Genetic factors driving multi-host infection in a core member of the root mycobiota

Ram-Sevak Raja-Kumar[1,*], Fantin Mesny[1,*], Arpan Kumar Basak[1,2,*], Jacy Newfeld[3], Guillaume Chesneau[1], Frederickson Entila[1], Tak Lee[1], Linda Rigerte[1], Stephanie Carvajal Acevedo[1], Bruno Hüttel[1,4], Pedro W. Crous[5,6,7], Jose G. Maciá-Vicente[8], Helen Stewart[9], Matthew Ryan[9], Ahmad M. Fakhoury[10], Soledad Sacristán[11,12], Isabelle Batisson[13], Stefano Dumontet[14], Wade H. Elmer[15], Jana Henzelyová[16], Joanna S. Kruszewska[17], Jessica M. Nelson[18], Cara M. Santelli[19], Markus Pauly[20], Antonio Molina[11,12], Kei Hiruma[3], Stéphane Hacquard[1,21#]

[1]Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany. [2]Malopolska Centre of Biotechnology, Jagiellonian University, Krakow, 30-387 Poland. [3]Department of Life Sciences, Multidisciplinary Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1, Komaba, Meguro-ku, Tokyo 153-8902, Japan. [4]Max Planck Genome Centre Cologne, Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany. [5]Westerdijk Fungal Biodiversity Institute, P.O. Box 85167, 3508 AD Utrecht, The Netherlands. [6]Microbiology, Department of Biology, Faculty of Science, Utrecht University, Padualaan 8, 3584 CT Utrecht, The Netherlands. [7]Department of Genetics, Biochemistry and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, 0002, South Africa. [8]Marine Sciences and Applied Biology, University of Alicante, PO Box 99, 03083 Alicante, Spain. [9]CABI, Imperial site at Silwood Park, Buckhurst Rd, Ascot SL5 7PY, United Kingdom. [10]School of Agricultural Sciences, Southern Illinois University, Carbondale, IL 62901, USA. [11]Centro de Biotecnología y Genómica de Plantas, Universidad Politécnica de Madrid (UPM)-Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA/CSIC), Campus Montegancedo-UPM, 28223-Pozuelo de Alarcón (Madrid), Spain. [12]Departamento de Biotecnología-Biología Vegetal, Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, 28040-Madrid, Spain. [13]Université Clermont-Auvergne, CNRS, Laboratoire Microorganismes: Génome et Environnement, F-63000 Clermont-Ferrand, France. [14]Department of Sciences and Technologies, University of Napoli 'Parthenope', Centro Direzionale, Isola C4, 80143 Napoli, Italy. [15]Department of Plant Pathology and Ecology, The Connecticut Agricultural Experiment Station, New Haven, Connecticut 06504, USA. [16]Department of Genetics, Institute of Biology and Ecology, Faculty of Science, Pavol Jozef Šafárik University in Košice, Mánesova 23, 04154, Košice, Slovakia. [17]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland. [18]Natural History Museum, University of Oslo, 0562 Oslo, Norway.

[19]Department of Earth and Environmental Sciences, University of Minnesota Twin Cities, Minneapolis, MN, USA. [20]Institute of Plant Cell Biology and Biotechnology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. [21]Cluster of Excellence on Plant Sciences (CEPLAS), Max Planck Institute for Plant Breeding Research, 50829 Cologne, Germany.

*First authors

[#]Corresponding author: S. Hacquard (hacquard@mpipz.mpg.de).

Abstract

Core members of the fungal root microbiota include pathogens capable of colonizing multiple hosts, yet the underlying genetic determinants remain unknown. We report that *Plectosphaerella cucumerina* is a core member of the *Arabidopsis thaliana* root microbiota displaying high pathogenic potential and multi-host colonization capabilities. Establishment of a *Plectosphaerella* reference culture collection, followed by whole-genome sequencing of 72 strains reveals subtle phenotypic and genotypic variation that associate with fungal phylogeny, but not host plant identity. Transcriptome profiling of a model *P. cucumerina* isolate in roots of multiple hosts identifies core and host-specific fungal processes linked to carbon catabolism and root cell wall deconstruction of the hosts. A fungal gene encoding a candidate β-1,3-glucanase (GH64) was identified as a key genetic factor driving infection and disease in plants that diverged 110 million years ago. The gene is enriched in plant-colonizing fungi and consistently functions as a disease determinant in the root pathogen *Colletotrichum incanum*. We conclude that diverse and tunable fungal repertoires of carbohydrate-active enzymes act as disease determinants and drive multi-host compatibility belowground.

Introduction

The root mycobiota – the community of fungi that associate with the roots of plants – include members that establish beneficial relationships but also isolates that conditionally cause diseases or hamper plant performance[1, 2, 3, 4, 5, 6]. Consistently, several core fungal taxa that reproducibly colonize roots of healthy plants in natural populations are often harmful under laboratory conditions[7, 8, 9], illustrating that the pathogenic lifestyle of several root mycobiota members is often kept at bay in nature. In *Arabidopsis thaliana*, it has been reported that the protective activity of bacterial root commensals is at least as important as the innate immune branch involving tryptophan-derived specialized metabolites for controlling fungi in roots[4, 9, 10, 11], thereby maintaining microbial interkingdom balance for plant health[12].

Given that soil and root-associated fungal communities are extensively shaped by the local climatic, edaphic and host conditions[13, 14, 15], high endemism in fungal communities has been repeatedly reported[16, 17], with very few core fungal species detected worldwide[18].

Widespread fungal root colonizers appear particularly successful due to their ability to efficiently disperse, infect multiple hosts, and plastically adapt to various environmental conditions[19]. However, the underlying molecular mechanisms enabling adaptation to a broad range of hosts and environments in root mycobiota members remains unknown.

Plant-associated fungi produce remarkable cocktails of carbohydrate-active enzymes (CAZymes) that target plant and/or microbial cell walls and outer layers. These enzymes facilitate fungal entry into plant tissues, promote carbohydrate extraction from host cell walls or modulate immune responses[8, 20, 21, 22, 23], but can also contribute to antagonistic microbe-microbe interactions[24, 25]. The repertoire of fungal CAZymes often explains fungal lifestyles or correlates with adaptation to specific hosts or organs[26, 27, 28, 29, 30, 31]. For example, transitions from saprotrophic fungi to ectomycorrhizal symbionts have occurred multiple times independently, involving significant losses of CAZymes involved in lignin and cellulose deconstruction[32, 33]. In contrast, endophytism in Arabidopsis root mycobiota members does not correlate with a reduction of saprotrophic traits[3, 8, 11, 34] but involves pectin-degrading enzymes that modulate fungal root colonization aggressiveness[8]. However, the extent to which the diversity, composition, and transcriptional activation of CAZyme repertoires contribute to fungal adaptation to multiple hosts and environments in root mycobiomes remains unclear.

We report that the fungus *Plectosphaerella cucumerina*[35] – a well-known pathogen of diverse plant species causing wilt disease, root rot disease or leaf necrosis[36, 37, 38, 39] – is also a prevalent species colonizing the roots of *A. thaliana* at a continental scale. The fungus (phylum Ascomycota, class Sordariomycetes, order Glomerellales, family Plectosphaerellaceae) exhibits broad host range colonization capabilities and acts as a root pathogen of the dicotyledonous plants *Solanum lycopersicum* (tomato) and *A. thaliana* but not of the monocotyledonous host *Hordeum vulgare* (barley). We assembled a large collection of 69 *P. cucumerina* isolates that were retrieved from various host plants worldwide as well as high quality – long read-based – genome assemblies. Multi-host recolonization experiments using the model *P. cucumerina* 0016 strain, followed by fungal transcriptome sequencing revealed 1) high plasticity in catabolic process activation that mirrors root cell wall composition of the cognate hosts, 2) the key contribution of a *P. cucumerina* β-1,3-glucanase (glycoside hydrolase familly 64, GH64) for driving multi-host infection and disease in host plants that diverged 110 million years ago (MyA) and 3) the evolutionary conservation of GH64 as a root infection determinant in the root pathogen *Colletotrichum incanum* (order Glomerellales, family Glomerellaceae). Our results indicate that adaptation to multiple hosts across various environments is conferred by a combination of genome-rich CAZyme repertoires and the plastic up-regulation of these genes in response to both conserved and host-specific cell wall cues.

Results

***P. cucumerina* is a core member of the *A. thaliana* root mycobiota and has a broad host range**. We hypothesized that specific root mycobiota members can reproducibly colonize roots of healthy Arabidopsis in nature, irrespective of differences in climate and soil conditions. To test for the presence of a core root mycobiota, we first used amplicon sequencing data from two previously published studies[9, 14] (18 sites, n = 291 root samples in total) and re-analyzed them at amplicon sequence variant (ASV)-level resolution (i.e., Zero-radius operational taxonomic unit[40], Zotu, see methods). From the 26,976,402 sequenced ITS reads, we identified 338 Zotus that were detected in *A. thaliana* roots in at least one of the 18 sites across Europe [mean relative abundance (RA) > 0.1%, **Fig. 1a** and **Supplementary Table 1**]. Notably we identified 5 Zotus that were reproducibly detected in more than 80% of the 291 root samples, thereby constituting a small – yet multispecies – core Arabidopsis root mycobiota (**Fig. 1a**). These corresponded to Zotu2 (*Fusarium* sp.), Zotu9 (*Plectosphaerella* sp.), Zotu1 (*Dactylonectria* sp.), Zotu6 (*Fusarium* sp.) and Zotu3 (*Fusarium* sp.) that have occurrence values of 1, 0.98, 0.92, 0.87 and 0.87, respectively (**Fig. 1a** and **Supplementary Table 1**). Cross referencing with 488 ITS sequences corresponding to cultured isolates established from four of these sites (see methods, **Supplementary Table 1**) confirmed that multiple (>10) representative isolates had 100% sequence match to these Zotus (**Supplementary Fig. 1a**). Particularly, 106 ITS sequences from cultured isolates matched Zotu9 and were unambiguously assigned to *Plectosphaerella cucumerina* (Phylum: Ascomycota; class: Sordariomycetes, family: Plectosphaerellaceae; based on the RDP classifier[41], **Supplementary Table 2, Supplementary Fig. 1b**), a well-known pathogen of several plant species, including Arabidopsis[37, 42]. We then used *P. cucumerina* as an input species in GlobalFungi[43] and detected this species across a variety of niches (**Fig. 1b**) and particularly in soil, mosses, topsoil, rhizosphere and root/rhizosphere samples (pairwise one-sided Fisher's exact tests, Benjamini-Hochberg correction, **Fig. 1c**), thereby confirming its presence at the soil-root interface. Root and/or rhizosphere samples in which *P. cucumerina* was detected were distributed worldwide across latitudes/longitudes and ecosystems (**Fig. 1d, Supplementary Fig. 2a,b**) and inspection of the corresponding plants revealed broad associations with a remarkable diversity of hosts, including monocots and dicots (**Fig. 1e**). The results indicate that *P. cucumerina* is a widespread root colonizer that has a broad range of hosts.

Genetic and phenotypic differentiation among *P. cucumerina* isolates is explained by phylogeny, but not by the host of isolation. We hypothesized that high genomic and phenotypic variation explains broad fungal distribution and host range. We assembled a collection of 72 *Plectosphaerella* strains that were isolated from diverse plant families

148 (including bryophytes[44]) and environments worldwide. Among these 69 belong to the *P.*
149 *cucumerina* species complex, and three belong to the closely related species *P. plurivora* (n =
150 2) and *P. delsorboi* (n = 1) (**Supplementary Table 3**, **Fig. 2a**). PacBio long-read genome
151 sequencing resulted in high-quality assemblies (sizes: 35.5 − 40.5 Mbp, median = 36.4 Mbp;
152 number of contigs: 9 – 188, median = 27.5; N50: 0.77 – 7.66 Mbp, median = 2.69 Mbp) that
153 contained similar sets of predicted genes (9,992 – 12,956 genes, median: 11,581.5)
154 (**Supplementary Fig. 3a-d**). Phylogeny reconstruction using 5,466 single copy orthologs
155 identified two *P. cucumerina* subspecies (subsp. 1: 39 strains, subsp. 2: 26 strains) and a *P.*
156 *cucumerina* outgroup consisting of four strains (**Fig. 2a**, **Supplementary Fig. 4a-d**). A burst in
157 transposon and retrotransposon elements was observed in subsp. 2 (**Supplementary Fig. 5a-**
158 **e**) but did not associate with expansion of CAZyme, protease or effector gene repertoires
159 (**Supplementary Fig. 6**). Comparative genomics across the 69 *P. cucumerina* isolates using
160 15,269 orthogroups identified a large core genome (> 95% of genes detected in 80-100% of
161 the strains) and a modest pan genome (< 5% of the genes detected in less than 80% of the
162 strains) (**Supplementary Fig. 7a,b**). Pan-genome encoded genes were enriched in candidate
163 effectors (**Supplementary Fig. 7c**) and more closely associated with transposons compared
164 to core genome-encoded genes (**Supplementary Fig. 7d**). Between-strain differentiation in
165 CAZyme, protease and effector repertoires was extensively explained by the *Plectosphaerella*
166 phylogeny (54%, 36%, and 45%, respectively, PERMANOVA, $P < 0.001$, **Fig. 2c**) but not by
167 the phylogeny of the host of isolation ($P > 0.05$) (**Fig. 2d**), thereby corroborating that the origin
168 of the host plant is not a strong driver of repertoire variation in *P. cucumerina.* Similar results
169 were observed for orthogroups (**Supplementary Fig. 8a,b**). We next tested whether these
170 genetic variations associate with phenotypic differentiation for two major traits, namely 1) effect
171 on *A. thaliana* development 28 days post inoculation in a gnotobiotic agar-based plate system
172 (**Fig. 2b** and **Supplementary Fig. 9**) and 2) carbon resource consumption using PM1-2
173 Biolog® Plates that contained 190 carbon sources (**Supplementary Fig. 10a**). Although these
174 strains were isolated from markedly different environments worldwide, they displayed little
175 phenotypic variation for these traits. Most *P. cucumerina* isolates were detrimental on
176 Arabidopsis, irrespective of their hosts of isolation (ANOVA $P = 2e\text{-}16$, Dunnett's post-hoc test
177 vs. mock, $P < 0.05$ for 44/69 *P. cucumerina* strains), contrasting with *P. plurivora* and *P.*
178 *delsorboi* isolates that were non-pathogenic under these conditions (**Fig. 2b** and
179 **Supplementary Fig. 9**). These isolates also showed remarkably similar growth phenotypes
180 when inoculated on the 190 metabolites used as sole carbon sources (coefficient of variation
181 < 6.5%, **Supplementary Fig. 10a,b**), and are well adapted to efficiently use plant cell wall-
182 derived monosaccharides as growth substrates (**Supplementary Fig. 10a)**. The data suggest
183 modest genetic variation across *P. cucumerina* isolates that is unlikely the sole factor driving
184 adaptation to a broad range of environmental and host conditions worldwide.

185

186 ***P. cucumerina* 0016 colonizes Arabidopsis, tomato and barley roots but only causes**
187 **disease in dicot hosts**. We next asked whether a single strain initially isolated from *A. thaliana*
188 roots (P16, hereafter named *P. cucumerina* 0016, subspecies 1) is able to colonize/cause
189 disease in tomato (*Solanum lycopersicum cv.* Micro-Tom) and barley (*Hordeum vulgare*) which
190 diverged from *A. thaliana* around 110, and 140 MyA, respectively[45, 46]. Re-colonization
191 experiments in a sterile vermiculite matrix using *P. cucumerina* 0016 and germ-free
192 Arabidopsis, tomato, and barley seedlings (see methods) revealed contrasting effects on plant
193 health and survival 21 days post fungal inoculation (i.e., Plant Performance Index: PPI =
194 survival rate x shoot fresh weight, **Fig. 3a**). Whilst the fungus was highly detrimental on both
195 tested dicotyledonous plants (19.5% and 10.7% reduction in PPI for Arabidopsis and tomato,
196 respectively) (two-sided t-test or Wilcoxon rank-sum test, *adj. P* < 0.05), it remained harmless
197 when inoculated on barley roots (*adj. P* > 0.05, **Fig. 3a**). Inspection of fungal load in roots of
198 these plants using quantitative PCR (qPCR) indicated that the fungus was able to colonize all
199 roots, including those of barley (two-sided t-test or Wilcoxon rank-sum test, *adj. P* < 0.05, **Fig.**
200 **3b**). We next stained infected roots using calcofluor white and WGA CF®633A for visualizing
201 the spatial distribution of fungal hyphae during infection. Consistent with our amplicon
202 sequencing data (**Fig. 1a**), numerous *P. cucumerina* 0016 extraradical hyphae were observed
203 throughout the entire root system of Arabidopsis, from the mature peridermal region to the root
204 tip, including lateral roots (**Fig. 3c**). Penetration of undifferentiated hyphae was repetitively
205 observed into or between peridermal (**Fig. 3c-i, iii, iv**) or cortical cell layers (**Fig. 3c-ii, v**),
206 thereby validating extensive endophytic propagation (**Fig. 3c**). In tomato roots, epidermal walls
207 were coated with extracellular hyphae and endophytic colonization was primarily confined to
208 the first epidermal cell layer (**Fig. 3d**). In contrast, despite extensive fungal proliferation at the
209 vicinity of barley root hairs, we did not observe endophytic proliferation in barley roots,
210 suggesting that the fungus is unable to invade and/or is stopped by the barley immune system
211 (**Fig. 3e**). Taken together, our data suggest host-specific differences in fungal endophytic
212 colonization that correlate with distinct outcomes on host performance (**Fig. 3a**).
213

214 ***P. cucumerina* 0016 transcriptional reprogramming *in planta* is modulated by host**
215 **identity.** Given the low intra-species genomic variation observed in *P. cucumerina*, we
216 hypothesized that fungal adaptation to a broad range of hosts is conferred by activation of
217 common gene sets driving "general" root infection, but also specific subsets responding to
218 variation in host physical and chemical cues. Using the same vermiculite system described
219 above, we sequenced the *P. cucumerina* 0016 transcriptome 21 days post inoculation on
220 Arabidopsis, tomato, and barley roots and compared this *in planta* transcriptome to that of a
221 control fungal transcriptome obtained from inoculated polyester membranes (Ctrl, *ex planta*,

222 see Methods). Mapping of Illumina short reads against corresponding reference genomes
223 (**Supplementary Fig. 11a,b**), followed by statistical analysis of differentially expressed genes
224 (DEGs*, in planta* vs Ctrl, DESeq2, Generalized Linear Model) identified 3,399 fungal genes
225 significantly regulated in response to the different hosts (**Supplementary Table 4**).
226 Multidimensional scaling (MDS) of Pearson distances computed on normalized read counts of
227 both expressed genes (**Supplementary Fig. 12**) and DEGs (**Fig. 4a**) identified fungal
228 transcriptome differentiation that is primarily shaped by *in planta* vs *ex planta* conditions (i.e.,
229 see x-axis explaining 60.85% of the variation, **Fig. 4a**), but also to a major extent by the host
230 plant identity (i.e., see y-axis explaining 30.63% of the variation; pairwise statistical analyses:
231 ANOVA-TukeyHSD, *adj. P* < 0.05, **Fig. 4a**). Comparison of all fungal DEGs identified for each
232 of the three comparisons (n = 1,862 induced, n = 1,537 repressed in total, *in planta* vs Ctrl)
233 revealed that 19.3% and 37.6% were consistently induced and repressed, respectively,
234 irrespective of the host plant (ATB sector, **Fig. 4b,c**). Consequently, most fungal DEGs
235 responded either to two plant species (AT, AB and TB sectors, 22.0% induced, 20.7%
236 repressed; **Fig. 4b,c**) or were host-specific (A, T and B sectors: induced: 58.7%; repressed:
237 41.7%; **Fig. 4b,c**), thereby illustrating high transcriptome plasticity in response to different
238 hosts (**Fig. 4d**). Notably host-specific DEGs prevailed during infection on Arabidopsis (A
239 sector) compared to tomato (T sector) or barley (B sector), potentially reflecting higher
240 adaptation to the host of isolation (**Fig. 4c,d**). GO-term enrichment analysis (*adj. P* < 0.05;
241 Benjamini-Hochberg method) for each of the above-mentioned DEG sectors (**Fig. 4c**) revealed
242 remarkable enrichment of terms related to catabolism that were exclusively observed in the
243 up-regulated gene sets (**Fig. 4e** and **Supplementary Fig. 13a,b**). Those included:
244 "Hydrolase", "Proteolysis", "Cellulase", "Catalytic activity", "Pectate lyase" or "Subtilase", that
245 were often identified as enriched in both host-specific (A, T, B) and shared (ATB) fungal DEGs
246 subsets. Taken together, our results revealed that *P. cucumerina* transcriptional
247 reprogramming is differently tailored depending on the colonized host and that catabolism of
248 host-derived products is likely key in driving infection across a broad range of hosts.
249
250 **Activation of CAZyme-encoding genes prevails during infection and mirrors root cell**
251 **wall composition.** We next hypothesized that *P. cucumerina* 0016-mediated cell wall
252 degradation prevails during root colonization and plastically adjust according to host cell wall
253 polysaccharide composition. Consistent with our GO-term enrichment analysis (**Fig. 4e**), we
254 observed that genes encoding carbohydrate active enzymes (CAZymes) were
255 disproportionally upregulated during root infection in all above-mentioned DEG sectors (A:
256 78/715 = 10.9%, T: 19/88 = 21.6%, B: 59/290 = 20.3%, AT: 17/104 = 16.3%, AB: 44/222 =
257 19.8%, TB: 17/84 = 20.2%, ATB: 101/359 = 28.1% of the DEGs, **Supplementary Fig. 14a-c**),
258 thereby contrasting with the expected proportion of CAZyme genes encoded in *P. cucumerina*

259 0016 genome (9.9%). Next, we constructed a force-directed network map (see methods),
260 where circular nodes represented plant cell wall-degrading CAZyme families (n = 83, see
261 methods) connected to their predicted plant substrates (as registered in the dbCAN-sub
262 database[47, 48] (n = 14, see methods). By mapping the mean Log2FC of CAZyme families
263 significantly overexpressed *in planta* (see circle size and grey gradient, Log2FC, Fisher's-exact
264 test, $P < 0.05$) onto the network backbone for each above-mentioned DEG sectors, we
265 predicted host cell wall components (square nodes, orange gradient, Log2FC, Fisher's-exact
266 test, $P < 0.05$) that were targeted by the fungus in response to a specific host (A, T, B, **Fig.**
267 **5a**), two hosts (AT, AB, TB, **Fig. 5a**) or all three hosts (ATB **Fig. 5a,b**). This revealed that
268 diverse CAZymes were predicted to target all major cell wall constituents (pectic
269 polysaccharides, cellulose, hemicellulose, lignin), irrespective of the host plant (ATB sector,
270 **Fig. 5a-c**). These CAZymes belong to polysaccharide lyase (PL) families 1_7, 3_2, 9_3, and
271 42, the carbohydrate binding module family 32, the glycoside hydrolase (GH) family 64, lytic
272 polysaccharide monooxygenases families AA9, AA16, and the carbohydrate esterase family 2
273 (**Fig. 5b**). Cell walls of dicotyledonous plants are more pectin-rich than those of monocots,
274 while monocot walls contain mainly structurally divers xylans and mixed-linkage glucans as
275 their dominant hemicelluloses[49].This is consistent with our targeted analysis of wall materials
276 in roots of Arabidopsis, tomato, and barley as pectic uronic acids were more prevalent in
277 Arabidopsis and tomato, while (arabino)-xylan and matrix polysaccharide based glucose were
278 more prevalent in barley (**Supplementary Fig. 15**). Interestingly, these differences correlated
279 with host-specific activation of fungal genes encoding pectin-degrading enzymes in the dicot
280 hosts (DEG sectors A, T, AT; PL3_2, PL 1_7, PL11_2 families), or (arabino)xylan-degrading
281 enzymes in the barley host (DEG sector B; GH51_2) (**Fig. 5a,c**). Our results suggest that
282 sensing and adequate transcriptional response to cell wall physicochemical cues contributes
283 to plastic infection of diverse hosts in *P. cucumerina*.
284
285 **GH64 CAZyme-encoding gene promotes multi-host infection and disease in *P.*
286 *cucumerina* 0016.** Based on our network map (ATB sector, **Fig. 5b**), we identified a single
287 copy gene encoding a secreted GH64 enzyme (jgi|Plecuc1|282412|) that displayed one of the
288 highest log2FC enrichment *in planta*, irrespective of the host plants (Arabidopsis: Log2FC =
289 7.4, tomato: Log2FC = 8.3 barley: Log2FC = 8.3, **Supplementary Table 4**). Following
290 AlphaFold protein structure prediction, a FoldSeek[50] search identified strong structural
291 similarity with the SpGlu64A enzyme secreted by *Streptomyces pratensis A* (**Fig. 6a,**
292 **Supplementary Fig. 16a**), a β-glucanase exhibiting strict substrate specificity toward β-1,3-
293 glucans[51]. These β-1,3-linked glucose units are constituents of plant cell walls (i.e. callose and
294 grass-specific hemicellulosic mixed-linkage glucans), but mainly represent immunogenic
295 molecules of fungal cell walls[52, 53, 54, 55]. By transfecting *P. cucumerina* 0016 protoplasts with a

296 ribonucleoprotein (RNP) complex consisting of the Cas9 protein and a single guide RNA
297 (sgRNA), we drove precise insertion of a phleomycin resistance DNA cassette into the coding
298 sequence of *GH64* via homologous recombination (method, **Supplementary Fig. 17a-d**).
299 Corresponding mutant spores inoculated on germ-free roots of Arabidopsis and tomato
300 seedlings showed 62%-77% reduction in pathogenicity (Kruskal-Wallis-Dunn Test, adjusted
301 with Bonferroni correction, $P < 0.0001$, **Fig. 6b**) and 18%-24% reduction in root colonization in
302 tomato and Arabidopsis plants (Kruskal-Wallis-Dunn Test, adjusted with Bonferroni correction,
303 $P < 0.0001$, **Fig. 6c**) compared to the WT strain. In contrast, no significant differences were
304 observed in the barley host ($P > 0.05$, **Fig. 6b,c**), nor in artificial growth media (**Fig. 6d**,
305 **Supplementary Fig. 17e,f**), indicating that the gene is a root colonization determinant of the
306 two tested dicot hosts. WT-like disease phenotypes were also observed for another mutant
307 (*Δpl11_2*, **Supplementary Fig. 18**), excluding the possibility that phleomycin resistance is
308 influencing these phenotypes. Notably, GH64 shows high sequence conservation across the
309 72 *Plectosphaerella* isolates (**Supplementary Fig. 16b**), has orthologs in specific fungal
310 lineages belonging to Sordariomycetes and Dothideomycetes and is significantly enriched in
311 genomes of plant saprotrophs, pathogens and endophytes (n = 2,534 fungal genomes, **Fig.
312 6e** and **Supplementary Fig. 16c,d**). Remarkably, inactivation of the corresponding GH64
313 orthologous gene in the root pathogen *Colletotrichum incanum*[4, 34, 56] (class Sordariomycetes,
314 family Glomerellaceae, jgi|Colin1|11312|, **Supplementary Fig. 19a**) resulted in 68.5%
315 reduction of fungal mutant load in *A. thaliana* roots and 10% increase in shoot fresh weight
316 compared to the WT *C. incanum* strain (24 days post infection, **Fig. 6f,g**). However, no
317 significant differences were noted in the Solanaceae host *Nicotiana benthmiana* (i.e., tobacco
318 **Supplementary Fig. 19b-d**). Taken together, our data indicate that *in planta* activation of this
319 gene in evolutionary-distant fungal lineages promotes root infection and disease.
320

321 Discussion

322 Our study reveals the remarkable occurrence of *P. cucumerina* in the roots of diverse host
323 plants, including the model plant species *A. thaliana*. *P. cucumerina* therefore represents one
324 of the very few core members of the Arabidopsis root mycobiota, based on both culture-
325 dependent and independent approaches. Despite this, our analysis revealed that *P.
326 cucumerina* isolates are reproducibly causing disease when inoculated on germ-free *A.
327 thaliana* roots, irrespective of their hosts of isolation. We previously showed that the most
328 prevalent fungal isolates colonizing *A. thaliana* roots in nature are often harmful in mono-
329 association experiments[8] and that these detrimental phenotypes are alleviated by the presence
330 of bacterial commensals that regulate fungal load and composition at the root interface, thereby
331 preventing fungal dysbiosis in nature[9, 10, 12]. It is therefore conceivable that *P. cucumerina*-
332 induced disease in nature relies on both the genetic determinants identified here for multi-host

333 infection but also on other genetic factors driving root microbiome manipulation, as previously
334 reported for closely related *Verticillium* soil-borne pathogens[57, 58, 59].
335 Our results indicate that the ability of *P. cucumerina* to colonise multiple hosts across
336 various environments is at least to some extent mediated by a remarkably large repertoire of
337 CAZymes acting on diverse plant cell wall substrates. In comparison with 93 other plant-
338 associated Sordariomycete fungi[8], *P. cucumerina* 0016 had a unusually diverse repertoire of
339 plant cell wall degrading enzymes (PCWDEs, 12th/94, n = 160, range 198-29), especially those
340 acting on cellulose (3rd/94, n = 92, range 94-6) and displayed the highest percentage of
341 PCWDEs within fungal CAZyme repertoires (1st/94, 76%, range 76%-33%)[8]. Our within-
342 species comparative genome analysis amongst 69 *P. cucumerina* isolates revealed
343 surprisingly little variation in gene repertoire categories, including CAZymes (range 1,212-
344 1,397, median = 1,291), which did not correlate with variation in fungal detrimental activity. We
345 propose that among-strain phenotypic variation in pathogenicity arose from differences in gene
346 expression, rather than in gene repertoire composition, as previously shown for closely related
347 fungal isolates with contrasted lifestyles[30, 34, 60, 61].
348 We also report the capacity of *P. cucumerina* to express differential sets of CAZymes
349 upon sensing of host-derived cues as a potential strategy driving multi-host compatibility.
350 Indeed, root-infecting fungi colonizing both dicot and monocot hosts have likely evolved
351 mechanisms to efficiently respond to markedly different root cell walls. Such sensing and
352 adequate transcriptional activation of fungal genes encoding PCWDEs was previously
353 reported for broad host range leaf pathogens[62, 63], and our data support that fungal sensing of
354 host-specific cell wall cues might contribute to host range expansion belowground. Therefore,
355 *P. cucumerina* dispersion success in nature is likely explained by its capability to infect multiple
356 hosts at any sites in order to avoid population bottleneck. Our results are consistent with the
357 idea that widespread soil-borne fungal pathogens often have a broad host range and
358 remarkably diverse CAZyme repertoires. Therefore, we conclude that both diverse and
359 tuneable repertoires of PCWDEs promotes multi-host infection.
360 We showed that infection of *A. thaliana* and *S. lycopersicum* roots by *P. cucumerina*
361 requires a candidate β-1,3 glucanase GH64, thereby promoting disease susceptibility. Given
362 that GH64-encoding gene is induced in response to a broad range of hosts, it is plausible that
363 the root signal is evolutionary conserved across monocots and dicots. However, because
364 endophytic fungal colonization was limited in barley roots and that pathogenicity was observed
365 in multiple dicot hosts that diverge at least 110 MyA (*A. thaliana* and *S. lycopersicum*), we
366 propose that this candidate β-1,3 glucanase promotes *P. cucumerina* proliferation in the root
367 endosphere. This is consistent with previous observations showing that endophytic
368 colonization aggressiveness and disease symptoms are often linked during root colonization
369 by mycobiota members[7, 8]. The differential enrichment of this gene in genomes of saprotrophs,

370 endophytes and pathogens also suggests a broader relevance for plant cell wall penetration
371 and/or degradation in diverse fungal lineages. Remarkably, mutation of the orthologous gene
372 in the root pathogen *C. incanum* also prevented root infection and disease in *A. thaliana*.
373 However, the gene was dispensable for tobacco root infection, consistent with its low
374 expression in the roots of this host. Irrespective of this, our data corroborate that GH64 is a
375 root infection determinant in at least two fungal pathogens that evolved independently in
376 distinct families within Sordariomycetes.

