Title: Symbiotic diversification relies on an ancestral gene network in plants

Authors: Leonardo Castanedo[1†], Katharina Melkonian[1†], Tatiana Vernié[1], Lucie Chauderon[1], Tifenn Pellen[1], Matheus E. Bianconi[1], Corinne Lefort[1], Cyril Libourel[1‡], Aurélie Le Ru[2], Simon Aziz[1], Cristina Latorre[1], Océane Thiercelin[1], Nathalie Rodde[3], Stephane Cauet[3], Caroline Callot[3], Laura Forrest[4], Tomoaki Nishiyama[5,6], David Wickell D[7], Fay-Wei Li[8,9], Jean Keller[1*], Pierre-Marc Delaux[1*]

Affiliations:

[1] LRSV, Université de Toulouse, CNRS, UPS, Toulouse INP, 31326 Castanet-Tolosan, France.

[2] Plateforme d'Imagerie FRAIB-TRI, Université de Toulouse, CNRS, Auzeville-Tolosane 31320, France.

[3] INRAE, CNRGV French Plant Genomic Resource Center, F-31320 Castanet-Tolosan, France.

[4] Royal Botanic Gardens Edinburgh, 20A Inverleith Row, Edinburgh, EH3 5LR, UK.

[5] Research Center for Experimental Modeling of Human Disease, Kanazawa University, Kanazawa, Ishikawa, Japan.

[6] School of Science, Academic Assembly, University of Toyama, Toyama, Toyama, Japan

[7] Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT, USA

[8] Boyce Thompson Institute, Ithaca, NY, USA

[9] Plant Biology Section, Cornell University, Ithaca, NY, USA

*Corresponding authors. Emails: pierre-marc.delaux@cnrs.fr jean.keller@cnrs.fr

† The authors contributed equally

‡ present address: Unité de Recherche Physiologie, Pathologie, et Génétique Végétales (PPGV), INP PURPAN, Université de Toulouse, Toulouse, France

Abstract:

Symbioses have been fundamental to colonization of terrestrial ecosystems by plants and their evolution. Emergence of the ancient arbuscular mycorrhizal symbiosis was followed by the diversification of alternative intracellular symbioses, such as the ericoid mycorrhizae (ErM). We aimed at understanding how these diversifications occurred. We sequenced the genomes of ErM-forming liverworts, and reconstituted symbiosis under laboratory conditions. We demonstrated the existence of a nutrient-regulated symbiotic state that enables ErM and underlies intracellular colonization of plant tissues. Comparative transcriptomic analyses identified an ancestral gene module associated with intracellular symbiosis beyond ErM. Genetic manipulations in the liverwort *Marchantia paleacea*, phylogenetics and transactivation assays demonstrated its essential function for intracellular symbiosis. We conclude that plant have maintained, and convergently recruited, an ancestral gene module for intracellular symbioses.

Main Text:

Symbiotic interactions have played a fundamental role in the colonization of terrestrial ecosystems by plants[1–3], influencing both population dynamics and biogeographical distribution, and shaping plant evolution[3]. The appearance of the first intracellularly-hosted fungal symbiont, the arbuscular mycorrhizal symbiosis (AMS), at the dawn of the land plants 450 million years ago is supported by the fossil record, evo-devo and phylogenetic studies[4–9]. The evolution of AMS in land plants was followed by the diversification of alternative intracellular symbioses and the origin of novel plant clades that colonized previously unoccupied terrestrial ecosystems[3]. For example, the evolution of ericoid mycorrhizae (ErM) around 140–75 Ma coincided with the origin and the appearance of the Ericaceae family of vascular plants[3,10,11]. ErM can also be found in bryophytes, within the leafy liverwort order Jungermanniales (Jungermanniopsida)[12,13]. It is estimated that ErM in Jungermanniales, alongside Ericaceae plants, contribute to the mobilization of nitrogen (N) and phosphate (P) in the habitats where these plants often co-exist[14–16].

The establishment of AMS relies on a set of four genes (*SYMRK*, *EPP1*, *CCaMK* and *CYCLOPS*), which constitute the Common Symbiosis Pathway (CSP)[17–19]. The CSP genes are both conserved and required for the establishment of AMS across land plants, including angiosperms and bryophytes[20,4,6,18,19]. Consistent with this, plant lineages that have lost the ability to form AMS also lost the CSP during evolution[4,21]. Very few exceptions to these correlated losses have been observed, and they always correspond to the occurrence of another type of intracellular symbiosis, such as the nitrogen-fixing root nodule symbiosis in *Lupinus*[5], or ErM in Ericaceae and Jungermanniales[5,19]. Altogether, this suggests that ErM, like other type of intracellular symbioses, have diversified from the

ancestral AMS. However, the cellular, molecular and genetic mechanisms underlying these diversification events remain unclear. Here, we aimed at exploring the mechanisms underlying symbiosis diversification by studying ErM in the leafy liverwort *Calypogeia fissa*.

Nutrient-dependent mutualistic intracellular symbiosis in *Calypogeia fissa*

We aimed at developing a model for studying ErM in controlled conditions. To do this, we selected the leafy liverwort *Calypogeia fissa* and generated axenic cultures. *C. fissa* was then mock-inoculated, or inoculated with the ErM-forming fungus *Hyaloscypha hepaticicola*, in a nutrient regime low in both phosphate (P) and nitrogen (N) (Fig. 1). We used this treatment to mimic the nutrient-limited conditions in which leafy liverworts can thrive[14]. Two parameters indicative of mutualism were scored: i) hyphal coils in colonized rhizoid-tips which are considered the symbiotic structures[12,22], and ii) plant growth (Fig. 1A-L). The growth of *H. hepaticicola* was discrete and highly localized around and inside swollen rhizoid tips of *C. fissa* where it formed intracellular coils (Fig. 1G-L). To assess whether colonization benefited *C. fissa* plants, we quantified plant growth in both *H. hepaticicola*-inoculated and mock-inoculated conditions over a 9-week time-course experiment (Fig. 1A-F). We observed that inoculated *C. fissa* plants produced significantly more biomass compared to mock-inoculated plants (Fig. 1C,F). Similar observations were made under extreme nutrient depletion (Fig. S1).

Given that nutrient availability regulates the establishment of AMS in angiosperms and thalloid liverworts, promoting a symbiotic permissive state[6,23,24], we hypothesized that P and N availability might regulate the establishment of ErM in *C. fissa*. We conducted inoculation experiments under a range of P supply that had no detrimental impact on the growth of *H. hepaticicola* cultures (Fig. S1). We applied these nutrient conditions to mock- or *H. hepaticicola*-inoculated *C. fissa* plants. Under conditions of low P supply, *C. fissa* inoculated with *H. hepaticicola* produced more biomass than mock-inoculated plants (Fig. 1D-F, and Fig. S1). This is in contrast with the minimal, or absence of, growth differences observed under high P supply (Fig. 1A-C, Fig. S1). Reducing N supply under high P led to increased mycorrhizal growth response (Fig. S1), indicating an interaction between the two nutrients, as in the case of AMS[25] Similar to the growth response, low P supply allowed the development of intracellular hyphae in swollen rhizoid tips, while high P abolished colonization (Fig. 1G-M, Fig. S1).

In addition to AMS, the nitrogen-fixing root nodule endosymbiosis formed by few angiosperm species and nitrogen-fixing bacteria, is also regulated by nutrient availability[26]. Here we provide experimental evidence that nutrient availability defines symbiotic permissiveness in a third intracellular symbiosis. We propose that the regulation of colonization by the plant nutrient status is a key driver in maintaining mutualism irrespective of the symbiont type.

Jungermanniales genomes

In order to provide genomic resources for Jungermanniopsida to further study ErM, we produced and annotated genome assemblies for *Calypogeia fissa* and *Solenostoma infuscum* (previously referred to as *Jungermannia infusca*). The *S. infuscum* genome was sequenced using the Illumina technology, yielding 248 million of paired-end reads assembled into 4,616 scaffolds (from 949bp to 1,4Mb, N50=244Kb) (Table S1). The *C. fissa* genome was sequenced using the PacBio long-read technology, resulting in an assembly comprising 82 scaffolds (from 46Kb to 66Mb, N50=16Mb, Table S1). As previously suggested by phylogenetic analyses conducted on transcriptome assemblies, the *C. fissa* and *S. infuscum* genomes both contain all four CSP genes *SYMRK*, *CCaMK*, *CYCLOPS* and *EPP1* (Fig. S2-5).

