.

560 Enzyme activity of the two candidate cellulases A and B was tested by combining in a 1.5 mL  
561 microcentrifuge tube 50  $\mu$ L enzyme; 100  $\mu$ L buffer A at either pH 4, 5, or 6; and 50  $\mu$ L either  
562 AZCL-HE-cellulose, AZCL- $\beta$ -glucan, or AZCL-xylan. Tubes were incubated at 4, 20, 37, 50,  
563 or 60°C for 48 h. To measure activity, samples were centrifuged at 13,000 rpm for 5 min to  
564 settle any debris, then 100  $\mu$ L supernatant was removed to a 96-well flat-bottom plate.  
565 Absorbance at 595 nm was measured by plate reader and blanked with a sample containing  
566 water instead of enzyme.

#### 567 **Identification of sugar- and sugar-alcohol transporters**

568

569 To identify putative sugar- and sugar-alcohol transporters we used Orthofinder 2.5.2 on all  
570 sequences from all genomes with Pfam annotation PF00083 (Sugar\_tr;  
571 [http://pfam.xfam.org/family/sugar\\_tr](http://pfam.xfam.org/family/sugar_tr)) combined with characterized sugar transporter



572 sequences from the PF00083 seed set. Additionally we added characterized cellodextrin  
573 (MH648002.1 (NCBI; from *Aspergillus niger*), S8AIR7 (UniProtKB; from *Penicillium*  
574 *oxalicum*)) and sugar alcohol transporters (AAX98668.1; from *Ambrosiozyma monospora*,  
575 CAR65543.1, CAG86001.1; from *Debaryomyces hansenii*, NP\_010036.1; from  
576 *Saccharomyces cerevisiae*). We used the presence of characterized sequences in the inferred  
577 orthogroups to identify orthologs in each genome included in this study. The number of  
578 orthologs of different transporters per genome were visualized using custom R scripts.

579

## 580 **Identification of class II peroxidases**

581 First we used diamond 0.9.22 to search all Ascomycota class II peroxidases downloaded from  
582 RedOxiBase (accessed Jul. 14, 2021) against the predicted proteins in all 83 genomes studied  
583 here. We then extracted all genomes which had a diamond hit to any of the downloaded  
584 sequences. Similar to the identification of sugar transporters we now used Orthofinder 2.5.2 to  
585 classify putative class II peroxidases present in our genomes based on the presence of  
586 downloaded genes in individual orthogroups. The number of orthologs per genome for different  
587 peroxidases were visualized using custom R scripts.

588

## 589 **Data Availability**

590 The complete analysis workflow used to acquire the results in this paper, including all custom  
591 python and R scripts, are available on Github ([https://github.com/reslp/LFS-cazy-](https://github.com/reslp/LFS-cazy-comparative)  
592 [comparative](https://github.com/reslp/LFS-cazy-comparative)). The workflow used to calculate phylogenomic trees is available on Github  
593 (<https://github.com/reslp/phylociraptor>). Genome assemblies and corresponding functional  
594 annotations will be deposited at the European Nucleotide Archive (ENA). Accession numbers  
595 are provided in supplementary material.

## 596 **Acknowledgments**

597 The early phases of this project were funded by the Austrian Science Fund (FWF grant P25237,  
598 “Evolution of Substrate Specificity in Lichens”) and carried out at the Institute of Plant  
599 Sciences (Uni Graz). PR would like to acknowledge the Theodor Körner Funds (Vienna,  
600 Austria) and Network of Biological Systematics Austria (NOBIS; Vienna, Austria) for funding  
601 that enabled genome sequencing of several species. SW received funding from the Icelandic  
602 Research Fund IRF (174307-051), from Deutsche Forschungsgemeinschaft DFG (WE 6443/1-  
603 1) and from LMU Munich (startup funds). MW received funding from the Swedish Research  
604 Council, grant VR 2016-03589. MW and MK would like to acknowledge support from the  
605 NRM Department of Bioinformatics and Genetics, the National Genomics Infrastructure in  
606 Stockholm, the SNIC/Uppsala Multidisciplinary Center for Advanced Computational Science  
607 and the UPPMAX computational infrastructure, and Linda Phillips for assistance with material.  
608 TS would like to acknowledge an NSERC Discovery Grant, a Canada Research Chair in  
609 Symbiosis, and the generosity of Susan Dalby and Eskild Petersen, who provided a place to  
610 work on this manuscript. Special thanks go to Sophie Dang at the Molecular Biology Service  
611 Unit, University of Alberta Department of Biological Sciences, for help in data acquisition, and  
612 members of the U of A Lichen Evolution Lab for reading and commenting on the manuscript.  
613 We are also grateful to Tanja Ernst and Andrea Brandl for laboratory assistance.

614

## 615 **References**

- 616 1. Schwendener, S. 1869. *Die Algentypen der Flechtengonidien*.  
617 Universitaetsbuchdruckerei, Basel.
- 618 2. Smith, D.C., Muscatine, L., Lewis, D. 1969. Carbohydrate movement from autotrophs to  
619 heterotrophs in parasitic and mutualistic symbiosis. *Biol. Rev.* 44, 17–90.

