

Plant material

Lunularia cruciata (*Lunularia*) culture was isolated in the Pyrennées mountains (France) and cultured *in vitro* on M medium (1). *Closterium peracerosum-strigosum-littorale* complex (strain NIES-67 - *Closterium*) was provided by the NIES (<http://mcc.nies.go.jp/02introduction-e.html>) and grown in MBL medium (2) at 20/24°C (16H light). *Spirogyra* sp. was cultivated in liquid 3N-BBM+V medium and kept shaking at 25 °C with a speed of 60 rpm as described previously (Delaux et al. 2012). *Medicago truncatula* wild-type and *ccamk-1* (TRV25) mutants, both in the Jemalong A17 background were used.

Treatment of *Lunularia* and *Closterium*

Lunularia cruciata young thallus produced from gemmae were inoculated with 100 spores of *Rhizophagus irregularis* (AM) or water (Mock) on M medium (3) and incubated at 20/24°C (16H light). Plants were collected after two month and frozen in liquid nitrogen before RNA extraction. Three weeks-old *Closterium* cultures were divided in 6 flasks (25ml culture in 125ml flasks) and treated for 24 hours with exudates of *Rhizophagus irregularis* (final concentration 400 spores/ml), generated as previously (4), or with water. Samples were then collected by centrifugation (1 min, 8,000g) and frozen in liquid nitrogen.

***Lunularia* and *Closterium* RNA extraction**

Frozen samples were grinded in liquid nitrogen and total RNA was extracted using RNeasy plant mini kit (Qiagen) following the manufacturer's recommendations. RNA quantification was performed by Nanodrop. RNA quality was assessed on Agilent Bioanalyzer. Samples with RNA integrity Number (RIN) higher than eight were selected.

Construction of RNA Seq Libraries for *Lunularia* and *Closterium*

Each RNA library was generated following Illumina TruSeq RNA Sample Preparation Guide and the Illumina TruSeq RNA Sample Preparation Kit (Illumina Inc., San Diego, California, USA). mRNA was purified from 1ug total RNA using poly-T oligo-attached magnetic beads. Subsequently, each mRNA sample was fragmented using divalent cations under elevated temperature, and purified. Double-stranded cDNAs were synthesized using SuperScript II (Invitrogen, Carlsbad, California, USA) Reverse Transcriptase and random primers for first strand cDNA synthesis followed by second strand synthesis using DNA Polymerase I and RNaseH for removal of mRNA. Double-stranded cDNA was purified using Agencourt AMPure XP beads (Qiagen, Valencia, California, USA) as recommended in the TruSeq RNA Sample Prep Guide. cDNAs were end-repaired by T4 DNA polymerase and Klenow DNA Polymerase, and phosphorylated by T4 polynucleotide kinase. The blunt ended cDNA was purified using Agencourt AMPure XP beads. The cDNA products were incubated with Klenow DNA Polymerase to add an 'A' base (Adenine) to the 3' end of the blunt phosphorylated DNA fragments and then purified using Agencourt AMPure XP beads. DNA fragments were ligated to Illumina adapters, which have a single 'T' base (Thymine) overhang at their 3'end. The adapter-ligated products were purified using Agencourt AMPure XP beads. Adapter ligated DNA was amplified in a Linker Mediated PCR reaction (LM-PCR) for 13 cycles using Phusion™ DNA Polymerase and Illumina's PE genomic DNA primer set followed by purification using Agencourt AMPure XP beads. Quality and quantity of finished libraries were assessed using an Agilent DNA1000 series chip assay (Agilent Technologies, Santa Clara, CA) and Invitrogen Qubit HS Kit (Invitrogen, Carlsbad, California, USA), respectively. Each library was standardized to 2µM. For samples of *Lunularia cruciata* colonized by *Rhizophagus irregularis*, twelve cycles of PCR were applied to amplify libraries, and size selection was performed on E-gel (Invitrogen). Libraries were quantified by qPCR using the KAPA Library Quantification Kit (PN11 KK4824).

RNA sequencing

Cluster generation was performed using a TruSeq Single Read Cluster Kit (v3) and the Illumina cBot, with libraries multiplexed for 1x100bp sequencing using the TruSeq 100bp SBS kit (v3) and HCS1.6 software, on an Illumina HiSeq2000. Images were analyzed using CASAVA. For *Lunularia cruciata* colonized by *Rhizophagus irregularis*, RNA-seq experiments were performed on an Illumina HiSeq2000, using a paired-end read length of 2 x100 bp with the Illumina TruSeq SBS sequencing kits v3 (PN FC-401-3001, HiSeq2000). Data have been generated from paired-end libraries (2x100 bp) with insert sizes ranging from 200 to 300 bp. Sequencing was performed at the GeT-PlaGe platform (<https://genomique.genotoul.fr>, Toulouse, France). In addition, 454 pyrosequencing was performed on colonized tissues. Approximately 0.5 µg of total RNA was used for construction of Roche 454 GS FLX sequencing library following the manufacturer's protocol. Libraries were sequenced using one-eighth of Pico Titer Plate. All sequenced reads have been submitted at the National Center for Biotechnology Information's Sequence Read Archive (accession nos. SRR1027885).

Closterium assembly

Raw data from 15 libraries containing a total of 404,421,355 raw Illumina 100bp single-end reads (~ 40 Gbp) were pooled for *de novo* transcriptome assembly. The fastx-toolkit v.0.0.13 (Hannon lab, Cold Spring Harbor) was used for quality assessment and TruSeq adapter trimming. Assembly was performed using the Trinity assembler version r20140413p1 with the following modified parameters: "--PasaFly --normalize_reads --min_kmer_cov 3 --trimmomatic --quality_trimming_params "LEADING:5 TRAILING:5 MINLEN: (5). The assembly produced 119,254 contigs representing 83,700 "gene units" (based on Trinity determination). The average contig length was 866 nt, with 65,495 contigs >= 500 nt and 10,506 contigs >= 2,000 nt. Unprocessed read data is deposited at the NCBI Short Read Archive (PRJNA296352).