Whereas the estimated size of the *S. infuscum* genome is *ca.* 400 Mb, the estimated genome size of *C. fissa* is *ca.* 800 Mb. Genome assembly completeness analyses indicated 95.3% and 96% completion for *S. infuscum* and *C. fissa*, respectively, with 94.6% of the annotated genes duplicated in *C. fissa* (Table S1). Gene structural annotation predicted 14,494 and 38,150 protein-coding genes in the genomes of *S. infuscum* and *C. fissa* respectively (Table S1). Overall, the genome assembly statistics supported the hypothesis that *C. fissa* underwent whole-genome duplication (WGD), as previously proposed[27] (Fig. 2A,B). This hypothesis is further supported by the distribution of synonymous substitutions per site (Ks) between pairs of paralogs in *C. fissa*, with a single peak near the origin (Ks = 0.05), which is indicative of a recent large-scale duplication event (Fig. 2A). Likewise, an intragenomic synteny analysis recovered 137 collinear blocks comprising 7,087 gene pairs (*i.e.*, syntelogs) representing 37.2% of the *C. fissa* genome (Fig. S6), with nearly all of these corresponding to the peak near the origin (Ks = 0.05) (Fig. 2A). In addition to paralogous Ks, orthologous Ks was estimated for *C. fissa* and *S. infuscum*. The resulting Ks plot of orthologous genes exhibited a peak around Ks = 0.35, consistent with a scenario of recent WGD in *C. fissa* (Fig. S6) after its divergence from *S. infuscum*. Intergenomic comparison of synteny between *C. fissa* and the distantly related species *M. polymorpha* identified relatively few collinear gene pairs between them, roughly 10% of genes in each species, likely owing to the species' deep phylogenetic divergence (Fig. S6). However, a clear 2:1 pattern of syntenic depth was observed among collinear genes (Fig. 2C), consistent with the apparent absence of WGD events in *M. polymorpha* (Fig. S6). The species tree inferred from orthologous genes recovered *C. fissa* and *S. infuscum* forming a clade (estimated divergence at 69.8 Ma, 95% HPD = 32 - 122 Ma) that is sister to the Marchantiopsida clade of complex thalloid liverworts (293.7 Ma, 95% HPD = 217 - 370 Ma) (Fig. 2D).

Here, we have generated two genome assemblies for the Jungermanniales which are distantly related to the model Marchantiopsida liverworts.

Induction of an ancestral gene module for intracellular symbioses

Availability of its genome assembly allowed us to conduct bulk RNA sequencing on mock-treated or *H. hepaticicola*-inoculated *C. fissa* plants, to test for the occurrence of a transcriptomic rewiring, as observed in other endosymbiosis[6,28–30]. For this, plants were grown in a permissive condition and harvested eight weeks after inoculation (Table S2). Differentially expressed genes were computed, identifying 787/743 genes up-/down-regulated (logFC threshold of |1|) in inoculated plants compared to the mock treatment in this condition permissive for ErM (Figure 5A, Supplementary Dataset S1). We hypothesized that comparing the transcriptomic response of *C. fissa* to the inoculation with *H. hepaticicola* in permissive and non-permissive symbiotic conditions would reveal intracellular-specific responses (Fig. 5A-B). To do so, we conducted RNA sequencing experiments with mock- or *H. hepaticicola*-inoculated *C. fissa* plants that were grown under a restrictive nutrient regime with high P supply for eight weeks. In this non-permissive condition, where no intracellular colonization was detected (Fig. 3A-C, 3G-H, and 3M), we computed 566/27 genes up/downregulated in the inoculated compared to mock-inoculated plants at a logFC of |1| (Fig. 5B, Supplementary Dataset S2). We compared the list of DEG in the permissive versus non-permissive conditions to identify the genes and processes associated with intracellular symbiosis during ErM, and conducted functional gene ontology enrichment in each condition (Fig. 5A-B, Tables S3-6). These comparisons revealed 82/1 genes up/downregulated in response to *H. hepaticicola* that were specifically responsive during the ErM-permissive state. Among the upregulated genes in permissive condition, we observed two highly enriched gene ontology categories, one encompassing a large number of enzymes associated with cell-wall remodeling and the other with nutrient transport (Fig. 5A-B, Table S3). Both categories might be involved in the reciprocal exchange of resources – carbon for nutrients – between *C. fissa* and the ErM fungus and for building the symbiotic interface.

Next, we hypothesized that, as illustrated by the phylogentic pattern of the CSP genes (Fig. 2, Fig. S3-8)[5], intracellular symbiosis might rely on shared gene modules, irrespective of the symbiont type and plant lineage. To test this hypothesis, we cross-referenced the specific transcriptomic signature of *C. fissa* in response to *H. hepaticicola* under the ErM-permissive condition to the transcriptomic response of the thalloid liverwort *Marchantia paleacea* colonized by the AMS-forming fungus *Rhizophagus irregularis*[6] (see Materials & Methods). This analysis identified a total of 20 Orthogroups in *C. fissa* upregulated during ErM that had ortholog genes upregulated during AMS in *M. paleacea*. Among them, we detected the LysM-RLK receptor-encoding gene *EPR3* (Calfis_g18879 & Calfis_g34255), and the two genes, *VAPYRIN* (Calfis_g4017 & Calfis_g10154) and *LIN* (Calfis_g31157 & Calfis_g22484), encoding required proteins in angiosperms for the development of the intracellular symbiosis-specific structure called the infectosome (Table S7) [31–35]. This suggests that the infectosome might be an ancestral feature induced during intracellular symbiosis in plants. To experimentally test this hypothesis, we used the genetically tractable liverwort *M. paleacea*. *M. paleacea* transformed with *MpaLINb$_{pro}$:GUS* constructs were mock- or *R. irregularis*-inoculated and the GUS expression pattern was monitored 6 weeks later. In mock-inoculated conditions, lines

transformed with the *MpaLINb$_{pro}$:GUS* construct showed faint GUS activity in the ventral epidermis (Fig. 3). This basal expression was maintained in the plants colonized by *R. irregularis.* However, colonization by *R. irregularis* triggered and additional expression pattern in the area of the parenchyma hosting the fungus (Fig. 3).

Altogether, these RNAseq data across species and the promoter:GUS lines demonstrate that the intracellular symbionts-mediated local induction of the infectosome defines an ancestral symbiotic feature in land plants.

ERS is essential for intracellular symbioses

To determine the mechanisms governing the induction of the infectosome, we searched for transcription factors in the 20 Orthogroups found up-regulated during both ErM and AMS and identified one GRAS (GRAS17) and one containing ERF transcription factors (Table S7). Although other GRAS proteins are involved in symbiotic processes[36–38], their ability to directly bind DNA remains debatable. Thus, we decided to focus on the ERF transcription factors. This Orthogroup contains one gene in *C. fissa* (Calfis_g17180) and two paralogs in *M. paleacea* (Marpal_utg000001g0000321, Marpal_utg000040g0076701) which are up-regulated during AMS (Table S7). Promoter GUS fusions revealed basal level of expression for both *M. paleacea* genes in the thallus and the rhizoids in non-inoculated condition (Fig. 4). Upon colonization by *R. irregularis,* an additional expression domain was observed in the inner parenchyma hosting the AM fungus, mirroring the *MpaLINb$_{pro}$:GUS* pattern (Fig 4). To test for the putative symbiotic function of this ERF Orthogroup in intracellular-symbiosis, we generated two single and two double mutants by CRISPR/Cas9 in *M. paleacea* (Fig. S9). The mutants and empty-vector control lines were inoculated with *R. irregularis* and harvested after 6 weeks. Thalli from the control showed accumulation of the AMS-induced pigment (Fig. 4), and displayed intracellular hyphae and arbuscules (Fig. 4). By contrast, not a single thallus of the two independent double mutant lines showed the accumulation of the pigment, or any colonization by *R. irregularis* (Fig. 4), mirroring the phenotypes of mutants in the CSP genes[18,19]. Given the complete lack of colonization of the double mutants, we named these ERFs *ERF Required for Symbiosis 1* and *2* (*ERS1* Marpal_utg000001g0000321 and *ERS2* Marpal_utg000040g0076701). Single mutants of one of the two paralogs, *MpaERS2,* displayed reduced colonization in three out of four independent experiments (Fig. S10), indicative of genetic redundancy between the two genes. To define the potential link of the ERS transcription factors with intracellular symbioses in other clades, we conducted a phylogenetic analysis on 127 plant species (Table S8). This analysis revealed 3 ERS clades. The first one is specific to non-flowering plants, including *MpaERS1* and *MpaERS2,* as well as the *C. fissa* ERS upregulated during ErM (Fig. 4). The other two clades likely originate from a duplication at the base of the angiosperms. One of these clades includes three genes from the legume *M. truncatula* – *ERN1*, *ERN2* and *ERN3* – which are known for their redundant symbiotic role during establishment of the nitrogen-fixing root nodule symbiosis in that species[39] (Fig. 4). The second angiosperm clade encompasses three other *M.*

truncatula genes including *ERM1*. *M. truncatula erm1* mutants show a quantitative defect in AM symbiosis[39]. We cross-referenced this phylogenetic data with transcriptomic data comparing mock-inoculated and symbiont-inoculated (root nodule or AM symbioses) plants available for 13 of the species represented on the tree. All tested species showed at least one ERS homolog up-regulated in response to AMS and/or RNS (Fig. 4, Table S9). Given this conserved symbiotic induction, and the phenotype of the *M. paleacea ers1ers2* mutant, we reasoned that ERS might regulate the expression of the infectosome. In angiosperm, Medicago ERN1 regulates gene expression via binding to the GCCGGC motif[40]. To determine whether this link might be conserved across land plants, we scanned the $VAPYRIN_{pro}$ and LIN_{pro} sequences across land plant species for the presence of this motif, and detected it in, respectively, 106 out of 168 promoters (~63%) and 147 out of 220 promoters (67%), indicating the conservation of the ERS – Infectosome logics between angiosperm and bryophytes (Table S10). To further test this link, we conducted *trans*-activation assays in *Nicotiana benthamiana* leaves with *MpaERS1* or *MpaERS2* on the $MpaLINb_{pro}$. While the $MpaLINb_{pro}$ did not show any basal expression (Fig. 4), co-expression with either of the two *MpaERS* led to its strong trans-activation (Fig. 4). Collectively, these data define *ERS* as a clade of transcription factors which have regulated the symbiotic induction of the infectosome across symbiotic types and plants lineages for the last 450 million years.