- 620 3. Drew, E.A., Smith, D.C. 1967. Studies in the physiology of lichens. VII. The physiology  
621 of the *Nostoc* symbiont of *Peltigera polydactyla* compared with cultured and free-living  
622 forms. *New Phytologist* 66: 379–388.
- 623 4. Richardson, D.L. Hill, D.J., Smith, D. 1968. Lichen Physiology - XI. The role of the alga  
624 in determining the pattern of carbohydrate movement between lichen symbionts. *New*  
625 *Phytologist* 67: 469–486.
- 626 5. Smith, D.C. 1980. Mechanisms of nutrient movement between the lichen symbionts. In:  
627 Cook, C.B., Pappas, P.W., Rudolph, E.D. (eds). *Cellular Interactions in Symbiosis and*  
628 *Parasitism*. Ohio State University Press. pp. 197–227.
- 629 6. Palmqvist, K., Dahlman, L., Jonsson, A., Nash, T. H. III. 2008. The carbon economy of  
630 lichens. In: Nash, T.H. III, *Lichen Biology*. 2nd edition. Cambridge University Press. pp.  
631 182–215.
- 632 7. Scott, G.D. 1973. Evolutionary aspects of symbiosis. In: Ahmadjian, V. & Hale, M.E.  
633 *The Lichens*. Academic Press, Inc., New York and London. pp. 581–598.
- 634 8. Wipf, D., Krajinski, F., van Tuinen, D., Recorbet, G., Courty, P.-E. 2019. Trading on the  
635 arbuscular mycorrhiza market: from arbuscules to common mycorrhizal networks. *New*  
636 *Phytologist* 223: 1127–1142.
- 637 9. Miyauchi, S., Kiss, E., Kuo, A. *et al.* 2020. Large-scale genome sequencing of  
638 mycorrhizal fungi provides insights into the early evolution of symbiotic traits. *Nature*  
639 *Communications* 11, 5125.
- 640 10. Ellers, J., Kiers, E.T., Currie, C.R., McDonald, B.R., Visser, B. 2012. Ecological  
641 interactions drive evolutionary loss of traits. *Ecology Letters* 15: 1071–1082.
- 642 11. Lutzoni, F., Pagel, M., Reeb, V. 2001. Major fungal lineages are derived from lichen  
643 symbiotic ancestors. *Nature* 411: 937–940.

- 644 12. Lumbsch, H.T., Schmitt, I., Döring, H., Wedin, M. 2001. ITS sequence data suggest  
645 variability of ascus types and support ontogenetic characters as phylogenetic  
646 discriminators in the *Agyriales* (Ascomycota). *Mycological Research* 105: 265–274.
- 647 13. Lumbsch, H.T., Schmitt, I., Lücking, R., Wiklund, E., Wedin, M. 2007. The  
648 phylogenetic placement of Ostropales within Lecanoromycetes (Ascomycota) revisited.  
649 *Mycological Research* 111: 257–267.
- 650 14. Thiyagaraja, V., Lücking, R., Ertz, D., Karunarathna, S.C., Wanasinghe, D.N.,  
651 Lumyong, S., Hyde, K.D. 2021. The evolution of life modes in Stictidaceae, with three  
652 novel taxa. *Journal of Fungi* 7(2): 105.
- 653 15. Wedin, M., Döring, H., Gilenstam, G. 2004. Saprotrophy and lichenization as options  
654 for the same fungal species on different substrata: environmental plasticity and fungal  
655 lifestyles in the *Stictis–Conotrema* complex. *New Phytologist* 164: 459–465.
- 656 16. Muggia, L., Baloch, E., Stabentheiner, E., Grube, M., Wedin, M. 2011. Photobiont  
657 association and genetic diversity of the optionally lichenized fungus *Schizoxylon*  
658 *albescens*. *FEMS Microbiology Ecology* 75: 255–272.
- 659 17. Brodo, I.M. 1973. Substrate Ecology. In Ahmadjian & Hale, *The Lichens*. Academic  
660 Press, Inc. pp. 401–441.
- 661 18. Spribille, T., Thor, G., Bunnell, F.L., Goward, T., Björk, C.R. 2008. Lichens on dead  
662 wood: species-substrate relationships in the epiphytic lichen floras of the Pacific  
663 Northwest and Fennoscandia. *Ecography* 31: 741–750.
- 664 19. Resl, P., Fernández-Mendoza, F., Mayrhofer, H., Spribille, T. 2018. The evolution of  
665 fungal substrate specificity in a widespread group of crustose lichens. *Proceedings of the*  
666 *Royal Academy of Sciences B* 285: 20180640.
- 667 20. Fahselt, D. 1994. Carbon metabolism in lichens. *Symbiosis* 17: 127–182.