***Lunularia* assembly**

To define the gene repertoire of *Lunularia*, a transcriptome assembly was generated by using the CLC Genomic Workbench suite (v7.0, CLC Bio, Denmark). We chose a strategy combining 500-bases-long reads from 454 pyrosequencing data and 100 base-long reads Illumina data. We mixed sequences obtained from *Lunularia* thallus associated or not to *R. irregularis*. Pyrosequencing data were performed on mycorrhizal thalli and generated 138,245 reads (51,920,771 nucleotides; mean 492 bases). Two optima of GC contents were observed, 28 and 48%, the first one being characteristic of *R. irregularis* (6). The fungal-originating reads were eliminated by mapping reads on the genomes of *R. irregularis* (Gloin1 and Rir sequences, (6, 7)) using CLC mapping tool. Illumina single reads (100 bp from Hiseq2000) were obtained from six libraries of RNA from non mycorrhizal thalli, leading to 173,377,946 reads (17.5 Gb). The resulting reads were trimmed based on the quality scores (limit: 0.05) and end ambiguity based on adaptor sequences used for libraries (maximum number of ambiguities: 2). The reads less than 50bp long after trimming were discarded. The trimmed and filtered reads from 454 and Illumina were de novo assembled into contigs using CLC software with default setting. The resulting assembly (Luc_v1; available at www.polebio.lrsv.ups-tlse.fr/Luc_v1/) is formed by 70,122 contigs (av. length 781; max length 22,117; N50= 1,275). Luc_v1 assembly was used for further annotation and gene expression analysis.

***Lunularia* gene expression analysis**

For gene expression analysis, we used additional triplicates Illumina libraries from mock and symbiotic samples (*Lunularia* and *Lunularia* + *R. irregularis*). Reads were mapped to the Luc_v1 assembly to calculate gene expression using CLC Genomic Workbench suite. To find genes significantly up-regulated during symbiosis, RPKM and proportion-based test statistics with a False Discovery Rate (FDR) correction for multiple testing were calculated (settings: minimum mapped read length fraction = 0.95; minimum similarity = 0.98). According to CLC recommendation, genes were considered as significantly up regulated when meeting the

requirements of "total difference reads mapped" >10, RPKM fold change>2 and FDR corrected p-value<0.05.

Spirogyra DNA extraction

Spirogyra sp. (AU1) was chosen for the genomic sequence. Genomic DNA was extracted from fresh materials with adapted CTAB method (8). *Spirogyra* filaments were collected by filtration and disrupted in the CTAB buffer with a bead beater.

Spirogyra sequencing and assembly

The genomic sequence was launched with Illumina MiSeq sequencer and HiSeq sequencer. Before loading the genomic DNA on the sequencer, fragmentation together with adaptor incorporation (tagmentation), limited-cycle PCR amplification, and size selection were performed to make size-specific DNA libraries compatible with Illumina sequencers. The sizes of insertions obtained with HiSeq and MiSeq methods were 426 bp and 502 bp respectively. FASTQC Toolkit version 0.10.1 was used to check and evaluate the raw data quality (9). The assembly process was performed with SOAPdenovo version 2.04 (10) following a length adjustment process with Flash (fast length adjustment of short reads) (11). Sequence assembly is deposited at the NCBI BioProject depository (Submission ID: SUB1029739 / BioProject ID: PRJNA290461)

Plasmid construction

To clone *Closterium peracerosum-strigosum-littorale* complex *CCaMK* (*CpCCaMK*) cDNA were generated (Fermentas) from 500ng of DNase-treated (Ambion) RNA extracted as described above. 1 µl was used in a PCR reaction using primers 5'-CACCATGGGCGACGTGGAGGAG-3' and 5'-TCACCACTGATGTTGCCGGAAT-3'. *Cp-CaCBD* was generated by PCR from *CpCCaMK*. First, two PCR were conducted in parallel; PCR1: primers *Cp-CaCBD*-GAT-F1 5'-CACCATGGGCGACGTGGAGGA-3' and *CaM-CaCBD*-R1 5'-

ACGCTGCGGAATTTCTTCTTCGCATTCAAGGTGAGCAGGCGCCGGCACA-3'. PCR2: Cp-CBD-GAT-R1 5'-TCACCACTGATGTTGCCGGAAT-3' and CaM-CaCBD-F1 5'-GAAATTCCGCAGCGTTGGTTTTCGCGGTGATCGTGAGCAACAAGCTCATGC-3'. After purification of the PCR products, one µl of each was used for PCR3, at 50°C for 5 cycles. Then, primers Cp-CaCBD-GAT-F1 and Cp-CBD-GAT-R1 were added to the PCR reaction and 30 more cycles at 62°C were performed. Gel purified PCR products were cloned in pENTR D TOPO (Invitrogen). After sequencing, pENTR CpCCaMK and CpCaCBD were transferred to pVP16 by LR reactions (Invitrogen) to generate "Maltose Binding Protein"(MBP)-CpCCaMK and MBP-Cp-CaCBD fusions for protein purification. pENTR CpCCaMK and pENTR Cp-CaCBD were then used to generate 35S:*CpCCaMK* and 35S:*Cp-CaCBD* by recombination in a pK7FWG2-RR vector. This vector allows constitutive expression of DsRED as a transformation marker. *Chlamydomonas reinhardtii* (CrCDPK1 - Cre33.g782750) and *Chlorella variabilis* (CvCDPK3 - Cva20239) CDPKs genes were synthesized by Biomatik with *Sall* and *NotI* restriction sites added on the 5' and 3' ends respectively. Both CDPKs were then digested with *Sall* and *NotI*, cloned in pENTR1A and transferred in pVP16 by LR recombination, resulting in MBP-CrCDPK and MBP-CvCDPK fusions. *Chlorokybus atmophyticus* and *Nephroselmis pyriformis* CCaMK were synthesized by Life Technology with *Bsal* sites and overhangs to use them as level-0 in GoldenGate cloning, accordingly to the standards for plant synthetic biology (12). First, level 1 modules were generated by fusing a 35S promoter and terminator to CaCCaMK and NpCCaMK, in the R2 backbone (13). Another level-1, with the fluorescent reporter DSred under control of the *Arabidopsis thaliana* Ubiquitin promoter and a 35S terminator was generated in a R1 level-1 backbone. The L1M-R2-p35S-CaCCaMK-t35S or L1M-R2-p35S-NpCCaMK-t35S modules were then cloned in position two of a Level-2 backbone with L1M-R1-pAtUBI-DSred-t35S in position 1, accordingly to patron (13).