Regulation of ERS by the CSP across plant symbioses

The CSP is a constitutively expressed pathway whose activation, mediated by the perception of symbiont-derived signals, leads to the phosphorylation of the transcription factor CYCLOPS by the kinase CCaMK. The CSP is phylogenetically linked with intracellular symbiosis[5] (Fig. 2). Our data indicate that the ERS - infectosome module is induced by diverse symbionts across plant lineages and allows – at the cellular level – the actual intracellular colonization to occur. We reasoned that the direct regulation of ERS by phosphorylated CYCLOPS would integrate these two seemingly independent pathways, connecting symbiont perception to the transcriptional induction of the infection program. In angiosperms, this theory is supported in the context of the derived nitrogen-fixing root nodule symbiosis with the direct regulation of *ERN1* by CYCLOPS via a *cis*-regulatory element (CYC-RE) in the $ERN1_{pro}$[41] To determine whether this logic was conserved in Jungermanniales, we first conducted transactivation assays with the *CYCLOPS* ortholog genes of *C. fissa* (*CfCYCLOPSa* and *CfCYCLOPSb*) and another leafy liverwort, *Scapania nemorea* (*SnCYCLOPS*). As a control, we used *CYCLOPS* from *Medicago truncatula*[18]. We expressed either *CfCYCLOPS* or *SnCYCLOPS* in combination with an autoactive version of CCaMK (MtCCaMK-Kin)[42] and the reporter construct $CYC\text{-}RE_{pro}:GUS$[18,43] in *N. benthamiana* leaves. As expected, no GUS activity was detected in *N. benthamiana* leaves transformed with $CYC\text{-}RE_{pro}:GUS$ alone (Fig. 4G). Strong GUS activity was observed upon co-transformation of the *MtCCaMK-Kin* (Fig. 4G) construct with either *CfCYCLOPS* (a and b), *SnCYCLOPS*, or with the positive *MtCYCLOPS* control construct. In addition, constructs mimicking CYCLOPS phosphorylation

(*CYCLOPS*^{DD44}, *CfCYCLOPSa*^{DD}, *CfCYCLOPSb*^{DD} and *MpaCYCLOPS*^{DD}, induced the *CYC-RE*$_{pro}$*:GUS* (Fig. S11).

Using an untargeted promoter analysis, we found a highly conserved motif which corresponds to the CYC-RE in the promoters of the *ERS* homologs across land plants, further supporting the link between the CSP and ERS (Fig. 4, fig. S12). We evaluated whether the activation of *CYC-RE*$_{pro}$*:GUS* by various CYCLOPS translated to the ERS promoters themselves. As for the *CYC-RE*$_{pro}$*:GUS* above, we conducted transactivation assays using constructs mimicking CYCLOPS phosphorylation (CYCLOPS^{DD}) from *M. paleacea* (*MpaCYCLOPS*^{DD}) and *C. fissa* (*CsCYCLOPSa*^{DD}, *CsCYCLOPSb*^{DD}) (Fig. 4H) combined with either *MpaERS1*$_{pro}$*:GUS, MpaERS2*$_{pro}$*:GUS or CfERS*$_{pro}$*:GUS.* The *MpaERS1*$_{pro}$*:GUS* displayed basal expression, while the other two ERS promoters alone did not display GUS signals (Fig. 4H). Co-expression with any of the *CYCLOPS*^{DD} constructs from *M. paleacea* or *C. fissa* led to the induction of both *CfERS*$_{pro}$*:GUS* and *MpaERS2*$_{pro}$*:GUS* (Fig. 4), confirming the link between the CSP and the ERS.

Altogether, these phylogenetic and experimental data demonstrate that transcription factors of the ERS clade have been acting as a transcriptional link between the CSP and the infectosome since the most recent common ancestor of the land plants. Diversification of symbiotic options has been associated with the convergent recruitment of the CSP – ERS – Infectosome programme. Conceptually, the diversification of receptors upstream the CSP may be sufficient for this recruitment. The discovery that a conserved infectosome core module underlies the intracellular accommodation of completely distinct microbial endophytes in alternative symbiotic types, and the identification of a single clade of transcription factors associated with infectosome induction, provides valuable insights for guiding efforts to develop synthetic intracellular symbioses in land plants.

References

1. Taylor, T. N. & Osborn, J. M. The importance of fungi in shaping the paleoecosystem. *Review of Palaeobotany and Palynology* **90**, 249–262 (1996).
2. Brundrett, M. C. Coevolution of roots and mycorrhizas of land plants. *New Phytologist* **154**, 275–304 (2002).
3. Brundrett, M. C. & Tedersoo, L. Evolutionary history of mycorrhizal symbioses and global host plant diversity. *New Phytologist* **220**, 1108–1115 (2018).
4. Delaux, P.-M., Radhakrishnan, G. V., Jayaraman, D., Cheema, J., Malbreil, M., Volkening, J. D., Sekimoto, H., Nishiyama, T., Melkonian, M., Pokorny, L., Rothfels, C. J., Sederoff, H. W., Stevenson, D. W., Surek, B., Zhang, Y., Sussman, M. R., Dunand, C., Morris, R. J., Roux, C., Wong, G. K.-S., Oldroyd, G. E. D. & Ané, J.-M. Algal ancestor of land plants was preadapted for symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 13390–13395 (2015).
5. Radhakrishnan, G. V., Keller, J., Rich, M. K., Vernié, T., Mbadinga Mbadinga, D. L., Vigneron, N., Cottret, L., Clemente, H. S., Libourel, C., Cheema, J., Linde, A.-M., Eklund, D. M., Cheng, S., Wong, G. K. S., Lagercrantz, U., Li, F.-W., Oldroyd, G. E. D. & Delaux, P.-M.

An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages. *Nat. Plants* **6**, 280–289 (2020).
6. Rich, M. K., Vigneron, N., Libourel, C., Keller, J., Xue, L., Hajheidari, M., Radhakrishnan, G. V., Le Ru, A., Diop, S. I., Potente, G., Conti, E., Duijsings, D., Batut, A., Le Faouder, P., Kodama, K., Kyozuka, J., Sallet, E., Bécard, G., Rodriguez-Franco, M., Ott, T., Bertrand-Michel, J., Oldroyd, G. E. D., Szövényi, P., Bucher, M. & Delaux, P.-M. Lipid exchanges drove the evolution of mutualism during plant terrestrialization. *Science* **372**, 864–868 (2021).
7. Redecker, D., Kodner, R. & Graham, L. E. Glomalean Fungi from the Ordovician. *Science* **289**, 1920–21 (2000).
8. Krings, M., Taylor, T. N., Hass, H., Kerp, H., Dotzler, N. & Hermsen, E. J. Fungal endophytes in a 400-million-yr-old land plant: infection pathways, spatial distribution, and host responses. *New Phytologist* **174**, 648–657 (2007).
9. Wellman, C. H. The nature and evolutionary relationships of the earliest land plants. *New Phytologist* **202**, 1–3 (2014).
10. Cullings, K. W. Single phylogenetic origin of ericoid mycorrhizae within the Ericaceae. *Can. J. Bot.* **74**, 1896–1909 (1996).
11. Schwery, O., Onstein, R. E., Bouchenak-Khelladi, Y., Xing, Y., Carter, R. J. & Linder, H. P. As old as the mountains: the radiations of the Ericaceae. *New Phytologist* **207**, 355–367 (2015).
12. Kowal, J., Pressel, S., Duckett, J. G., Bidartondo, M. I. & Field, K. J. From rhizoids to roots? Experimental evidence of mutualism between liverworts and ascomycete fungi. *Annals of Botany* **121**, 221–227 (2018).
13. Pressel, S., Bidartondo, M. I., Field, K. J. & Duckett, J. G. Advances in understanding of mycorrhizal-like associations in bryophytes. *BDE* **43**, (2021).
14. Read, D. J., Duckett, J. G., Francis, R., Ligrone, R. & Russell, A. Symbiotic fungal associations in 'lower' land plants. *Phil. Trans. R. Soc. Lond. B* **355**, 815–831 (2000).
15. Read, D. J., Leake, J. R. & Perez-Moreno, J. Mycorrhizal fungi as drivers of ecosystem processes in heathland and boreal forest biomes. *Can. J. Bot.* **82**, 1243–1263 (2004).
16. Ward, E. B., Duguid, M. C., Kuebbing, S. E., Lendemer, J. C. & Bradford, M. A. The functional role of ericoid mycorrhizal plants and fungi on carbon and nitrogen dynamics in forests. *New Phytologist* **235**, 1701–1718 (2022).
17. Parniske, M. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nat Rev Microbiol* **6**, 763–775 (2008).
18. Vernié, T., Rich, M., Pellen, T., Teyssier, E., Garrigues, V., Chauderon, L., Medioni, L., Van Beveren, F., Libourel, C., Keller, J., Girou, C., Lefort, C., Le Ru, A., Martinez, Y., Reinhardt, D., Kodama, K., Shimazaki, S., Morel, P., Kyozuka, J., Mbengue, M., Vandenbussche, M. & Delaux, P.-M. Conservation of symbiotic signaling since the most recent common ancestor of land plants. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2408539121 (2025).
19. Rich, M., Vernié, T., Tiwari, M., Chauderon, L., Causse, J., Boussaroque, A., Bianconi, M. E., Vandenbussche, M., Valdes-Lopez, O., Mbengue, M., Ané, J.-M. & Delaux, P.-M. EPP1 is an ancestral component of the plant Common Symbiosis Pathway. (2025).
20. Wang, B., Yeun, L. H., Xue, J., Liu, Y., Ané, J. & Qiu, Y. Presence of three mycorrhizal genes in the common ancestor of land plants suggests a key role of mycorrhizas in the colonization of land by plants. *New Phytologist* **186**, 514–525 (2010).