377 β-Glucans represent a diverse group of polysaccharides that are found in both plant
378 and fungal cell walls[64]. These glucans can be recognized as MAMPs (Microbe-Associated
379 Molecular Patterns) or DAMPs (Damage-Associated Molecular Patterns) and initiate plant
380 immune responses[54, 65, 66]. To suppress β-glucan-mediated immune activation, plant-colonizing
381 fungi have been shown to use various strategies[67], including down-regulation of β-glucan
382 biosynthesis genes *in planta*[68], secretion of lectins that bind β-glucans[69] or of molecules that
383 inhibit plant endoglucanase[70]. *Magnaporthe oryzae* was recently shown to secrete a GH17
384 exo-β-1,3-glucanase (Ebg1) that hydrolyzes β-1,3-glucans into glucose, thereby preventing
385 plant immune responses[21]. Similarly, a fungal insect pathogen was shown to deploy a host-
386 inducible β-1,3-glucanase called BbEng1 in order to deplete immunogenic β-1,3-glucans from
387 fungal cell walls, thereby evading host immune surveillance and promoting successful fungal
388 mycosis[71]. Therefore, suppression of β-1,3-glucan-induced immune activation could represent
389 one strategy deployed by *P. cucumerina* to promote endophytic infection in roots of
390 Arabiodopsis and tomato. Furthermore, *P. cucumerina* could also deploy this GH64 enzyme
391 *in planta* to break down callose, a plant cell wall polysaccharide composed of β-1,3-linked
392 glucose units[72, 73]. Given that callose deposition is well known to occur at pathogen infection
393 sites and in plasmodesmata to block pathogen invasion[74, 75, 76], it is tempting to speculate that
394 *P. cucumerina* could also have evolved this strategy to overcome callose-mediated restriction
395 of fungal propagation in roots. Irrespective of the mode of action, GH64 represents a target of
396 choice to efficiently control destructive root pathogens that have a broad host range.
397

398 # Methods

399 **Re-analysis of *A. thaliana* mycobiota profiling data**

400 Data from a European transect study profiling the root mycobiota of *A. thaliana* at 17 sites were
401 re-analysed[14], together with profiling data of *A. thaliana* roots grown in Cologne agricultural
402 soil[9]. Paired sequencing reads were joined and demultiplexed using QIIME[77] and its scripts
403 *join paired reads* and *split libraries fastq*. Quality filtering step was performed with phred score
404 of 30 or higher. For fungal reads that could not be merged, the respective forward reads were
405 kept for further processing. Filtered and demultiplexed reads were trimmed to an equal length
406 of 220bp using USEARCH with option -fastx truncate[78]. Those reads were dereplicated and

407 forwarded to the UNOISE3 pipeline[40]. Amplicon sequence variants (ASVs) were determined
408 using the USEARCH unoise3 command, where reads with sequencing errors are corrected
409 and possible chimeras are removed, leaving a set of correct amplicon sequences. ASVs reads
410 were checked against an ITS sequences database (full-length ITS sequences from the
411 National Center for Biotechnology Information (NCBI)) to remove non-fungal reads. Fungal
412 ASVs were classified using the Warcup database[79]. To receive count-tables, ASV sequences
413 were then mapped back against the whole read dataset at a 97% similarity cut-off (using
414 USEARCH -usearch global).
415
416 **Analysis of *P. cucumerina* niches, hosts and prevalence in soils worldwide**
417 To identify which plant hosts *P. cucumerina* colonizes in nature, we accessed the GlobalFungi
418 online database[43] (in November 2024). Samples in which *P. cucumerina* (SH1002406.10FU)
419 was detected were compared to total samples registered in the database, focusing on sample
420 niches. Overrepresentation of a niche in samples in which *P. cucumerina* was detected was
421 tested with a one-sided Fisher's exact test, using the formula
422 stats.fisher_exact(alternative='greater') in Scipy[80], followed by Benjamini-Hochberg *P* value
423 correction. To identify plant species that host *P. cucumerina* in their roots and rhizosphere,
424 "root", "rhizosphere" and "root+rhizosphere" samples in which *P. cucumerina* was detected
425 were browsed and the "dominant plant species" metadata associated with each sample were
426 collected. These species were grouped by phylogenetic classes, according to the NCBI
427 Taxonomy database[81]. The plant phylogeny presented on Figure 1 was obtained using
428 Taxallnomy[82] based on NCBI Taxonomy data.
429 To analyse the prevalence of *P. cucumerina* in soils, previously curated ITS1-5.8S-
430 ITS2 rDNA sequences of species[35] were downloaded and ITS1 sequences were extracted
431 from them using ITSx[83]. We then linked these ITS1 sequences to ASVs from the Global
432 Mycobiome data set[84] using blastn with a percentage of identity threshold of 100%. These
433 ASVs were detected in 780/3,194 soil samples (relative abundance > 0). Metadata of sites at
434 which ASVs were detected were then analysed, focusing on latitude, longitude and biome
435 types. We tested the enrichment of specific biomes in soils containing *P. cucumerina* using a
436 random sampling approach. We randomly sampled 780 soil samples 99,999 times, and each
437 time counted the number of soil samples associated to each biome. We used a Fisher's exact
438 test to compare whether the set of soils in which *P. cucumerina* was detected contains
439 significantly more/less samples associated to each biome than the randomly sampled sets of
440 soils. Statistical testing was performed with Scipy 1.5.3 and the function
441 stats.fisher_exact(alternative='two-sided')[80].
442
443 **Assembly of a culture collection of *Plectosphaerella* isolates**

444 We previously isolated 7 strains of *P. cucumerina* and 1 strain of *P. plurivora* from surface-
445 sterilized *A. thaliana* roots sampled across Europe[9]. We imported 64 additional strains from 17
446 collaborators or culture collections, sampled from different hosts and environments across
447 different continents. See **Supplementary Table 3** for details.
448
449 ***A. thaliana* recolonization experiments with spores of *Plectosphaerella* isolates**
450 We performed plant recolonization experiments with individual *Plectosphaerella* strains in an
451 agar-based system. First, to prepare fungal spore stock solutions, fungi were grown on oatmeal
452 medium. This medium was prepared by crushing 5 g of oatmeal in liquid nitrogen with a mortar
453 and pestle until obtaining a powder. This powder was boiled for 30 min in 200 mL of water on
454 a heating plate at 180°C. Then, 4 g of Difco Agar (ref. 214530) was added to the mixture prior
455 to autoclaving. Fungi were revived from 30% glycerol stocks and grown on this medium for 10-
456 14 days. Sterile water was poured on the grown mycelium, pipetted then filtered on Miracloth,
457 then centrifuged at 4000 x g for 10min to concentrate spores. Spores were counted on a
458 Malassez cell, and their concentration was adjusted to 10^7 spores/mL in 25% glycerol. These
459 stock solutions were flash-frozen in liquid nitrogen, then kept at -80°C. *A. thaliana* Col-0 seeds
460 were sterilized 15 min in 70% ethanol, then 5 min in 8% sodium hypochlorite. They were then
461 washed six times in sterile double-distilled water and one time in 10 mM $MgCl_2$. They were
462 kept 5 days at 4°C in the dark for stratification. Culture medium was prepared by mixing 2.2 g
463 Murashige and Skooge (Duchefa Biochemie, ref. M0222) and 0.5 g MES buffer (Roth, ref.
464 4256.5) powders in 1 L of double-distilled water. Then, pH was adjusted to 5.7, and 10 g of
465 Difco Agar (ref. 214530, 1% final concentration) was added. After autoclaving, 120x120mm
466 petri dishes were filled each with 65 mL of medium. The top 2 cm of agar were cut out, and 7
467 seeds were deposited where the cut was made. Plates were sealed with micropore tape and
468 their bottoms were wrapped (up to the seeds level) in paper to hide roots from the light. Plants
469 were grown for 10 days at 21°C (10h with light intensity 4 at 19°C, and 14 h in the dark) in
470 growth chambers. After 10 days, culture plates were re-opened and only the four biggest plants
471 in each plate were kept. To prepare a spore inoculum, 5 µL of defrosted spore stock solution
472 was pipetted and mixed with 995 µL of sterile water. This mixture was centrifuged at 9,000 g
473 for 10 min. The top 950µL in the tube were replaced with fresh sterile water. With this
474 procedure, glycerol was removed and the concentration of spores in the tube was set to 500
475 spores/10 µL. Ten-days old seedlings were inoculated by pipetting down 10 µL of inoculum
476 (500 spores) on their primary roots. Plates were then re-sealed and put back in growth
477 chambers, with settings mentioned above. After 28 more days in culture, plants were
478 phenotyped by measuring their shoot fresh weight.
479
480 **Biolog Plate Experiment**

481 Mycelia of individual *Plectosphaerella* isolates (n = 72) were harvested using sterile tips and
482 transferred into 1 mL of 10 mM $MgCl_2$ containing one stainless-steel bead (Ø = 2.85–3.45 mm;
483 Carl Roth) and ten smaller beads (Ø = 0.75–1 mm; Carl Roth). Samples were homogenized
484 using a paint shaker (SK450, Fast & Fluid Management, Sassenheim, Netherlands) for 10
485 minutes. The resulting fungal suspensions were diluted to a final biomass concentration of
486 0.2 mg/mL in a 1:1 mixture of Biolog FF and IFO inoculating fluids (Biolog Inc.), supplemented
487 with 1% Dye Mix A (Biolog Inc.) to enable colorimetric detection. For each isolate, 100 µL of
488 the inoculum was loaded into Biolog PM1 and PM2A MicroPlates™, which together contain
489 190 distinct carbon sources. Plates were incubated at 25 degrees without shaking and
490 respiration was monitored every 20 minutes for 5 days using the OmniLog® instrument (Biolog
491 Inc.), which records OmniLog Respiration Units (ORU). The area under the curve (AUC) of
492 ORU values was calculated for each well over the 5-day period and log-transformed (log AUC)
493 for normalization prior to data visualization and statistical analysis.
494
495 **Genome sequencing**
496 Two strains in our collection (P16 and P117) were previously sequenced with the PacBio
497 technology and assembled[8]. Their genomes made available on database GenBank:
498 bioprojects PRJNA371204 and PRJNA570880. While we used previously published genome
499 assembly JAGPXD000000000 for P16, we used an improved version of JAGSXJ000000000
500 for P117, which was downloaded from the Mycocosm portal[85] with the identifier Plecucu2. For
501 P30, we used a genome assembled from Illumina short reads made available on the
502 Mycocosm portal with identifier Plecu1. Other genomes used in this study were newly
503 sequenced with PacBio sequencing. Individual strains were revived on Potato Growth Agar
504 medium (Roth, ref. CP74.1), and cultured for 7 days. Six strains (P43, P143, P340, P421,
505 P455, P612) were sequenced with PacBio CLR technology. Their gDNA was extracted
506 according to a previously introduced cetyltrimethylammonium bromide (CTAB) protocol[32]. RNA
507 of these strains was sequenced with Illumina short reads to support gene prediction in the
508 assemblies. All the other strains (n = 63) were sequenced with PacBio CCS technology (HiFi
509 reads). Their gDNA was extracted according to a previously introduced CTAB protocol[32] ,
510 without the described purification step of the DNA extracts. Native fungal DNA was found in
511 many cases to be recalcitrant to sequencing. Therefore, PacBio barcoded ultra-low libraries
512 were prepared according to the protocol "Procedure Checklist - Preparing HiFi SMRTbell®
513 Libraries from Ultra-Low DNA Input" starting with DNA fragmentation using g-Tubes (Covaris).
514 Typically, five libraries were pooled and then sequenced on a Sequel IIe SMRT cell with 8
515 million ZMWs with Binding kit 2.0 or Binding kit 2.2 and the Sequel II Sequencing Kit 2.0 for 30
516 h. If necessary gDNA or final libraries were additionally size-selected on a Blue Pippin (Sage
517 Sciences) device to remove smaller DNA fragments. After sequencing, the HiFi data was

518 demultiplexed with SMRTlink10.0, and adapter sequences were trimmed away with cutadapt
519 v3.5[86].
520
521 **Genome assembly and gene annotation**
522 Genomes were assembled and polished from PacBio reads using Flye v2.9-b1768[87] with
523 parameter *–genome-size 40m*, decided based on the size of previously sequenced *P.*
524 *cucumerina* genomes[8, 30]. Parameter –pacbio-hifi was used for genomes sequenced with
525 PacBio CCS and –pacbio-raw was used for genomes sequenced with PacBio CLR. Gene
526 prediction was carried out in all the genome assemblies except for P16 and P117, for which
527 we used the gene predictions from the JGI annotation pipeline[85]. We used FGENESH v8.0.0[88]
528 with similarity matrix *Torrubiella hemipterigena* and parameters *-skip bad prom -skip bad term*.
529 CAZyme-encoding genes were annotated using dbcan v4.0.0[89]. Candidate effector-encoding
530 genes were annotated by first predicting secreted proteins with SignalP 6.0[90], then by using
531 EffectorP v3.0[91] (in fungal mode) on predicted secreted proteins. To annotate protease-
532 enconding genes, the annotation tool emapper v2.1.5[92] based on eggNOG orthology data[93]
533 was run using DIAMOND as a similarity search algorithm[94], then we considered as proteases
534 all the proteins linked by emapper to a PFAM identifier associated to a MEROPS family[95].
535 Transposons were annotated in genome assemblies with reasonaTE v1.0[96] with parameters -
536 *mode annotate -tool all*.
537
538 **Comparative genomics**
539 Orthology prediction was performed on the total set of proteins predicted in our data set using
540 OrthoFinder v2.5.4[97] with parameter -S blast. Orthogroups were classified into intra-species
541 conservation levels. Those represented in 100% *P. cucumerina* genomes were called *core*.
542 The ones represented in at least 75% *P. cucumerina* genomes were considered *broadly*
543 *conserved*. The ones represented in less than 20% *P. cucumerina* genomes were considered
544 to have low conservation. We considered as *subspecies-associated* the orthogroups
545 represented in more than 20% of one *P. cucumerina* subspecies, while absent from the other.
546 Remaining orthogroups were considered as *sparsely distributed*. To test for overrepresentation
547 of these conservation levels among orthogroups encoding CAZymes, effectors and proteases,
548 Fisher's exact tests were performed using the function stats.fisher_exact(alternative='two-
549 sided') from Scipy[80]. Principal components analysis (PCA) revealing similarities of CAZyme,
550 protease and effector catalogues were performed by counting the number of annotated genes
551 per family (CAZy and MEROPS families for CAZymes and proteases respectively, and
552 orthogroups for effectors), then calculating pairwise Jaccard distances between strains.
553 Jaccard dissimilarity matrices were calculated with R package Vegan v2.6-4
554 (https://github.com/jarioksa/vegan), using function vegdist(method="jaccard"). PCA plots were

555 then reconstructed from these matrices using function prcomp in R. PERMANOVA tests were
556 computed to test for effects of strain phylogeny and the host of isolation on CAZyme, effector
557 and protease catalogues, using the function
558 adonis2(JaccardMatrix~PhylogeneticGroup+HostOfIsolation, permutations = 9999). Following
559 PERMANOVA, given the significant effect of *Plectosphaerella* phylogenetic groups on
560 genomic compositions in gene repertoires, pairwise comparisons aiming to detect significant
561 differences between phylogenetic groups were computed with the function
562 pairwise.perm.manova() from R package RVAideMemoire v 0.9-83-12, with default Benjamini-
563 Hochberg correction applied on *P* values.
564
565 **Phylogeny reconstruction**
566 We used four different methods to reconstruct the phylogeny of our *Plectosphaerella* collection,
567 based on genome sequences. First two trees were built from 5,466 single copy orthologs,
568 identified in our data set by OrthoFinder[97]. Sequences of these 5,466 gene families were
569 aligned independently using MAFFT v7.407 with parameter –auto[98], then trimmed with Trimal
570 v1.4.rev22 using default settings[99]. These trimmed alignments were used as input for IQ-TREE
571 v2.1.1[100]. The algorithm ModelFinder[101] implemented in IQ- TREE identified maximum-
572 likelihood model JTT+F+I+G4 to be the most adapted one considering the input data. We
573 therefore reconstructed phylogenies with IQ-TREE and RAxML-NG v1.1.0[102] using this model.
574 While prior alignment concatenation using AMAS[103] was necessary in the case of RAxML-NG,
575 the IQ-TREE software carries out concatenation before phylogeny reconstruction. Additionally,
576 we reconstructed phylogenies with two different coalescent approaches. First, we used method
577 STAG[97], as implemented in OrthoFinder. STAG built individual gene family trees (n=13,694),
578 then reconstructed a coalescent species tree from them. Finally, we used Astral v5.7.1[104] with
579 default parameters, to reconstruct a coalescent species tree from all the single copy ortholog
580 gene trees generated by OrthoFinder (n = 5,466). The phylogenetic tree used as a reference
581 in this study and presented on multiple figures is the one generated by IQ-TREE.
582
583 **Multi-host re-colonization experiments with *P. cucumerina* 0016**
584 Arabidopsis and tomato seeds were sterilized for 15 min in 70% ethanol, then 5 min in 8%
585 sodium hypochlorite. Subsequently, the seeds were rinsed five times with sterile double-
586 distilled water and once with a 10 mM Magnesium Chloride ($MgCl_2$) solution. They were then
587 stratified in darkness at 4°C for a period of 2 to 3 days. Barley seeds were sterilized for 1 hour
588 in 6% sodium hypochlorite and then five washes in sterile double-distilled water. Before
589 inoculation, Arabidopsis and tomato seedlings were grown for 14 days and barley seedlings
590 were grown for 7 days in ½ Murashige and Skoog (MS) media (without carbon source, pH 5.7).
591 The conditions in the growth chamber are 21°C, for 10h with light (intensity 4) at 19°C and 14h

592 in the dark in growth chambers. Inoculation in the roots with *P. cucumerina* 0016 was carried
593 out by harvesting the spores from mycelium grown for 8 days on Potato extract Glucose Agar
594 medium (PGA) using sterile double-distilled water. Harvested spores were filtered through
595 Miracloth (Merck, 475855-1R) centrifuged for 5 min at 600 x g (Eppendorf- centrifuge 5810R)
596 and then pellets were resuspended in 1 mL of sterile double-distilled water. Spores were
597 counted using Hemocytometer and the concentration was adjusted to 2.5 x 10^6 spores/mL.
598 These spores were inoculated onto the roots of Arabidopsis, tomato, and barley seedlings.
599 After 30 min of incubation, the seedlings were transferred to sterile Magenta boxes containing
600 40 g vermiculite with 95 mL of ½ MS media (without carbon source, pH 5.7). After transferring
601 four plants into this "vermiculite system", the boxes were sealed using Micropore by placing an
602 empty magenta box upside down on top. Plants (Arabidopsis, tomato, and barley) were grown
603 for 21 days post inoculation (dpi) in the same growth conditions as described above. As a
604 control condition for RNA-seq experiment, sterile Nucleopore Track-Etched polyester
605 membranes ("FungusOnMembrane") were deposited on "Vermiculite System", then 10 µL
606 drops of fungal spores (2.5 x 10^6 spores/mL of *P. cucumerina* 0016 in sterile double-distilled
607 water) were placed on each one. The membranes were collected and processed as the root
608 samples of our test condition. After 21 dpi, the roots were harvested and flash-frozen (if they
609 were for *in planta* RNA seq experiment), shoot fresh weight (SFW) was measured for each
610 plant. For analysis, the plant performance index was calculated as the product of shoot fresh
611 weight (SFW) and survival rate of the plants[8]. In further correlation analyses, we used log
612 transformed PPI normalized to mock controls using the Hedges' g method. The harvested
613 roots were stored at -80°C until further use.
614
615 **High-throughput DNA extraction and quantification of fungal load at the roots.**
616 Total DNA was isolated from Arabidopsis, tomato and barley frozen roots samples (one root
617 per tube) using a modified high-throughput version of the FastDNA SPIN kit for Soil (MP
618 Biomedicals)[105]. Root samples were homogenized in a Precellys 24 TissueLyser (Bertin
619 Technologies) for 2 x 30 s at 6,200 rpm with 15 s intervals. 450 µL of sodium phosphate buffer
620 (MP Biomedicals) and 50 µL of MT buffer (MP Biomedicals) was added and then homogenized
621 as described before. Subsequently, the centrifugation was performed for 15 min at 16,000 x
622 g (Eppendorf- centrifuge 5424R) and 150 µL of the supernatant were transferred to a 96-well
623 Acroprep Advance filter plate (with 0.2 µm Supor filter; Pall). Once full, the filter plate was
624 positioned on a PCR plate filled with 50 µL Binding Matrix (MP Biomedicals) per well and
625 centrifuged for 15 min to remove residual soil particles. This and all subsequent centrifugation
626 steps were carried out in a swing out centrifuge at 200 x g (Eppendorf- centrifuge 5424R). The
627 PCR plate was sealed and shaken for 3 min to allow binding of the DNA to the Binding Matrix.
628 The suspension was pipetted onto a second filter plate of the same kind, positioned on a

629 collection plate, and centrifuged for 15 min. The flowthrough was discarded. Then, 200 µL
630 SEWS-M washing buffer (MP Biomedicals) were pipetted into each well of the filter plate and
631 centrifuged for 5 min. This washing step was carried out a second time. The flowthrough was
632 discarded and followed by centrifugation for 5 min to remove residual SEWS-M buffer. Finally,
633 30 µL nuclease-free water were added to each well and left to incubate at room temperature
634 for 3 min. Subsequent centrifugation for 5 min eluted the DNA into a clean PCR plate. Fungal
635 colonization of these root samples was then measured by quantitative PCR (qPCR) without
636 prior adjustment of DNA concentrations. For each Arabidopsis root sample, two reactions were
637 conducted with one pair of primers ITS1F (5'-CTTGGTCATTTAGAGGAAGTAA-3') and ITS2
638 (5'-GCTGCGTTCTTCATCGATGC-3') which target the fungal ITS1 sequence, and two pair of
639 primers specific to plant samples: UBQ10F (5'-TGTTTCCGTTCCTGTTATCT-3') and UBQ10R
640 (5-ATGTTCAAGCCATCCTTAGA-3) that target the *Ubiquitin10* Arabidopsis gene, nAP256F
641 (5'-CTACCACTGGTCACTTGATC-3') nAP257R (5'-CCAGGCATACTTGAATGACC-3') that
642 target the *EF1-alpha* tomato gene, UBQF (5'-ACCCTCGCCGACTACAACAT-3') and UBQR
643 (5-CAGTAGTGGCGGTCGAAGTG -3') that target the *Ubiquitin* barley gene[8, 106]. Each reaction
644 was performed by mixing 5 µL of iQ™ SYBR Green Supermix with 0.3 µL of 10 µM forward
645 primer, 0.3 µL of 10 µM reverse primer, 3.4 µL of sterile double-distilled water and 1 µL of
646 template DNA. A BioRad CFX Connect Real-Time system was used with the following
647 program: 3 min of denaturation at 95°C, followed by 39 cycles of 15 s at 95°C, 30 s at 60°C
648 and 30 s at 72°C. We then calculated a single colonization index for each sample using the
649 following formula: Index = 2−(Cq(ITS1)/Cq(UBQ10))[8].
650
651 **Confocal microscopy of root colonization by fungi.**
652 Roots of Arabidopsis (5 weeks old), tomato (5 weeks old), and barley (4 weeks old) grown for
653 21 days in mono-association with *P. cucumerina* 0016 were harvested and conserved in 70%
654 ethanol. During the process, the roots were incubated in 20% of Potassium Hydroxide (KOH)
655 at room temperature with orbital shaking at 60 rpm overnight. At the end of the incubation, the
656 KOH was removed and the roots were rinsed twice with 1x Phosphate-Buffered Saline (PBS)
657 at pH 7.4. They were then double-stained with 0.1% Calcofluor white (Sigma, cat: 18909) and
658 25 µg/ mL (for Arabidopsis roots)- 50 µg/mL (for tomato and barley roots) of Wheat Germ
659 Agglutinin conjugated to fluorophore Biotium CF633 (WGA-CF633) buffered at pH 7.4 in PBS.
660 After 30 mins of incubation, samples were then washed thrice in PBS and imaged with a
661 Confocal LSM880 from Zeiss and the associated software ZEN v2.3 SP1.
662
663 ***In planta* fungal transcriptome analysis.**
664 PolyA enrichment was carried out on the RNA extracts, then RNAseq library was prepared
665 with the NEBNext Ultra™ II Directional RNA Library Prep Kit for Illumina (New England