Medicago root transformation and Mycorrhization assays

Roots of *Medicago truncatula* A17 or *ccamk-1* mutant (TRV25) in the A17 background (14) were transformed using *Agrobacterium rhizogenes* AR1193 following (15). Three weeks after transformation, roots were checked for red fluorescence under a dissecting scope. Plants with positive roots were transferred to 30ml pots inoculated with 1/15 of Solrize Pro (Agrauxine, France) in a mix of sand and clay. After 5 weeks of incubation at 21°C (16H light) plants were collected and transgenic roots (*i.e.* showing red fluorescence) were harvested under a dissecting scope (Figure S21A). Mycorrhization level was quantified using the grid intersect method.

RNA extraction and RT-PCR

Roots were collected as described previously and immediately frozen in liquid nitrogen after collection under microscope. RNA was extracted using the Qiagen RNAeasy extraction kit. RNA were treated with DNase (Ambion) and 1 ug used to generate cDNA. cDNA were generated with Qiagen superscriptIII. For PCR, 0.5 ul of cDNA were used for each sample with primers MtPT4-F 5'-GACACGAGGCGCTTTCATAGCAGC-3' and MtPT4-R 5'-GTCATCGCAGCTGGAACAGCACCG-3' for *MtPT4* and MtUbi-F 5'-GCCGAAAACAGCTAGAAGA-3' and MtUbi-R 5'-GGAGACGGAGAACAAGGTGA-3' for *MtUBIQUITIN*. The same cDNA were used to test for expression of the transgenes using primers in Figure S21B, C.

Protein extraction and purification

pVP16 vectors containing *CpCCaMK*, *CrCDPK*, *CvCDPK* or *MtCDPK* were transformed in *E. coli* Rosetta strain. Single colonies were grown overnight in LB supplemented with Ampicillin (Sigma). 50 ml of LB in a 250 ml flask were then inoculated with 1.25ml of the overnight cultures. Once OD_{600nm} reached 0.5, 0.5 mM isopropyl-b-D-1-thiogalactopyranoside was added. 4 hours after induction the cultures were pelleted at 4°C and suspended in 5 ml of extraction buffer (20 mM HEPES pH7.4, 300 mM NaCl, 5 mM MgCl₂, 10% Glycerol, 2mM Mercaptoethano) in 50 ml

conical tubes. The solutions were sonicated for 15 seconds three times and 500 μ l of Triton X100 added. After 15 minutes centrifugation at 4°C, the supernatant was mixed with 500 μ l of 50% slurry of rinsed (1X PBS) Amylose resin (New-England Biolab) and incubated for at least two hours at 4°C with slow shaking. Beads were then washed three times with PBS 1X and eluted with 300 μ l 10mM Maltose in 10mM Tris-HCl pH 7.2 (1000G for 5 minutes). Protein extracts were stored at -80°C.

Calmodulin binding assays

Calmodulin from bovine testis (Sigma-Aldrich) was labeled with biotin using ECL Protein Biotinylation Module (GE Healthcare). A polyvinylidene difluoride membrane (Immun-Blot, Rio-Rad) was rinsed with Methanol for two minutes and incubated in TBS 1X (50 mM Tris-HCl, 150 mM NaCl) for 20 minutes. 5 μ l of purified CCaMK and CDPK proteins at 100 ng/ μ l were spotted and the membrane was blocked in Blocking solution (TBS 1X, 0.1% Tween 20, 5% skim milk) at room temperature for 1 hour. The membrane was then incubated overnight with 100 ng/mL biotinylated Calmodulin in Binding buffer (TBS 1X containing 1 mM CaCl_2 or 5 mM EGTA) at 4°C. After several washes in Rinsing solution (TBS 1X, 0.1% Tween 20, 1 mM CaCl_2 or 5 mM EGTA) the membrane was incubated with streptavidin-conjugated horseradish peroxidase in binding buffer at room temperature for 1 h. Binding of biotinylated CaM to CCaMK was detected by ECL Plus Western Blotting Detection Reagents (GE Healthcare).

***In vitro* kinase assays**

Kinase assays were performed as described in (16). 5 μ g MBP-CCaMK protein (0.5 μ g for assays with Calmodulin) were incubated in Kinase buffer (50 mM HEPES, pH 7.5, 1 mM DTT, 200 μ M ATP, and 5 μ Ci [$g\text{-}^{32}\text{P}$]ATP, 10 mM MgCl_2) 20 μ L reaction at 30°C for five or ten minutes, in presence of 0.2 mM CaCl_2 (+ Ca^{2+}) or 2.5 mM EGTA (+EGTA). For assays with Calmodulin 0.5 μ M bovine Calmodulin (Millipore) was added. The reactions were terminated by

boiling at 95°C for 5 min. Samples were then separated by 10% SDS-PAGE. The radioactivity was quantified by Molecular Dynamics Storm® 860 phosphorimager (Sunnyvale, CA) and data were analyzed using the Molecular Dynamics ImageQuant® software.

Extended dataset

The detailed list of genomes and transcriptomes used for the phylogenetic analysis is provided in Dataset S1. Except for *Lunularia* and *Closterium* transcriptomes all the transcriptomes used were generated by the One Thousand Plant Project (<http://onekp.com/project.html>). Details for RNA extraction, sequencing and assemblies are available in (17).