21. Bravo, A., York, T., Pumplin, N., Mueller, L. A. & Harrison, M. J. Genes conserved for arbuscular mycorrhizal symbiosis identified through phylogenomics. *Nature Plants* **2**, 15208 (2016).
22. Smith, S. E. & Read, D. J. *Mycorrhizal Symbiosis*. (2010).
23. Li, X.-R., Sun, J., Albinsky, D., Zarrabian, D., Hull, R., Lee, T., Jarratt-Barnham, E., Chiu, C. H., Jacobsen, A., Soumpourou, E., Albanese, A., Kohlen, W., Luginbuehl, L. H., Guillotin, B., Lawrensen, T., Lin, H., Murray, J., Wallington, E., Harwood, W., Choi, J., Paszkowski, U. & Oldroyd, G. E. D. Nutrient regulation of lipochitooligosaccharide recognition in plants via NSP1 and NSP2. *Nat Commun* **13**, 6421 (2022).
24. Shi, J., Wang, X. & Wang, E. Mycorrhizal Symbiosis in Plant Growth and Stress Adaptation: From Genes to Ecosystems. *Annu. Rev. Plant Biol.* **74**, 569–607 (2023).
25. Breuillin-Sessoms, F., Floss, D. S., Gomez, S. K., Pumplin, N., Ding, Y., Levesque-Tremblay, V., Noar, R. D., Daniels, D. A., Bravo, A., Eaglesham, J. B., Benedito, V. A., Udvardi, M. K. & Harrison, M. J. Suppression of Arbuscule Degeneration in *Medicago truncatula phosphate transporter4* Mutants Is Dependent on the Ammonium Transporter 2 Family Protein AMT2;3. *The Plant Cell* **27**, 1352–1366 (2015).
26. Zhao, B., Jia, X., Yu, N., Murray, J. D., Yi, K. & Wang, E. Microbe-dependent and independent nitrogen and phosphate acquisition and regulation in plants. *New Phytol* **242**, 1507–22 (2024).
27. Villarreal A., J. C., Crandall-Stotler, B. J., Hart, M. L., Long, D. G. & Forrest, L. L. Divergence times and the evolution of morphological complexity in an early land plant lineage (Marchantiopsida) with a slow molecular rate. *New Phytologist* **209**, 1734–1746 (2016).
28. Luginbuehl, L. H. & Oldroyd, G. E. D. Understanding the Arbuscule at the Heart of Endomycorrhizal Symbioses in Plants. *Current Biology* **27**, R952–R963 (2017).
29. Schiessl, K., Lilley, J. L. S., Lee, T., Tamvakis, I., Kohlen, W., Bailey, P. C., Thomas, A., Luptak, J., Ramakrishnan, K., Carpenter, M. D., Mysore, K. S., Wen, J., Ahnert, S., Grieneisen, V. A. & Oldroyd, G. E. D. NODULE INCEPTION Recruits the Lateral Root Developmental Program for Symbiotic Nodule Organogenesis in Medicago truncatula. *Current Biology* **29**, 3657-3668.e5 (2019).
30. Li, F.-W., Nishiyama, T., Waller, M., Frangedakis, E., Keller, J., Li, Z., Fernandez-Pozo, N., Barker, M. S., Bennett, T., Blázquez, M. A., Cheng, S., Cuming, A. C., De Vries, J., De Vries, S., Delaux, P.-M., Diop, I. S., Harrison, C. J., Hauser, D., Hernández-García, J., Kirbis, A., Meeks, J. C., Monte, I., Mutte, S. K., Neubauer, A., Quandt, D., Robison, T., Shimamura, M., Rensing, S. A., Villarreal, J. C., Weijers, D., Wicke, S., Wong, G. K.-S., Sakakibara, K. & Szövényi, P. Anthoceros genomes illuminate the origin of land plants and the unique biology of hornworts. *Nat. Plants* **6**, 259–272 (2020).
31. Liu, C.-W., Breakspear, A., Stacey, N., Findlay, K., Nakashima, J., Ramakrishnan, K., Liu, M., Xie, F., Endre, G., De Carvalho-Niebel, F., Oldroyd, G. E. D., Udvardi, M. K., Fournier, J. & Murray, J. D. A protein complex required for polar growth of rhizobial infection threads. *Nat Commun* **10**, 2848 (2019).
32. Feddermann, N., Duvvuru Muni, R. R., Zeier, T., Stuurman, J., Ercolin, F., Schorderet, M. & Reinhardt, D. The PAM1 gene of petunia, required for intracellular accommodation and morphogenesis of arbuscular mycorrhizal fungi, encodes a homologue of VAPYRIN: Intracellular accommodation of AM fungi. *The Plant Journal* **64**, 470–481 (2010).

33. Pumplin, N., Mondo, S. J., Topp, S., Starker, C. G., Gantt, J. S. & Harrison, M. J. *Medicago truncatula* Vapyrin is a novel protein required for arbuscular mycorrhizal symbiosis. *The Plant Journal* **61**, 482–494 (2010).

34. Murray, J. D., Muni, R. R. D., Torres-Jerez, I., Tang, Y., Allen, S., Andriankaja, M., Li, G., Laxmi, A., Cheng, X., Wen, J., Vaughan, D., Schultze, M., Sun, J., Charpentier, M., Oldroyd, G., Tadege, M., Ratet, P., Mysore, K. S., Chen, R. & Udvardi, M. K. *Vapyrin* , a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of *Medicago truncatula*. *The Plant Journal* **65**, 244–252 (2011).

35. Liu, M., Jia, N., Li, X., Liu, R., Xie, Q., Murray, J. D., Downie, J. A. & Xie, F. CERBERUS is critical for stabilization of VAPYRIN during rhizobial infection in *Lotus japonicus*. *New Phytologist* **229**, 1684–1700 (2021).

36. Liu, W., Kohlen, W., Lillo, A., Op Den Camp, R., Ivanov, S., Hartog, M., Limpens, E., Jamil, M., Smaczniak, C., Kaufmann, K., Yang, W.-C., Hooiveld, G. J. E. J., Charnikhova, T., Bouwmeester, H. J., Bisseling, T. & Geurts, R. Strigolactone Biosynthesis in *Medicago truncatula* and Rice Requires the Symbiotic GRAS-Type Transcription Factors NSP1 and NSP2. *The Plant Cell* **23**, 3853–3865 (2011).

37. Gobbato, E., Marsh, J. F., Vernié, T., Wang, E., Maillet, F., Kim, J., Miller, J. B., Sun, J., Bano, S. A., Ratet, P., Mysore, K. S., Dénarié, J., Schultze, M. & Oldroyd, G. E. D. A GRAS-Type Transcription Factor with a Specific Function in Mycorrhizal Signaling. *Current Biology* **22**, 2236–2241 (2012).

38. Xue, L., Cui, H., Buer, B., Vijayakumar, V., Delaux, P.-M., Junkermann, S. & Bucher, M. Network of GRAS Transcription Factors Involved in the Control of Arbuscule Development in *Lotus japonicus*. *Plant Physiology* **167**, 854–871 (2015).

39. Middleton, P. H., Jakab, J., Penmetsa, R. V., Starker, C. G., Doll, J., Kaló, P., Prabhu, R., Marsh, J. F., Mitra, R. M., Kereszt, A., Dudas, B., VandenBosch, K., Long, S. R., Cook, D. R., Kiss, G. B. & Oldroyd, G. E. D. An ERF Transcription Factor in *Medicago truncatula* That Is Essential for Nod Factor Signal Transduction. *The Plant Cell* **19**, 1221–1234 (2007).

40. Andriankaja, A., Boisson-Dernier, A., Frances, L., Sauviac, L., Jauneau, A., Barker, D. G. & De Carvalho-Niebel, F. AP2-ERF Transcription Factors Mediate Nod Factor–Dependent Mt *ENOD11* Activation in Root Hairs via a Novel *cis* -Regulatory Motif. *The Plant Cell* **19**, 2866–2885 (2007).

41. Cerri, M. R., Wang, Q., Stolz, P., Folgmann, J., Frances, L., Katzer, K., Li, X., Heckmann, A. B., Wang, T. L., Downie, J. A., Klingl, A., De Carvalho-Niebel, F., Xie, F. & Parniske, M. The *ERN 1* transcription factor gene is a target of the CC a MK / CYCLOPS complex and controls rhizobial infection in *Lotus japonicus*. *New Phytologist* **215**, 323–337 (2017).

42. Gleason, C., Chaudhuri, S., Yang, T., Muñoz, A., Poovaiah, B. W. & Oldroyd, G. E. D. Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. *Nature* **441**, 1149–1152 (2006).