666 Biolabs). Sequencing was then performed in single read mode on a HiSeq 3000 system. RNA-
667 seq reads from Agar System were trimmed using Trimmomatic[107] v0.38 and parameters
668 TRAILING:20 AVGQUAL:20 HEADCROP: 10 MINLEN:100. For RNA-seq reads from
669 Vermiculite System, trim-galore v0.6.5 pipeline was used[108]. Subsequently, these reads were
670 employed for the quantification of transcripts in each sample. A unified transcriptome was
671 constructed by concatenating the transcriptomes of plants and *P. cucumerina*
672 (Plecuc1_GeneCatalog_CDS_20171208 from JGI Mycocosm[8, 85]) for various hosts. The
673 indexing for Arabidopsis was based on TAIR11[109] from JGI- Phytozome[110], for tomato was
674 based on TMCSv1.2.1[111], for barley was based on Hvulgare_462_r1[112]. The creation of an
675 index for the concatenated transcriptome was accomplished using Salmon v1.1.0[113]. The
676 quantification of transcripts was performed with Salmon quant to determine the read counts for
677 each sample. These datasets were then imported into R utilizing the tximport v1.12.3[114].
678 Bioconductor package to compute the Transcript Per Million (TPM) and Reads Per Kilobase
679 per Million mapped reads (RPKM) values. Package *DESeq2* v1.24.0[115] was used for
680 quantification of differential expression of genes (DEGs) and $\log_2$ fold-change (Log2FC) was
681 calculated using "FungusOnMembrane" as control group. Log2FC values were corrected by
682 shrinkage with the algorithm in *apeglm* v1.6.0 package[116]. Fungal DEGs obtained on infecting
683 different host conditions were used for further analysis. Subsequently, a matrix was created by
684 incorporating the $\log_{10}$ (counts per million, CPM) values for individual samples (columns) along
685 with exclusively the fungal DEGs (rows). Computation of $\log_{10}$ CPM values was based on the
686 normalized read counts.
687
688 **Statistical analysis of differentially expressed genes.**
689 Using the computed CPM matrix, the Pearson correlation among samples was determined and
690 was used as the basis for multi-dimensional scaling (MDS) analysis. MDS was carried out
691 utilizing the distance matrix, and pairwise statistical analyses were conducted using Dim1 and
692 Dim2 for each sample group to assess the significance of impact across conditions, employing
693 the ANOVA-TukeyHSD function in R. This matrix was also employed for clustering the DEGs
694 using *k*-means clustering algorithm and the optimum *k* (ranging from 1 to 100) was determined
695 using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). A
696 heatmap was generated using the Log2FC to demonstrate expression patterns in comparison
697 to the control. Visualization of DEG within the groups was illustrated through VennDiagram[117]
698 and heatmap[118] separately. Gene annotations were assigned by matching the respective
699 gene/transcript/protein IDs to available annotations in JGI Plecuc1 (*Plectosphaerella*
700 *cucumerina* RP01 v1.0[8]) that includes Gene Ontology (GO), INTERPRO, EffectorP[91], and
701 CAZyme[119] terms associated with the fungal genes. For each functional term, gene set

702 enrichment analysis was performed using Fisher's Exact test. The *P* values were adjusted for
703 each condition using the Benjamini-Hochberg method.
704
705 **Construction of CAZyme network plots targeting predicted Cell Wall Substrates.**
706 CAZyme database was utilized for CAZyme enrichment analysis and visualization. Substrates
707 were categorized manually based on the class of Cell Wall (CW) (string parsing
708 *"pectin|cellulose|lignin|arabinan|galactan|glycan|glucan|xylan"* using *tidyverse*) components,
709 and redundant substrates were grouped together to simplify visualization. A network graph
710 was created using an adjacency matrix consisting of CAZyme families (rows) and possible
711 substrate families (columns) (CAZyDB vCAZyDB.09242021 from dbCAN2 repository)[47]. A
712 value of 1 was assigned to CAZyme families targeting potential substrates, with the rest being
713 designated as 0. Using this matrix network map was constructed, where the edges indicated
714 the number of enzymes (circle nodes) targeting potential substrates (square nodes). *P* values
715 derived from the enrichment analysis of the corresponding CAZyme families were employed
716 to represent the nodes in the network graphs, mean Log2FC of the CAZymes was employed
717 to illustrate the color of the edges. By integrating the network structure and the significantly
718 enriched CAZymes, the activation of CW components under specific conditions was visualized
719 using the igraph[120] package in R (v3.6.1 & v4.3.1).
720
721 **Monosaccharide compositional analysis and crystalline cellulose measurement.**
722 Cellulose content and matrix polysaccharide monosaccharide composition of the samples
723 were analysed as per previously described method[121]. In short, Alcohol insoluble residue (AIR)
724 was prepared from freeze-dried root tissues and split into two samples. One half of the samples
725 were treated with a weak acid (4% sulfuric acid) to release matrix polysaccharide-derived
726 sugars, while the other half of the samples were treated first with a strong acid (72% sulfuric
727 acid) to swell cellulose. Then the sulfuric acid concentration was diluted to 4% to yield
728 monosaccharides both derived from cellulose and the matrix polymers. Subtraction of the two
729 values allows for the quantification of crystalline cellulose. Monosaccharides of all fractions
730 were quantified using a Professional IC Vario high-performance anion-exchange
731 chromatography system 1068 (Metrohm, Herisau, Switzerland) equipped with a CarboPac
732 PA20 column (Thermo Fisher Scientific, Waltham, MA, USA) and an amperometric detector
733 (Metrohm). The used sodium hydroxide gradient is specified as described[121]. Lignin content of
734 the AIR samples was determined as previously described[122].
735
736
737 **GH64 occurrences in the fungal kingdom**

738 The CAZYme family GH64 was procured from the JGI Mycocosm[85] and from the CAZyme
739 database[119] respectively. These datasets were merged to preserve the taxonomic
740 categorization for the fungal kingdom that were present in the CAZyme database based on the
741 portal ID (genomic identifiers). The number of GH64 occurrences was ascertained by cross-
742 referencing the merged dataset. Finally, the FunGuild database was integrated into the
743 concatenated dataset through species or genus identifiers. Collectively, three tables were
744 merged and summarised using a phylogenetic tree of the fungal kingdom (constructed from
745 the *taxonomizr* R package[123]). The count data for the number of GH64 in genomes (from JGI
746 Mycocosm[85]) and the ecological information (FunGuild[124]) were incorporated into the
747 phylogenetic tree utilizing iTOL v6.0[125]. Due to redundancy in the annotation of ecological
748 niches (GUILD category in the FunGuild repository[124]), we classified those annotations that
749 contain the string "endophyte" as endophyte while others were classified as "non-endophyte."
750 Statistical analysis for Phylum was executed employing the pairwise Wilcoxon test with
751 Bonferroni *P* value adjustment, and for enrichment analysis within Ascomycota members (at
752 the class level) and within the ecological group, Fisher's exact test was performed with *P* value
753 adjustment using Bonferroni correction.
754
755 **Structural analysis of GH64**
756 The structure of the GH64 protein from *P. cucumerina* 0016 was predicted with AlphaFold3,
757 after signal peptide removal with SignalP v6.0[90]. The best predicted model was then used as
758 query for Foldseek[50] search (online, accessed December 2024), looking for structural similarity
759 with experimentally resolved protein structures registered in the PDB database. Structure
760 alignment between the GH64 enzyme of P16 and the best hit (SpGlu64A, protein from
761 *Streptomyces* pratensis[51], PDB code: 7VTK) was performed using TM-align[126], as
762 implemented in iCn3d[127].
763
764 **CRISPR/Cas9-mediated genome editing in *P. cucumerina* 0016**
765 *Design of synthetic guide RNA (sgRNA):* The structure of the specific sgRNA oligo is T7
766 promoter- Start of transcription- Protospacer- Constant part of sgRNA.
767 AAGCTAATACGACTCACTATAG(G)NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAA
768 ATAGCAAG. sgRNAs (**Supplementary Table 5**) were designed with the sgRNA designer
769 tools CRISPick (https://portals.broadinstitute.org/gppx/crispick/public) and CCTop[128]. An
770 sgRNA with a single, unique genomic hit to the target locus was identified using CCTop, and
771 its on-target score was retrieved from CRISPick.
772 *Synthesis of sgRNA:* For the preparation of gRNA synthesis, specific and constant gRNA
773 oligos were mixed by combining 1 µL of the specific protospacer oligo (100 pmol stock), 1 µL
774 of the constant gRNA oligo (100 pmol stock), and 8 µL of H2O, totaling 10 µL. 95°C for 5 min,

775 cooling from 95°C to 85°C at 2°C/sec, and further cooling from 85°C to 25°C at -0.1°C/sec,
776 with a total time of approximately 15 min. A T4 DNA mix was prepared to fill in overhangs,
777 consisting of 2.5 µL dNTPs (10 mM stock), 2 µL NEB buffer 2.1 (10x buffer stock), 5 µL H2O,
778 and 0.5 µL T4 DNA polymerase (NEB, M0203L), totaling 10 µL. This T4 DNA mix was added
779 to the annealed oligos, and the mixture was incubated at 12°C for 20 min. The product was
780 purified using a PCR column (Macherey-Nagel NucleoSpin Gel and PCR Clean-up Kit) and
781 eluted with 30 µL H2O. The concentration was measured with a Nanodrop, ensuring it was
782 around 130 ng/µL, and the band size was verified on a 3% agarose gel, expecting a band of
783 approximately 120 bp. For RNA synthesis, the HiScribe T7 High Yield RNA Synthesis Kit (Neb,
784 E2040S) was used, following the protocol for small RNAs. The reaction mixture included 6 µL
785 NTPs (25 mM stock), 1.5 µL T7 Buffer (10x buffer stock), 1.5 µL T7 Polymerase, 11 µL
786 Template (1-2 µg DNA) and, totaling 20 µL (overnight incubation at 37°C). After that, 14 µL of
787 nuclease free H2O, 4 µL of DNase 1 buffer (10x), and 2 µL of DNase 1 were added, and the
788 setup was incubated for 30 min at 37°C.
789 *Purification of sgRNA:* For sgRNA purification, RNA Clean & Concentrator 25 kit (Zymo
790 Research, Catalog No R1017 & R1018) was used. The eluted samples were immediately
791 stored at -80°C, with a concentration greater than 1 µg/µL. The integrity of the gRNA was
792 checked on 3-4% agarose gel electrophoresis. If a single band was observed, it was suitable
793 for further steps. If a smear was present, the sgRNA synthesis needed to be repeated. If a
794 double band was observed, double the amount of sgRNA than recommended was used for
795 transformation.
796 *Design of the Donor DNA:* Transgenes (donor DNAs) were designed to include a homology
797 arm on the 5' site of the Cas9 cutting site (60 bp) (**Supplementary Table 5**), the promoter
798 Pcpc1, the phleomycin resistance gene *BleoR*[129], the transcription terminator of tryptophan
799 synthase (TtrpC), and a homology arm on the 3' site of the Cas9 cutting site (60 bp). This
800 construct was ordered from the company Biocat (https://www.biocat.com/gensynthese) in
801 plasmid pUC57.
802 *Amplification of Donor DNA using PCR:* Once the construct was received, donor DNA
803 amplification was performed using PCR. Prior to this, primers (**Supplementary Table 5**; P5-
804 P16 used in pairs) were designed to amplify the donor DNA (*Bleo* cassette with the flanking
805 region, as described) from the vector. The reaction conditions included 5 µL of Buffer (5X), 0.4
806 µL of dNTPs (25 mM), 1 µL of DNA Template, 1 µL of Forward Primer (100 µM), 1 µL of
807 Reverse Primer (100 µM), 1 µL of Phusion DNA polymerase, and 41.6 µL of Sterile water. The
808 PCR conditions were set as follows: 98°C for 30 s, followed by 35 cycles of denaturation at
809 98°C for 8 s, annealing at required temperature for 30 s, and extension at 72°C for 45 s, with
810 a final extension at 72°C for 10 min. A 1% agarose gel electrophoresis with a 1 kb ladder
811 (GeneRuler, Thermo Fisher Scientific) was run to check the integrity of the DNA, ensuring a

812 single clean band was obtained. The DNA band was cut from the gel, and gel purification was
813 performed using the Macherey Nagel kit. The concentration was checked using a Nanodrop,
814 ensuring it was greater than 200 ng/µL.
815 *Protoplast isolation:* Spores at a concentration of 1.5 x 10^7 cells/ mL were inoculated into 100
816 mL of PGB medium. The flask was incubated at 25°C with orbital shaking at 180 rpm for 20 h.
817 Next day, the mycelia were collected by passing the entire culture through a 40 µm filter, and
818 the residues were washed vigorously using 20 ml of 0.7M NaCl to completely remove the
819 spores from the mycelia[130]. Mycelia (350 mg) were transferred to 10 mL of enzyme solution
820 (20 mg driselase in 10 mL of 0.7 M Nacl) and the mixture was incubated in the enzyme solution
821 for 2 h at 28°C with orbital agitation at 60 rpm. After incubation, the solution was filtered (40
822 µm filter). The filtrate was centrifuged at 250 x g for 5 min, with acceleration and deceleration
823 adjusted to 3 in cg (Eppendorf- centrifuge 5810R). The pellet was resuspended in 1 ml of STC
824 (Sucrose 20% (w/v), TrisHcl (1M, pH 8)- 10 mM, Cacl2 (1M)- 50 mM) buffer. The protoplasts
825 were counted using a hemocytometer, and the volume was adjusted to 5 x 10^6 cells/mL.
826 *PEG mediated transformation:* For the PEG-mediated transformation, prebound Cas9-gRNA
827 complex (RNP- Ribonucleoprotein complex)[131] was prepared by mixing Cas9 and sgRNA as
828 follows: 0.5 µL of 10X NEBuffer r3.1, 1 µL of 20 µM EnGen Spy Cas9 HF (NEB, M0667M),
829 and 1.5 µL of sgRNA 2 (approximately 1.5 µg, with the desired amount being 20-40 pmol, or
830 0.8-1.6 µg for 121 nucleotides). This whole mixture was incubated at 25°C for 30 min and
831 further kept on ice until use. For protoplast transformation, 200 µL volume of protoplasts were
832 taken in a 50 mL Falcon tube and kept on ice. The pre-bound Cas9-gRNA complex (~20 pmol)
833 and donor DNA (x µL containing 2.5 µg of DNA) were added to the same tube and gently
834 mixed, the total volume should be as low as possible. The tube with mixture was incubated on
835 ice for 30 min. Then, 1.5 mL of PEG-STC (6 g of PEG 8000 in 14 mL of STC buffer) was added
836 to the tube, which was rotated horizontally to softly mix the reagent. This was followed by
837 incubation at room temperature for 15 min. Subsequently, 5 mL of CM media (containing 6 g
838 of yeast extract, 6 g casein acid hydrolysate, and 10 g sucrose in 1 L H_20) was added and
839 mixed well by horizontal rotation. The tube was incubated in the dark at 25°C for 18 h before
840 being centrifuged at 3,000 x g for 5 min. The pelleted fungi were visible on the side of the tube,
841 and the supernatant was removed. Regenerated protoplasts were resuspended together in
842 100 µL of sterile water the entire solution was plated on square plates containing PGA with
843 200 µg/mL *Bleo* antibiotic and incubated in a 23°C incubator. After 2 days, these colonies were
844 transferred to new plates (PGA with *Bleo*), marked with numbers, and allowed to grow for 4-5
845 days before proceeding with screening the mutants. For control, untransformed protoplasts
846 were plated on PGA and used for screening steps to maintain uniformity for the wild type
847 samples and transformed colonies, especially in the colony PCR step. To screen and to verify
848 the integration of *BleoR* gene cassette at the right site, we used external primers that anneal

849 to the genomic sequences immediately outside the 5′ and 3′ homology arms (**Supplementary**
850 **Table 5)**. For colony PCR, the Phire Plant Direct PCR Master Mix was used (Thermofisher,
851 F170S). Mycelium was picked with a tip and it was transferred into 30 µL of the dilution buffer.
852 This was incubated at 95°C for 15 min and then overnight at 4°C. The next day, the template
853 was ready. The reaction mixture as follows: 10 µL of Master Mix, 1 µL of 10 µM forward primer,
854 1 µL of 10 µM reverse primer, 0.5 µL of Template, and 7.5 µL of nuclease free water. Once
855 the PCR was done, agarose gel electrophoresis was run to check the results and once a
856 positive colony was identified, DNA extraction was performed using the Qiagen DNeasy Plant
857 Mini kit, and PCR was conducted again with Phusion polymerase to double-check the results.
858

859 ***In vitro* growth assay**

860 To assess the growth rates of mutant *P. cucumerina* 0016 strains relative to the wild type, an
861 *in vitro* growth assay was conducted. Spores were initially collected from mycelia grown on
862 PGA. Using a hemocytometer, the spores were counted, and their concentration was adjusted
863 to 2.5 x 10^6 spores/mL. These spores were then inoculated into 30 mL of PGB and incubated
864 at 23°C with orbital shaking at 180 rpm for 24 hours. After incubation, a 1 mL sample was
865 taken to measure the optical density (OD) at 600 nm with a spectrophotometer. The remaining
866 culture was used to harvest and recount spores using a hemocytometer. This experiment
867 included three biological replicates. *In vitro* growth dynamics of *P. cucumerina* 0016 wild-type
868 (WT) and Δgh64 mutant strains in PGB medium were performed in 96 well-plates and
869 measured using spectrophotometer. For WT and Δgh64 conditions, 180 µl PGB was
870 dispensed per well whereas for mock 200 µl PGB was dispensed per well. The outer wells
871 were filled with water to minimize edge effects. Mycelia were harvested from PGA and then
872 they were washed with water to remove residual agar. Further, they were resuspended in water
873 with the concentration adjusted to 2 mg/ml. Later, they were homogenized using a Precellys
874 24 TissueLyser (Bertin Technologies) at 6,200 rpm for 2 × 30 s with a 15 s interval.
875 Subsequently, 20 µl of mycelial suspension was added to wells containing 180 µl PGB. Each
876 condition was represented by six wells per experiment and three independent biological
877 experiments were conducted. Optical density at 600 nm (OD600) was measured at 0, 24, 48,
878 72, and 96 h using microplate reader (Tecan, Infinite M Nano). For the agar plate assay, 20 µl
879 of mycelial suspension (2 mg/ml) was placed at the centre of PGA plates for each condition,
880 and mycelial growth were recorded after 10 days. For both the experiment, the samples were
881 incubated at 23°C.
882

883 ***Experiments with C. incanum***

884 *Targeted gene replacement: Agrobacterium*-mediated transformation was used based on
885 Hiruma et al., 2023[60]. Briefly, the upstream and downstream (~1500 bp each) regions

886 surrounding the *GH64* gene in *Colletotrichum incanum* MAFF 238704 (*Ci*) were amplified by
887 polymerase chain reaction (PCR) using the PrimeSTAR HS DNA polymerase (TaKaRa). PCR
888 products were electrophoresed on a 1% (w/v) agarose gel, then DNA was extracted from the
889 cut gel using the FastGene™ Gel/PCR Extraction Kit (NIPPON Genetics). The purified
890 upstream and downstream regions were mixed with SalI-treated pBIG4MRHrev using the In-
891 Fusion® HD Cloning Kit (TaKaRa). The generated plasmid was subsequently introduced to
892 *Agrobacterium tumefaciens* C58C1 by electroporation (1800 V, 5 ms) using the Eppendorf
893 Eporator. *Ci* spores were prepared by culturing agar plugs from Mathur's agar (3% w/v) plates
894 into Mathur's agar flasks, incubating for 2 weeks, then harvested by flooding the flask with GI
895 broth supplemented with 200 μM acetosyringone before washing twice with GI broth
896 supplemented with acetosyringone. Transformed *Agrobacterium* (OD_{600} = 0.4) was mixed 1:1
897 with Ci spores (10^7 spores/mL) on sterile paper filter discs on GI agar (1.5% w/v) supplemented
898 with 200 μM acetosyringone. Two days after transformation, filter paper discs were flipped onto
899 Mathur's agar plates supplemented with hygromycin (200 μg/mL), cefotaxime (100 μg/mL),
900 and spectinomycin (100 μg/mL) and incubated for 2 days at room temperature (~25°C) before
901 filter papers were removed. Colonies were allowed to merge, then were re-streaked onto fresh
902 Mathur's plates supplemented with hygromycin, cefotaxime, and spectinomycin (same
903 concentrations). Colony PCR was performed to validate the success of the gene replacement
904 using primers listed in **Supplementary Table 5**.
905 *Seed inoculation*: Agar plugs from fungal cultures streaked on Mathur's agar plates were
906 transferred to fresh Mathur's agar plates and incubated at 25°C for 7 days under continuous
907 light (~10 μmol m^{-2} s^{-1}). *A. thaliana* Col-0 or *Nicotiana benthamiana* seeds were surface
908 sterilized with 70% ethanol for 30 seconds twice, followed by 6% sodium hypochlorite with
909 0.01% Triton X for 5 minutes. After being washed 5 times with sterilized double-distilled water
910 (DW), seeds were sowed on half-strength Murashige and Skoog (MS) agar (1.3% w/v) plates
911 without sucrose. Spores from Mathur's plates were collected by flooding the fungal culture
912 plate with 5 mL DW, washed once in DW, then counted by hemocytometry in a light
913 microscope. Spores were diluted to 5,000 spores/mL prior to inoculating 3 μL of the spore
914 suspension 3 cm beneath seeds. Plates were incubated vertically in a plant growth chamber
915 at 22°C under fluorescent light (~100 μmol m^{-2} s^{-1}) with a 16-hour photoperiod at ~40%
916 humidity. Seedlings were harvested at 24 dpi for shoot fresh weight measurement. Roots were
917 sampled, flash frozen in liquid nitrogen, and stored at -80°C until RNA was extracted using the
918 Nucleospin® RNA kit (Macherey-Nagel) according to the manufacturer's instructions. RNA was
919 immediately used for cDNA synthesis, then stored at -80°C.
920 *Quantitative real-time polymerase chain reaction (qRT-PCR)*: cDNA was synthesized from
921 ~100 ng total RNA using the PrimeScript™ RT Master Mix (TaKaRa) in a volume of 10 μL
922 following manufacturer's instructions. qRT-PCR was performed using the THUNDERBIRD®

Next SYBR qPCR Mix (TOYOBO) with 250 nM primers using CFX Opus 384 Real-Time PCR System (BIORAD) in a volume of 10 µL for 48 cycles. Primers are listed in **Supplementary Table 5**.

Statistics

For the graphs and statistical analysis, the R programming environment (R version 4.4.0) was used for data analysis and visualization (https://www.r-project.org/). The normality of the data was assessed with the Shapiro–Wilk test using stats package (V4.4.0) and homogeneity of variances with Levene's test using car package (V3.1-2). The two-sided T-tests and two-sided non-parametric Wilcox rank-sum test were performed using t.test and wilcox.test from the stats package (V4.4.0) in R. ANOVA-TukeyHSD tests were conducted using emmeans package (V1.10.3). Non-parametric Kruskal-Wallis tests were carried out with stats package (V4.4.0), and Dunn's tests were performed using FSA (V0.9.5) and rcompanion packages (V2.4.36). Reported *P* values were Bonferroni adjusted using stats package (V4.4.0).

Data availability

RNASeq raw data are available at the European Nucleotide Archive (ENA) under the study ID PRJEB89666. The raw genomic data of the *Plectosphaerella* isolates have been deposited in the ENA database under the project IDs PRJEB61839 (P143), PRJEB50298 (P340, P421, P455) and PRJEB90544 (all other isolates). See **Supplementary Table 3** for individual ENA sample IDs.

Code availability

Scripts and raw data used to generate figures are available at https://github.com/fantin-mesny/Scripts_comparison_of_Plectosphaerella_strains for WGS, https://github.com/arpankbasak/Plecuc1_Repo. For RNAseq, and https://github.com/RamSevakRK/Genetic-factors-driving-multi-host-infection-in-a-core-member-of-the-root-mycobiota for all the other experiments.

Acknowledgements

This work was supported by funds to S.H. from European Research Council consolidator grants MICRORULES (758003) and MICROBIOSIS (101089198), the Cluster of Excellence on Plant Sciences (CEPLAS) under Germany's Excellence Strategy-EXC 2048/1-Project ID: 390686111, and the Priority Programme: Deconstruction and Reconstruction of the Plant Microbiota (SPP DECRyPT 2125; HA 8169/2-2), both funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). A.K.B. was funded by the

Polish National Agency for Academic Exchange NAWA (BPN/BEK/2023/1/00315). J.N. obtained funding from a JST SPRING grant (JPMJSP2108) and K.H. from a JST grant (JPMJAN23D4). J.G.M-V. is funded by the Beatriz Galindo program (BG23/00164) of the Spanish Ministry of Science, Innovation and Universities. We would like to thank Khushi Joshi (MPIPZ), Thorsten Thiergart (MPIPZ), and Katharina Grosche (Heinrich-Heine University) for the excellent technical assistance and Bart Thomma (University of Cologne) for providing protocols for fungal genome editing.

Author contributions

Conceptualization: S.H., F.M., and R-S.R-K. Supervision: S.H. Methodology and Investigation: R-S.R-K., A.B.K, F.M., S.H., J.N., G.C., F.E., L.R., S.C.A., M.P., and K.H. Analysis Support: T.L. and T.T. Sequencing Support: B.H. Contribution to biological material: P.W.C., J.G.M-V., H.S., M.R., H.R., A.M.F., S.S., A.M., I.B., S.D., W.H.E., J.H., J.S.K., J.M.N., and C.M.S. Visualization: R-S.R-K., A.B.K, F.M. and S.H. Writing original draft: S.H and R-S.R-K.

References

1. Bonfante P, Venice F, Lanfranco L. The mycobiota: fungi take their place between plants and bacteria. *Curr Opin Microbiol* **49**, 18-25 (2019).
2. Zeilinger S*, et al.* Friends or foes? Emerging insights from fungal interactions with plants. *FEMS Microbiol Rev* **40**, 182-207 (2016).
3. Almario J*, et al.* Root-associated fungal microbiota of nonmycorrhizal Arabis alpina and its contribution to plant phosphorus nutrition. *Proc Natl Acad Sci U S A* **114**, E9403-E9412 (2017).
4. Hiruma K*, et al.* Root Endophyte Colletotrichum tofieldiae Confers Plant Fitness Benefits that Are Phosphate Status Dependent. *Cell* **165**, 464-474 (2016).
5. Redkar A, Sabale M, Zuccaro A, Di Pietro A. Determinants of endophytic and pathogenic lifestyle in root colonizing fungi. *Curr Opin Plant Biol* **67**, 102226 (2022).
6. Netherway T*, et al.* Pervasive associations between dark septate endophytic fungi with tree root and soil microbiomes across Europe. *Nature Communications* **15**, (2024).
7. Kia SH, Glynou K, Nau T, Thines M, Piepenbring M, Macia-Vicente JG. Influence of phylogenetic conservatism and trait convergence on the interactions between fungal root endophytes and plants. *ISME J* **11**, 777-790 (2017).
8. Mesny F*, et al.* Genetic determinants of endophytism in the Arabidopsis root mycobiome. *Nat Commun* **12**, 7227 (2021).
9. Duran P*, et al.* Microbial Interkingdom Interactions in Roots Promote Arabidopsis Survival. *Cell* **175**, 973-983 e914 (2018).