Sequence collection

Prediction of the amino acid sequences for the new assemblies of *Closterium*, *Lunularia* and *Spirogyra sp.* and for the recently published assemblies of *Nitella mirabilis*, *Mesostigma viride*, *Spirogyra sp* and *Coleochaete orbicularis* was done using the AUGUSTUS package using the parameter set defined for *Arabidopsis thaliana* (18). Using the combined amino acid dataset from these sources, putative homologs of the various genes used in this study were extracted by using the HMMER package (<http://hmmer.org>) and the respective PFAM domains (<http://pfam.xfam.org>) present in each of the families to which these proteins belong (19, 20). The domains in each of these families used for the searches were: LysM-RLKs [PF01476, PF00069], GRAS [PF03514], DMI1 [PF06241], CDPKs [EF-hand, PF00069], VAPYRIN [PF00635, CL0465], and RAM2 [PF12710, PF01553]. Sequences were then de-duplicated and used for further analyses. For the GRAS family, sequences that had a GRAS domain were further filtered by checking for the presence of the LHRII, PFYRE and SAW motifs characteristic of all known GRAS proteins to avoid the use of fragmented sequences. Presence of the domains was confirmed for all the proteins using InterProScan (<http://www.ebi.ac.uk/interpro/>). For the phosphate transporter, ABCG and the H⁺-ATPase families, BLASTp search was conducted

using the full-length protein from *Medicago* as query to reduce the number of sequences. The best equal hits were selected. For IPD3, BLASTp search was conducted using the highly conserved C-terminal domain. For each gene, size distribution of the collected was determined and short outliers removed. These outliers were blasted on the *Medicago* genome to determine if they could belong to the clade of interest (reciprocal blast, with initial query in the first equal hits). This was the case for *Lunularia* RAM1 and *Amborella* PT4.

Phylogenetic analysis

All the potential homologs were aligned using MAFFT (<http://mafft.cbrc.jp/alignment/server/>) and the alignment curated using BioEdit. For each gene family, Maximum likelihood trees were constructed with RAxML version 8.1.22 (21, 22). First substitution models were determined using the program parameter “PROTGAMMAAUTO”, optimal models were: LG for CDPK, DMI1, GPAT, H⁺-ATPase, PT, ABCG and Vapyrin; JTT for GRAS and IPD3; WAG for LysM-RLK and MLD-RLK. We used 10 different starting trees on a given alignment using the parsimonator algorithm of RAxML with 10 different random seeds. We estimated 10 maximum likelihood analyses on different starting trees. Next, we conducted the standard bootstraps and performed ML analysis on 100 replicates. Finally, we selected the tree with best likelihood score and drew bootstrap supports onto each selected tree. The newick-formatted trees were read and processed using the *ETE2* library (23) and the scripts were written in Python 2.7.9 for bioinformatics analysis. The clades of interest were then extracted from the larger tree, re-aligned and a Maximum-likelihood tree built using MEGA6 (24) and the substitution models define by RaxML. 100 bootstraps were conducted. Short sequences removed before building the initial phylogeny were included in these targeted analyses.

References

1. Delaux PM, *et al.* (2012) Origin of strigolactones in the green lineage. *New Phytol.* 195(4):857-871

2. Nichols H (1973) *Handbook of Phycological Methods* (Cambridge University Press).
3. Bécard G & Piché Y (1989) Fungal Growth Stimulation by CO₂ and Root Exudates in Vesicular-Arbuscular Mycorrhizal Symbiosis. *Appl Environ Microbiol* 55(9):2320-2325.
4. Mukherjee A & Ané JM (2011) Germinating spore exudates from arbuscular mycorrhizal fungi: molecular and developmental responses in plants and their regulation by ethylene. *Mol Plant Microbe Interact* 24(2):260-270.
5. Grabherr MG, *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29(7):644-652.
6. Tisserant E, *et al.* (2013) Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proc Natl Acad Sci U S A* 110(50):20117-20122.
7. Lin K, *et al.* (2014) Single nucleus genome sequencing reveals high similarity among nuclei of an endomycorrhizal fungus. *PLoS Genet* 10(1):e1004078.
8. Winnepenninckx B, Backeljau T, & De Wachter R (1993) Extraction of high molecular weight DNA from molluscs. *Trends Genet* 9(12):407.
9. Patel RK & Jain M (2012) NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7(2):e30619.
10. Luo R, *et al.* (2012) SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* 1(1):18.
11. Magoč T & Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* 27(21):2957-2963.
12. Patron NJ (2014) DNA assembly for plant biology: techniques and tools. *Curr Opin Plant Biol* 19:14-19.
13. Engler C, *et al.* (2014) A golden gate modular cloning toolbox for plants. *ACS Synth Biol* 3(11):839-843.
14. Lévy J, *et al.* (2004) A putative Ca²⁺ and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science* 303(5662):1361-1364.
15. Boisson-Dernier A, *et al.* (2001) Agrobacterium rhizogenes-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol Plant Microbe Interact* 14(6):695-700.
16. Miller JB, *et al.* (2013) Calcium/Calmodulin-dependent protein kinase is negatively and positively regulated by calcium, providing a mechanism for decoding calcium responses during symbiosis signaling. *Plant Cell* 25(12):5053-5066.
17. Johnson MT, *et al.* (2012) Evaluating methods for isolating total RNA and predicting the success of sequencing phylogenetically diverse plant transcriptomes. *PLoS One* 7(11):e50226.
18. Stanke M, Steinkamp R, Waack S, & Morgenstern B (2004) AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res* 32(Web Server issue):W309-312.
19. Finn RD, Clements J, & Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39(Web Server issue):W29-37.
20. Finn RD, *et al.* (2014) Pfam: the protein families database. *Nucleic Acids Res* 42(Database issue):D222-230.
21. Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312-1313.
22. Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22(21):2688-2690.
23. Huerta-Cepas J, Dopazo J, & Gabaldón T (2010) ETE: a python Environment for Tree Exploration. *BMC Bioinformatics* 11:24.
24. Tamura K, Stecher G, Peterson D, Filipowski A, & Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30(12):2725-2729.

Supplementary Figures Legends

Appendix S1. Detailed Material and method

Figure S1. Maximum-likelihood tree of the LysM-domain containing receptor-like kinases identified in the extended dataset.