43. Singh, S., Katzer, K., Lambert, J., Cerri, M. & Parniske, M. CYCLOPS, A DNA-Binding Transcriptional Activator, Orchestrates Symbiotic Root Nodule Development. *Cell Host & Microbe* **15**, 139–152 (2014).

44. Singh, S., Katzer, K., Lambert, J., Cerri, M. & Parniske, M. CYCLOPS, A DNA-Binding Transcriptional Activator, Orchestrates Symbiotic Root Nodule Development. *Cell Host & Microbe* **15**, 139–152 (2014).

45. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods* **18**, 170–175 (2021).
46. Uliano-Silva, M. MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio High Fidelity reads. (2022).
47. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
48. Baril, T., Galbraith, J. & Hayward, A. Earl Grey: A Fully Automated User-Friendly Transposable Element Annotation and Analysis Pipeline. *Molecular Biology and Evolution* **41**, msae068 (2024).
49. Gabriel, L., Brůna, T., Hoff, K. J., Ebel, M., Lomsadze, A., Borodovsky, M. & Stanke, M. BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. *Genome Res.* **34**, 769–777 (2024).
50. Kim, D. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* **37**, (2019).
51. Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M. & Li, H. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
52. Li, H. Protein-to-genome alignment with miniprot. *Bioinformatics* **39**, btad014 (2023).
53. Brůna, T., Lomsadze, A. & Borodovsky, M. GeneMark-ETP significantly improves the accuracy of automatic annotation of large eukaryotic genomes. *Genome Res.* **34**, 757–768 (2024).
54. Iwata, H. & Gotoh, O. Benchmarking spliced alignment programs including Spaln2, an extended version of Spaln that incorporates additional species-specific features. *Nucleic Acids Research* **40**, e161–e161 (2012).
55. Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12**, 59–60 (2015).
56. Kovaka, S., Zimin, A. V., Pertea, G. M., Razaghi, R., Salzberg, S. L. & Pertea, M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol* **20**, 278 (2019).
57. Pertea, G. M. & Pertea, M. GFF Utilities: GffRead and GffCompare. (2020).
58. Stanke, M., Schöffmann, O., Morgenstern, B. & Waack, S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* **7**, 62 (2006).
59. Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics* **24**, 637–644 (2008).
60. Gabriel, L., Hoff, K. J., Brůna, T., Borodovsky, M. & Stanke, M. TSEBRA: transcript selector for BRAKER. *BMC Bioinformatics* **22**, 566 (2021).
61. Huang, N. & Li, H. compleasm: a faster and more accurate reimplementation of BUSCO. *Bioinformatics* **39**, btad595 (2023).
62. Dainat, J. AGAT: Another Gff Analysis Toolkit to handle annotations in any GTF/GFF format. (2022).
63. Blum, M., Andreeva, A., Florentino, L. C., Chuguransky, S. R., Grego, T., Hobbs, E., Pinto, B. L., Orr, A., Paysan-Lafosse, T., Ponamareva, I., Salazar, G. A., Bordin, N., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunic, I., Llinares-López, F., Marchler-Bauer, A., Meng-Papaxanthos, L., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P.,

Piovesan, D., Rivoire, C., Sigrist, C. J. A., Thanki, N., Thibaud-Nissen, F., Thomas, P. D., Tosatto, S. C. E., Wu, C. H. & Bateman, A. InterPro: the protein sequence classification resource in 2025. *Nucleic Acids Research* **53**, D444–D456 (2025).
64. Lee, E., Helt, G. A., Reese, J. T., Munoz-Torres, M. C., Childers, C. P., Buels, R. M., Stein, L., Holmes, I. H., Elsik, C. G. & Lewis, S. E. Web Apollo: a web-based genomic annotation editing platform. *Genome Biol* **14**, R93 (2013).
65. Chen, H. & Zwaenepoel, A. Inference of Ancient Polyploidy from Genomic Data. in *Polyploidy* (ed. Van De Peer, Y.) vol. 2545 3–18 (Springer US, New York, NY, 2023).
66. Tang, H., Krishnakumar, V., Zeng, X., Xu, Z., Taranto, A., Lomas, J. S., Zhang, Y., Huang, Y., Wang, Y., Yim, W. C., Zhang, J. & Zhang, X. JCVI: A versatile toolkit for comparative genomics analysis. *iMeta* **3**, e211 (2024).
67. Smith, G. S., Johnston, C. M. & Cornforth, I. S. COMPARISON OF NUTRIENT SOLUTIONS FOR GROWTH OF PLANTS IN SAND CULTURE. *New Phytologist* **94**, 537–548 (1983).
68. Atlas, M. R. *Handbook of Microbiological Media*. (CRC Press, 2010).
69. Abràmoff, D. M. D. Image Processing with ImageJ. *Biophotonics International* **11**, 36–42 (2004).
70. Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E. & Notredame, C. Nextflow enables reproducible computational workflows. *Nat Biotechnol* **35**, 316–319 (2017).
71. Ewels, P. A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M. U., Di Tommaso, P. & Nahnsen, S. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol* **38**, 276–278 (2020).
72. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
73. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
74. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
75. Sayols, S., Scherzinger, D. & Klein, H. dupRadar: a Bioconductor package for the assessment of PCR artifacts in RNA-Seq data. *BMC Bioinformatics* **17**, 428 (2016).
76. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. (2011).
77. Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Batut, P., Chaisson, M. & Gingeras, T. R. STAR: ultrafast universal RNA-seq aligner. (2012).
78. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294 (2016).
79. Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012).
80. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417–419 (2017).
81. Morgan, M., Obenchain, V., Hester, J. & Pagès, H. SummarizedExperiment: A container (S4 class) for matrix-like assays. (2024).
82. Love, M. I., Soneson, C., Hickey, P. F., Johnson, L. K., Pierce, N. T., Shepherd, L., Morgan, M. & Patro, R. Tximeta: Reference sequence checksums for provenance identification in RNA-seq. *PLoS Comput Biol* **16**, e1007664 (2020).

83. Chen, Y., Chen, L., Lun, A. T. L. & Baldoni, P. L. edgeR 4.0: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. (2024).
84. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X. & Yu, G. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *The Innovation* **2**, 100141 (2021).
85. Libourel, C., Keller, J., Brichet, L., Cazalé, A.-C., Carrère, S., Vernié, T., Couzigou, J.-M., Callot, C., Dufau, I., Cauet, S., Marande, W., Bulach, T., Suin, A., Masson-Boivin, C., Remigi, P., Delaux, P.-M. & Capela, D. Comparative phylotranscriptomics reveals ancestral and derived root nodule symbiosis programmes. *Nat. Plants* **9**, 1067–1080 (2023).
86. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* **20**, 238 (2019).
87. Emms, D. M. & Kelly, S. STAG: Species Tree Inference from All Genes. Preprint at https://doi.org/10.1101/267914 (2018).
88. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. & Madden, T. L. BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
89. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* **30**, 772–780 (2013).
90. Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972–1973 (2009).
91. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS ONE* **5**, e9490 (2010).
92. Edgar, R. C. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat Commun* **13**, 6968 (2022).
93. Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A. & Lanfear, R. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**, 1530–1534 (2020).
94. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. & Noble, W. S. MEME Suite: Tools for motif discovery and searching. *Nucleic Acids Research* **37**, 202–208 (2009).
95. Cerri, M. R., Frances, L., Kelner, A., Fournier, J., Middleton, P. H., Auriac, M.-C., Mysore, K. S., Wen, J., Erard, M., Barker, D. G., Oldroyd, G. E. & De Carvalho-Niebel, F. The Symbiosis-Related ERN Transcription Factors Act in Concert to Coordinate Rhizobial Host Root Infection. *Plant Physiol.* pp.00230.2016 (2016) doi:10.1104/pp.16.00230.
96. Mai, U. & Mirarab, S. TreeShrink: fast and accurate detection of outlier long branches in collections of phylogenetic trees. *BMC Genomics* **19**, 272 (2018).
97. Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N., Matschiner, M., Mendes, F. K., Müller, N. F., Ogilvie, H. A., Du Plessis, L., Popinga, A., Rambaut, A., Rasmussen, D., Siveroni, I., Suchard, M. A., Wu, C.-H., Xie, D., Zhang, C., Stadler, T. & Drummond, A. J. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* **15**, e1006650 (2019).
98. Douglas, J., Zhang, R. & Bouckaert, R. Adaptive dating and fast proposals: Revisiting the phylogenetic relaxed clock model. *PLoS Comput Biol* **17**, e1008322 (2021).
99. Harris, B. J., Clark, J. W., Schrempf, D., Szöllősi, G. J., Donoghue, P. C. J., Hetherington, A. M. & Williams, T. A. Divergent evolutionary trajectories of bryophytes and