10. Wolinska KW, *et al.* Tryptophan metabolism and bacterial commensals prevent fungal dysbiosis in Arabidopsis roots. *Proc Natl Acad Sci U S A* **118**, (2021).
11. Lahrmann U, *et al.* Mutualistic root endophytism is not associated with the reduction of saprotrophic traits and requires a noncompromised plant innate immunity. *New Phytol* **207**, 841-857 (2015).
12. Getzke F, Thiergart T, Hacquard S. Contribution of bacterial-fungal balance to plant and animal health. *Curr Opin Microbiol* **49**, 66-72 (2019).
13. Vetrovsky T, *et al.* A meta-analysis of global fungal distribution reveals climate-driven patterns. *Nat Commun* **10**, 5142 (2019).
14. Thiergart T, *et al.* Root microbiota assembly and adaptive differentiation among European Arabidopsis populations. *Nat Ecol Evol* **4**, 122-131 (2020).
15. Mikryukov V, *et al.* Connecting the multiple dimensions of global soil fungal diversity. *Sci Adv* **9**, eadj8016 (2023).
16. Talbot JM, *et al.* Endemism and functional convergence across the North American soil mycobiome. *Proc Natl Acad Sci U S A* **111**, 6341-6346 (2014).
17. Meiser A, Balint M, Schmitt I. Meta-analysis of deep-sequenced fungal communities indicates limited taxon sharing between studies and the presence of biogeographic patterns. *New Phytol* **201**, 623-635 (2014).
18. Egidi E, *et al.* A few Ascomycota taxa dominate soil fungal communities worldwide. *Nat Commun* **10**, 2369 (2019).
19. Fisher MC, *et al.* Emerging fungal threats to animal, plant and ecosystem health. *Nature* **484**, 186-194 (2012).
20. Gu XY, Cao ZY, Li ZQ, Yu HY, Liu WD. Plant immunity suppression by an β-1,3-glucanase of the maize anthracnose pathogen. *Bmc Plant Biol* **24**, (2024).
21. Liu H, *et al.* Plant immunity suppression by an exo-beta-1,3-glucanase and an elongation factor 1alpha of the rice blast fungus. *Nat Commun* **14**, 5491 (2023).
22. Tamburrini KC, *et al.* The disordered C-terminal tail of fungal LPMOs from phytopathogens mediates protein dimerization and impacts plant penetration. *Proc Natl Acad Sci U S A* **121**, e2319998121 (2024).
23. Maiti IB, Kolattukudy PE. Prevention of Fungal Infection of Plants by Specific-Inhibition of Cutinase. *Science* **205**, 507-508 (1979).
24. Eichfeld R, *et al.* Transcriptomics reveal a mechanism of niche defense: two beneficial root endophytes deploy an antimicrobial GH18-CBM5 chitinase to protect their hosts. *New Phytologist* **244**, 980-996 (2024).
25. Eitzen K, Sengupta P, Kroll S, Kemen E, Doehlemann G. A fungal member of the Arabidopsis thaliana phyllosphere antagonizes Albugo laibachii via a GH25 lysozyme. *Elife* **10**, (2021).

1033 26. Lebreton A, Zeng QC, Miyauchi S, Kohler A, Dai YC, Martin FM. Evolution of the Mode
1034 of Nutrition in Symbiotic and Saprotrophic Fungi in Forest Ecosystems. *Annu Rev Ecol*
1035 *Evol S* **52**, 385-404 (2021).
1036 27. Zhao ZT, Liu HQ, Wang CF, Xu JR. Comparative analysis of fungal genomes reveals
1037 different plant cell wall degrading capacity in fungi (vol 14, pg 274, 2013). *Bmc*
1038 *Genomics* **15**, (2014).
1039 28. Kameshwar AKS, Ramos LP, Qin WS. CAZymes-based ranking of fungi (CBRF): an
1040 interactive web database for identifying fungi with extrinsic plant biomass degrading
1041 abilities. *Bioresour Bioprocess* **6**, (2019).
1042 29. O'Connell RJ*, et al.* Lifestyle transitions in plant pathogenic Colletotrichum fungi
1043 deciphered by genome and transcriptome analyses. *Nat Genet* **44**, 1060-1065 (2012).
1044 30. Muñoz-Barrios A*, et al.* Differential Expression of Fungal Genes Determines the
1045 Lifestyle of *Plectosphaerella cucumerina* Strains During *Arabidopsis thaliana*
1046 Colonization. *Mol Plant Microbe In* **33**, 1299-1314 (2020).
1047 31. King BC, Waxman KD, Nenni NV, Walker LP, Bergstrom GC, Gibson DM. Arsenal of
1048 plant cell wall degrading enzymes reflects host preference among plant pathogenic
1049 fungi. *Biotechnol Biofuels* **4**, (2011).
1050 32. Kohler A*, et al.* Convergent losses of decay mechanisms and rapid turnover of
1051 symbiosis genes in mycorrhizal mutualists (vol 47, pg 410, 2015). *Nature Genetics* **47**,
1052 (2015).
1053 33. Miyauchi S*, et al.* Large-scale genome sequencing of mycorrhizal fungi provides
1054 insights into the early evolution of symbiotic traits. *Nature Communications* **11**, (2020).
1055 34. Hacquard S*, et al.* Survival trade-offs in plant roots during colonization by closely
1056 related beneficial and pathogenic fungi. *Nat Commun* **7**, 11362 (2016).
1057 35. Giraldo A, Crous PW. Inside Plectosphaerellaceae. *Stud Mycol* **92**, 227-286 (2019).
1058 36. Sanchez-Vallet A*, et al.* Disruption of abscisic acid signaling constitutively activates
1059 Arabidopsis resistance to the necrotrophic fungus *Plectosphaerella cucumerina*. *Plant*
1060 *Physiol* **160**, 2109-2124 (2012).
1061 37. Sanchez-Vallet A*, et al.* Tryptophan-derived secondary metabolites in Arabidopsis
1062 thaliana confer non-host resistance to necrotrophic Plectosphaerella cucumerina fungi.
1063 *Plant J* **63**, 115-127 (2010).
1064 38. Carlucci A, Raimondo ML, Santos J, Phillips AJ. Plectosphaerella species associated
1065 with root and collar rots of horticultural crops in southern Italy. *Persoonia* **28**, 34-48
1066 (2012).
1067 39. Ramos B, González-Melendi P, Sánchez-Vallet A, Sánchez-Rodríguez C, López G,
1068 Molina A. Functional genomics tools to decipher the pathogenicity mechanisms of the

1069 necrotrophic fungus *Plectosphaerella cucumerina* in *Arabidopsis thaliana*. *Mol Plant*
1070 *Pathol* **14**, 44-57 (2013).
1071 40. C. ER. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon
1072 sequencing. *bioRxiv*, (2016).
1073 41. Wang Q, Cole JR. Updated RDP taxonomy and RDP Classifier for more accurate
1074 taxonomic classification. *Microbiol Resour Ann* **13**, (2024).
1075 42. Molina A*, et al.* Arabidopsis cell wall composition determines disease resistance
1076 specificity and fitness. *Proc Natl Acad Sci U S A* **118**, (2021).
1077 43. Vetrovsky T*, et al.* GlobalFungi, a global database of fungal occurrences from high-
1078 throughput-sequencing metabarcoding studies. *Sci Data* **7**, 228 (2020).
1079 44. Nelson JM, Hauser DA, Hinson R, Shaw AJ. A novel experimental system using the
1080 liverwort Marchantia polymorpha and its fungal endophytes reveals diverse and
1081 context-dependent effects. *New Phytol* **218**, 1217-1232 (2018).
1082 45. Ku HM, Vision T, Liu JP, Tanksley SD. Comparing sequenced segments of the tomato
1083 and *Arabidopsis* genomes: Large-scale duplication followed by selective gene loss
1084 creates a network of synteny. *P Natl Acad Sci USA* **97**, 9121-9126 (2000).
1085 46. Chaw SM, Chang CC, Chen HL, Li WH. Dating the monocot-dicot divergence and the
1086 origin of core eudicots using whole chloroplast genomes. *J Mol Evol* **58**, 424-441
1087 (2004).
1088 47. Zheng JF, Ge QW, Yan YC, Zhang XP, Huang L, Yin YB. dbCAN3: automated
1089 carbohydrate-active enzyme and substrate annotation. *Nucleic Acids Res* **51**, W115-
1090 W121 (2023).
1091 48. Zhang H*, et al.* dbCAN2: a meta server for automated carbohydrate-active enzyme
1092 annotation. *Nucleic Acids Res* **46**, W95-W101 (2018).
1093 49. Vogel J. Unique aspects of the grass cell wall. *Current Opinion in Plant Biology* **11**,
1094 301-307 (2008).
1095 50. van Kempen M*, et al.* Fast and accurate protein structure search with Foldseek. *Nat*
1096 *Biotechnol* **42**, 243-246 (2024).
1097 51. Ma JW, Jiang ZQ, Yan QJ, Lv A, Li YX, Yang SQ. Structural and functional analysis of
1098 SpGlu64A: a novel glycoside hydrolase family 64 laminaripentaose-producing β-1,3-
1099 glucanase from. *Febs J* **291**, 2009-2022 (2024).
1100 52. Wanke A*, et al.* Plant species-specific recognition of long and short β-1,3-linked
1101 glucans is mediated by different receptor systems. *Plant Journal* **102**, 1142-1156
1102 (2020).
1103 53. Melida H*, et al.* Non-branched beta-1,3-glucan oligosaccharides trigger immune
1104 responses in Arabidopsis. *Plant J* **93**, 34-49 (2018).

1105 54. Rebaque D*, et al.* Cell wall-derived mixed-linked β-1,3/1,4-glucans trigger immune
1106 responses and disease resistance in plants. *Plant Journal* **106**, 601-615 (2021).
1107 55. Barghahn S, Arnal G, Jain N, Petutschnig E, Brumer H, Lipka V. Mixed Linkage β-
1108 1,3/1,4-Glucan Oligosaccharides Induce Defense Responses in *Hordeum vulgare* and
1109 *Arabidopsis thaliana*. *Frontiers in Plant Science* **12**, (2021).
1110 56. Sato TM, T.; Imamura, Y.; Nojima, H.; Moriwaki, J.; Yaguchi, Y. . Anthracnose of
1111 Japanese radish caused by Colletotrichum dematium. *J Gen Plant Pathol* **71**, 380-383
1112 (2005).
1113 57. Snelders NC, Petti GC, van den Berg GCM, Seidl MF, Thomma B. An ancient
1114 antimicrobial protein co-opted by a fungal plant pathogen for in planta mycobiome
1115 manipulation. *Proc Natl Acad Sci U S A* **118**, (2021).
1116 58. Snelders NC*, et al.* Microbiome manipulation by a soil-borne fungal plant pathogen
1117 using effector proteins. *Nat Plants* **6**, 1365-1374 (2020).
1118 59. Snelders NC, Rovenich H, Thomma B. Microbiota manipulation through the secretion
1119 of effector proteins is fundamental to the wealth of lifestyles in the fungal kingdom.
1120 *FEMS Microbiol Rev* **46**, (2022).
1121 60. Hiruma K*, et al.* A fungal sesquiterpene biosynthesis gene cluster critical for mutualist-
1122 pathogen transition in. *Nature Communications* **14**, (2023).
1123 61. Ren Ujimatsu JT, Seishiro Aoki, Masami Nakamura, Hiromi Haba, Atsushi Minami, Kei
1124 Hiruma A fungal transcription factor converts a beneficial root endophyte into an
1125 anthracnose leaf pathogen. *Current Biology* **35**, 1989-2005.e1986 (2025).
1126 62. Blanco-Ulate B, Morales-Cruz A, Amrine KC, Labavitch JM, Powell AL, Cantu D.
1127 Genome-wide transcriptional profiling of Botrytis cinerea genes targeting plant cell
1128 walls during infections of different hosts. *Front Plant Sci* **5**, 435 (2014).
1129 63. Allan J, Regmi R, Denton-Giles M, Kamphuis LG, Derbyshire MC. The host generalist
1130 phytopathogenic fungus Sclerotinia sclerotiorum differentially expresses multiple
1131 metabolic enzymes on two different plant hosts. *Sci Rep* **9**, 19966 (2019).
1132 64. Fesel PH, Zuccaro A. beta-glucan: Crucial component of the fungal cell wall and elusive
1133 MAMP in plants. *Fungal Genet Biol* **90**, 53-60 (2016).
1134 65. Rovenich H, Zuccaro A, Thomma BP. Convergent evolution of filamentous microbes
1135 towards evasion of glycan-triggered immunity. *New Phytol* **212**, 896-901 (2016).
1136 66. Molina A, Sanchez-Vallet A, Jorda L, Carrasco-Lopez C, Rodriguez-Herva JJ, Lopez-
1137 Solanilla E. Plant cell walls: source of carbohydrate-based signals in plant-pathogen
1138 interactions. *Curr Opin Plant Biol* **82**, 102630 (2024).
1139 67. van Boerdonk S, Saake P, Wanke A, Neumann U, Zuccaro A. beta-Glucan-binding
1140 proteins are key modulators of immunity and symbiosis in mutualistic plant-microbe
1141 interactions. *Curr Opin Plant Biol* **81**, 102610 (2024).

1142 68. Oliveira-Garcia E, Deising HB. Attenuation of PAMP-triggered immunity in maize
1143 requires down-regulation of the key beta-1,6-glucan synthesis genes KRE5 and KRE6
1144 in biotrophic hyphae of Colletotrichum graminicola. *Plant J* **87**, 355-375 (2016).
1145 69. Wawra S*, et al.* The fungal-specific beta-glucan-binding lectin FGB1 alters cell-wall
1146 composition and suppresses glucan-triggered immunity in plants. *Nat Commun* **7**,
1147 13188 (2016).
1148 70. Rose JK, Ham KS, Darvill AG, Albersheim P. Molecular cloning and characterization of
1149 glucanase inhibitor proteins: coevolution of a counterdefense mechanism by plant
1150 pathogens. *Plant Cell* **14**, 1329-1345 (2002).
1151 71. Wang HF*, et al.* Insect fungal pathogens secrete a cell wall-associated glucanase that
1152 acts to help avoid recognition by the host immune system. *Plos Pathogens* **19**, (2023).
1153 72. Chen XY, Kim JY. Callose synthesis in higher plants. *Plant Signal Behav* **4**, 489-492
1154 (2009).
1155 73. Wang B, Andargie M, Fang RQ. The function and biosynthesis of callose in high plants.
1156 *Heliyon* **8**, (2022).
1157 74. Nishimura MT, Stein M, Hou BH, Vogel JP, Edwards H, Somerville SC. Loss of a
1158 callose synthase results in salicylic acid-dependent disease resistance. *Science* **301**,
1159 969-972 (2003).
1160 75. Millet YA*, et al.* Innate immune responses activated in Arabidopsis roots by microbe-
1161 associated molecular patterns. *Plant Cell* **22**, 973-990 (2010).
1162 76. Wu SW, Kumar R, Iswanto ABB, Kim JY. Callose balancing at plasmodesmata. *J Exp*
1163 *Bot* **69**, 5325-5339 (2018).
1164 77. Caporaso JG*, et al.* QIIME allows analysis of high-throughput community sequencing
1165 data. *Nat Methods* **7**, 335-336 (2010).
1166 78. Edgar RC. Search and clustering orders of magnitude faster than BLAST.
1167 *Bioinformatics* **26**, 2460-2461 (2010).
1168 79. Deshpande V*, et al.* Fungal identification using a Bayesian classifier and the Warcup
1169 training set of internal transcribed spacer sequences. *Mycologia* **108**, 1-5 (2016).
1170 80. Virtanen P*, et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python.
1171 *Nat Methods* **17**, 261-272 (2020).
1172 81. Schoch CL*, et al.* NCBI Taxonomy: a comprehensive update on curation, resources
1173 and tools. *Database (Oxford)* **2020**, (2020).
1174 82. Sakamoto T, Ortega JM. Taxallnomy: an extension of NCBI Taxonomy that produces
1175 a hierarchically complete taxonomic tree. *BMC Bioinformatics* **22**, 388 (2021).
1176 83. Bengtsson-Palme J*, et al.* Improved software detection and extraction of ITS1 and ITS2
1177 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of
1178 environmental sequencing data. *Methods Ecol Evol* **4**, 914-919 (2013).

1179 84. Tedersoo L*, et al.* The Global Soil Mycobiome consortium dataset for boosting fungal
1180 diversity research. *Fungal Divers* **111**, 573-588 (2021).
1181 85. Grigoriev IV*, et al.* MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic*
1182 *Acids Res* **42**, D699-D704 (2014).
1183 86. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing
1184 reads. *EMBnetjournal* **17**, 10–12 (2011).
1185 87. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using
1186 repeat graphs. *Nat Biotechnol* **37**, 540-546 (2019).
1187 88. Solovyev V, Kosarev P, Seledsov I, Vorobyev D. Automatic annotation of eukaryotic
1188 genes, pseudogenes and promoters. *Genome Biol* **7 Suppl 1**, S10 11-12 (2006).
1189 89. Huang L*, et al.* dbCAN-seq: a database of carbohydrate-active enzyme (CAZyme)
1190 sequence and annotation. *Nucleic Acids Res* **46**, D516-D521 (2018).
1191 90. Teufel F*, et al.* SignalP 6.0 predicts all five types of signal peptides using protein
1192 language models. *Nature Biotechnology* **40**, 1023-+ (2022).
1193 91. Sperschneider J, Dodds PN. EffectorP 3.0: Prediction of Apoplastic and Cytoplasmic
1194 Effectors in Fungi and Oomycetes. *Mol Plant Microbe Interact* **35**, 146-156 (2022).
1195 92. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-
1196 mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at
1197 the Metagenomic Scale. *Mol Biol Evol* **38**, 5825-5829 (2021).
1198 93. Huerta-Cepas J*, et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically
1199 annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic*
1200 *Acids Res* **47**, D309-D314 (2019).
1201 94. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND.
1202 *Nat Methods* **12**, 59-60 (2015).
1203 95. Rawlings ND, Barrett AJ, Finn R. Twenty years of the MEROPS database of proteolytic
1204 enzymes, their substrates and inhibitors. *Nucleic Acids Res* **44**, D343-350 (2016).
1205 96. Riehl K, Riccio C, Miska EA, Hemberg M. TransposonUltimate: software for transposon
1206 classification, annotation and detection. *Nucleic Acids Res*, (2022).
1207 97. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative
1208 genomics. *Genome Biol* **20**, 238 (2019).
1209 98. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7:
1210 improvements in performance and usability. *Mol Biol Evol* **30**, 772-780 (2013).
1211 99. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated
1212 alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973
1213 (2009).
1214 100. Minh BQ*, et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic
1215 Inference in the Genomic Era. *Mol Biol Evol* **37**, 1530-1534 (2020).

1216 101. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder:
1217 fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**, 587-589
1218 (2017).
1219 102. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable
1220 and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*
1221 **35**, 4453-4455 (2019).
1222 103. Borowiec ML. AMAS: a fast tool for alignment manipulation and computing of summary
1223 statistics. *PeerJ* **4**, e1660 (2016).
1224 104. Zhang C, Sayyari E, Mirarab S. ASTRAL-III: Increased Scalability and Impacts of
1225 Contracting Low Support Branches. *Lect N Bioinformat* **10562**, 53-75 (2017).
1226 105. Getzke F*, et al.* Cofunctioning of bacterial exometabolites drives root microbiota
1227 establishment. *Proc Natl Acad Sci U S A* **120**, e2221508120 (2023).
1228 106. Sarkar D*, et al.* The inconspicuous gatekeeper: endophytic Serendipita vermifera acts
1229 as extended plant protection barrier in the rhizosphere. *New Phytol* **224**, 886-901
1230 (2019).
1231 107. Sewe SO, Silva G, Sicat P, Seal SE, Visendi P. Trimming and Validation of Illumina
1232 Short Reads Using Trimmomatic, Trinity Assembly, and Assessment of RNA-Seq Data.
1233 *Methods Mol Biol* **2443**, 211-232 (2022).
1234 108. Krueger Fea. FelixKrueger/TrimGalore: Version 0.6.8. (Jan. 28, 2023). *Zenodo*.
1235 109. Cheng CY, Krishnakumar V, Chan AP, Thibaud-Nissen F, Schobel S, Town CD.
1236 Araport11: a complete reannotation of the Arabidopsis thaliana reference genome.
1237 *Plant J* **89**, 789-804 (2017).
1238 110. Goodstein DM*, et al.* Phytozome: a comparative platform for green plant genomics.
1239 *Nucleic Acids Res* **40**, D1178-1186 (2012).
1240 111. Aoki K*, et al.* Large-scale analysis of full-length cDNAs from the tomato (Solanum
1241 lycopersicum) cultivar Micro-Tom, a reference system for the Solanaceae genomics.
1242 *Bmc Genomics* **11**, 210 (2010).
1243 112. Beier S*, et al.* Construction of a map-based reference genome sequence for barley,
1244 Hordeum vulgare L. *Sci Data* **4**, 170044 (2017).
1245 113. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-
1246 aware quantification of transcript expression. *Nat Methods* **14**, 417-419 (2017).
1247 114. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level
1248 estimates improve gene-level inferences. *F1000Res* **4**, 1521 (2015).
1249 115. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
1250 RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).

1251 116. Zhu A, Ibrahim JG, Love MI. Heavy-tailed prior distributions for sequence count data:
1252 removing the noise and preserving large differences. *Bioinformatics* **35**, 2084-2092
1253 (2019).
1254 117. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-
1255 customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 35 (2011).
1256 118. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in
1257 multidimensional genomic data. *Bioinformatics* **32**, 2847-2849 (2016).
1258 119. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The
1259 Carbohydrate-Active EnZymes database (CAZy): an expert resource for
1260 Glycogenomics. *Nucleic Acids Res* **37**, D233-238 (2009).
1261 120. Csárdi G NT, Traag V, Horvát S, Zanini F, Noom D, Müller K igraph: Network Analysis
1262 and Visualization in R. doi:10.5281/zenodo.7682609, R package version 2.1.4,
1263 https://CRAN.R-project.org/package=igraph. . (2025).
1264 121. Yeats TV, T.; Sorek, N.; Ibáñez, A.B.; Bauer, S. Rapid Determination of Cellulose,
1265 Neutral Sugars, and Uronic Acids from Plant Cell Walls by One-step Two-step
1266 Hydrolysis and HPAEC-PAD. *Bio-protocol* **6**, (2016).
1267 122. Foster CE, Martin TM, Pauly M. Comprehensive compositional analysis of plant cell
1268 walls (lignocellulosic biomass) part II: carbohydrates. *J Vis Exp*, (2010).
1269 123. Sherrill-Mix S. taxonomizr: Functions to Work with NCBI Accessions and Taxonomy. R
1270 package version 0.11.1, https://cran.r-project.org/web/packages/taxonomizr. . (2017).
1271 124. Nguyen NH*, et al.* FUNGuild: An open annotation tool for parsing fungal community
1272 datasets by ecological guild. *Fungal Ecol* **20**, 241-248 (2016).
1273 125. Letunic I, Bork P. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic
1274 tree display and annotation tool. *Nucleic Acids Res* **52**, W78-W82 (2024).
1275 126. Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the
1276 TM-score. *Nucleic Acids Res* **33**, 2302-2309 (2005).
1277 127. Wang J*, et al.* iCn3D: From Web-Based 3D Viewer to Structural Analysis Tool in Batch
1278 Mode. *Front Mol Biosci* **9**, 831740 (2022).
1279 128. Stemmer M, Thumberger T, Del Sol Keyer M, Wittbrodt J, Mateo JL. CCTop: An
1280 Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One* **10**,
1281 e0124633 (2015).
1282 129. Mozsik L, Hoekzema M, de Kok NAW, Bovenberg RAL, Nygard Y, Driessen AJM.
1283 CRISPR-based transcriptional activation tool for silent genes in filamentous fungi. *Sci*
1284 *Rep* **11**, 1118 (2021).
1285 130. Rehman L, Su XF, Guo HM, Qi XL, Cheng HM. Protoplast transformation as a potential
1286 platform for exploring gene function in. *Bmc Biotechnol* **16**, (2016).

131. Leisen T*, et al.* CRISPR/Cas with ribonucleoprotein complexes and transiently selected telomere vectors allows highly efficient marker-free and multiple genome editing in. *Plos Pathogens* **16**, (2020).
132. Abramson J*, et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).
133. Emms DM, & Kelly, S. . STAG: Species Tree Inference from All Genes. . *bioRxiv, doi: https://doiorg/101101/267914* (2018).
134. Corpet F. Multiple Sequence Alignment with Hierarchical-Clustering. *Nucleic Acids Res* **16**, 10881-10890 (1988).

Main Figures

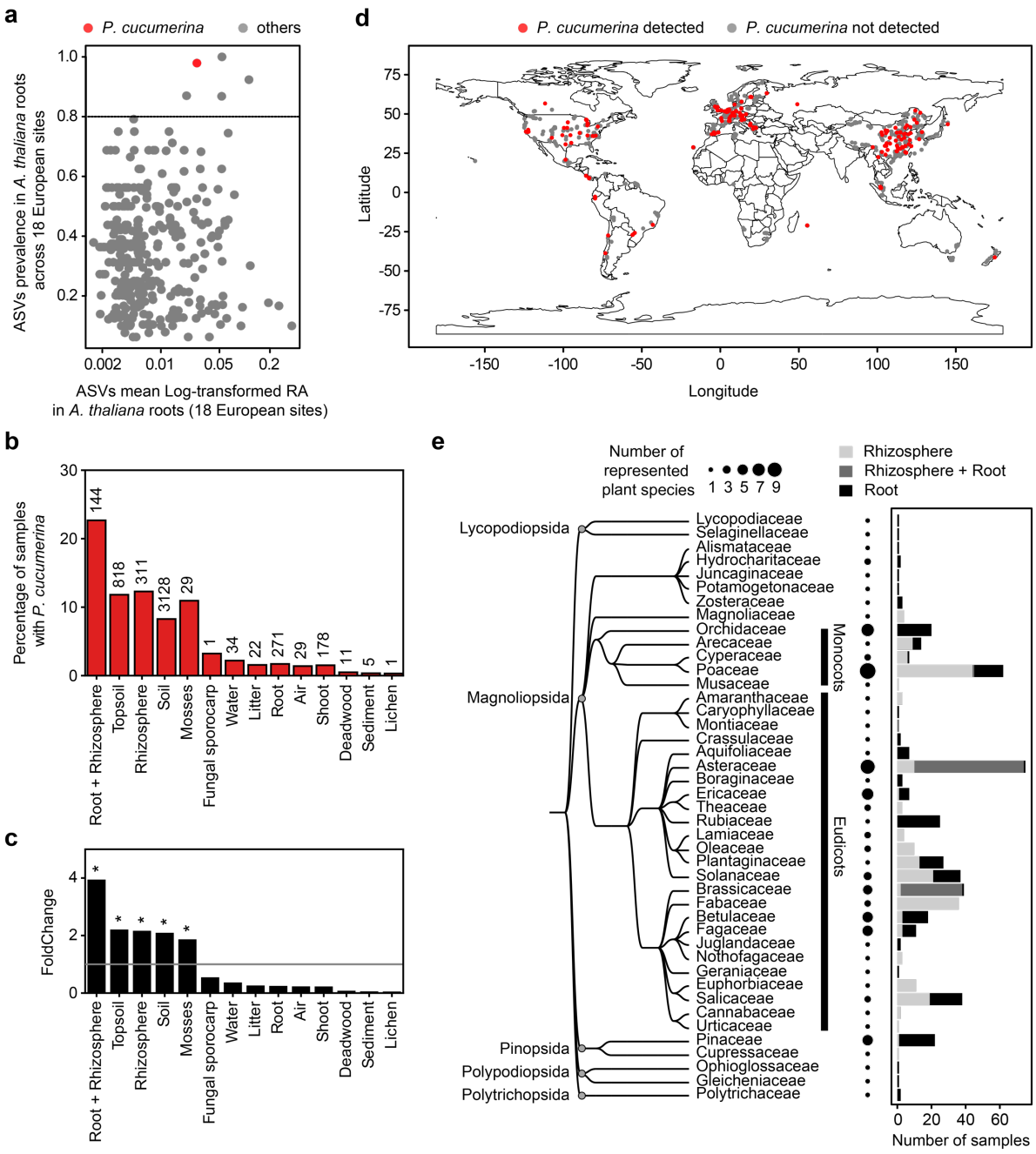


Fig. 1: *P. cucumerina* is a core member of the *A. thaliana* root mycobiota and a widespread multi-host root colonizer. a, Re-analysis of amplicon sequence variants (i.e., zero radius OTUs, Zotus) sequenced from *A. thaliana* root samples (n = 281) from two previous studies[9, 14]. Zotus are plotted according to their prevalence and mean log-transformed relative abundance (RA) in roots across all samples. RA was calculated across samples where Zotus were detected (minimum 0.1% per sample). Only samples having more than 1,000 reads were considered for the analysis. The five Zotus above the dashed line corresponds to the most widespread root-colonizing taxa. *P. cucumerina* zOTU is highlighted in red. **b**, Analysis of GlobalFungi data[43] showing detection of *P. cucumerina* in a variety of niches (n = 4,982/84,844 samples). Percentage of samples in which the *P. cucumerina* species was detected are plotted

for each niche, with the corresponding numbers of samples annotated on the barplot. **c**, Fold-change values corresponding to the proportion of samples from a given niche in which *P. cucumerina* was detected divided by the proportion of samples in all other niches in which *P. cucumerina* was detected. Asterisks highlight the niches in which detection of *P. cucumerina* is significantly overrepresented (pairwise one-sided Fisher's exact tests + Benjamini-Hochberg correction) **d**, Geographic origin of root and rhizosphere samples of the GlobalFungi database (n = 19,030 samples), with samples in which species *P. cucumerina* was detected in red (n = 726 samples). **e**, Phylogeny (family level) of the plant hosts associated to the GlobalFungi root and rhizosphere samples in which *P. cucumerina* was detected (n = 726 samples). The phylogenetic tree was assembled based on species names and NCBI Taxonomy data using Taxallnomy[82]. A bubble plot shows the number of species represented in each family. The stacked barplot highlights the number of samples associated to this plant host family in which *P. cucumerina* was detected, distinguishing root, rhizosphere and root + rhizosphere samples.