- 668 21. Beckett, R.P., Zavarzina, A.G., Liers, C. 2013. Oxidoreductases and cellulases in  
669 lichens: Possible roles in lichen biology and soil organic matter turnover. *Fungal Biology*  
670 117: 431–443.
- 671 22. Palmqvist, K. 2000. Carbon economy in lichens. *New Phytologist* 148: 11–36.
- 672 23. Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B.  
673 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for  
674 glycogenomics. *Nucleic Acids Research* 37(Database issue): D233–D238.
- 675 24. Kohler, A., Kuo, A., Nagy, L. *et al.* 2015. Convergent losses of decay mechanisms and  
676 rapid turnover of symbiosis genes in mycorrhizal mutualists. *Nature Genetics* 47: 410–  
677 415.
- 678 25. Lutzoni, F., Nowak, M.D., Alfaro, M.E. *et al.* 2018. Contemporaneous radiations of  
679 fungi and plants linked to symbiosis. *Nature Communications* 9, 5451.
- 680 26. Li, Y., J. L. Steenwyk, Y. Chang, Y. Wang, T. Y. James, J. E. Stajich, J. W. Spatafora,  
681 M. Groenewald, C. W. Dunn, C. T. Hittinger, X.-X. Shen, & A. Rokas 2021. A genome-  
682 scale phylogeny of the kingdom Fungi. *Current Biology* 31: 1653–1665.
- 683 27. Miadlikowska, J., Kauff, F., Högnabba, F. *et al.* 2014. A multigene phylogenetic  
684 synthesis for the class Lecanoromycetes (Ascomycota): 1307 fungi representing 1139  
685 infrageneric taxa, 317 genera and 66 families. *Molecular Phylogenetics and Evolution*  
686 79: 132–168.
- 687 28. Nelsen, M.P., Lücking, R., Boyce, C.K., Lumbsch, H.T., Ree, R.H. 2020. The  
688 macroevolutionary dynamics of symbiotic and phenotypic diversification in lichens.  
689 *Proceedings of the National Academy of Sciences* 117: 21495–21503.
- 690 29. Nakamura, M.A., Nascimento, A.S. & Polikarpov I. 2017. Structural diversity of  
691 carbohydrate esterases. *Biotechnology Research and Innovation* 1(1):35–51.

- 692 30. Plett, J.M., Martin, F.M. 2018. Know your enemy, embrace your friend: using omics to  
693 understand how plants respond differently to pathogenic and mutualistic  
694 microorganisms. *The Plant Journal* 93: 729–746.
- 695 31. Yagüe, E., Estévez, M.P. 1990. Cellobiose induces extracellular  $\beta$ -1,4-glucanase and  $\beta$ -  
696 glucosidase in the epiphytic lichen *Evernia prunastri*. *Plant Physiology and Biochemistry*  
697 28: 203–207.
- 698 32. Yagüe, E., Estévez, M.P. 1989.  $\beta$ -1,4-glucanase (cellulase) location in the symbionts of  
699 *Evernia prunastri*. *Lichenologist* 21: 147–151.
- 700 33. Honegger, R. 2012. The symbiotic phenotype of lichen-forming ascomycetes and their  
701 endo- and epibionts. In Hock, B. (ed.) *Fungal Associations, 2nd Edition*. The Mycota IX:  
702 287–338. Springer-Verlag.
- 703 34. Elenkin, A. 1902. K voprosu o „vnutrennem“ saprofitizme” („endosaprofitizm”) u  
704 lishaynikov’ [On the question of “inner saprotrophy” (“endosaprophytism”) in lichens].  
705 *Bulletin du Jardin Impérial Botanique de St.-Pétersbourg* 2: 65–84.
- 706 35. Kono, M., Kon, Y., Ohmura, Y., Satta, Y., Terai, Y. 2020. In vitro resynthesis of  
707 lichenization reveals the genetic background of symbiosis-specific fungal-algal  
708 interaction in *Usnea hakonensis*. *BMC Genomics* 21: 671
- 709 36. Domozych, D.S., Ciancia, M., Fangel, J.U., Mikkelsen, M.D., Ulvskov, P., Willats,  
710 W.G.T. 2012. The cell walls of green algae: a journey through evolution and diversity.  
711 *Frontiers in Plant Science* 3, Article 82: 1–7.r
- 712 37. Yagüe E, Orus M.I., Estévez, M.P. 1984 . Extracellular polysaccharidases synthesized  
713 by the epiphytic lichen *Evernia prunastri* (L.) Ach. *Planta* 160: 212–216.
- 714 38. Guerrero, J., Yagüe, E., Estévez, M.P. 1992. Cellulase production by lichens. *Journal of*  
715 *Plant Physiology* 140: 508–510.