tracheophytes from a complex common ancestor of land plants. *Nat Ecol Evol* **6**, 1634–1643 (2022).
100. Hu, R., Li, X., Hu, Y., Zhang, R., Lv, Q., Zhang, M., Sheng, X., Zhao, F., Chen, Z., Ding, Y., Yuan, H., Wu, X., Xing, S., Yan, X., Bao, F., Wan, P., Xiao, L., Wang, X., Xiao, W., Decker, E. L., Van Gessel, N., Renault, H., Wiedemann, G., Horst, N. A., Haas, F. B., Wilhelmsson, P. K. I., Ullrich, K. K., Neumann, E., Lv, B., Liang, C., Du, H., Lu, H., Gao, Q., Cheng, Z., You, H., Xin, P., Chu, J., Huang, C.-H., Liu, Y., Dong, S., Zhang, L., Chen, F., Deng, L., Duan, F., Zhao, W., Li, K., Li, Z., Li, X., Cui, H., Zhang, Y. E., Ma, C., Zhu, R., Jia, Y., Wang, M., Hasebe, M., Fu, J., Goffinet, B., Ma, H., Rensing, S. A., Reski, R. & He, Y. Adaptive evolution of the enigmatic Takakia now facing climate change in Tibet. *Cell* **186**, 3558-3576.e17 (2023).
101. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* **67**, 901–904 (2018).
102. Engler, C. & Marillonnet, S. Golden Gate Cloning. in *DNA Cloning and Assembly Methods* (eds Valla, S. & Lale, R.) vol. 1116 119–131 (Humana Press, Totowa, NJ, 2014).
103. Patron, N. J., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., Raitskin, O., Leveau, A., Farré, G., Rogers, C., Smith, A., Hibberd, J., Webb, A. A. R., Locke, J., Schornack, S., Ajioka, J., Baulcombe, D. C., Zipfel, C., Kamoun, S., Jones, J. D. G., Kuhn, H., Robatzek, S., Van Esse, H. P., Sanders, D., Oldroyd, G., Martin, C., Field, R., O'Connor, S., Fox, S., Wulff, B., Miller, B., Breakspear, A., Radhakrishnan, G., Delaux, P., Loqué, D., Granell, A., Tissier, A., Shih, P., Brutnell, T. P., Quick, W. P., Rischer, H., Fraser, P. D., Aharoni, A., Raines, C., South, P. F., Ané, J., Hamberger, B. R., Langdale, J., Stougaard, J., Bouwmeester, H., Udvardi, M., Murray, J. A. H., Ntoukakis, V., Schäfer, P., Denby, K., Edwards, K. J., Osbourn, A. & Haseloff, J. Standards for plant synthetic biology: a common syntax for exchange of DNA parts. *New Phytologist* **208**, 13–19 (2015).
104. Andriankaja, A., Boisson-Dernier, A., Frances, L., Sauviac, L., Jauneau, A., Barker, D. G. & De Carvalho-Niebel, F. AP2-ERF Transcription Factors Mediate Nod Factor–Dependent Mt *ENOD11* Activation in Root Hairs via a Novel *cis* -Regulatory Motif. *The Plant Cell* **19**, 2866–2885 (2007).
105. Harris, L. W. & Davies, T. J. A Complete Fossil-Calibrated Phylogeny of Seed Plant Families as a Tool for Comparative Analyses: Testing the 'Time for Speciation' Hypothesis. *PLoS ONE* **11**, e0162907 (2016).

Acknowledgments

We are grateful to Silvia Perotto (University of Turin) for kindly providing the *Hyaloscypha hepaticicola* (syn. *Rhizoscyphus ericae* UAMH 6735) strain used in this study. We thank the genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul, https://doi.org/10.15454/1.5572369328961167E12), the NIG supercomputer at ROIS National Institute of Genetics, and the Data Integration and Analysis Facility at the National Institute for Basic Biology for providing computing resources. We acknowledge the TRI-FRAIB imaging facility, member of the national infrastructure France-BioImaging supported by the French National Research Agency (ANR-10-INBS-04). Celophane was provided by Futamura Chemical, Co., LTD. We thank Dr. Elizabeth Jamet (LRSV, CNRS) for kindly allowing us to use the camera for phenotyping. We are grateful to Dr.

M. Shimamura for letting us know the change in taxonomy. **Funding:** Research performed at LRSV was also supported by the "Laboratoires d'Excellence (LABEX)' TULIP (ANR-10-LABX-41) and the "École Universitaire de Recherche (EUR)" TULIP-GS (ANR-18-EURE-0019). This work was supported by the project Engineering Nitrogen Symbiosis for Africa (ENSA) funded through a grant to the University of Cambridge by the Bill and Melinda Gates Foundation (OPP1172165) and the UK Foreign, Commonwealth and Development Office as Engineering Nitrogen Symbiosis for Africa (OPP1172165); the German Research Foundation (DFG Walter Benjamin fellowship project number 536856410 to KM); the Horizon Europe programme (MSCA-PF grant 101105838 'SYMBIOLOSS' to MEB); and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no.101001675-ORIGINS) to P.-M.D. MEXT and JSPS Kakenhi JP22128008 and JP23657063 to TN. *S. infuscum* culture to library construction were performed by K. Yamada. Illumina sequencing of S. infuscum was performed by Dr. Katushi Yamaguchi and Dr. Shuji Shigenobu at the Core facility of NIBB through NIBB Collaborative Research Program (12-716) to TN. **Author contributions:** Conceptualization: PMD, TV, KM, LCa, JK ; Methodology: ALR, CC, CLe, CLi, LC, LCh, MEB, OT, SA, TP, TV, KM, CRL, DW ; Investigation: CLi, LCa, LCh, MEB, SA,TP, TV, KM, DW, LF ; Visualization: LCa, MEB, SA, TV, KM, DW, SA ; Funding acquisition: PMD, KM, MEB ; Project administration: PMD ; Supervision: PMD, TV, KM, JK, FWL ; Writing – original draft: LCa,PMD, TV, KM, JK ; Writing – review & editing: all authors

Competing interests: Some findings in this manuscript are considered for patent application.

Data and materials availability: The *Solenostoma infuscum* genome assembly with annotation and raw reads are available on INSDC under the BioProject accession PRJDB19613. The *Calypogeia fissa* genome assembly and annotation is availablee on ENA-EMBL, with the reference number PRJEB86056. All DEG results, functional annotation, orthogroups and template scripts have been deposited in the repository 10.57745/MD77E8. An interactive Genome Browser can be found and queried here: https://bbric-pipelines.toulouse.inra.fr/myGenomeBrowser?browse=1&portalname=Calypogeia_fissa&owner=jean.keller@cnrs.fr&key=DKW4edLL. The raw RNAseq read of *Calypogeia fissa* are available on the SRA database, with the reference number PRJNA1186220. All read counts, orthogroups, differentially expressed genes and functional annotation of *C. fissa* can be accessed at: https://doi.org/10.57745/MD77E8.

Supplementary Materials

Figure S1-S12

Tables S1 to S12

References (*45–105*)

Figure captions

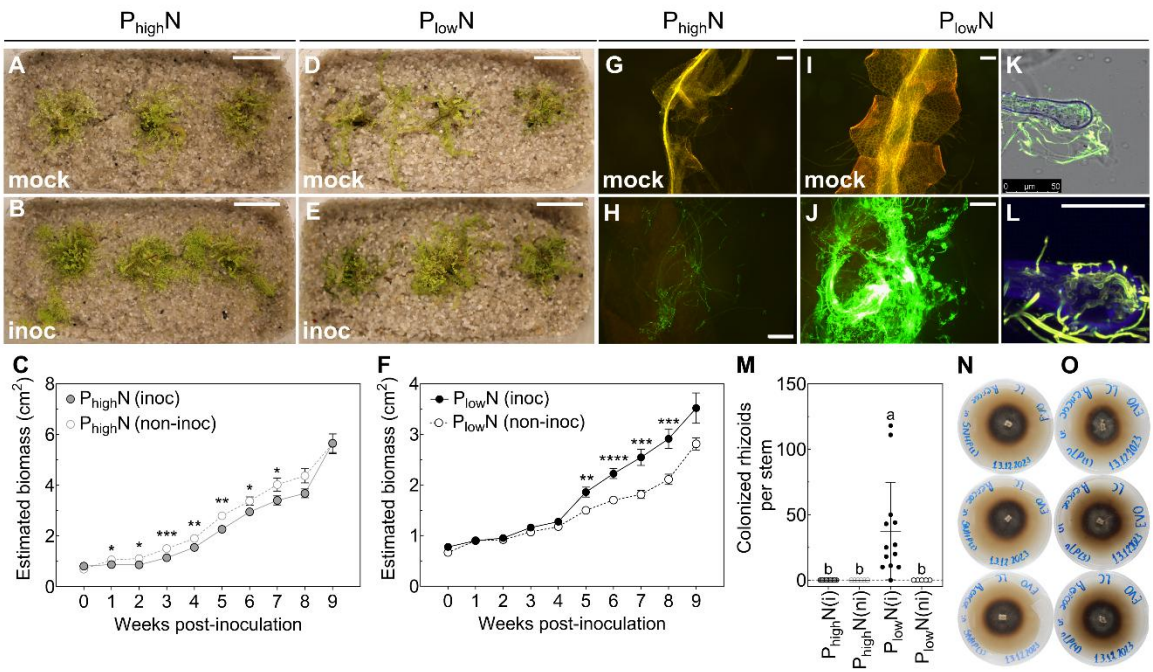

Figure 1. Establishment of Ericoid mycorrhizae in *Calypogeia fissa* relies on a nutrient starvation-induced permissive state.