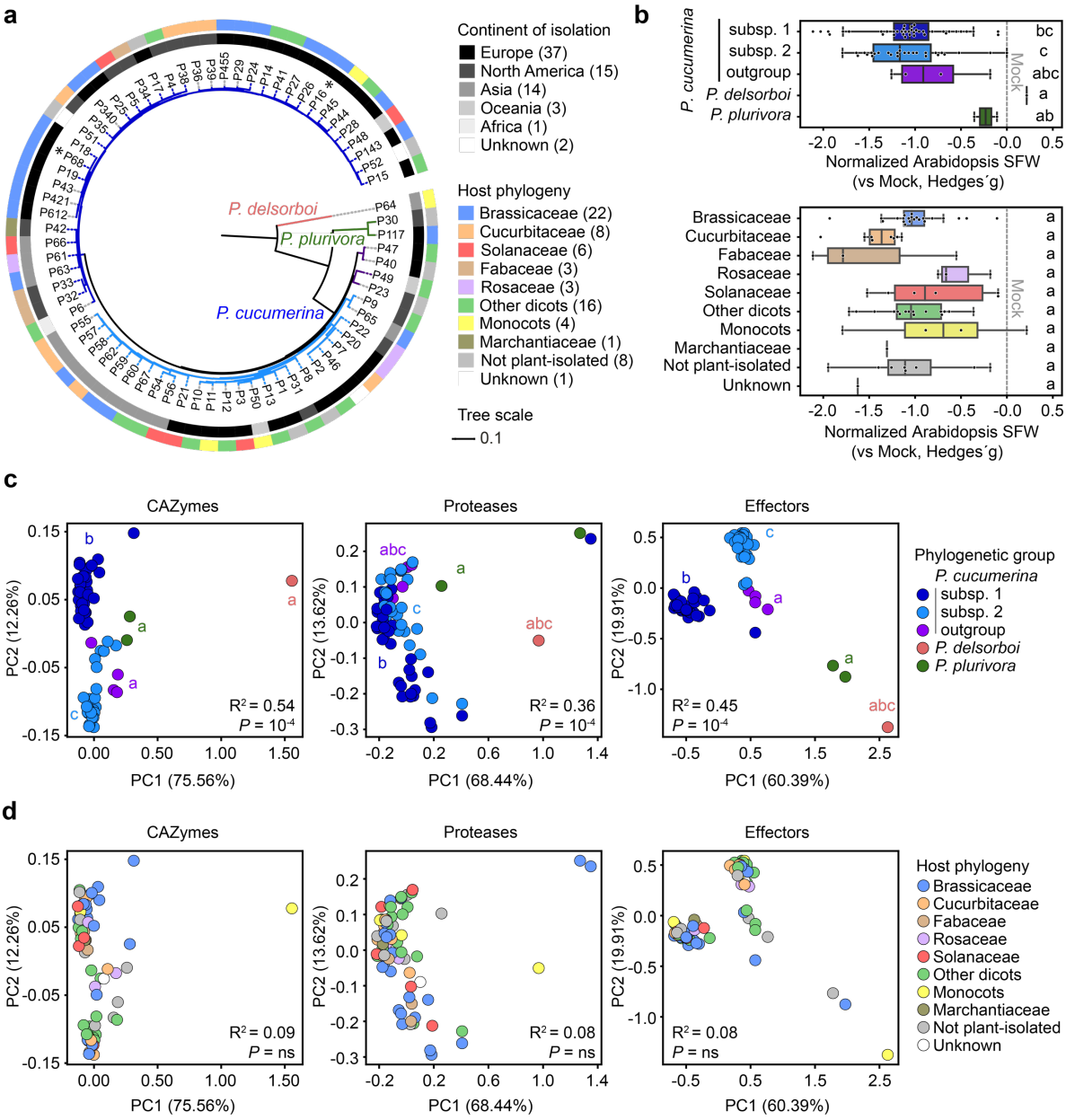


Fig. 2: *P. cucumerina* strains isolated from phylogenetically diverse plants do not show host preference signature. **a**, Phylogeny of the 72 *Plectosphaerella* isolates, annotated with their continent of isolation (inner circle) and the phylogenetic group of their host (outer circle). The phylogenetic tree was reconstructed from the concatenated alignments of 5,466 single copy orthologs (see Methods). Two *P. cucumerina* subspecies can be distinguished and are depicted in two shades of blue whereas four strains do not belong to any of these subspecies and are depicted in violet (later referred to as "outgroup"). Isolates P68 and P16 that respectively correspond to *P. cucumerina* BMM (PcBMM, see previous work[39]) and *P. cucumerina* 0016 (see this work) are highlighted with an asterisk. **b**, Individual fungal effects on *A. thaliana* Col-0 plants organized according to the phylogenetic group of *Plectosphaerella* strains (top), or family names of the hosts of isolation (bottom). Standard effect sizes (Hedges'

g) were calculated to compare the shoot fresh weights of 28-day old inoculated and mock-treated plants (agar-based gnotobiotic system, n = 12-20 plants per fungal conditions). Letters describe differences between categories according to a ANOVA-TukeyHSD post-hoc test, performed after ANOVA testing. **c-d**, Principal Coordinate Analyses profiling inter-strain differences in genomic repertoires of CAZymes, proteases and candidate effectors across all *Plectosphaerella* isolates (n = 72 strains). The PCoA plots are color-coded according to *Plectosphaerella* phylogenetic group (**c**) or family names of the hosts of isolation (**d**). R^2 and *P* resulting from PERMANOVA tests assessing the effect of these two categories on the three genomic repertoires are shown for each graph. Letters indicate significant differences between groups based on pairwise permutation MANOVAs followed by *P* value adjustment using the Benjamini-Hochberg method.

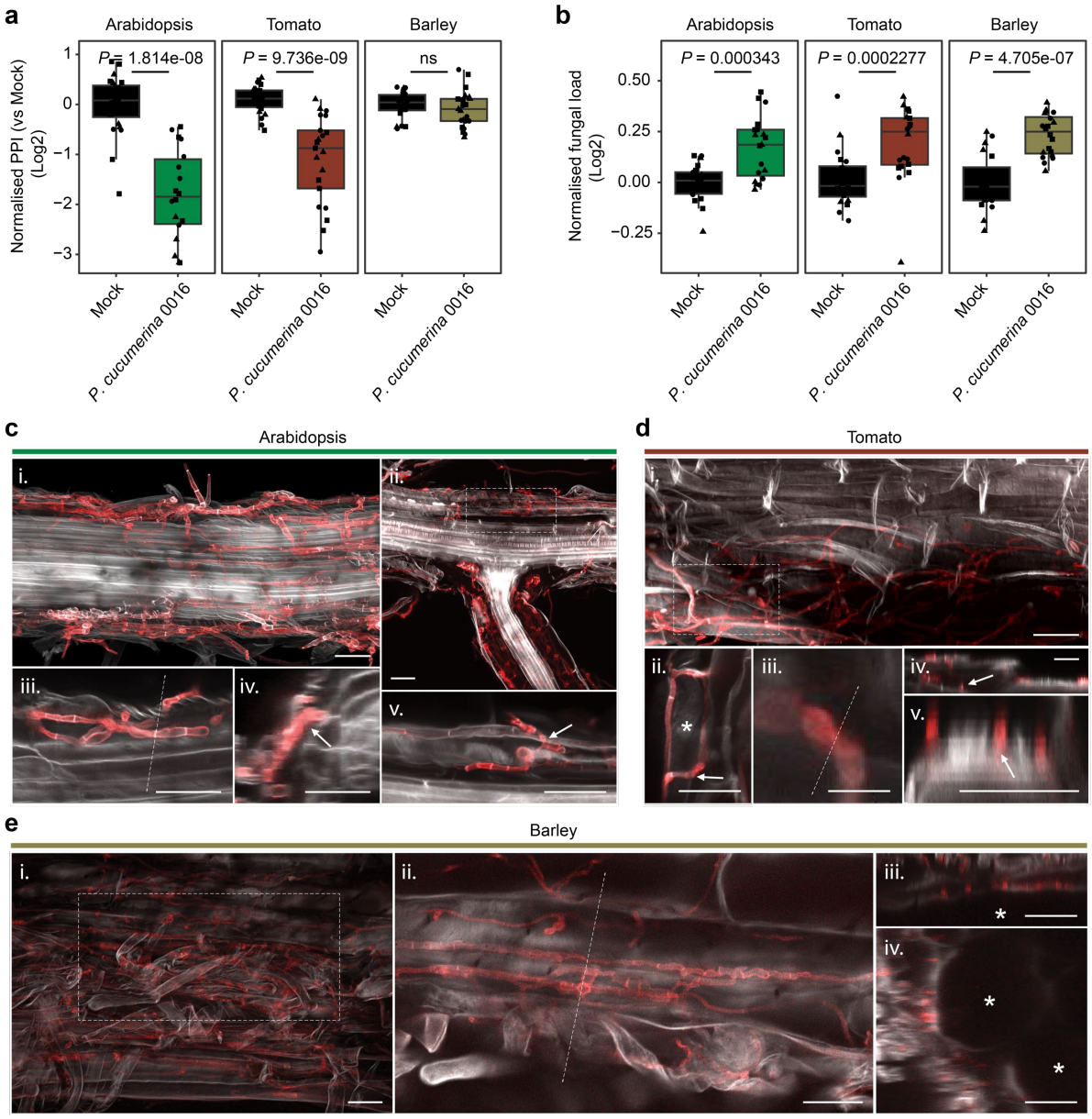
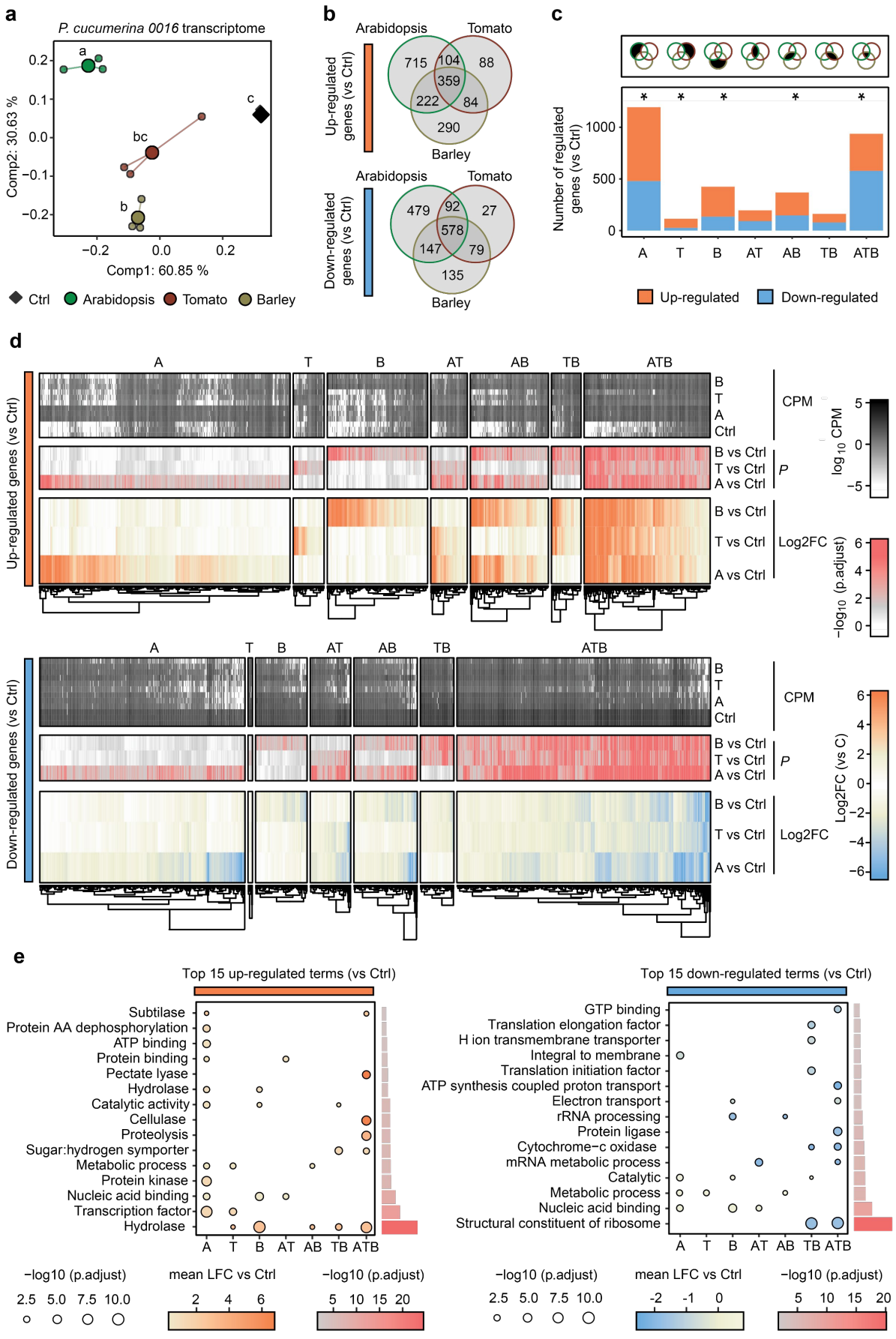


Fig. 3: Colonization of *P. cucumerina* 0016 in Arabidopsis, tomato and barley roots and impact on host performance: a, Plant Performance Indices (PPI) of Arabidopsis, tomato, and barley colonized by *P. cucumerina* 0016 were normalized to mock-treated plants and log2 transformed. PPI values were calculated by multiplying the survival rate by the shoot fresh weight of the plants. **b**, Relative abundance of *P. cucumerina* 0016 in the roots of Arabidopsis, tomato, and barley measured by qPCR-based quantification of the fungal ITS *vs* plant *UBQ* gene. **a-b**, Pre-grown seedlings (2-week-old Arabidopsis and tomato and a week-old barley) were inoculated with *P. cucumerina* 0016 and harvested 21 days post fungal inoculation (vermiculite-based gnotobiotic system). Three independent biological replicates are depicted (circles, triangles, squares), each consisting of 4 – 8 plants (n = 12 – 24 plants per condition). Pairwise comparisons of plant performance between mock-treated and *P. cucumerina* 0016 treated plants were performed using either the two-sided t-test or Wilcoxon rank-sum test ($P \leq$

0.05). **c**-**e**, Confocal imaging of Arabidopsis (**c**), tomato (**d**), barley (**e**) roots colonized by *P. cucumerina* 0016 21 days post fungal inoculation. Roots were double-stained with Calcoflour white and Wheat Germ Agglutinin coupled to fluorophore CF633A (WGACF633; Biotium) and imaged using Zeiss LSM 880 confocal microscopy. **c**, i. Maximum intensity projection of Z-stack images of an Arabidopsis primary root extensively colonized. ii. Maximum intensity projection of middle section of the primary root from which a lateral root emerges. iii. Enlargement of one optical section from (i), iv. Orthogonal projection of (iii) showing undifferentiated hyphae penetrating peridermal cell layer (arrow). v. Enlargement of one optical section shown in (ii) (dashed square) depicting an intracellular hypha colonizing the root cortex (arrow). **d**, i. Maximum intensity projection of a middle section of a primary root. ii. Enlargement and Z-stack image of (i) (dashed square) showing intercellular hyphae (arrow) surrounding an epidermal cell (star). iii. Enlargement of one optical section shown in (ii) (arrow). iv and v. Orthogonal projection of (ii). Undifferentiated hyphae adhere to the epidermal cell (arrows). **e**, i. Maximum intensity projection of the middle section of the root. ii. Z-stack slice of (i) (dashed square) showing the colonization of fungal hyphae in the epidermal layer of the barley root. iii and iv. Orthogonal projection of (ii) showing hyphae adhering to epidermal cells (stars).


a
P. cucumerina 0016 transcriptome
Comp2: 30.63 %
Comp1: 60.85 %
Ctrl
Arabidopsis
Tomato
Barley
b
Arabidopsis
Tomato
Barley
Up-regulated genes (vs Ctrl)
Down-regulated genes (vs Ctrl)
c
Number of regulated genes (vs Ctrl)
A
T
B
AT
AB
TB
ATB
Up-regulated
Down-regulated
d
Up-regulated genes (vs Ctrl)
Down-regulated genes (vs Ctrl)
CPM
P
Log2FC
B vs Ctrl
T vs Ctrl
A vs Ctrl
log10 CPM
−log10 (p.adjust)
Log2FC (vs C)
e
Top 15 up-regulated terms (vs Ctrl)
Subtilase
Protein AA dephosphorylation
ATP binding
Protein binding
Pectate lyase
Hydrolase
Catalytic activity
Cellulase
Proteolysis
Sugar:hydrogen symporter
Metabolic process
Protein kinase
Nucleic acid binding
Transcription factor
Hydrolase
Top 15 down-regulated terms (vs Ctrl)
GTP binding
Translation elongation factor
H ion transmembrane transporter
Integral to membrane
Translation initiation factor
ATP synthesis coupled proton transport
Electron transport
rRNA processing
Protein ligase
Cytochrome−c oxidase
mRNA metabolic process
Catalytic
Metabolic process
Nucleic acid binding
Structural constituent of ribosome
−log10 (p.adjust)
mean LFC vs Ctrl

1470 **Fig. 4: *In planta* transcriptome of P. *cucumerina* 0016 is modulated by host plant identity.**
1471 **a**, Multidimensional Scaling (MDS) plot of Pearson distances computed on normalized read
1472 counts of differentially expressed genes (DEGs, n = 3,399, *in planta* vs *ex planta*). The different
1473 plant hosts are shown in green (Arabidopsis), red (tomato), and khaki (barley). *Ex planta*
1474 control condition is depicted in black (Ctrl, membrane). Large data points represent the mean
1475 values of the respective datasets in MDS1 and MDS2, whereas small data points represent
1476 the three biological replicates. Multiple statistical comparisons across conditions were
1477 performed using TukeyHSD test ($P \leq 0.05$). **b**, Venn diagram illustrating the number of DEGs
1478 (|log2FC| ≥ 1 and FDR ≤ 0.05) that are up-regulated (orange) and down-regulated (blue) *in*
1479 *planta vs* Ctrl condition. **c**, Stacked bar graph of fungal DEGs that are up- and down-regulated
1480 *in planta.* The categories A, T, and B denote fungal genes uniquely up- or down-regulated in
1481 Arabidopsis, tomato, or barley roots, respectively. AT, AB, and TB denote genes commonly
1482 up- or down-regulated in two hosts (Arabidopsis and tomato, Arabidopsis and barley, or tomato
1483 and barley, respectively), while ATB denotes genes consistently up- or down-regulated across
1484 all hosts. Categories that are enriched in up- or down-regulated DEGs *in-planta* are marked
1485 with an asterisk (*, proportion test with Pearson's chi-square). **d**, Heatmaps depicting DEGs
1486 that are up-regulated (orange) and down-regulated (blue), respectively. Bottom heatmaps
1487 depict log2FC of DEGs that are hierarchically clustered along the y-axis, with corresponding
1488 conditions displayed along the x-axis. Middle heatmaps depict *adj. P* of DEGs (Benjamini-
1489 Hochberg correction, -log10). Top heatmaps show counts per million of DEGs (log10 scale).
1490 **e**, Gene Ontology (GO) term enrichment analysis. Top 15 significantly enriched GO terms that
1491 are up (left) and down regulated (right), respectively, *in planta vs* Ctrl. The x-axis represents
1492 A, T, B, AT, AB, TB, and ATB sectors (see above) and the y-axis the GO terms. The color of
1493 each sphere represents the mean log2FC of DEGs, while the size of the sphere indicates the
1494 mean *adj. P* within each GO category for the respective conditions. The bar graph illustrates
1495 the cumulative *adj. P* for the respective GO. These values were computed using the
1496 hypergeometric test with Benjamini-Hochberg correction ($P \leq 0.05$).
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506

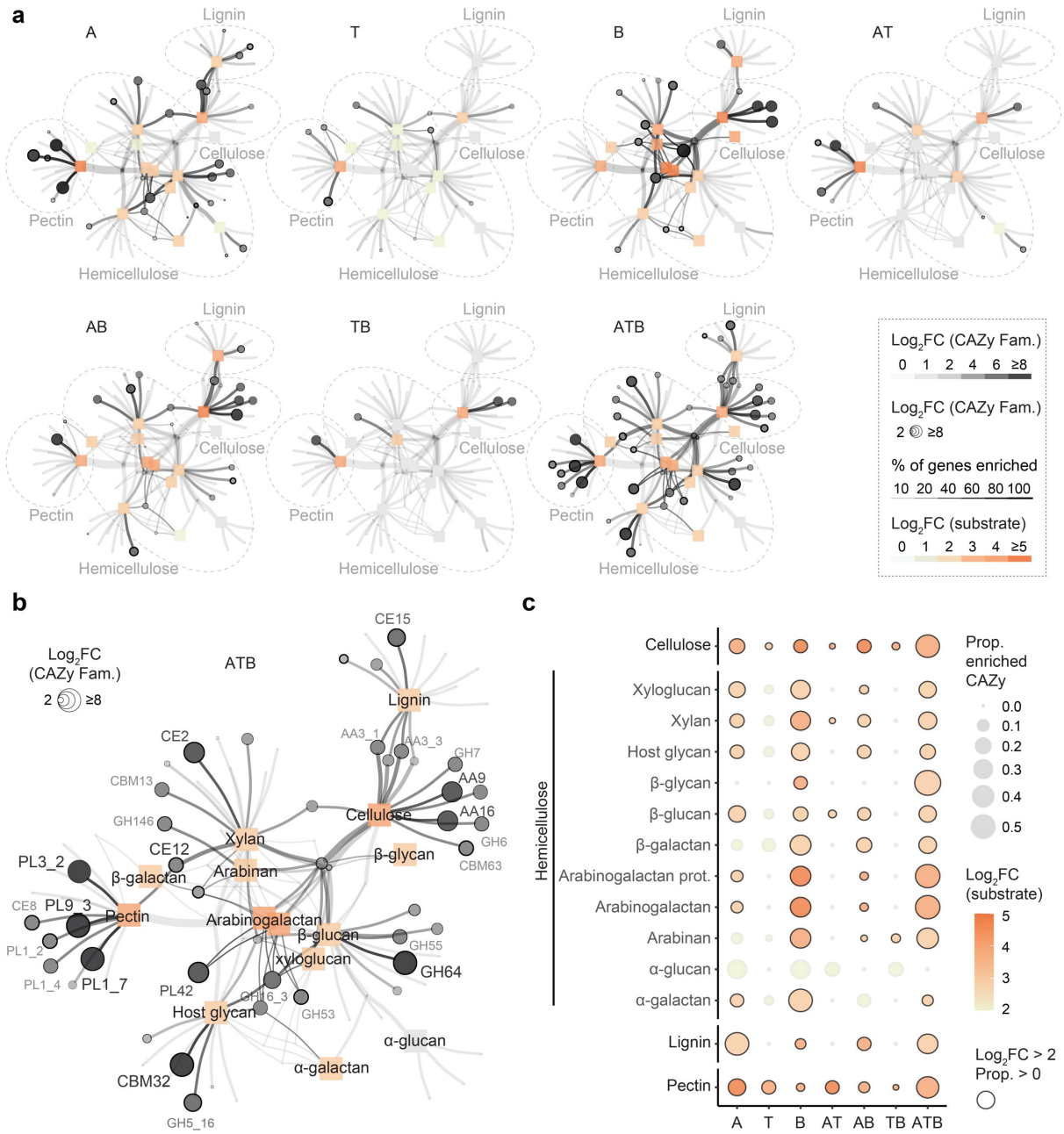


Fig. 5: Host-specific and core sets of *P. cucumerina* 0016 CAZyme-encoding genes induced during root colonisation of different hosts. a-b, The networks illustrate CAZyme families acting on plant and plant/fungal cell walls (circles) alongside with their potential plant cell wall substrates (n = 14 substrates, squares) defined based on the CAZyDB[119] and dbCAN-sub databases[47]. The color gradient of the nodes and edges (white to black) and the size of the circles reflect the mean Log2FC of statistically enriched CAZyme family *in planta* vs *ex planta* (Fisher's exact test) for each DEG subset (**a,** A, T, B, AT, AB, TB, ATB **b**, enlargement of the ATB network, see **Fig. 4c**). For each CAZyme family, the thickness of the circles indicates the proportion of genes induced among the DEGs across conditions. Predicted substrates targeted by CAZymes are shown using a grey-to-orange gradient that depicts the

1519 mean Log2FC of the enriched CAZymes targeting specific substrates across conditions.
1520 Substrates represented in grey squares are classified as non-significant (with a proportion of
1521 enriched CAZymes equals to 0). **c,** Balloon plot illustrating the utilization of substrates through
1522 the differential expression of fungal CAZymes based on their log2FC *in planta* vs *ex planta*.
1523 The colors within the circles (yellow-to-orange) depicts the mean log2FC of the CAZymes
1524 predicted to target primary cell wall substrates such as cellulose, diverse hemicellulose
1525 components, lignin, or pectin (see y-axis) for the different DEGs subsets (see x-axis, A, T, B,
1526 AT, AB, TB, and ATB). The size of the circles corresponds to the proportion of DEGs within
1527 the CAZyme family that target the different cell wall substrates (y-axis). Circles marked in black
1528 represent groups that show a mean log2FC along with a proportion of enriched CAZymes > 0
1529 and in grey where the proportion of enriched CAZymes equal to 0, as determined by Fisher's
1530 exact test.
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555