- 716 39. de los Ríos, A., Ramírez, R., Estévez, P. 1997. Production of several isoforms of  $\beta$ -1,4-  
717 glucanase by the cyanolichen *Peltigera canina*. *Physiologia Plantarum* 100: 159–164.
- 718 40. Tolpysheva, T.Y. 2007. Dependence of the cellulolytic activity in the lichen  
719 *Hypogymnia physodes* (L.) Nyl. on the substrate. *Moscow University Biological Sciences*  
720 *Bulletin* 62: 65–68.
- 721 41. Beckett, R.P., Ntombela, N., Scott, E., Gurjanov, O.P., Minibayeva, F.V., Liers, C. 2015.  
722 Role of laccases and peroxidases in saprotrophic activities in the lichen *Usnea undulata*.  
723 *Fungal Ecology* 14: 71–78.
- 724 42. Liers, C., Ullrich, R., Hofrichter, M., Minibayeva, F.V., Beckett, R.P. 2011. A heme  
725 peroxidase of the ascomyceteous lichen *Leptogium saturninum* oxidizes high-redox  
726 potential substrates. *Fungal Genetics and Biology* 48: 1139–1145.
- 727 43. Baloch, E., Lücking, R., Lumbsch, H.T., Wedin, M. 2010. Major clades and  
728 phylogenetic relationships between lichenized and non-lichenized lineages in Ostropales  
729 (Ascomycota: Lecanoromycetes). *Taxon* 59: 1483–1494
- 730 44. Yamamoto, Y., Miura, Y., Higuchi, M., Kinoshita, Y., Yoshimura, I. 1993. Using lichen  
731 tissue cultures in modern biology. *Bryologist* 96: 384–393.
- 732 45. Parrent, J.L., James, T.Y., Vasaitis, R. *et al.* 2009. Friend or foe? Evolutionary history  
733 of glycoside hydrolase family 32 genes encoding for sucrolytic activity in fungi and its  
734 implications for plant-fungal symbioses. *BMC Evol Biol* 9, 148.  
735 <https://doi.org/10.1186/1471-2148-9-148>
- 736 46. Frey, E. 1932. Die Spezifität der Flechtengonidien. Alte und neue Probleme. *Berichte*  
737 *der schweizerischen botanischen Gesellschaft* **41(2)**: 180–198.
- 738 47. Armstrong, R.A., Smith, S.N. 2009. Carbohydrates in the hypothallus and areolae of the  
739 crustose lichen *Rhizocarpon geographicum* (L.) DC.. *Symbiosis* 49, 95.  
740 <https://doi.org/10.1007/s13199-009-0016-z>

- 741 48. Hammer, S. 1996. Prothallus structure in *Cladonia*. *Bryologist* 99: 212–217.
- 742 49. Orus, M.I., Ascaso, C. 1982. Localización de hifas liquénicas en los tejidos conductores  
743 de *Quercus rotundifolia* Lam. *Collectanea Botanica* (Barcelona) 13: 325–338.
- 744 50. Greenhalgh, G. N., & Anglesea, D. 1979. The distribution of algal cells in lichen thalli.  
745 *Lichenologist* 11: 283–292. doi:10.1017/s0024282979000335
- 746 51. Lücking, R., Hodkinson B.P., Leavitt, S.D. 2017. The 2016 classification of lichenized  
747 fungi in the Ascomycota and Basidiomycota – Approaching one thousand genera.  
748 *Bryologist* 119: 361–416. <https://doi.org/10.1639/0007-2745-119.4.361>
- 749 52. Smith, D.C., Douglas, A.E. 1987. *The Biology of Symbiosis*. Edward Arnold, London.  
750 302 pp.
- 751
- 752 53. Smith, D.C. 1978. Is a lichen a good model of biological interactions in nutrient-poor  
753 environments? In: Shilo, M. (ed) *Strategies of Microbial Life in Extreme Environments*.  
754 Dahlem Konferenzen, Life Sciences Research Report 13. Verlag Chemie. pp. 291–303.
- 755 54. Smith, D.C. 1992. The symbiotic condition. *Symbiosis* 14: 3–15.
- 756 55. Honegger, R. 1986. Ultrastructural studies in lichens. I. Haustorial types and their  
757 frequencies in a range of lichens with trebouxioid photobionts. *New Phytologist* 103:  
758 785–795.
- 759 56. Poelt, J. 1989. Die Entstehung einer Strauchflechte aus einem Formenkreis krustiger  
760 Verwandter. *Flora* 183: 65–72.
- 761 57. Spribille, T., Tagirdzhanova, G., Goyette, S., Tuovinen, V., Case, R., Zandberg, W.  
762 2020. 3D biofilms: in search of the polysaccharides holding together lichen symbioses.  
763 *FEMS Microbiology Letters* DOI: 10.1093/femsle/fnaa023.
- 764 58. Hale, M.E. 1973. Fine structure of the cortex in the lichen family Parmeliaceae viewed  
765 with the scanning electron microscope. *Smithsonian Contr Bot.* 10: 1–92.