Figure S2. Maximum likelihood tree of the malectin-like ecto-domain (MLD) receptor like kinases

Figure S3. A) DMI2 is composed by a malectin-like ecto-domain (MLD) and a kinase. DMI2 evolved via duplication in land-plants. Maximum-likelihood tree of the DMI2 clade is presented. The entire MLD-Kinase tree is available in Figure S2. B) The orthologs of Medicago DMI2 found in liverworts display the same domain organization.

Figure S4. Maximum-likelihood tree of the DMI1 homologs. The colored boxes on the right indicate the conservation of the filter domain (ADAGNHA in Medicago DMI1). Black, blue and white squares indicate identical, homologous and different amino acids respectively.

Figure S5. Maximum-likelihood tree of the Calcium dependent protein kinases identified in the extended dataset. The colored boxes on the right indicate the conservation of the calmodulin binding domain domain (QMDPEIVSRLQSFNARRKLRAAAIASVWS in Medicago CCaMK). Black, blue and white squares indicate identical, homologous and different amino acids respectively.

Figure S6. A) Maximum-likelihood tree of IPD3. IPD3 appeared in advanced charophytes and is conserved as a single copy gene in land plants. B) IPD3 is composed by two main domains, a DNA-binding domain located at the N-terminus and an activation domain. Among these two domains, the DNA binding domain is the most conserved, especially in charophytes.

Figure S7. Charophyte proteins ortholog to the land plant LysM-RLK display a similar domain organization.

Figure S8. A) CpCCaMK display the same organization than MtCCaMK with a kinase domain at the N-terminus and EF-hands on the C-terminus. B) The auto-phosphorylation site, analogous to T271 in MtCCaMK, and to T265 in *Lotus japonicas* CCaMK, is conserved (S242) C) *Top left:* A structural model was predicted as previously for MtCCaMK (Miller et al. 2013). *Top right:* In absence of auto-phosphorylation a single hydrogen-bond is formed with E291. *Bottom left:* Mimicking a phosphorylation of S242 results in the formation of a hydrogen-bond with R245, similar to what is observed for MtCCaMK. *Bottom right:* Furthermore, mimicking a phosphorylation of S243 improve this network by forming three hydrogen-bonds with R295 and C215. D) Autophosphorylation assay of CpCCaMK. Like MtCCaMK, autophosphorylation of CpCCaMK is promoted by the presence of calcium (Ca^{++}). E) Effect of calmodulin on CpCCaMK auto-phosphorylation. Similarly to MtCCaMK, auto-phosphorylation of CpCCaMK is reduced in presence of calmodulin.

Figure S9. A) The calmodulin binding domain of *Closterium* CCaMK (CpCCaMK) and *Chlorokybus atmophyticus* CCaMK (CaCCaMK) are divergent. A chimeric protein Cp-CaCBD corresponding to CpCCaMK with 8 mutations mimicking the calmodulin binding domain of CaCCaMK was generated by PCR and called Cp-CaCBD (see Appendix S1). B) The ability of Cp-CaCBD to bind calmodulin was strongly reduced but not abolished (the original plot is available Figure S20). C) Trans-complementation assays of *Mtccamk-1* TRV25 mutant. Letters indicate statistically supported groups according to ANOVA and TuckeyHSD tests (P-values < 0.01). *Ccamk-1* roots expressing Cp-CaCBD were colonized to a level not distinguishable from the wild-type transformed with an empty vector (EV). D) Representative picture of WGA-FITC stained AM fungi colonizing plant roots of *ccamk-1* mutant expressing Cp-CaCBD. Bar= 100 μm . E) Expression of the symbiotic phosphate transporter *MtPT4* was observed in colonized roots of *ccamk-1* transformed with Cp-CaCBD. F) Hyphopodium formed on *ccamk-1* roots expressing Cp-CaCBD resulted in a normal level of internal hyphae and arbuscules.

Figure S10. A) The auto-phosphorylation site analogous to MtCCaMK T271, and *Lotus japonicas* CCaMK T265, is conserved in angiosperms (red vertical line), lycophytes (yellow), mosses (light-green), liverworts (dark-green) and most advanced charophytes (dark-blue). Few advanced charophytes display a Cysteine instead, the impact of this mutation has not been investigated. Amino-acids that cannot be phosphorylated are present in basal-charophytes (dark-blue light-blue) and in chlorophytes (light-blue). B) *Top left:* A structural model was predicted for *Nitella mirabilis* CCaMK as previously for MtCCaMK (Miller et al. 2013) and CpCCaMK (Figure S8). *Top right:* In absence of auto-phosphorylation two hydrogen-bonds are formed with E287. *Bottom left:* Mimicking a phosphorylation of S238 results in the formation of a third hydrogen-bond with R291. *Bottom right:* mimicking a phosphorylation of the neighbouring T239 resulted in a single H-bond between S238 and R291. C) and D) Structural modelling of *Chlorokybus atmophyticus* and *Mesostigma viride* CCaMK confirmed the absence of H-bond network similar to *bona fide* CCaMK.

Figure S11. Maximum-likelihood tree of the GRAS transcription factors.

Figure S12. Maximum-likelihood tree of the ABC transporter, subfamily G, identified in the extended dataset.

Figure S13. A) Maximum-likelihood phylogenetic tree of STR and STR2 extracted from the ABC, subfamily G, transporter tree on the left. High-resolution image of the ABCG tree is available in Figure S12. STR and STR2 evolved from duplication in land plants. B) Expression of the *Lunularia cruciate* orthologs, *LcSTR* and *LcSTR2*, is up-regulated during mycorrhization.

Figure S14. Maximum-likelihood tree of the Glycerol-3-phosphate acyltransferases identified in the extended dataset.

Figure S15. A) Maximum-likelihood phylogenetic tree of RAM2 extracted from the GPAT tree on the left. High-resolution image of the GPAT tree is available in Figure S14. RAM2 originates in

land-plants. Two paralogs are present in liverworts. B) Expression of both paralogs is up-regulated during mycorrhization in the liverwort *Lunularia cruciata*. C) RAM2 is composed by two domains, a haloacid dehydrogenase domain and an acetyltransferase domain. These two types of domains are present in other eukaryotes, including chlorophytes and charophytes, but their combination is unique to GPAT.