(A-F) Representative images (A, B, D, E) and estimated growth (C, F) of *Calypogeia fissa* plants in mock-inoculated (non-inoc) and inoculated (inoc) treatments. Scale bars: 2 cm. Data points in (C, F) indicate the mean ± SEM (n = 27) of technical replicates per time-point from one representative experiment of a total of three independent experiments. **(G-L)** Representative fluorescence and confocal microscopy images of *C. fissa* stems harvested from non-inoc (A, D) and inoc (B, E) treatments, and stained with WGA. Plants were photographed and harvested after 8 weeks of growth. Scale bars: 200 µm Confocal microscopy images of samples shown in (K, L), were further stained with Congo red (blue). Scale bar: 25 µm. **(M)** Number of swollen and colonized rhizoid tips per stem in non-inoc (ni) and inoc (i) treatments. Data points indicate the mean ± SD (n = 4 to 13) of technical stem replicates from one experiment. Plants were grown under moderate supply of nitrogen [N], and either high phosphate [P] (A-C,G,H), or low [P] (D-F,I-L). *Hyaloscypha hepaticicola* was used in all inoculation experiments. Different characters indicate differences (P < 0.05) between treatments, based on a Welch's (double-tailed) t-test (C, F), or based on a Kruskal-Wallis test with post-hoc Dunn's test (M). **(N, O)** Representative images of *H. hepaticicola* grown under moderate [N] supply, and high [P] (N), or low [P] (O).

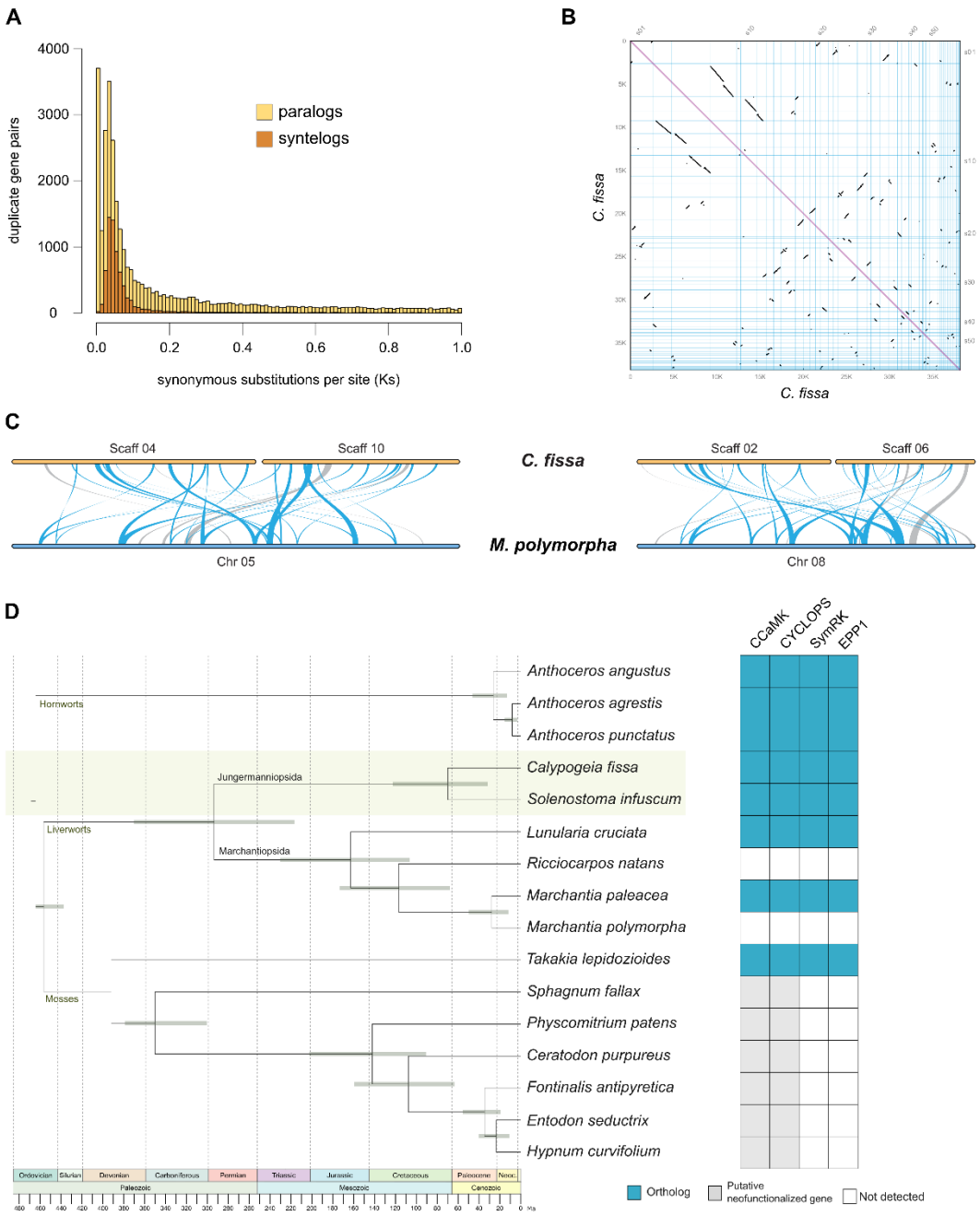

Figure 2. A recent whole-genome duplication in the leafy liverwort *Calypogeia fissa*. **(A)** Distribution of synonymous substitutions per site (Ks) between paralog pairs and syntelogs in the *C. fissa* genome. **(B)** Intragenomic synteny analysis identifies conserved collinear gene arrangements within the *C. fissa* genome, revealing 137 collinear blocks comprising 7,087 syntelogs and covering 37.2% of the *C. fissa* genome. **(C)** Synteny between *C. fissa* and *Marchantia polymorpha*. **(D)** Time-calibrated phylogeny of bryophyte species. Node bars represent the 95% highest posterior density (HPD) intervals. The root of the tree and the crown

node of mosses were fixed to the mean ages from [2105] (see Methods). Presence of selected symbiosis-related genes is shown on the right panel.