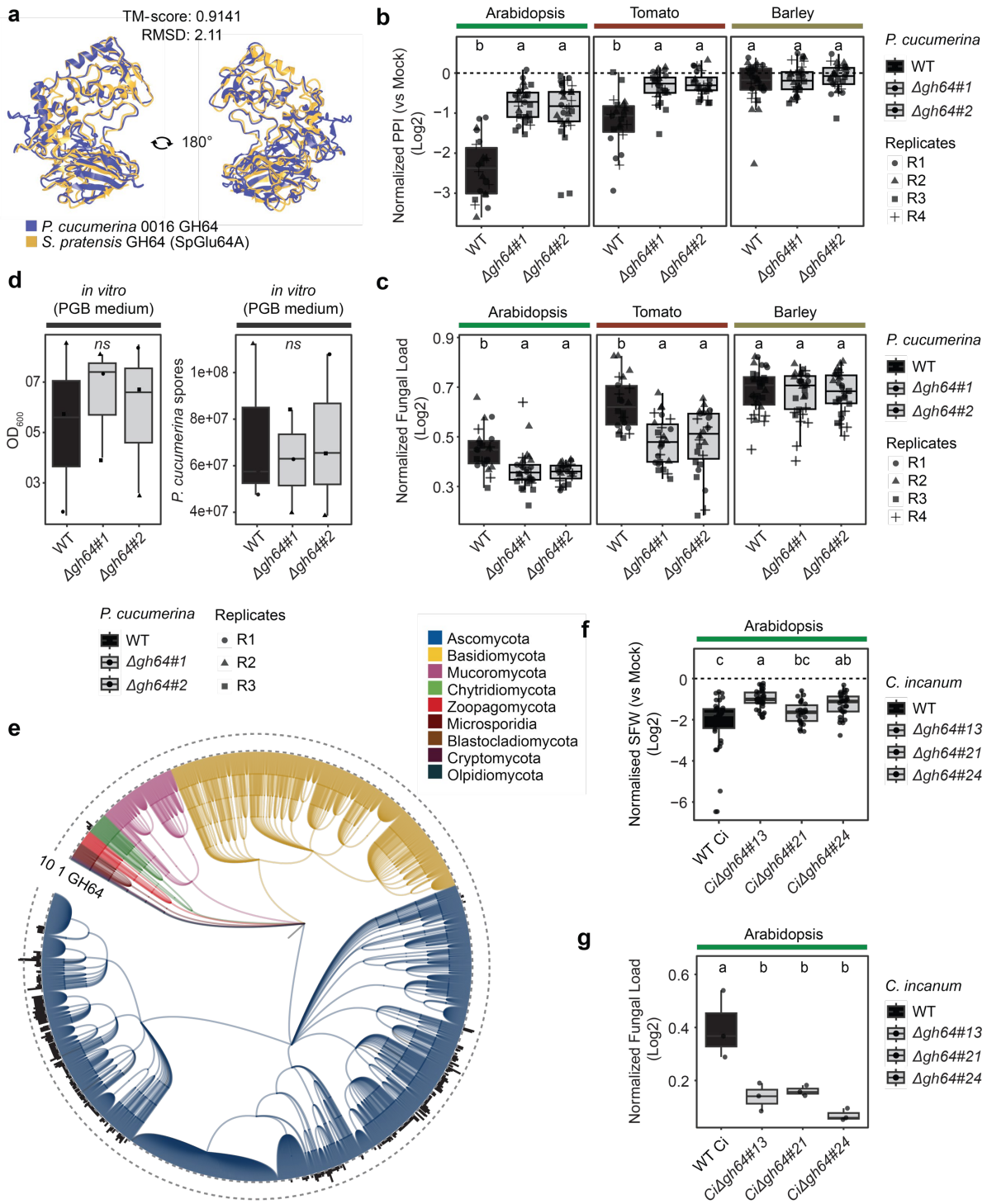
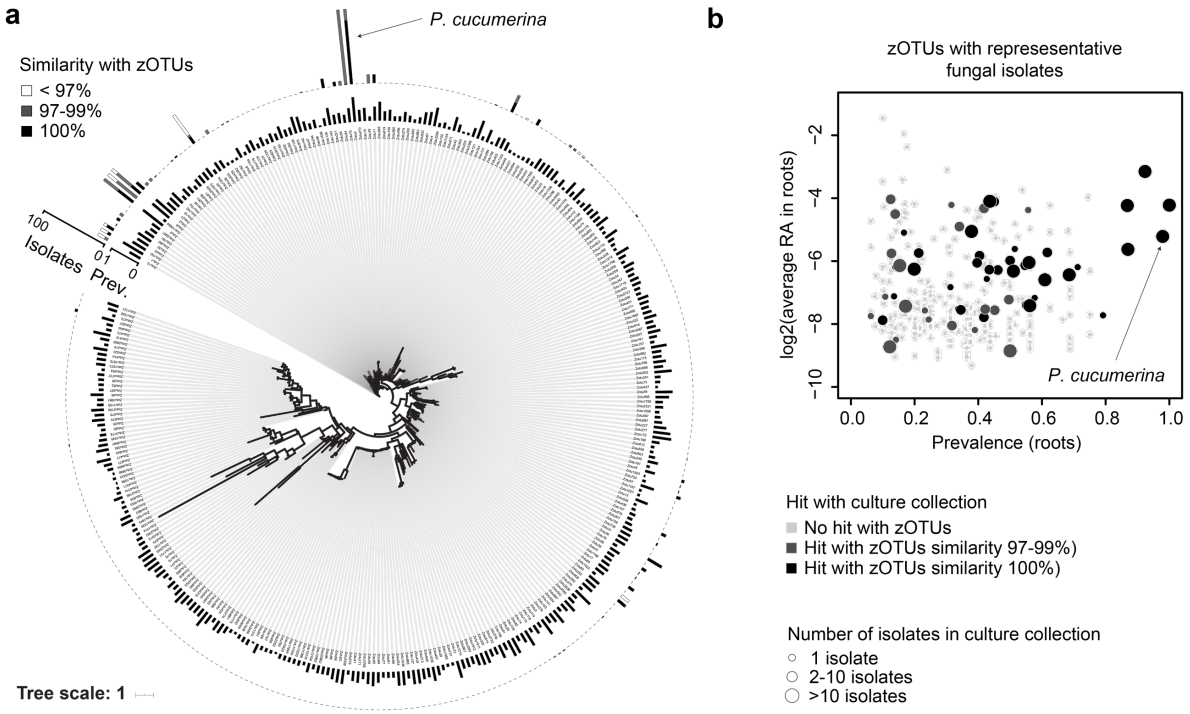


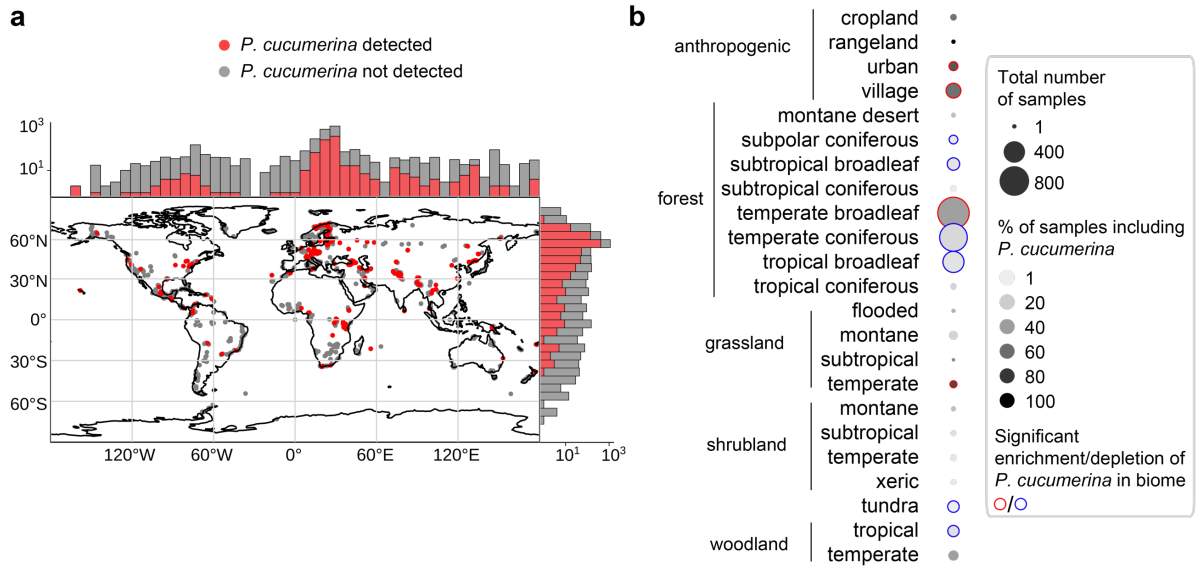
Fig. 6 Phenotype assessment of *P. cucumerina* 0016 *and C. incanum Δgh64* mutant strains: a, AlphaFold3-predicted structure[132] of *P. cucumerina* 0016 GH64 (in blue; pTM=0.92) showing high similarity to the previously resolved structure of the GH64 enzyme SpGlu64A of *Streptomyces pratensis*[51] (in orange), according to FoldSeek search analysis[50]. **b,** Plant Performance Indices (PPI, see method) of Arabidopsis, tomato, and barley colonized by *P. cucumerina* 0016 were normalized to mock-treated plants and log2 transformed. PPI values were calculated by multiplying the germination rate by the shoot fresh weight of the plants. **c,**

1565 Relative abundance of *P. cucumerina* 0016 in the roots of Arabidopsis, tomato, and barley
1566 measured by qPCR-based detection of the fungal ITS vs. plant *UBQ10* gene. **b-c**, Pre-grown
1567 seedlings (2-week-old Arabidopsis and tomato and a week-old barley) were inoculated with *P.*
1568 *cucumerina* 0016 and harvested 21 days post fungal inoculation. Four independent biological
1569 replicates are depicted (circles, triangles, squares, and plus), each consisting of 5 – 8 plants
1570 (n = 20 – 32 plants per condition). Multiple pairwise comparisons of plant performance between
1571 mock-treated and *P. cucumerina* 0016 treated plants were performed using Kruskal-Wallis-
1572 Dunn Test, adjusted with Bonferroni correction ($P \leq 0.05$). **d,** In vitro growth assessment of *P.*
1573 *cucumerina* 0016 WT and *Δgh64* mutant strains. An initial spore inoculum of 8x10^4 spores/mL
1574 was used and the fungus was inoculated on PGB medium for 24 hours. OD_{600} was measured
1575 after 24 hours using a spectrophotometer (left) and yield of *P. cucumerina* 0016 spores were
1576 counted after 24 hours using a hemocytometer. **e,** Phylogenetic tree constructed using 2,534
1577 fungal genomes from the MycoCosm[85] (JGI) with NCBI taxonomy indices. The colors of the
1578 clade indicate the fungal phylum. The barplot at the leaf nodes denotes the number of GH64
1579 occurrences that are identified in the genomes (CAZyme family-GH64 from JGI mycocosm). **f,**
1580 Shoot fresh weight (SFW) of *A. thaliana* plants colonized by *Colletotrichum incanum* wild type
1581 (WT *Ci*) and three independent mutant lines (*CiΔgh64#13*, *#21*, *#24*) normalized to mock-
1582 treated plants (dashed line) and log2 transformed. Plants (n = 26 - 36 per condition) were
1583 harvested at 24 dpi. Statistical significance was determined by ANOVA-TukeyHSD, $P < 0.05$.
1584 **g,** Relative load of WT *Ci* and mutant strains in the roots of *A. thaliana* measured by qRT-PCR-
1585 based detection of *C. incanum* ACTIN transcripts normalized to Arabidopsis ACTIN transcripts
1586 at 24 dpi (n = 3 replicates per conditions, 8 - 12 plants were pooled per replicate). Statistical
1587 significance was determined by ANOVA-TukeyHSD, $P < 0.05$.
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601

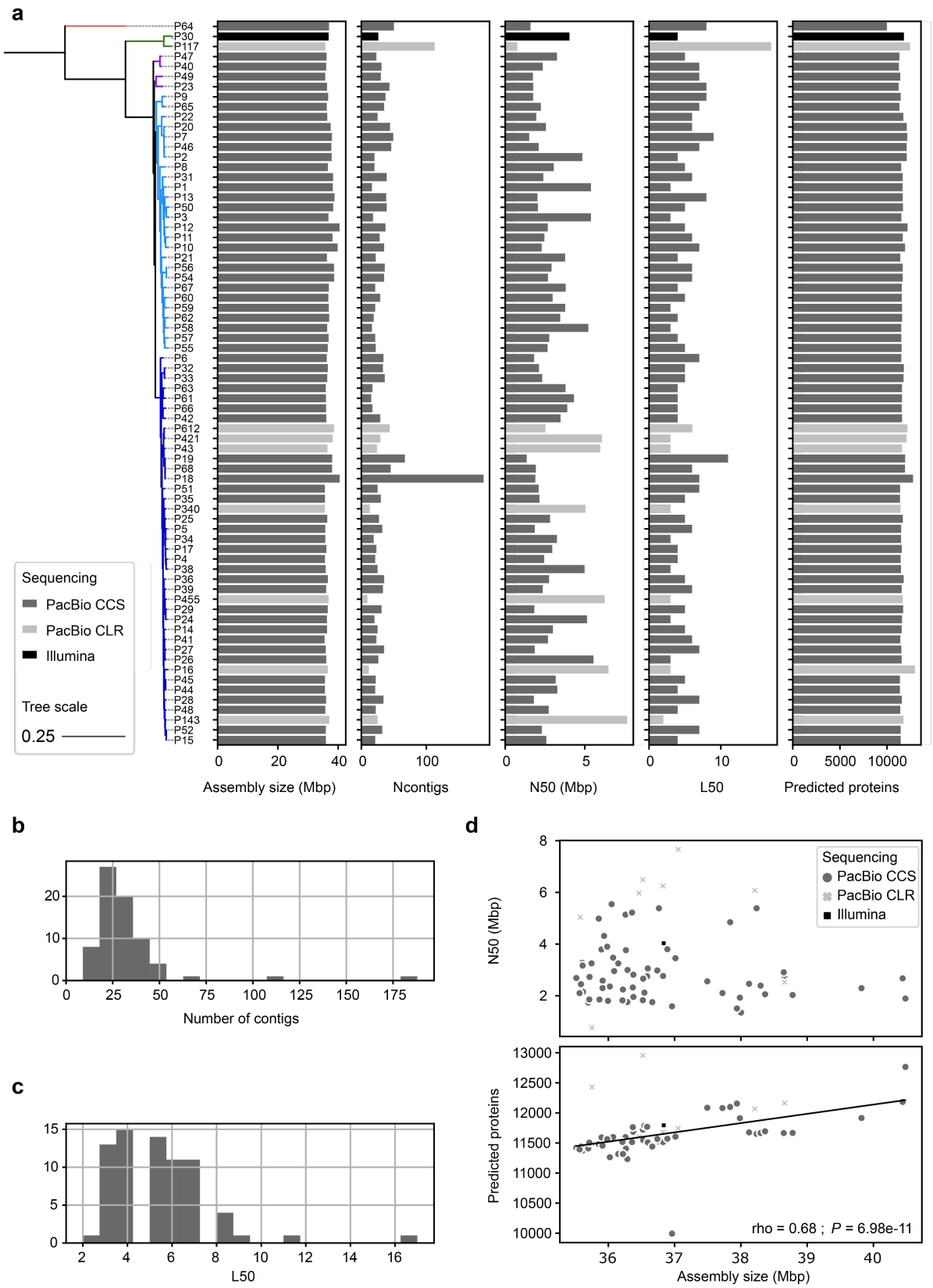
Supplementary Figures



Supplementary Fig. 1. Core fungal taxa detected in Arabidopsis roots across Europe and cross-reference with fungal culture collections. a, Maximum likelihood tree based on ITS1 sequences of 338 zero-radius OTUs (Zotus) detected in root samples retrieved from two previous studies (n = 291 samples). This includes *A. thaliana* root samples harvested from 17 European sites and from the Cologne agricultural soil. The inner circle depicts zOTU prevalence across18 sites. The outer barplot depicts Zotus for which representative isolates match culture collection isolates (n = 488 strains, deriving from 4 sites) based on three sequence similarity thresholds (< 97%, 97-99% and 100%). **b,** Zotus are plotted according their mean log2 relative abundance (RA) across all root samples (n = 291 samples), relative to their prevalence across all sample sites (n = 18). RA is only calculated across samples where Zotus were present (i.e., representing at least 0.1% RA per samples). Only root samples with > 1,000 reads were considered for this analysis. Zotus that have hits against the ITS of cultured fungal isolates are colored (dark grey: 97-99% sequence similarity, black: 100% sequence similarity). The size of the circle depicts the number of hits (ranging from 1 isolate to more than 10 isolates).

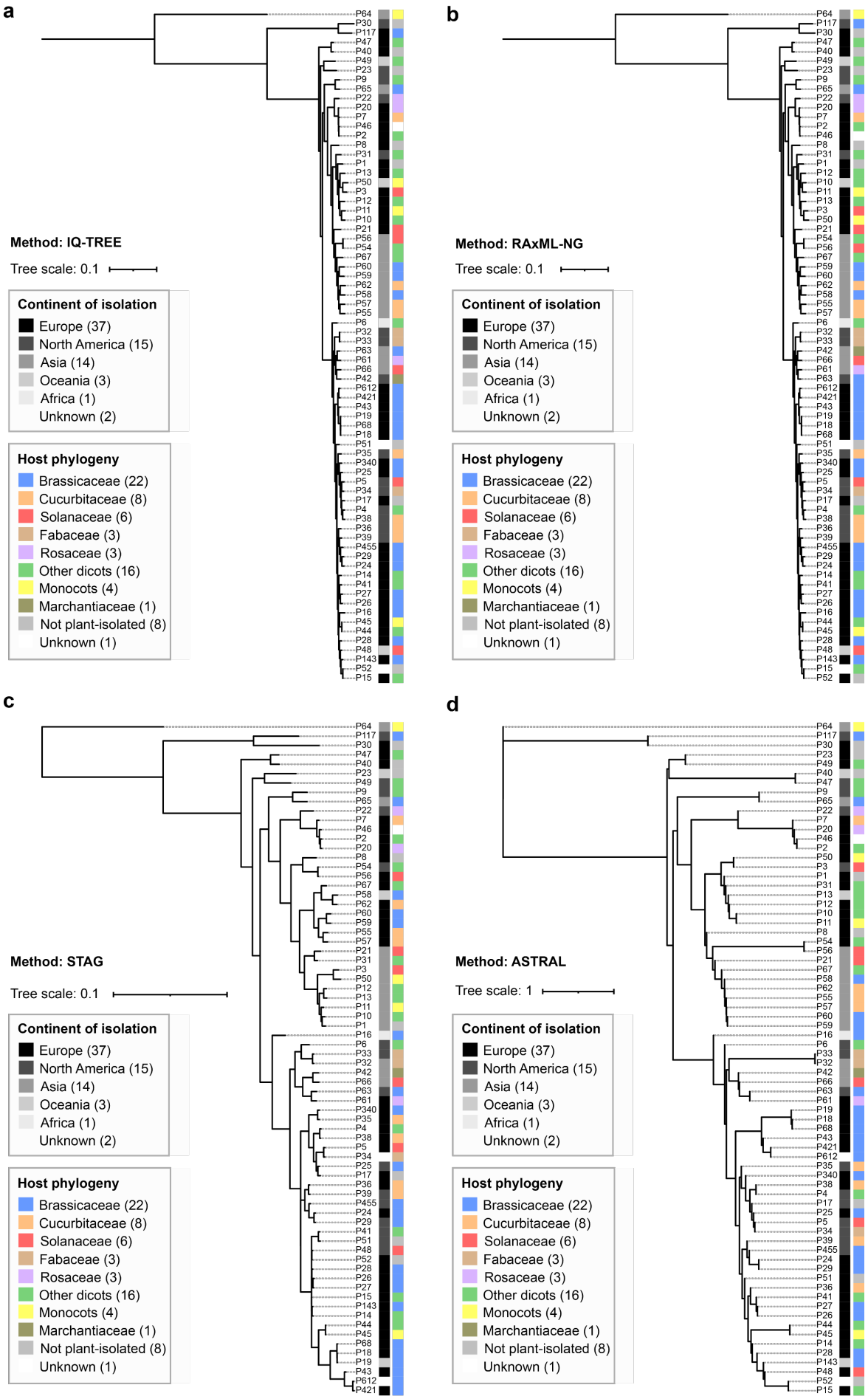


Supplementary Fig. 2: *P. cucumerina* is a ubiquitous soil-borne fungus **a,** Detection of *P. cucumerina* (presence of curated ITS1 sequences[35]) in soil samples originating from all continents (n = 3,194 samples based on Global Soil Mycobiome dataset[84]). **b,** Soil biome enrichment/depletion in samples containing *P. cucumerina*, as tested by Fisher's exact test (see Methods) using the Global Soil Mycobiome dataset and the metadata associated to soil samples.



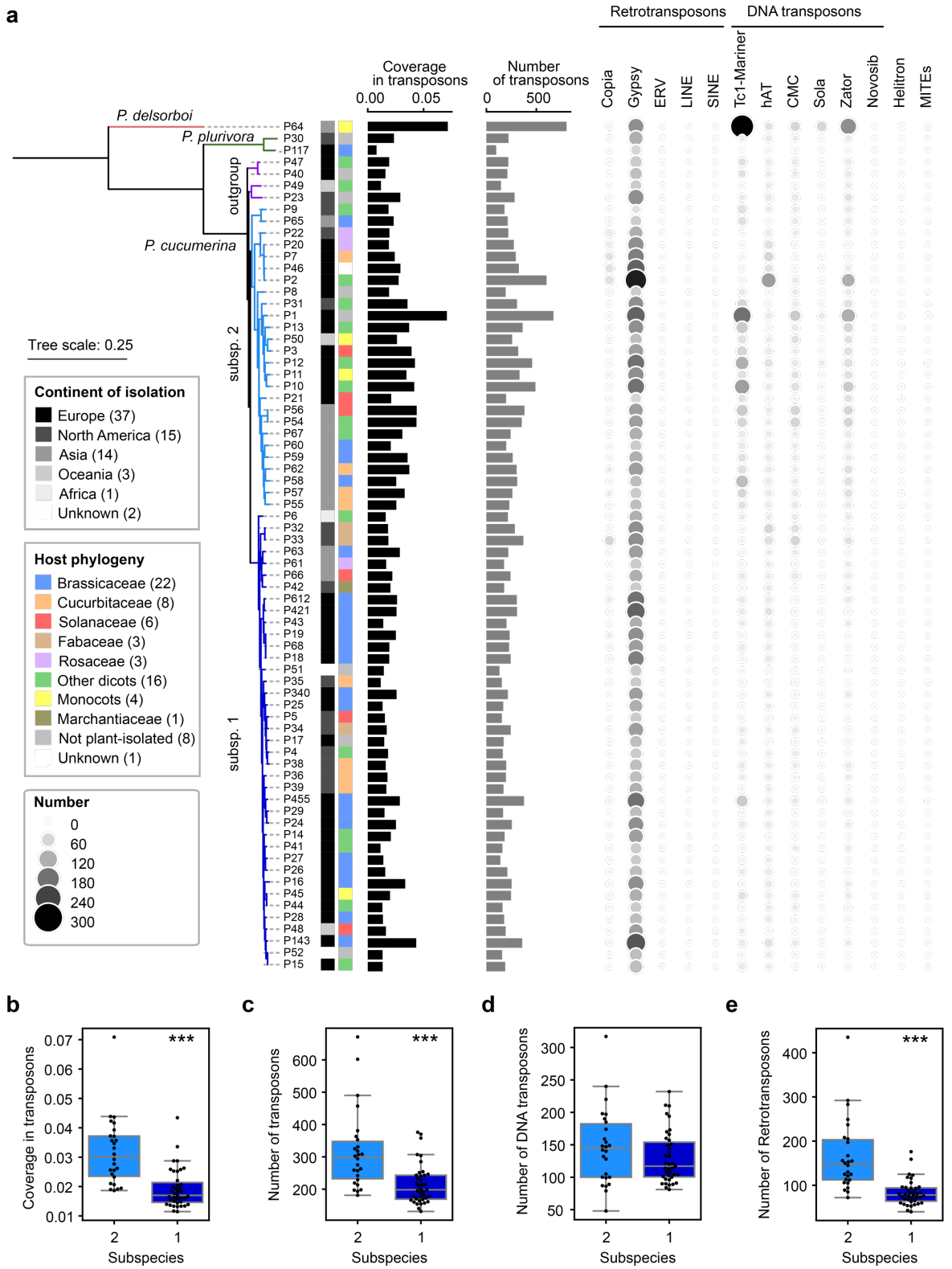
Supplementary Fig. 3: Assembly size and quality for 72 *Plectosphaerella* genomes. a, For each genome in our genomic dataset, the following values are plotted: assembly size, number of contigs in the assembly, N50 and L50 values, number of predicted genes in the

assembly. These values are represented along the *Plectosphaerella* collection phylogeny reconstructed with IQ-TREE (See Methods). Note that isolates P68 and P16 correspond respectively to *P. cucumerina* BMM (PcBMM) and *P. cucumerina* 0016 (see this work). **b,** Distribution of the number of contigs in the genomic data set (n = 72, one value per genome assembly). **c,** Distribution of the L50 values in the genomic data set (n = 72). **d,** N50 values (top) and numbers of predicted genes (bottom), plotted in function of assembly size. The number of predicted genes is significantly correlated to the assembly size according to Spearman's rank correlation.


a
Method: IQ-TREE
Tree scale: 0.1
Continent of isolation
Europe (37)
North America (15)
Asia (14)
Oceania (3)
Africa (1)
Unknown (2)
Host phylogeny
Brassicaceae (22)
Cucurbitaceae (8)
Solanaceae (6)
Fabaceae (3)
Rosaceae (3)
Other dicots (16)
Monocots (4)
Marchantiaceae (1)
Not plant-isolated (8)
Unknown (1)
b
Method: RAxML-NG
Tree scale: 0.1
Continent of isolation
Europe (37)
North America (15)
Asia (14)
Oceania (3)
Africa (1)
Unknown (2)
Host phylogeny
Brassicaceae (22)
Cucurbitaceae (8)
Solanaceae (6)
Fabaceae (3)
Rosaceae (3)
Other dicots (16)
Monocots (4)
Marchantiaceae (1)
Not plant-isolated (8)
Unknown (1)
c
Method: STAG
Tree scale: 0.1
Continent of isolation
Europe (37)
North America (15)
Asia (14)
Oceania (3)
Africa (1)
Unknown (2)
Host phylogeny
Brassicaceae (22)
Cucurbitaceae (8)
Solanaceae (6)
Fabaceae (3)
Rosaceae (3)
Other dicots (16)
Monocots (4)
Marchantiaceae (1)
Not plant-isolated (8)
Unknown (1)
d
Method: ASTRAL
Tree scale: 1
Continent of isolation
Europe (37)
North America (15)
Asia (14)
Oceania (3)
Africa (1)
Unknown (2)
Host phylogeny
Brassicaceae (22)
Cucurbitaceae (8)
Solanaceae (6)
Fabaceae (3)
Rosaceae (3)
Other dicots (16)
Monocots (4)
Marchantiaceae (1)
Not plant-isolated (8)
Unknown (1)

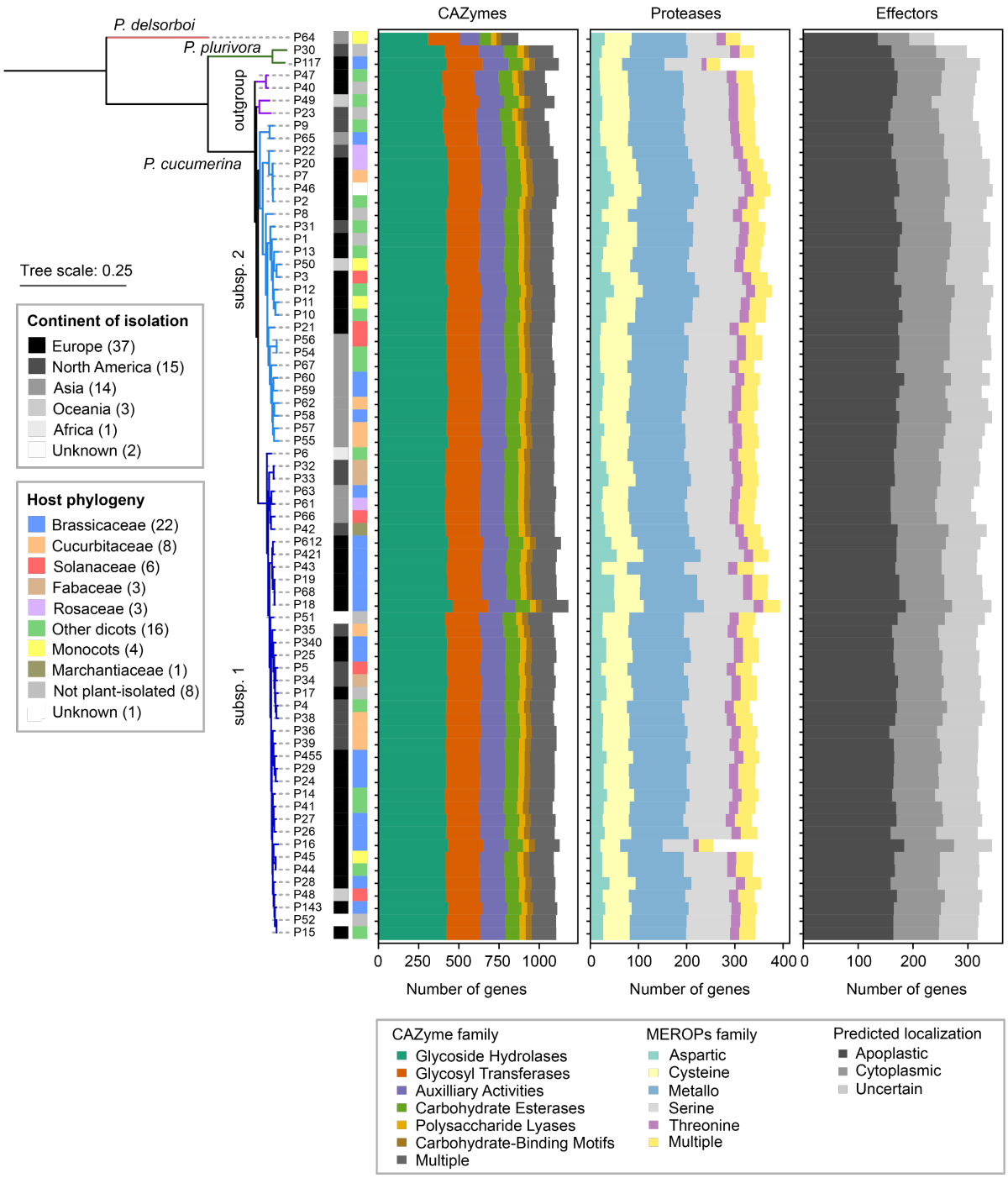
1694

1695 **Supplementary Fig. 4: Phylogenomic trees of the 72 *Plectosphaerella* strains,**
1696 **reconstructed with four different methods**. **a.** Phylogenetic tree reconstructed with IQ-
1697 TREE[100] (model *JTT+F+I+G4*, as defined by ModelFinder[101]) from the trimmed concatenated
1698 alignments of 5,466 single copy orthologs identified by orthology prediction with OrthoFinder[97].
1699 **b.** Phylogenetic tree reconstructed with RAxML-NG[102] (model *JTT+F+I+G4*, as defined by
1700 ModelFinder) from the trimmed concatenated alignments of 5,466 single copy orthologs
1701 identified by orthology prediction with OrthoFinder. **c.** Coalescent phylogenetic tree
1702 reconstructed from total gene family trees generated by OrthoFinder, with built-in method
1703 STAG[133]. **d.** Coalescent phylogenetic tree reconstructed with ASTRAL[104] from the OrthoFinder-
1704 reconstructed gene family trees of 5,466 single copy orthologs. See Methods for further details.
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731