Figure S16. A) Maximum-likelihood tree of VAPYRIN. VAPYRIN appeared in land plants and is conserved as a single copy gene. B) VAPYRIN is composed by two domains: a motile-sperm domain combined with Ankyrin repeats. These two types of domains are present in other eukaryotes, including chlorophytes and charophytes, but their combination is unique to VAPYRIN.

Figure S17. Maximum-likelihood tree of the H⁺-ATPase identified in the extended dataset.

Figure S18. Maximum-likelihood tree of the Phosphate transporters identified in the extended dataset.

Figure S19. Maximum-likelihood tree of the Phosphate transporter of the PT4 and LcPT clades. High-resolution image of the phosphate transporter tree is available in Figure S18. B) Expression of *LcPT*, from *Lunularia cruciata* is up-regulated during mycorrhization (here contig_26970, the other contig spanning *LcPT*, contig_49319, shows a 406-fold up-regulation).

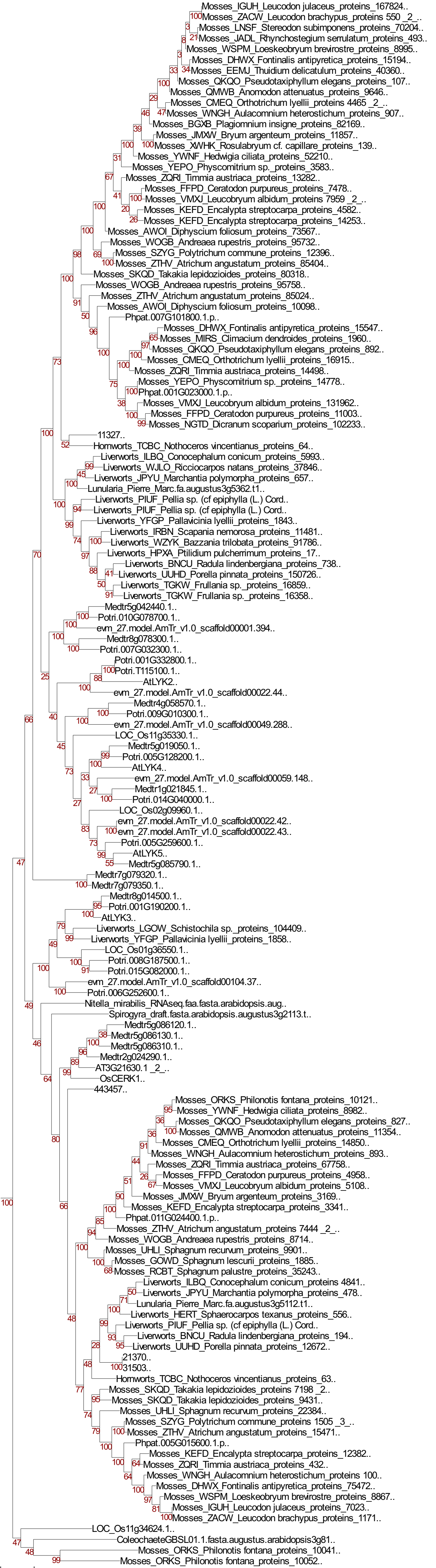
Figure S20. Drops corresponding to 500, 200 and 100 ng of purified proteins were plotted on a membrane and incubated in presence of calcium and biotinylated calmodulin or with biotinylated calmodulin only (see Appendix S1 for details). *Left:* blot revealed with Poncea S staining. *Right:* gel revealed using streptavidin-conjugated horseradish peroxidase and ECL western blotting reagent. For each protein at least three assays have been conducted including two with all the proteins on the same membrane. One of these experiments is presented here. In absence of calcium no binding was detected.

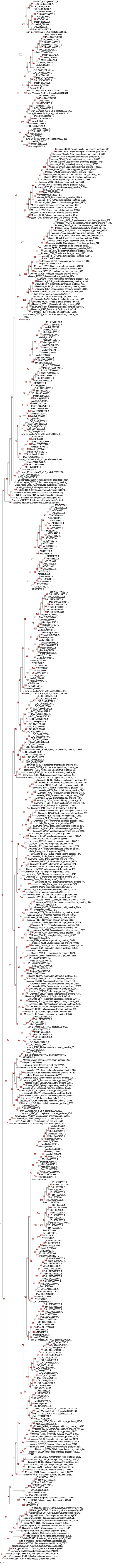
Figure S21. A) Transgenic roots of A17 (WT) or *ccamk-1* mutant transformed with the constructs used for trans-complementation assays. These constructs contain a constitutively expressed red fluorescent protein (DSRED) as a transformation marker. This marker was used to select transgenic roots for all the experiments. B) RNA samples were extracted from collected transgenic roots and the expression of the transgenes tested by RT-PCR. In parallel, expression of *DSRED* was confirmed. Expression of the transgenes and *DSRED* was confirmed on a total of four plants per construct coming from two independent experiments. C) Primers used to validate the expression of the transgenes and *DSRED*.

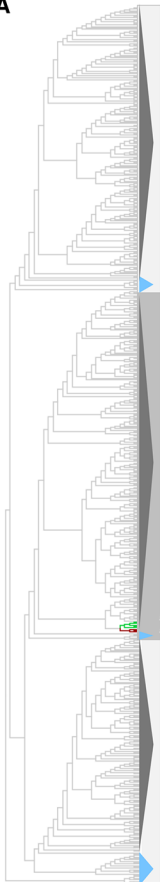
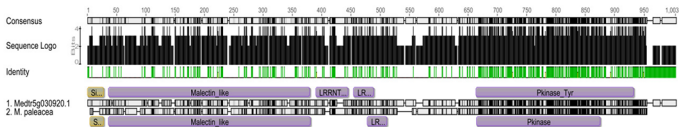
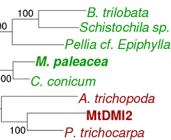
Dataset S1. Extended dataset.

Dataset S2. *Lunularia cruciata* genes up-regulated upon colonization by *Rhizophagus irregularis*.