- 766 59. Honegger, R. 1991. Functional aspects of the lichen symbiosis. *Ann. Rev. Plant Physiol.*  
767 *Plant Mol. Biol.* 42: 553–578.
- 768 60. Tapper, R. 1981. Glucose uptake by *Trebouxia* and associated fungal symbiont in the  
769 lichen symbiosis. *FEMS Microbiology Letters* 10: 103–106.
- 770 61. Campbell, J., Bengtson, J., Fredeen, A.L., Coxson, D.S., Prescott, C.E. 2013. Does  
771 exogenous carbon extend the realized niche of canopy lichens? Evidence from sub-boreal  
772 forests in British Columbia. *Ecology* 94: 1186–1195.
- 773 62. Jahn, L., Schafhauser, T., Pan, S., Weber, T., Wohlleben, W., Fewer, D., Sivonen, K.,  
774 Flor, L. van Pée, K.–H., Caradec, T., Jacques, P., Huijbers, M.M.E., van Berkel, W.J.H.,  
775 Ludwig-Müller, J. 2017. *Cyanodermella asteris* sp. nov. from the inflorescence of *Aster*  
776 *tataricus*. *Mycotaxon* 132: 107–123
- 777 63. Tagirdzhanova, G., Saary, P., Tingley, J., Díaz-Escandón, D., Abbott, W., Finn, R.,  
778 Spribille, T. 2021. Predicted input of uncultured fungal symbionts to a lichen symbiosis  
779 from metagenome-assembled genomes. *Genome Biology and Evolution* 13: evab047.

780

781

782

783

784

785

786

787 **Figure captions**

788

789 **Figure 1.** Distribution and ancestral states of CAZymes and selected sugar transporters across  
790 the evolution of *Lecanoromycetes* and related classes of *Ascomycota* projected onto a

791 maximum likelihood phylogenomic tree based on 1310 loci. Symbols beside tree tips refer to  
792 life history traits and phototrophic partners of LFS under study. Heatmaps with shades of red  
793 indicate the number of genes in different CAZyme classes or involved in degrading complex  
794 PCW components. Columns from left to right: AA - Auxiliary Activities, CBM - Carbohydrate  
795 binding module, CE - Carbohydrate Esterases, GH - Glycoside Hydrolases, GT - Glycosyl  
796 Transferases, PL - Polysaccharide Lyases. cell - Number of genes in 35 CAZyme families  
797 involved in cellulose and hemicellulose breakdown. pec - Number of genes in 11 CAZyme  
798 families involved in pectin breakdown. lign - Number of genes in 3 CAZyme families and class  
799 II peroxidases involved in lignin breakdown. Selection of CAZyme sets follows (9, 24). celldex  
800 - Number of cellodextrin transporters. malt - number of maltose transporters. total - Total  
801 number of sugar transporters. Below the heatmap are ancestral sizes of CAZyme families  
802 involved in (hemi-)cellulose, pectin and lignin degradation. Colored circles on tree branches  
803 indicate significantly expanded CAZyme families. The size of the circles indicates the number  
804 of individual CAFE runs in which a family was found to be significantly expanded.

805

806

807 **Figure 2.** Similarity of CAZyme sets involved in the breakdown of different complex plant-  
808 based polysaccharides based on phylogenetically corrected Principal Components Analysis.  
809 Different colors indicate taxonomic groups. A - Similarity of CAZyme families involved in  
810 cellulose and hemicellulose breakdown. B - Similarity of CAZyme families involved in lignin  
811 breakdown, C - Similarity of CAZyme families involved in pectin breakdown. Displayed  
812 below are representative members of the two subclasses *Ostropomycetidae* (with light green  
813 border) and *Lecanoromycetidae* (dark green border). D - *Graphis scripta*, E - *Icmadophila*  
814 *ericetorum*, F - *Mycoblastus sanguinarius*, G - *Peltigera leucophlebia*, H - *Xylographa*  
815 *carneopallida*, I - *Agyrium rufum*, J - *Evernia prunastri*, K - *Cladonia macilenta*. Image credits:

816 *Agyrium rufum*: Paul Cannon (fungi.myspecies.info); Creative Commons: BY-NC 4.0.  
817 *Peltigera leucophlebia*: Jason Hollinger, uploaded by Amada44, CC-BY 2.0,  
818 <https://commons.wikimedia.org/w/index.php?curid=24213606>. *Evernia prunastri*: by Jason  
819 Hollinger, CC-BY 2.0, <https://commons.wikimedia.org/w/index.php?curid=50595319>.  
820 *Cladonia macilenta*: Bruce McCune & Sunia Yang - Lichen, CC-BY 4.0-NC,  
821 [https://lichens.twinferntech.net/pnw/species/Cladonia\\_macilenta.shtml](https://lichens.twinferntech.net/pnw/species/Cladonia_macilenta.shtml); other images by the  
822 authors.