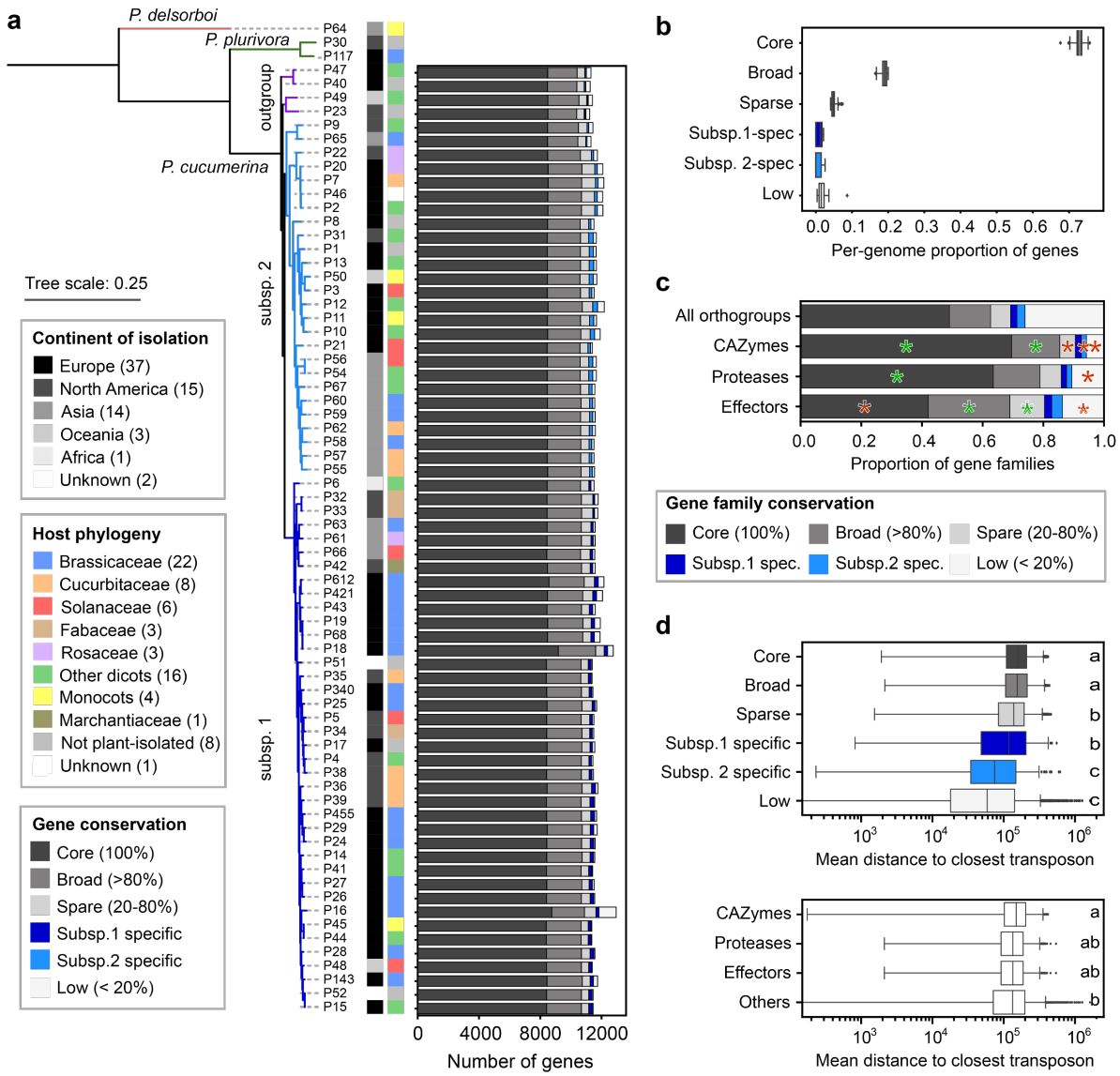


Supplementary Fig. 5: Transposon annotation of the 72 *Plectosphaerella* genome assemblies. a, Results of transposon annotation with the pipeline TransposonUltimate[96]. Along the *Plectosphaerella* collection phylogeny, barplots show the percentage of each genome covered by transposons ("coverage") and the total numbers of transposons annotated. On the right, a bubble plot depicts the number of transposons annotated per family.

b, Boxplot displaying the difference of *P. cucumerina* genomes in subspecies 1 and 2 in coverage in transposons. Mann-Whitney test revealed a significant difference (statistic = 134.0, $P = 6.10 \times 10^{-7}$). **c,** Boxplot displaying the difference of *P. cucumerina* genomes in subspecies 1 and 2 in transposon number. Mann-Whitney test revealed a significant difference (statistic = 194.5, $P = 2.94 \times 10^{-5}$). **d,** Boxplot displaying the difference of *P. cucumerina* genomes in subspecies 1 and 2 in DNA transposon number. Mann-Whitney test revealed no significant difference (statistic = 437.0, $P = 0.35$). **e,** Boxplot displaying the difference of *P. cucumerina* genomes in subspecies 1 and 2 in retrotransposon number. Mann-Whitney test revealed no significant difference (statistic = 108.0, $P = 9.46 \times 10^{-8}$).

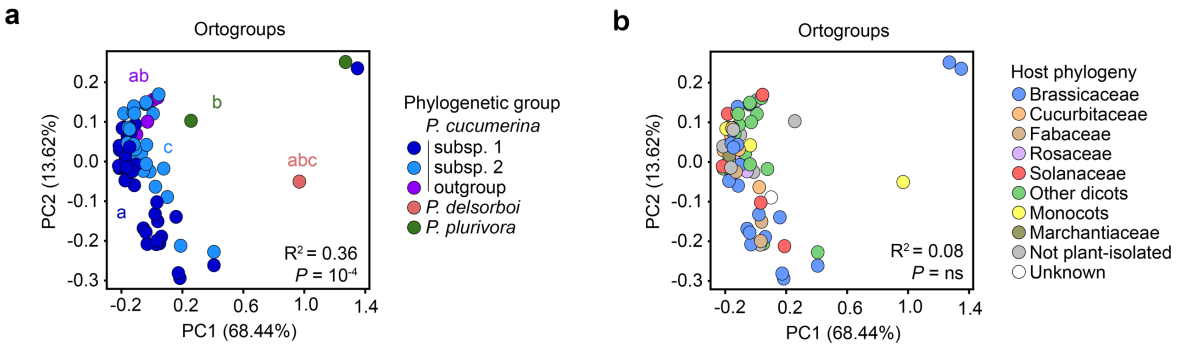


Supplementary Fig. 6: Genome annotation of the 72 *Plectosphaerella* isolates in CAZyme-, protease- and effector-encoding genes. Number of CAZyme- (left), protease- (center) and effector- (right) encoding genes annotated in each of the 72 *Plectosphaerella* genomes, using methods dbcan[47], emapper[92] and EffectorP[91], respectively. Different colors divide the barplots to highlight different functional families.

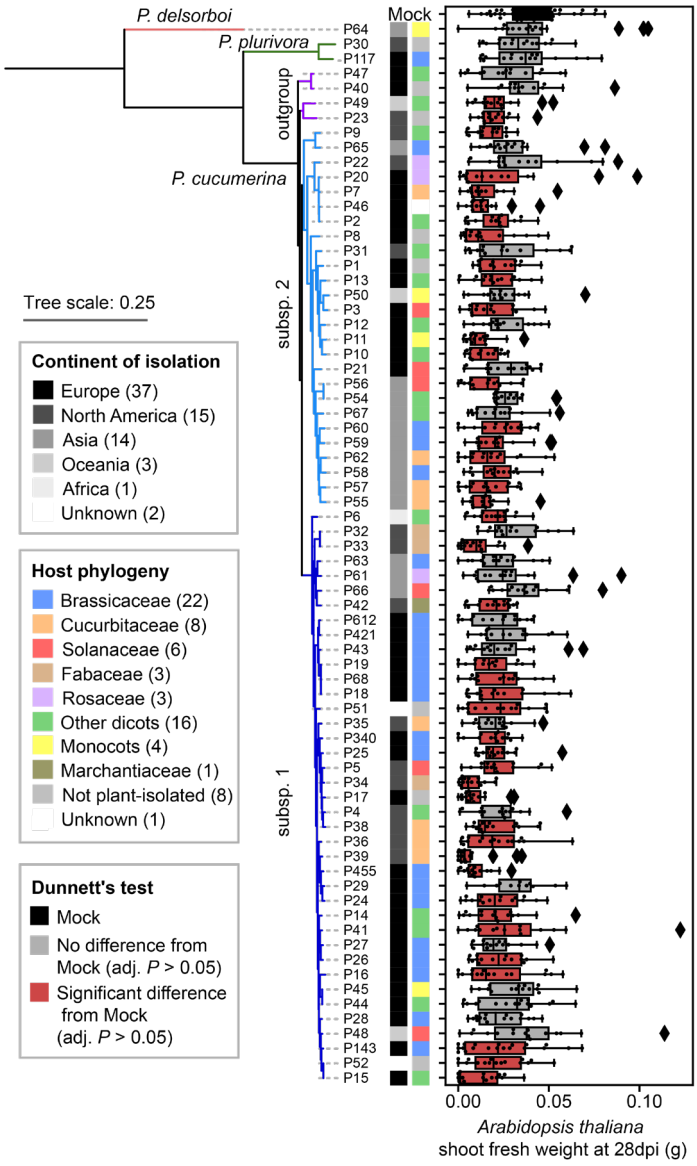


Supplementary Fig. 7: Conservation of gene families across the genomic dataset. a, Following orthology prediction (OrthoFinder), gene families (i.e., n = 15,269 orthogroups) were analysed for their conservation across the 69 *P. cucumerina* isolates and classified as core (represented in 100% of *P. cucumerina* isolates), broadly conserved (represented in > 80% of *P. cucumerina* isolates), sparsely conserved (in 20-80% of *P. cucumerina* isolates), subspecies-associated (in one subspecies' members while absent in the other subspecies) and low-conservation (in < 20% *P. cucumerina* isolates). The number of genes in each category is shown in a stacked barplot, along the *Plectosphaerella* phylogeny. **b,** Per-genome proportions of genes in each one of these categories are shown in a boxplot, revealing that about 75% of *P. cucumerina* gene families are "core" (i.e. conserved across the full collection). **c,** Proportions of total gene families, as well as gene families encoding CAZymes, proteases and effectors, in the different intra-species conservation levels. Fisher's exact tests comparing proportions of CAZyme, protease and effector gene families to total orthogroups were

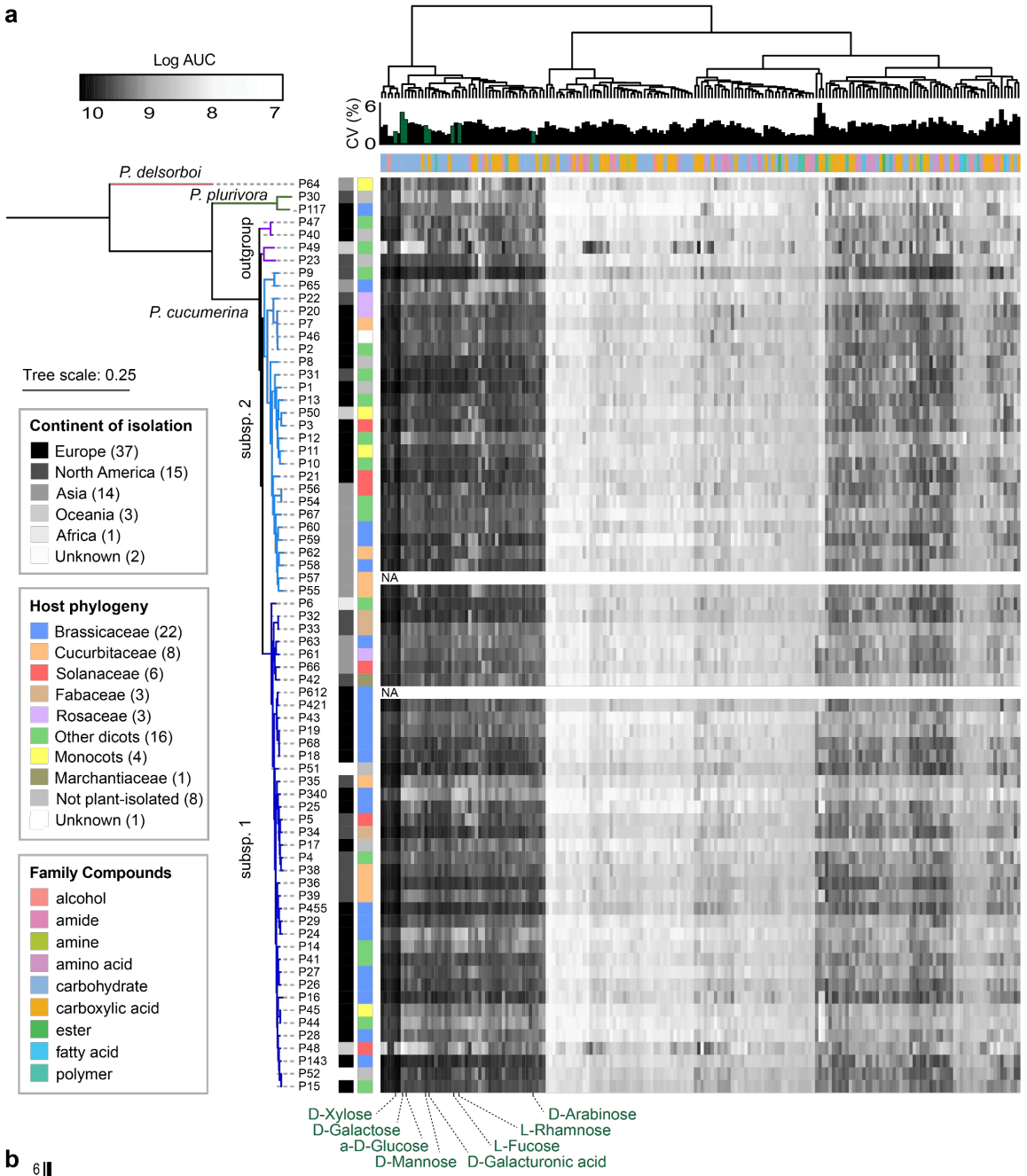
computed for each conservation level. Significant enrichments/depletions are highlighted with asterisks (green: enrichments, red: depletion). **d,** Mean distance to the closest transposon (in the same contig) were plotted for each gene family, grouped by conservation level (top) or by gene function (bottom). Letters reflect significant differences between groups according to Kruskal-Wallis tests ($P < 0.05$) followed by post-hoc Dunn test.

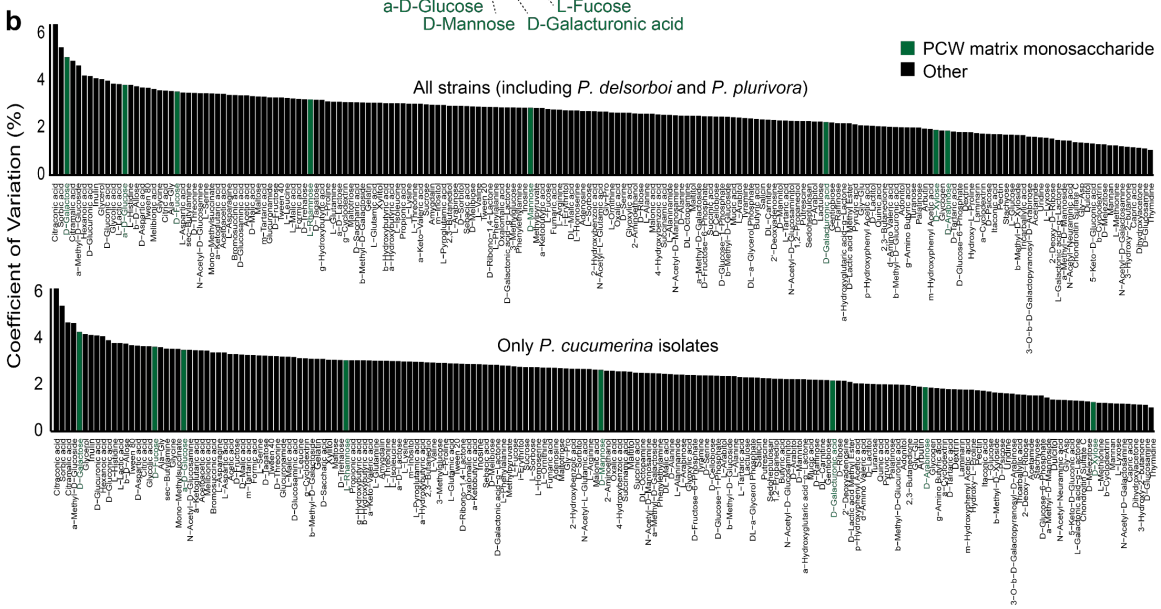


Supplementary Fig. 8: Effect of phylogeny and the host of isolation on inter-strain differences in gene family repertoire. a,b, Principal Coordinate Analysis profiling inter-strain differences in total gene family repertoires (= 15,269 orthogroups, OrthoFinder). This PCoA is colored according to strain phylogenetic group (**a**) or host phylogeny (**b**). R^2 and *P* value resulting from PERMANOVA test assessing the effect of these two categories on the three genomic repertoires are shown at the bottom of each graph.



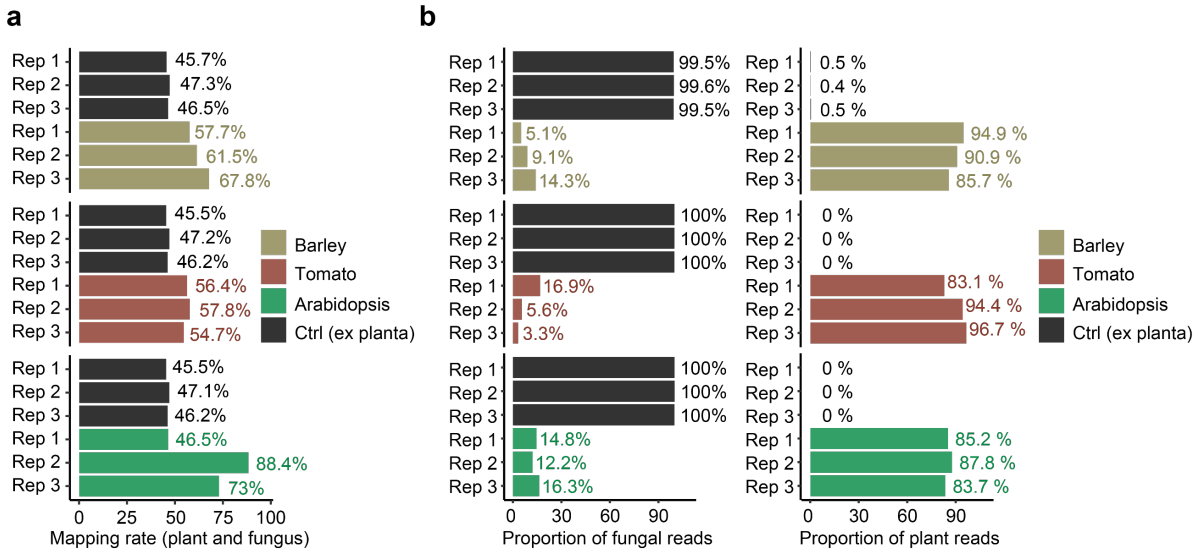
Supplementary Fig. 9: Individual fungal effects of *Plectosphaerella* isolates on *A. thaliana* performance in gnotobiotic system. Shoot fresh weight (SFW) of plants which roots were inoculated with 500 spores of each strain, after 28 days in culture in an agar-based gnotobiotic system (for each fungus: 4 independent biological replicates, 4 plants per replicate). These weight values reflect individual fungal effects on plants. Boxes colored in red reveal a significant difference between the fungus-inoculated and mock-treated plants, according to ANOVA and Dunnett's test ($P < 0.05$). Boxes are represented along the strain phylogeny, reconstructed with IQ-TREE. Black diamonds represent outlier values. Note that isolates P68 and P16 correspond respectively to *P. cucumerina* BMM (PcBMM) and *P. cucumerina* 0016 (see this work).


a
Log AUC
10 9 8 7
CV (%)
6
0
P. delsorboi
P. plurivora
P. cucumerina
outgroup
subsp. 2
subsp. 1
Tree scale: 0.25
Continent of isolation
Europe (37)
North America (15)
Asia (14)
Oceania (3)
Africa (1)
Unknown (2)
Host phylogeny
Brassicaceae (22)
Cucurbitaceae (8)
Solanaceae (6)
Fabaceae (3)
Rosaceae (3)
Other dicots (16)
Monocots (4)
Marchantiaceae (1)
Not plant-isolated (8)
Unknown (1)
Family Compounds
alcohol
amide
amine
amino acid
carbohydrate
carboxylic acid
ester
fatty acid
polymer
NA
NA
D-Xylose
D-Galactose
a-D-Glucose
D-Mannose
D-Galacturonic acid
L-Fucose
L-Rhamnose
D-Arabinose

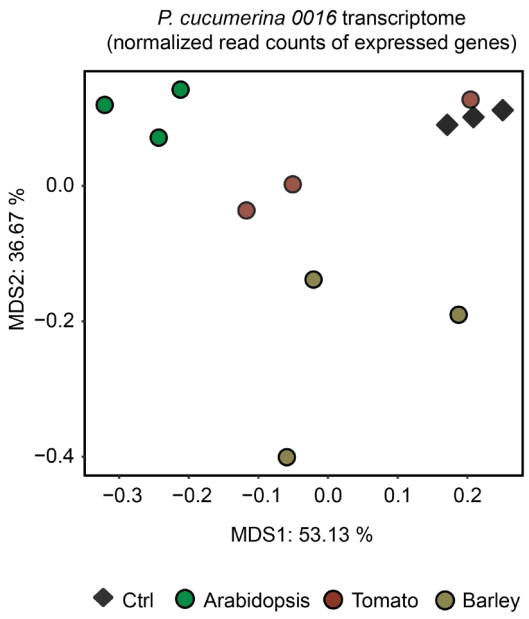

b
PCW matrix monosaccharide
Other
All strains (including P. delsorboi and P. plurivora)
Only P. cucumerina isolates
Coefficient of Variation (%)
6
4
2
0

1882

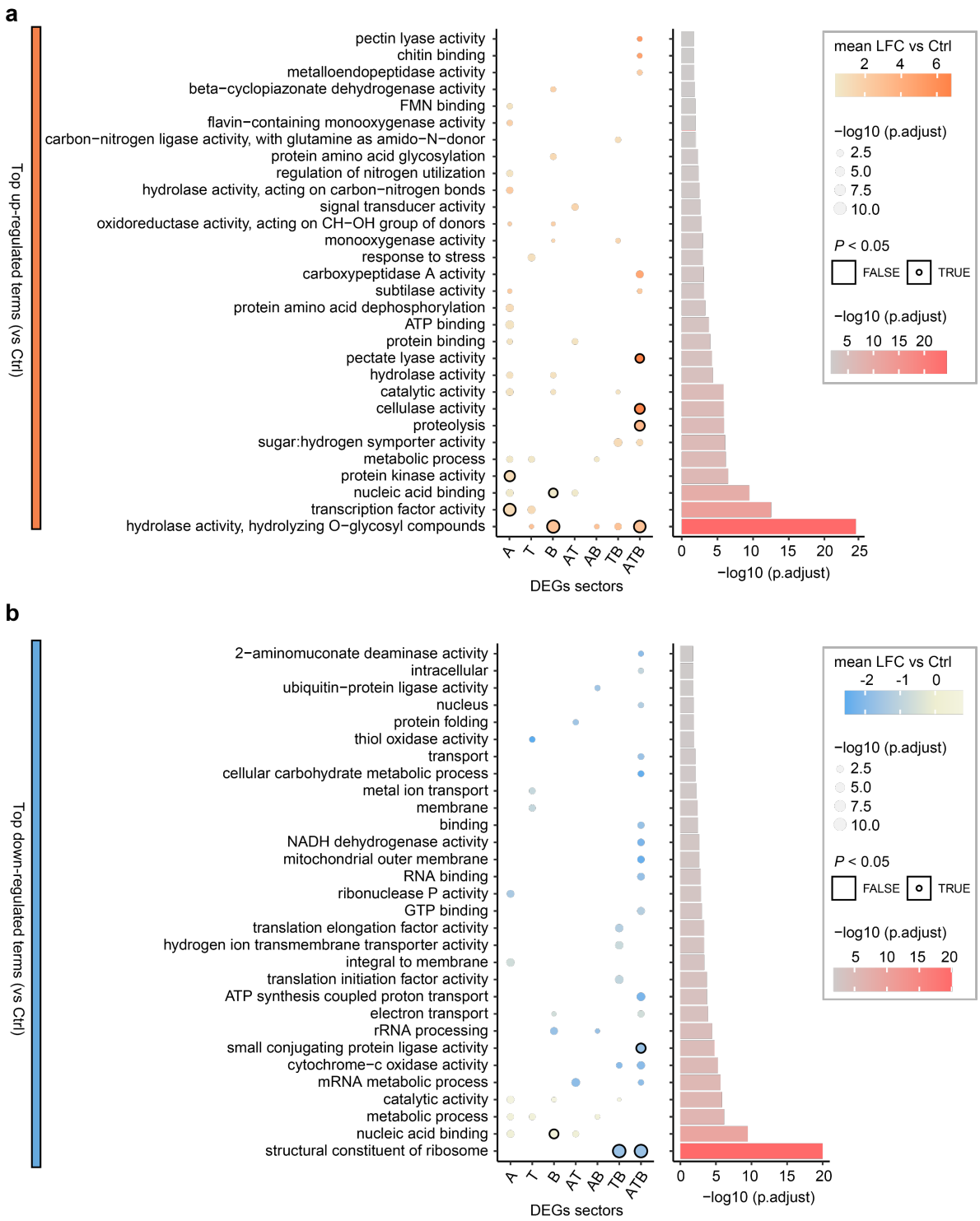
1883 **Supplementary Fig. 10: Carbon resource consumption profiles of *Plectosphaerella***
1884 **isolates. a,** Heatmap showing the respiration profiles of *Plectosphaerella* strains (n = 72, y-
1885 axis) over 5 days when grown on individual carbon sources (x-axis, n = 190) using Biolog PM1
1886 and PM2 plates. The color intensity represents the log-transformed area under the curve (log
1887 AUC) of respiration. Carbon sources are clustered based on their utilization patterns across all
1888 *Plectosphaerella* strains. Strains are ordered according to their phylogeny. Plant cell wall
1889 (PCW) matrix monosaccharides (**Supplementary Fig. 15)** are highlighted in dark green. **b,**
1890 Bar plots showing the coefficient of variation (y-axis) in respiration across isolates for each
1891 carbon source (x-axis). The top panel includes all strains, while the bottom panel focuses on
1892 *P. cucumerina* isolates only. PCW matrix monosaccharides (**Supplementary Fig. 15**) are
1893 highlighted in dark green.
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919