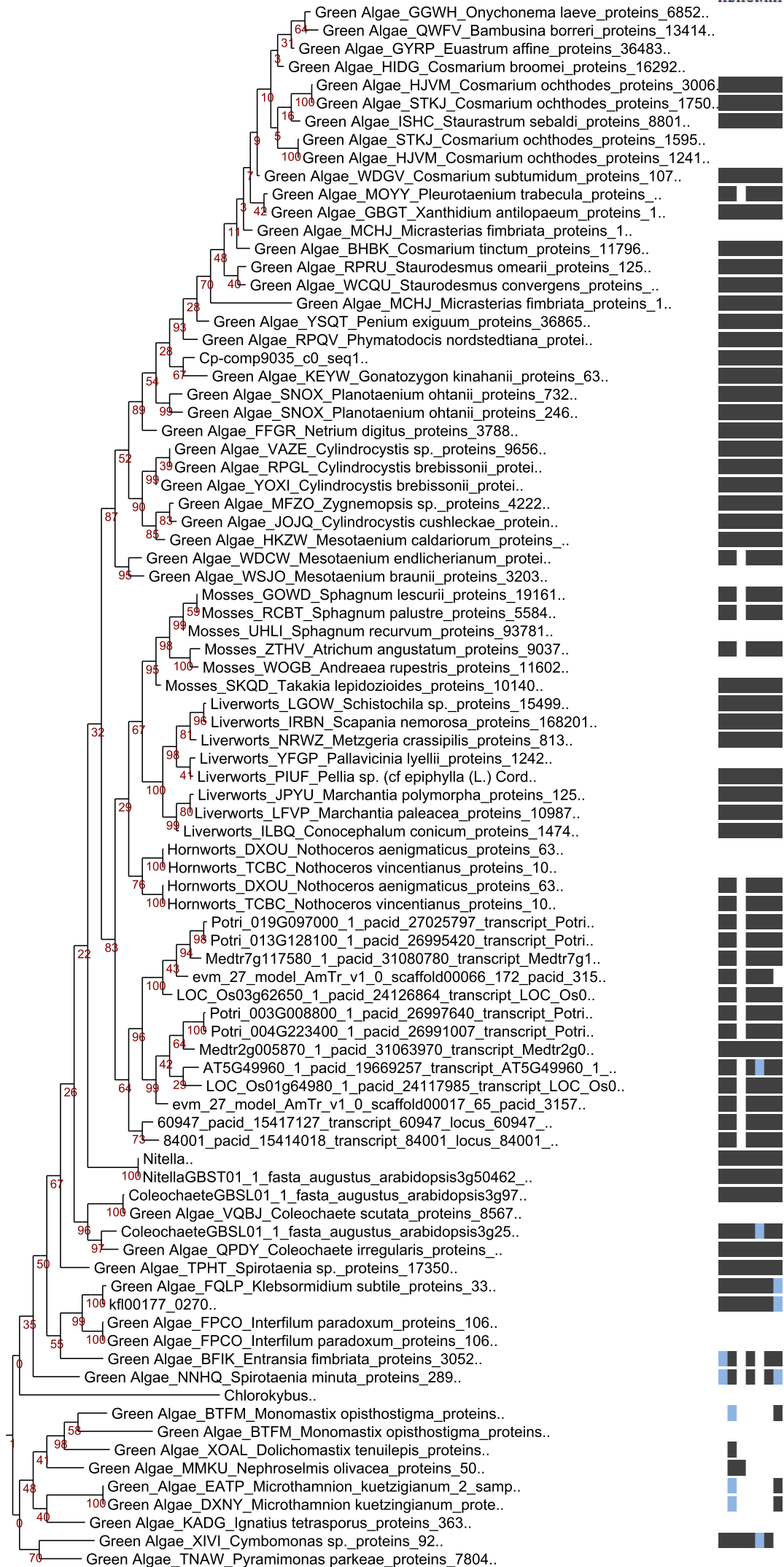
Dataset S3. Sequences used to generate the phylogenetic trees.

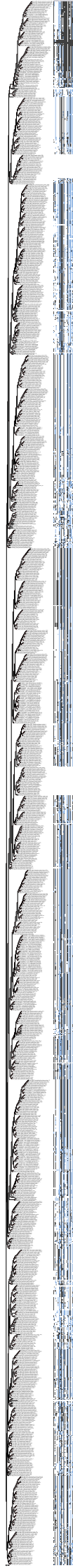


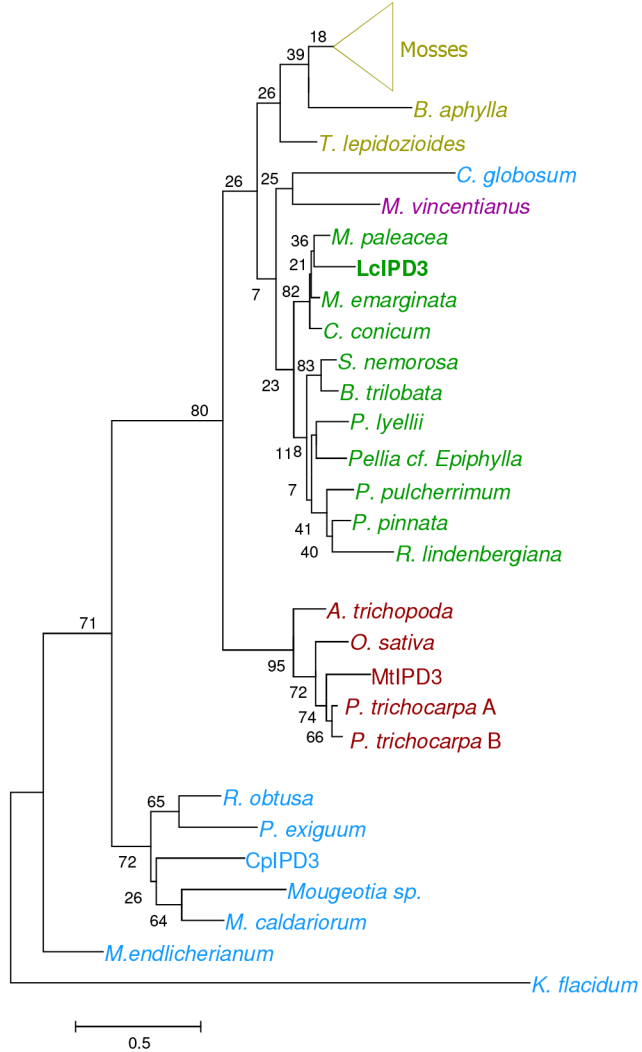
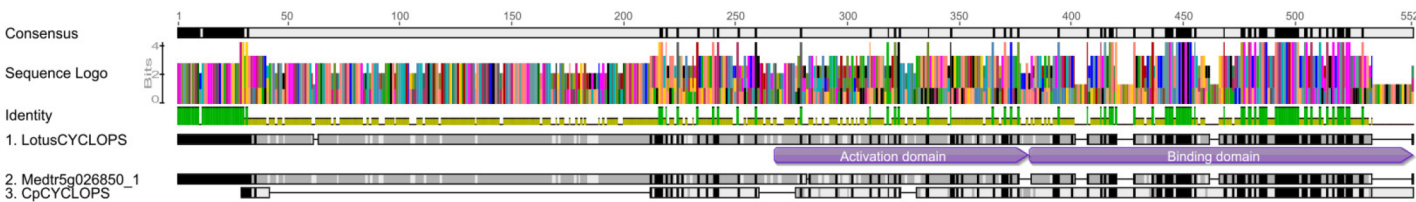