823

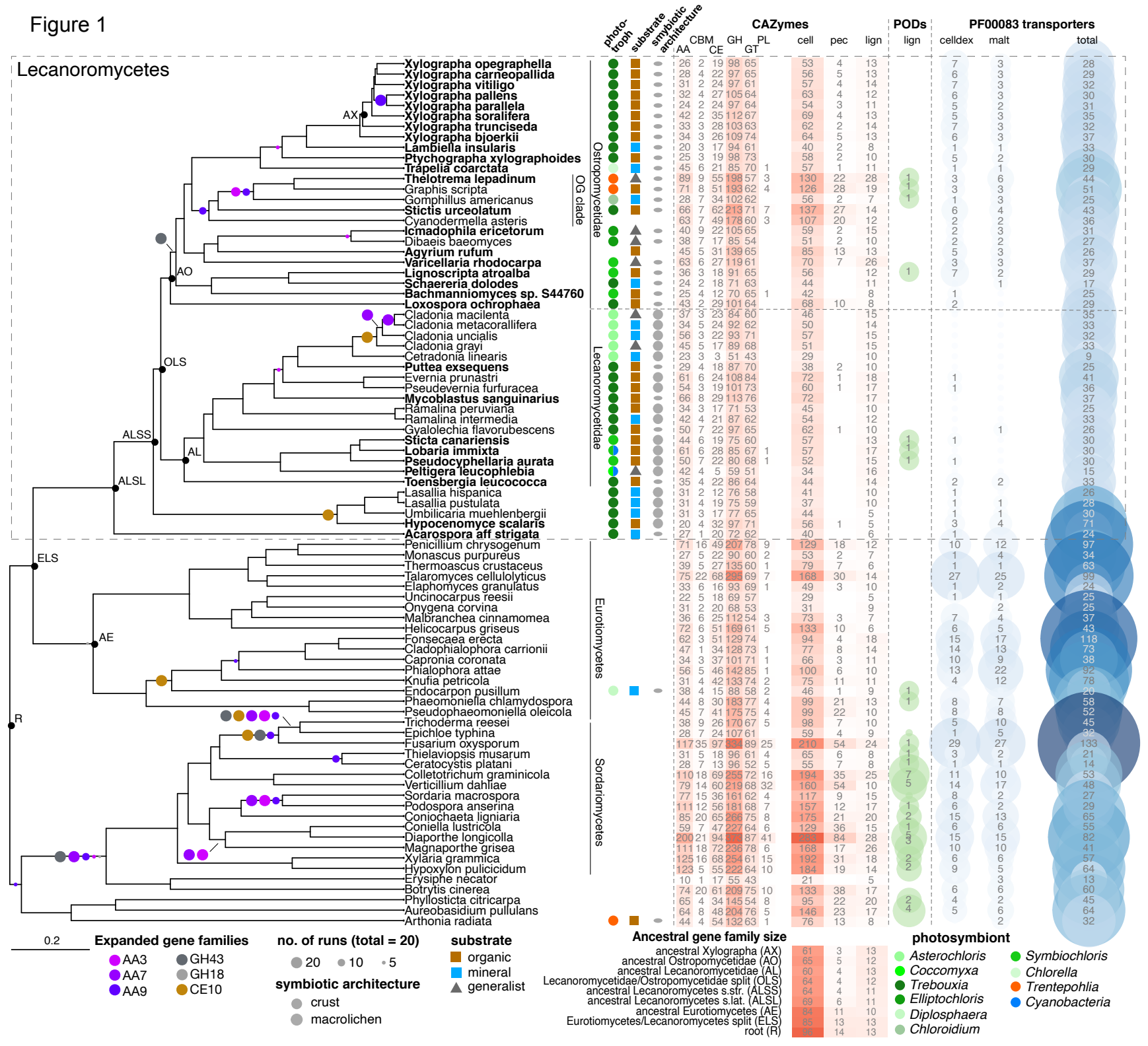
824

825 **Figure 3.** Gene trees of two CAZyme families involved in cellulose (GH5) and hemicellulose  
826 (GH43) breakdown. Each tree includes all experimentally characterized sequences combined  
827 with all sequences from the 83 genomes studied here. CAZyme Subfamilies are labeled with  
828 numbers and grey rectangles. Three columns along tree tips display additional information of  
829 corresponding sequences when available. EC - Enzyme Code of experimentally characterized  
830 sequences downloaded from [cazy.org](http://cazy.org). Sub Loc - Predicted subcellular location of Enzyme with  
831 DeepLoc. Tax - Taxonomic assignment of organisms from which the sequence comes from.  
832 For larger subfamilies, functions of characterized sequences based in Enzyme Code numbers  
833 are given as colored squares. Sequences used for heterologous expression experiments are  
834 marked with red arrows in GH5 subfamily 5.