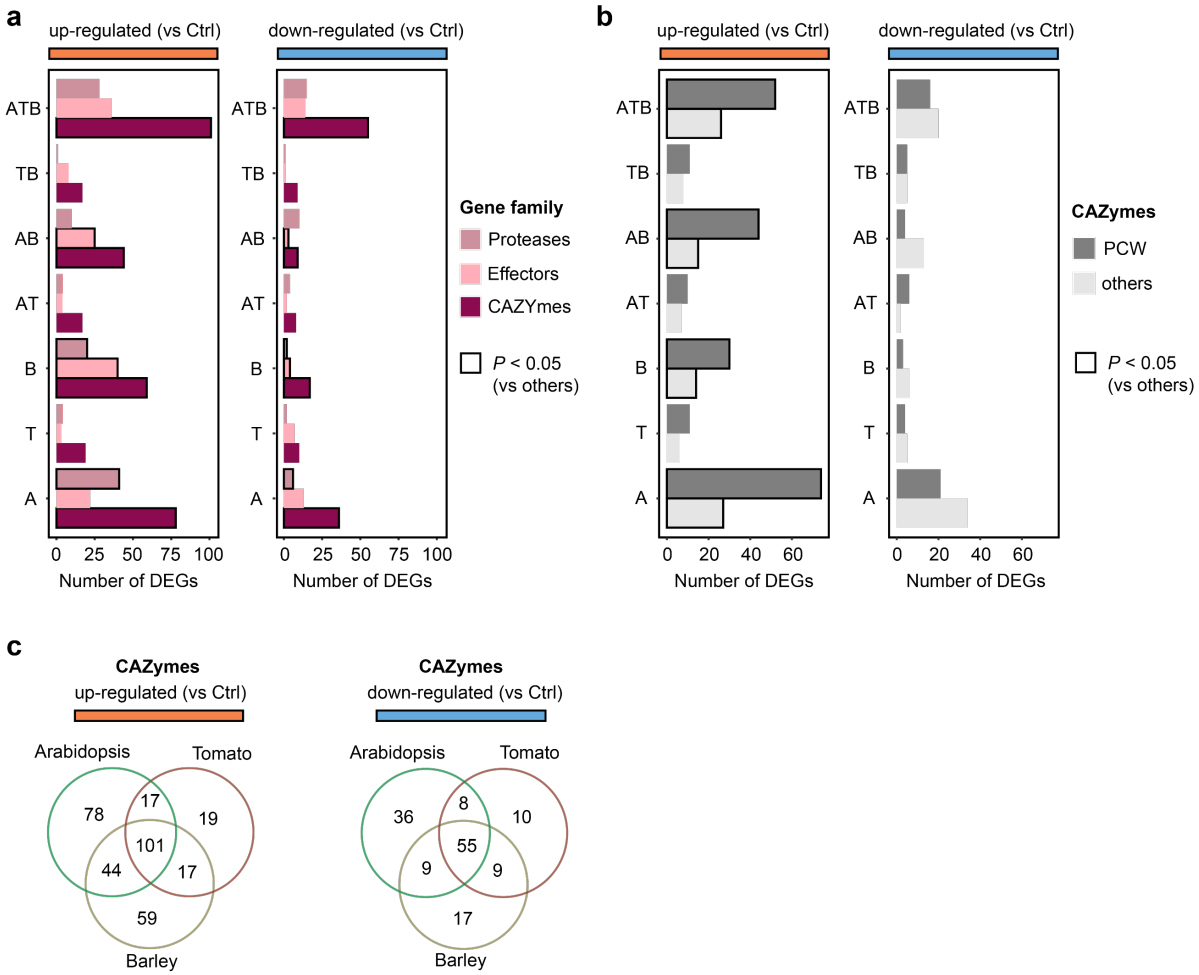
Supplementary Fig. 11: Mapping rates of plant and fungal RNAseq reads against corresponding reference genomes. a, The bar chart depicts for each sample the proportion of sequencing reads (i.e., mapping rate) that aligned against the concatenated reference transcriptome of the plant host (either *A. thaliana* Col-0 - TAIR11[109], Micro-Tom - TMCSv1.2.1[111], barley - Hvulgare_462_r1[112] and *P. cucumerina* 0016 utilizing Salmon[113]. **b**, The proportion of fungal reads (left) and plant reads (right) mapped against the corresponding reference transcriptomes. The colors represent the mapping rate of fungal and host reads in roots of barley (khaki), tomato (deep red) and Arabidopsis (green) and of fungal reads in the absence of the host (Ctrl, black).



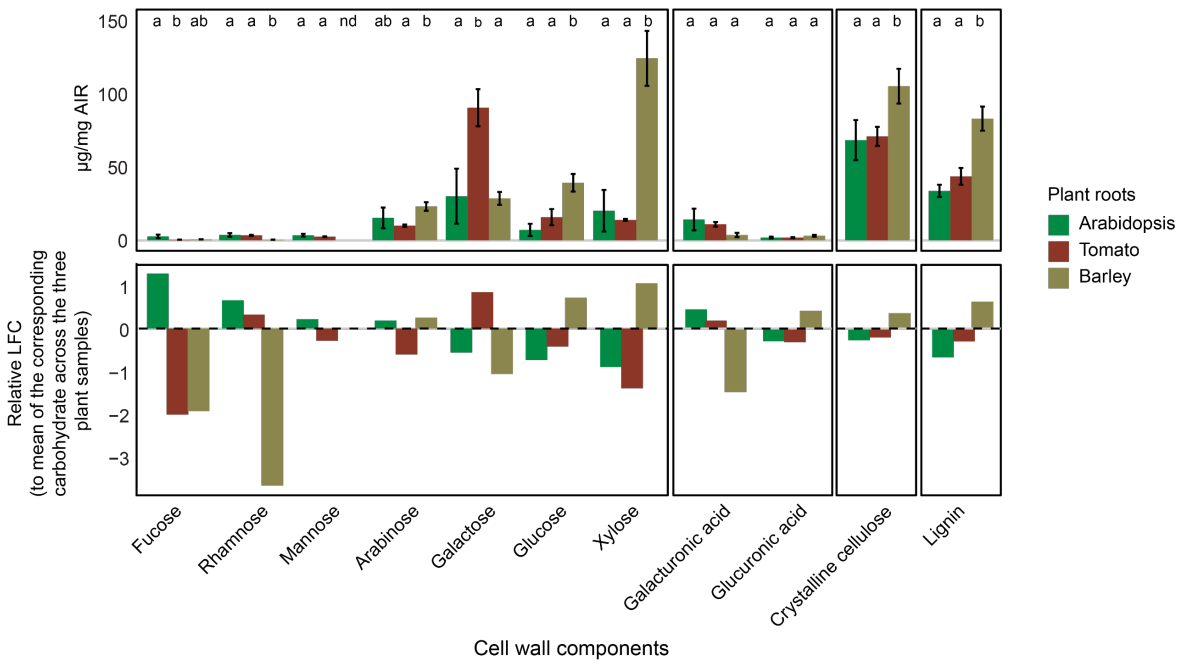
Supplementary Fig. 12. Multidimensional Scaling (MDS) plot of Pearson distances computed on normalized read counts of expressed fungal genes. Fungal transcriptome in roots of Arabidopsis (Green), tomato (Red), and barley (Khaki), as well as in control membrane (Ctrl, *ex planta*, Black). Data points indicate three biological replicates for each condition.



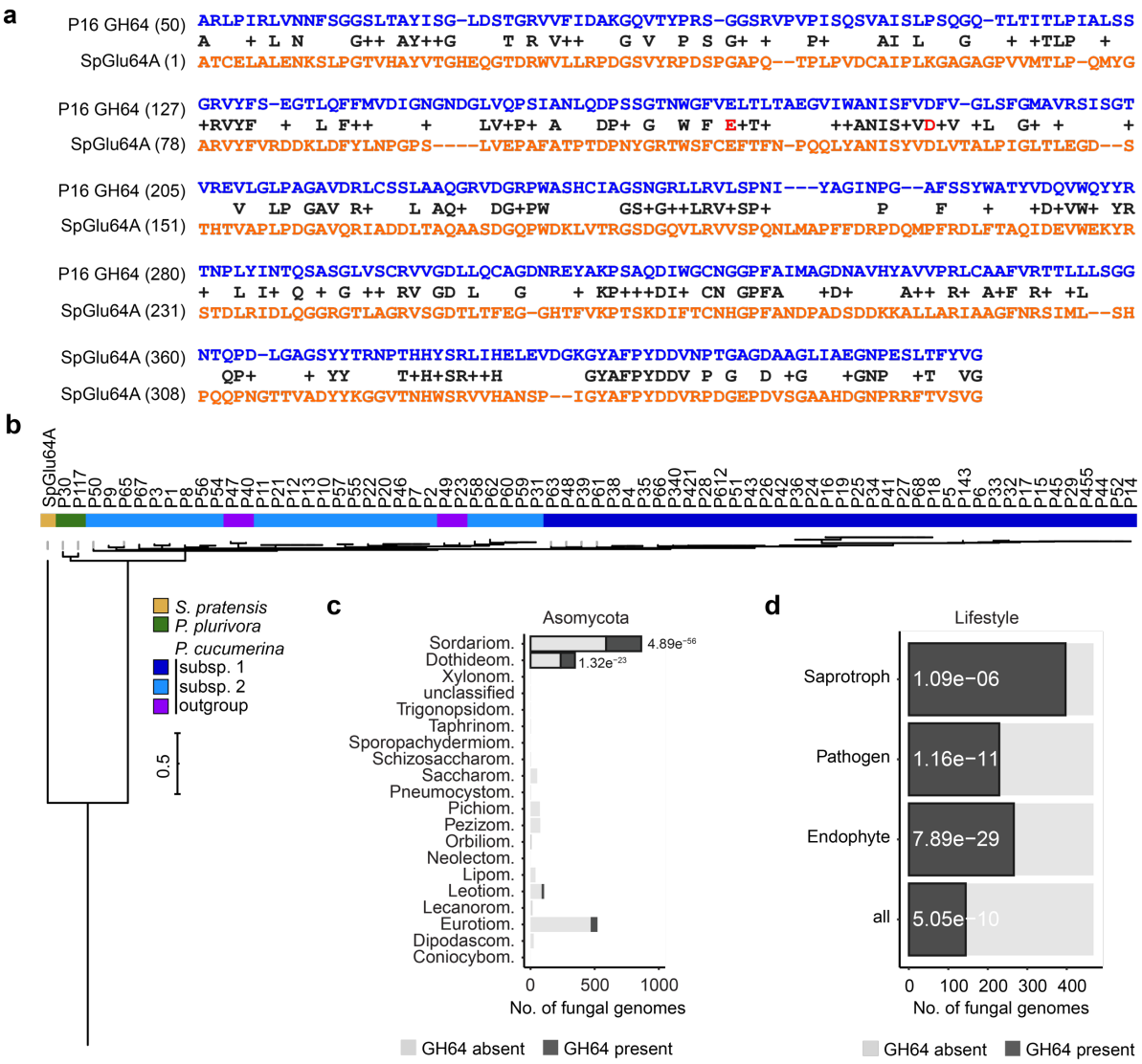
Supplementary Fig. 13: Fungal GO terms significantly regulated during root colonization. a,b, The heatmap illustrates the average $\log_2$ fold-change values for the categories (x-axis), marked circles signify $P \leq 0.05$ and the size of the circle indicates $-\log_{10}$ *P*-adjusted values (FDR, Fisher's Exact Test). The corresponding barplot shows the cumulative $-\log_{10}$ *adj. P* values (x-axis and colors grey-red) for the 30 most common GO-terms in y-axes. GO terms that are up-regulated are represented in orange (**a**) and those that are down-regulated are represented in blue (**b**).



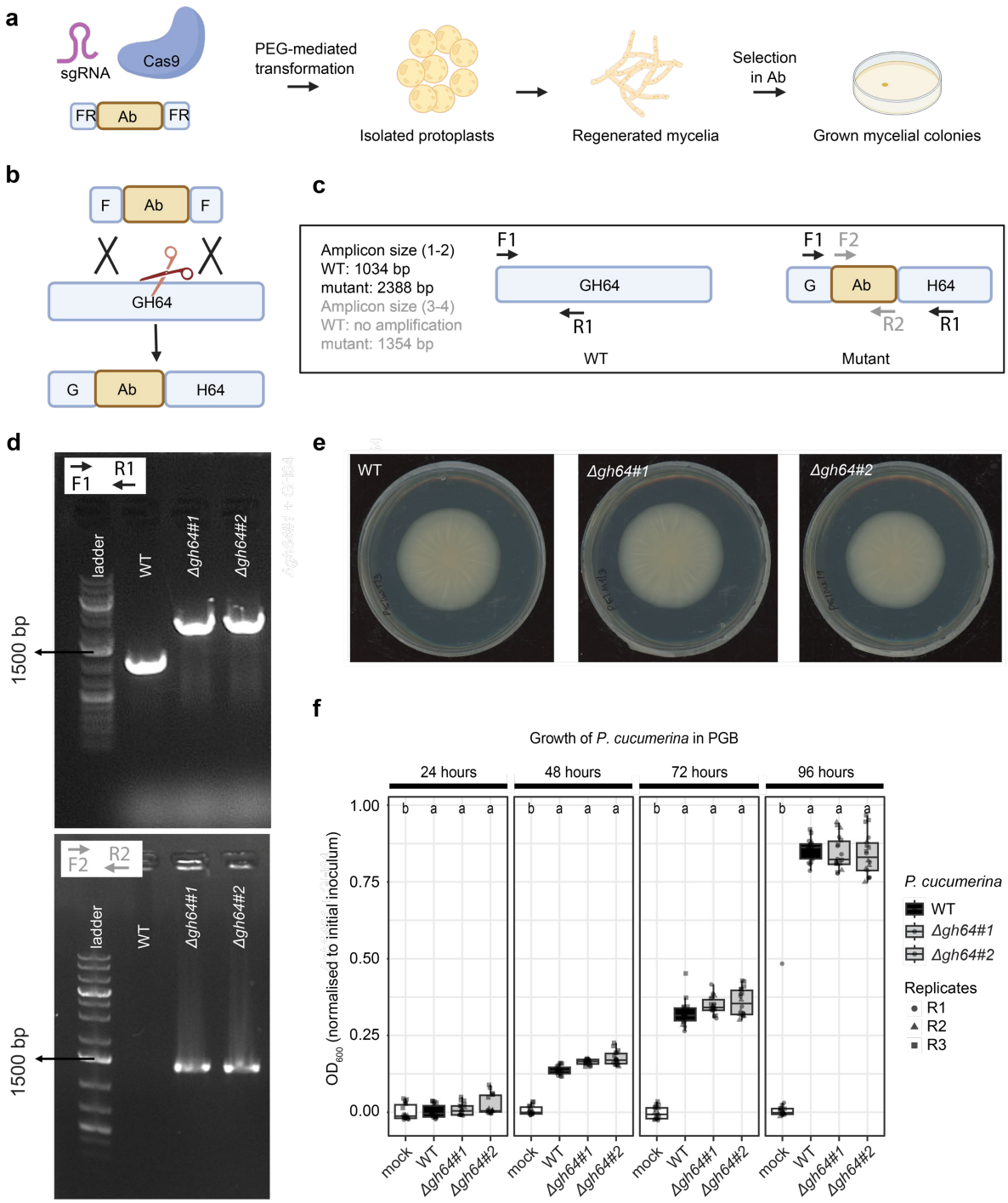
Supplementary Fig. 14: Regulation of CAZyme-encoding genes during root colonization. a,b, The combination of barplot illustrates the number of differentially expressed genes (DEGs, x-axes) that are either upregulated (left, denoted in orange) or downregulated (right, denoted in blue) for the shared sectors on the y-axes. Marked bars signify those sectors that are $P \leq 0.05$ (prop.test, chi-square statistics) indicating the ratio of gene family to be elevated among the total DEGs in the dataset. The colors of the barplot elucidate the predetermined gene families Proteases (deeppink), Effectors (lightpink) and CAZYmes (magenta) (**a**) and in grey plant-cell-wall components (PCW; mined "pectin, cellulose, lignin, arabinan, galactan, glycan, glucan, xylan, galacturon) CAZY-Substrate (dbCAN[47]) (**b**). **c**, Venn diagram illustrating the number of DEGs encoding CAZymes (|log2FC| ≥ 1 and FDR ≤ 0.05) that are up-regulated (orange) and down-regulated (blue) *in planta*, respectively.



Supplementary Fig. 15: Root cell wall composition analysis of Arabidopsis, tomato, and barley: The upper bar plot illustrates the concentration of wall matrix monosaccharides, including uronic acids [in µg/mg alcohol insoluble residue AIR, y axis], cellulose, and lignin detected in roots of 5-week-old *A. thaliana* (n = 36 plants) and tomato (n = 18 plants), and 4-week-old barley (n = 6 plants). The cell wall components were arranged based on the different assays used and in descending order based on their abundance in Arabidopsis. Multiple pairwise comparisons between the three plant species for each cell wall component were performed using Kruskal-Wallis-Dunn Test, adjusted with Bonferroni correction ($P \leq 0.05$). The lower bar plot illustrates the relative enrichment of carbohydrates (log2FC) for each plant with respect to the mean value measured across all three plants. Data derived from three biological replicates.

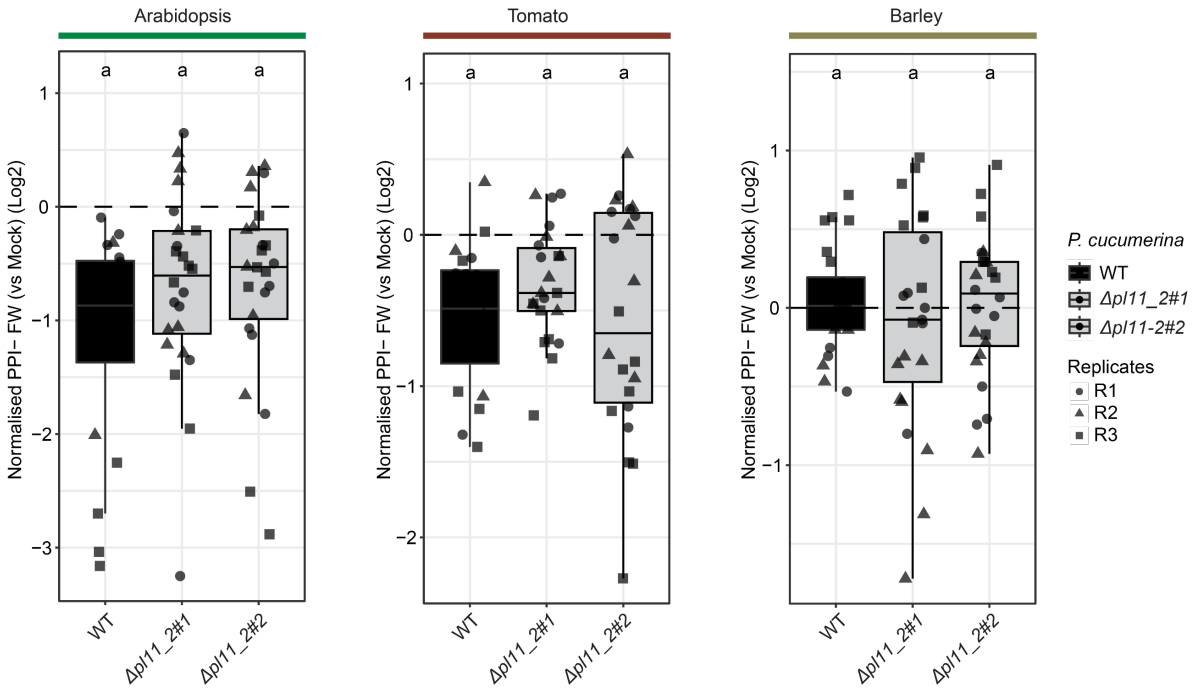


Supplementary Fig. 16: **Function and evolutionary conservation of GH64 in the fungal kingdom**. **a,** Structure-aware sequence alignment of *P. cucumerina* 0016 GH64 (in blue) with SpGlu64A (in orange) shows conservation of catalytic residues (in red), despite overall low sequence similarity (sequence identity=34.3%). **b,** Maximum-likelihood phylogenetic tree (model JTT+R2 of iqtree[100]) of the orthologs of *P. cucumerina* 0016 GH64 in the *Plectosphaerella* genome collection, together with SpGlu64A. Labels refer to the strain encoding the protein, while color strips highlight the strain phylogenetic group. Note: except for *P. delsorboi* which does not encode any GH64 enzyme, a single GH64-encoding gene was annotated in every *Plectosphaerella* genome. **c,d,** Stacked bar plot depicting the presence/absence of GH64-encoding genes in 2,534 fungal genomes (x-axes) across distinct classes of Ascomycota members (**c**) and fungal lifestyles (**d**) (y-axes). Highlighted in dark grey are those where *GH64* family genes are encoded in the genome, and those with thick black borders are enriched within the group (Fisher's exact test followed by *P* value adjustment using Bonferroni correction).

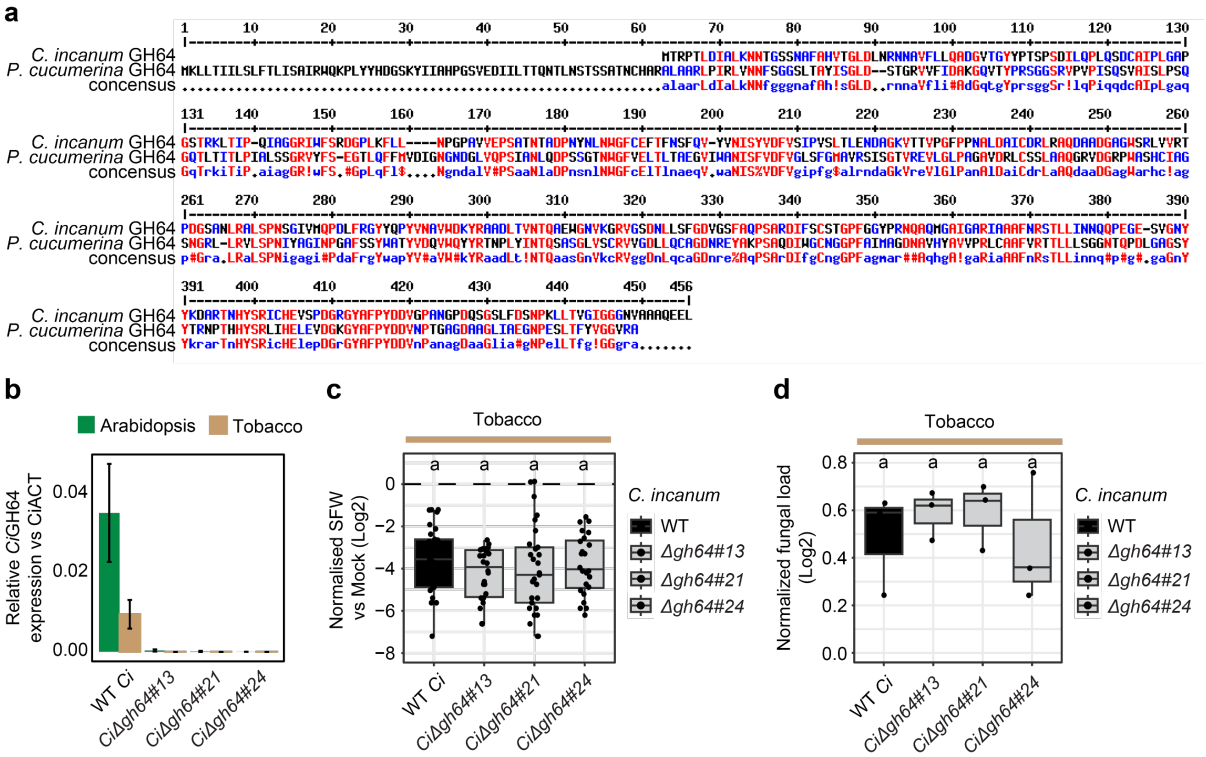


Supplementary Fig. 17: Protoplasts, regenerated Protoplasts and *P. cucumerina* 0016 mutants validation: a, Schematic diagram illustrating PEG-mediated transformation and CRISPR/Cas9- mediated knockout method. (PEG-mediated transformation method: 1. Isolation of protoplasts, 2. transformation of Donor DNA and Cas9-sgRNA complex into the protoplasts. 3. Regeneration of protoplasts. 4. Selection of potential colonies in media containing Phleomycin.) **b,** The CRISPR/Cas9 system enhances HR-mediated integration by introducing a double-stranded break in the target DNA, allowing for effective integration with short homology flanks of 60bp. **c**, Schematic representation of the *gh64* gene region in WT

and in mutant fungi. Primer pairs used in panel d are depicted. **d**, Targeted disruption of the GH64 gene. Top: Agarose gel electrophoresis results using primer pair F1, R1. The wild type shows a band at 1034 bp, while mutants display a band shift to 2388 bp due to the insertion of the antibiotic resistance sequence. Lane 1: Ladder; Lane 2: Wild type (WT) control; Lane 3, 4: Mutant 1, 2; Bottom: Amplification of the phleomycin antibiotic resistance cassette using primer pair F2, R2. The gel schema is identical to the top panel, confirming the presence of the phleomycin cassette in the mutant lines. **e,** *In vitro* growth of *P. cucumerina* 0016 WT *versus Δgh64* mutant strains in PGA plate. **f**, Dynamics of *in vitro* growth of *P. cucumerina* 0016 WT *versus Δgh64* mutant strains in PGB medium.



Supplementary Fig. 18: Phenotype assessment of P. cucumerina 0016 Δpl11_2 mutant strains. Plant Performance Indices (PPI) of Arabidopsis, tomato, and barley colonized by *P. cucumerina* 0016 WT and mutant strains (*Δpl11_2#1*, *Δpl11_2#2*) were normalized to mock-treated plants and log2 transformed. PPI values were calculated by multiplying the germination rate by the shoot fresh weight of the plants. Pre-grown seedlings (2-week-old Arabidopsis and tomato and a week-old barley) were inoculated with *P. cucumerina* 0016 and harvested 21 days post fungal inoculation. Three independent biological replicates are depicted (circles, triangles, squares), each consisting of 5-8 plants (n = 15-24 plants per condition). Multiple pairwise comparisons of plant performance between mock-treated and *P. cucumerina* 0016 treated plants were performed using Kruskal-Wallis-Dunn Test, adj. with Bonferroni correction ($P \leq 0.05$).



Supplementary Fig. 19: *C. incanum* GH64 sequence conservation, expression and functional relevance for tobacco root infection. a, Protein sequence alignment of *P. cucumerina* 0016 GH64 (bottom) with *C. incanum* GH64 (up) obtained from Multalin[134]. Note the absence of signal peptide in *C. incanum* GH64 sequence. **b**, *GH64* transcript quantification in roots of Tobacco and Arabidopsis. The relative expression of the *C. incanum GH64* gene was measured by qRT-PCR on tobacco or Arabidopsis roots inoculated with *C. incanum* WT (WT *Ci*) and *Δgh64* mutant lines and normalized to the expression of *C. incanum ACTIN* gene at 24 dpi. Statistical significance was evaluated by Tukey's post-hoc test indicated by different letters, $P < 0.05$. **c**, Shoot fresh weight (SFW) of tobacco plants inoculated with WT *Ci* or *Δgh64* mutants and normalized to mock-treated plants (24 dpi). Statistical significance was evaluated by Tukey's post-hoc test indicated by different letters ($P < 0.05$). **d**, Relative fungal load quantification of WT and mutant lines in tobacco roots was estimated by qRT-PCR based quantification of *C. incanum ACTIN* transcripts normalized to tobacco *ACTIN* transcripts (24 dpi). Statistical significance was evaluated by Tukey's post-hoc test indicated by different letters, $P < 0.05$.