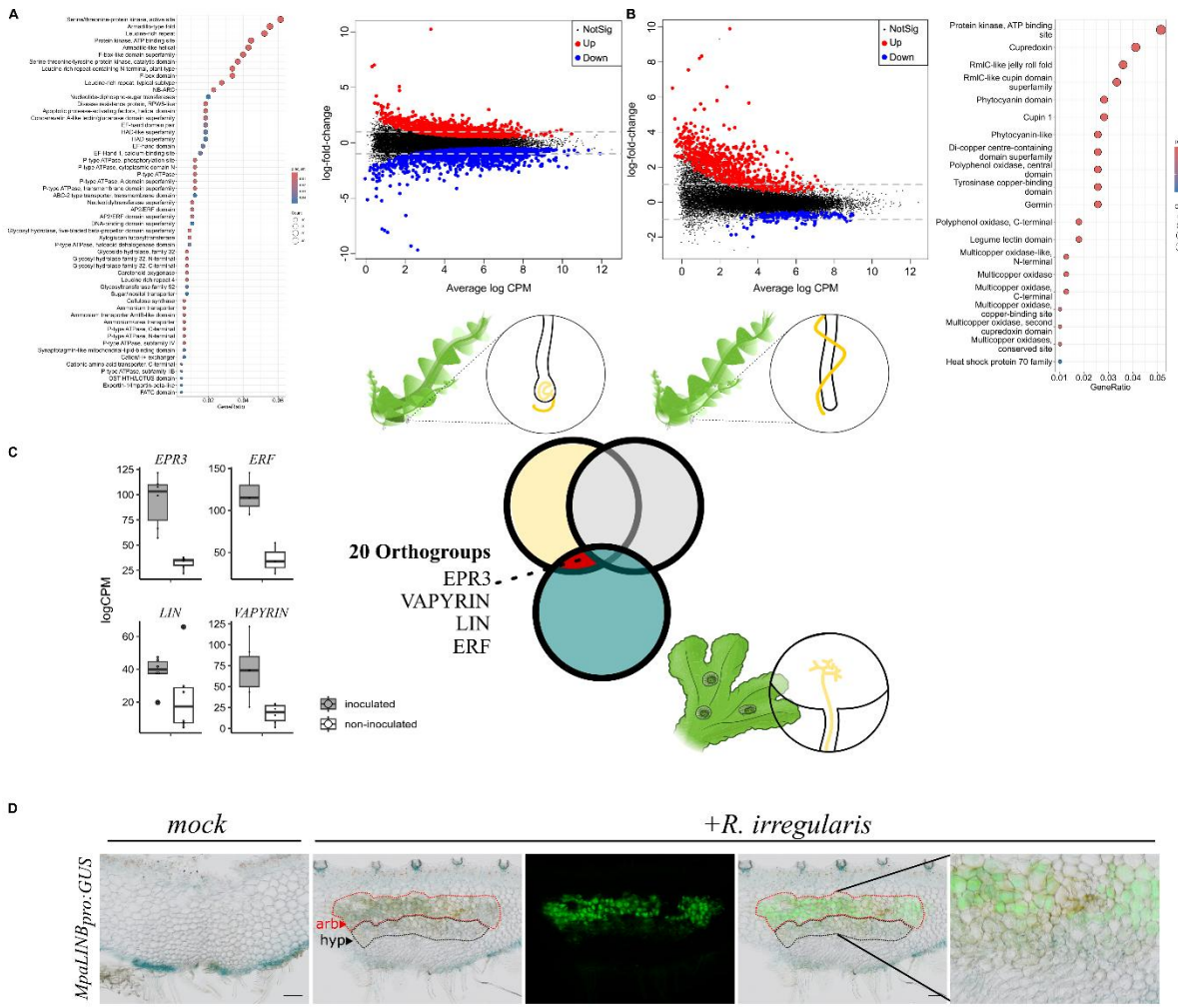

Figure 3. Transcriptional reprogramming during intracellular symbiosis. (A) Differential gene expression analysis comparing mock-treated and *Hyaloscypha hepaticicola*-inoculated *Calypogeia fissa* plants under ericoid mycorrhiza-permissive conditions. A total of 787 genes were upregulated, and 743 genes were downregulated in inoculated plants relative to mock-treated controls. **(B)** Differential gene expression analysis under ericoid mycorrhiza-permissive restrictive conditions, with 566/27 genes up/downregulated in inoculated plants compared to mock-treated controls. **(C)** Comparative transcriptomic analysis of ericoid mycorrhiza-permissive and restrictive conditions alongside arbuscular mycorrhiza-permissive conditions. The orange-overlapping area in the Venn diagram depicts the conserved expressed gene cluster that was found to be upregulated during both ericoid and arbuscular mycorrhizae. **(D)** Expression bar plots displaying expression of *C. fissa EPR3*, *ERF*, *LIN* and *VAPYRIN*. **(E)** Promoter activity of the *Marchantia paleacea LINb* gene monitored using *MpaLINBpro:GUS* at 6 weeks post inoculation. From left to right, mock inoculated (bright field), *R.irregularis*-inoculated (bright field), *R.irregularis*-inoculated (WGA-Alexafluor 488), *R.irregularis*-inoculated (merged), *R.irregularis*-inoculated (merged magnified). scale bar= 100 µM.

Arrowheads indicate the the hyphae- containing area (hyp, black- dotted line), and the arbuscule area (arb, red- dotted line).

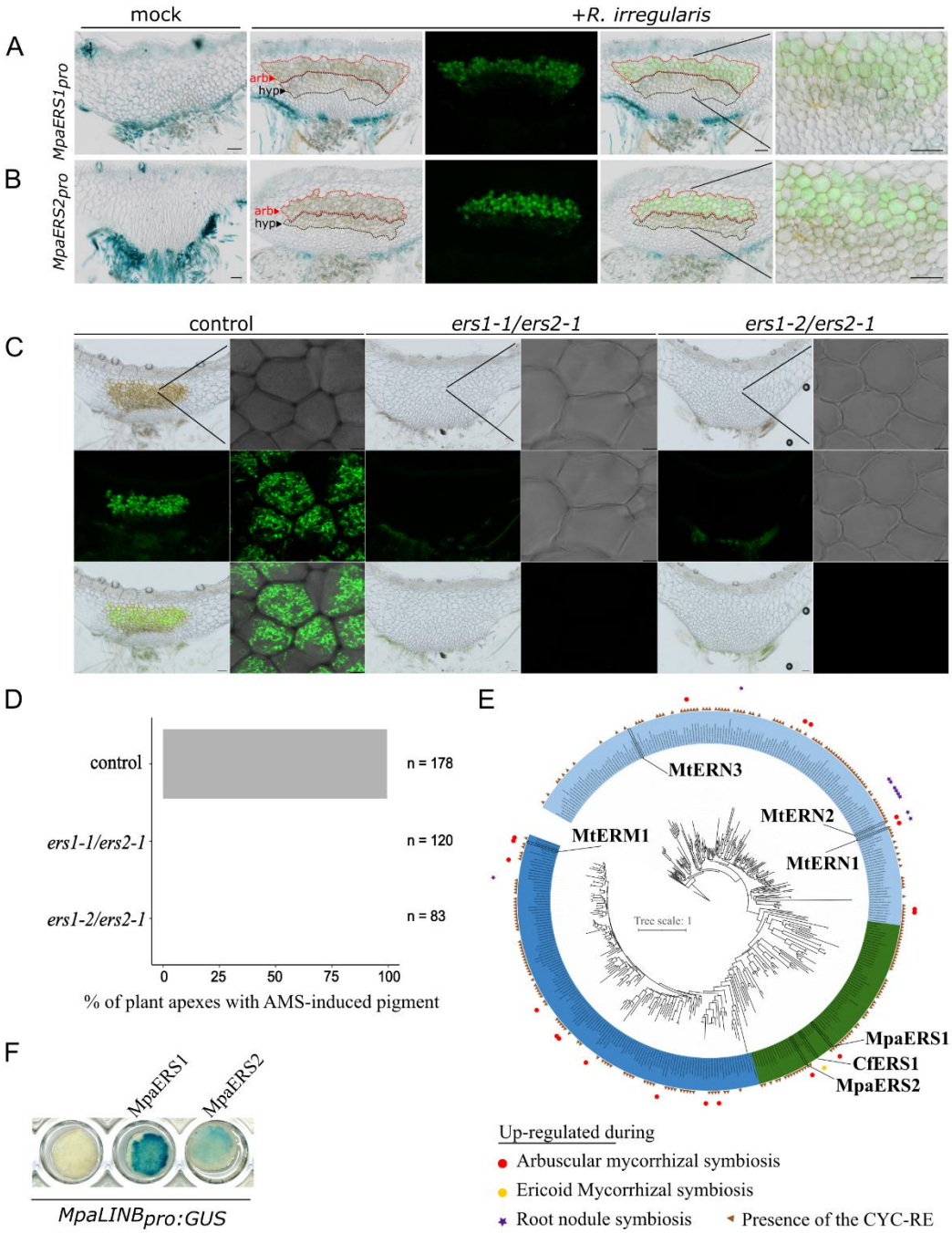
A
mock
+R. irregularis
MpaERS1pro
arb
hyp
B
MpaERS2pro
C
control
ers1-1/ers2-1
ers1-2/ers2-1
D
control
ers1-1/ers2-1
ers1-2/ers2-1
n = 178
n = 120
n = 83
0
25
50
75
100
% of plant apexes with AMS-induced pigment
E
MtERN3
MtERN2
MtERM1
MtERN1
Tree scale: 1
MpaERS1
CfERS1
MpaERS2
Up-regulated during
Arbuscular mycorrhizal symbiosis
Ericoid Mycorrhizal symbiosis
Root nodule symbiosis
Presence of the CYC-RE
F
MpaERS1
MpaERS2
MpaLINBpro:GUS
G
MtCYCLOPS
CfCYCLOPSa
CfCYCLOPSb
SnCYCLOPS
+ MtCCaMK-K
CYC-REpro:GUS
H
MpaCYCLOPSDD
CfCYCLOPSaDD
CfCYCLOPSbDD
MpaERS1pro:GUS
CfERSpro:GUS
MpaERS2pro:GUS

Figure 4. ERS bridge the CSP and the infectosome. (A-B) Promoter activity of the *Marchantia paleacea ERS1* and *ERS2* genes monitored using *MpaERS1pro:GUS* and *MpaERS2pro:GUS* at 6 weeks posr inoculation. From left to right, mock inoculated (bright field), *R.irregularis*-inoculated (bright field), *R.irregularis*-inoculated (WGA-Alexafluor 488), *R.irregularis*-inoculated (merged), *R.irregularis*-inoculated (merged magnified). scale bar= 100 µM. Arrowheads indicate the the hyphae- containing area (hyp, black- dotted line), and the arbuscule area (arb, red- dotted line). **(C)** Representative images of control plants transformed with an empty vector and *Mpaers* double mutants inoculated with *R. irregularis* at 6 weeks post inoculation. Green channel: WGA-Alexafluor 488 staining of fungal structures. Scale bar 100 µM or 10 µM. **(D)** Scoring of plant apexes with AMS-induced pigment at 6 weeks post inoculation with *R. irregularis* in control plants transformed with an empty vector and *Mpaers* double mutants. Combined results of 4 independent experiments are shown. **(E)** Reconstructed phylogeny of ERS orthologs across angiosperms (2 clades in blue) and bryopythes (clade in green), with the *CYC-RE* presence and up-regulated genes during different symbioses mapped on it. **(F)** *Nicotiana benthamiana* leaves expressing $MpaLINB_{pro}$*:GUS* alone or together with *MpaERS1* or *MpaERS2*. **(G)** *N. benthamiana* leaves expressing $CYC\text{-}RE_{pro}$*:GUS* alone or together with *MtCCAMK-K* and *CYCLOPS* from *Medicago truncatula* (*MtCYCLOPS*)**,** or *Calypogeia fissa* (*CsCYCLOPSa and b*) or *Scapania nemorea* (*SnCYCLOPS*)**. (H)** *N. benthamiana* leaves expressing ERS_{pro}:*GUS* fusions from *Marchantia paleacea* ($MpaERS1_{pro}$*:GUS*, $MpaERS2_{pro}$*:GUS) and C. fissa* ($CsERS_{pro}$*:GUS*) alone or together with constructs mimicking CYCLOPS phosphorylation ($CYCLOPS^{DD}$) from *M. paleacea* ($MpaCYCLOPS^{DD}$)**,** or *Calypogeia fissa* ($CsCYCLOPSa^{DD}$ and $CsCYCLOPSb^{DD}$). GUS staining was performed at 48HPI and is shown in blue.