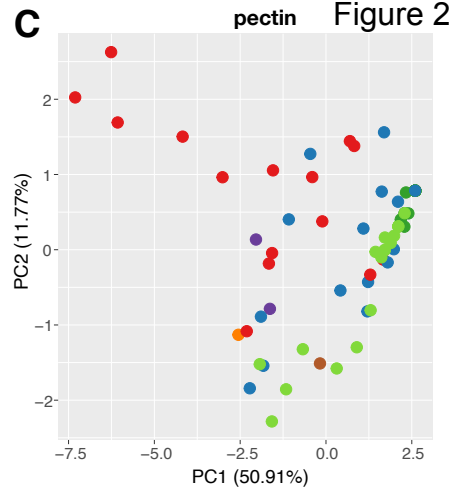
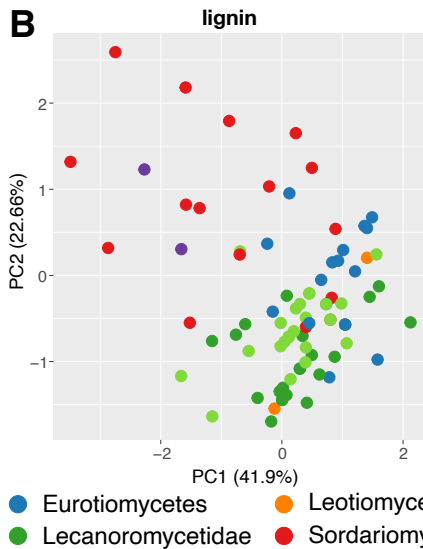
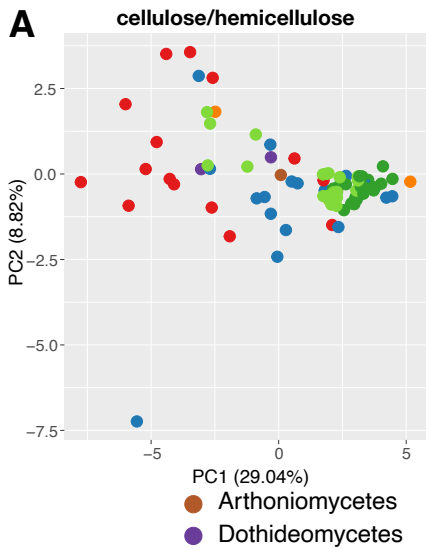
835

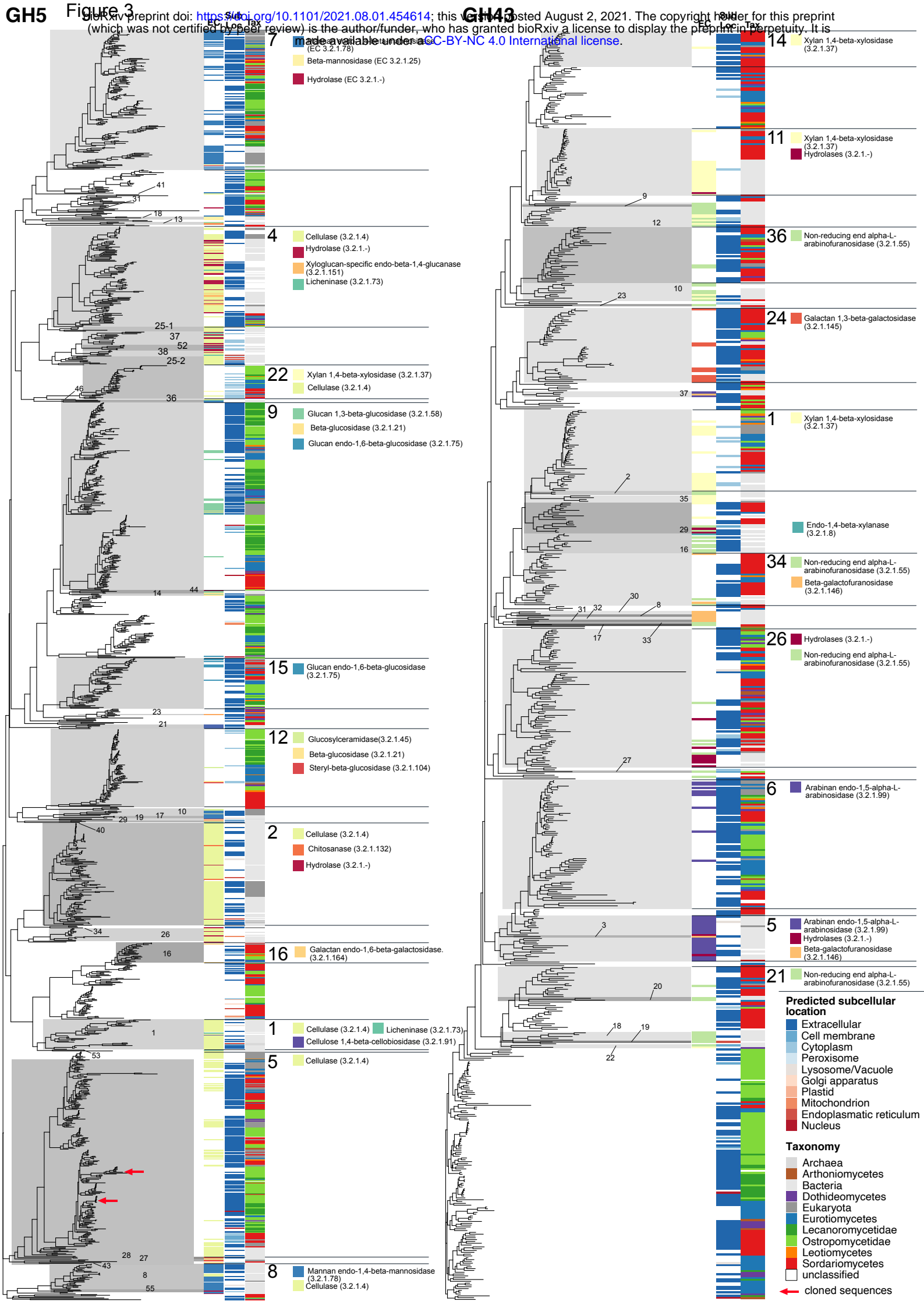
836 **Figure 4.** Enzymatic activity of two putative cellulases A and B from *Xylographa bjoerkii* and  
837 orthologues of these genes in 83 ascomycete genomes. A - Activity of cellulase A and B at  
838 different pH conditions. B - Activity of cellulase A and B at different temperatures. C - Activity  
839 of cellulase A and B on different substrates. D - Number of orthologs of cellulase A and B in  
840 different genomes.

Figure 1









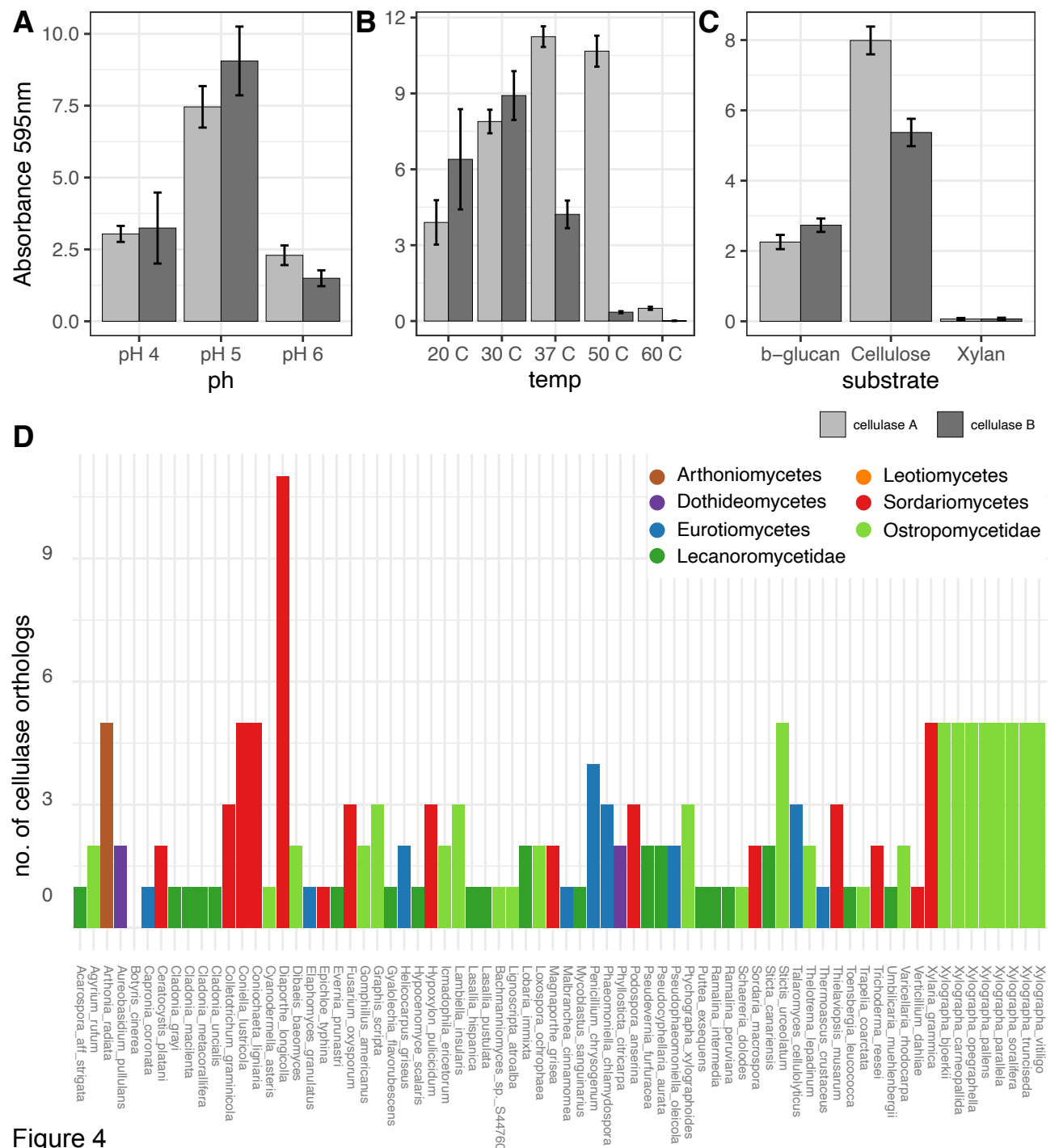


Figure 4