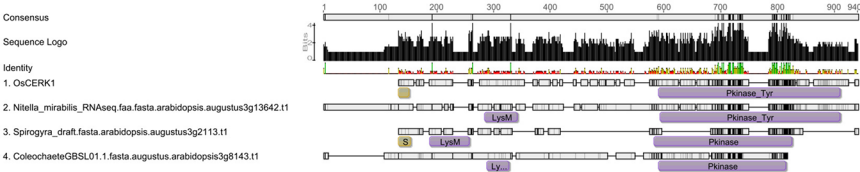
A**B****DMI2**

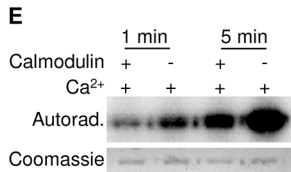
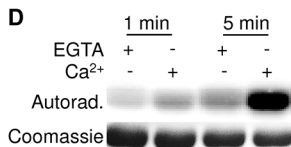
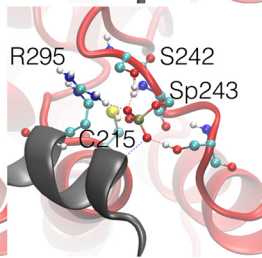
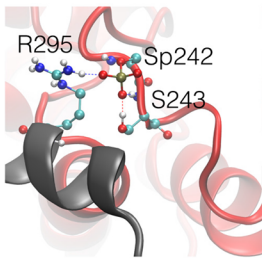
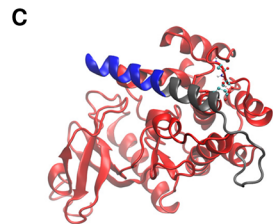
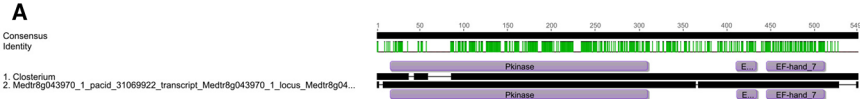
- ▶ Land plant clade
- ▶ Charophyte clade
- Angiosperms
- Liverworts

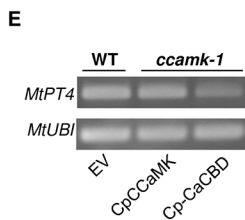
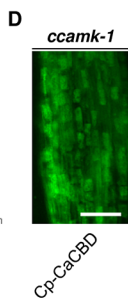
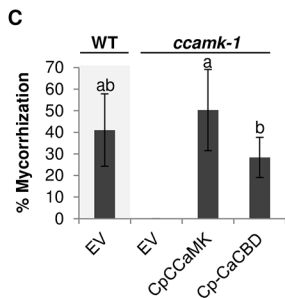
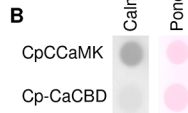




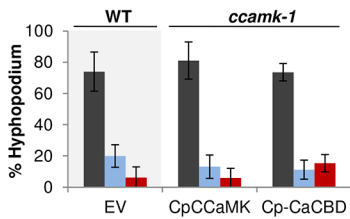
A**B**







F ■ Arbuscules ■ Internal hyphae ■ Hyphopodium only



A

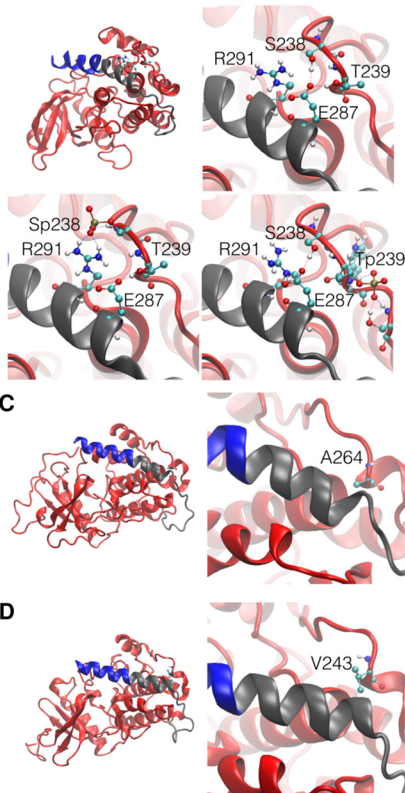
Medr8g043970_1_pacid_31069922
 Potri_008G011400_1_pacid_27039
 Potri_010G247400_1_pacid_26979
 LOC_Os05g41090_1_pacid_2415217
 evm_27_model_AmTr_v1_0_scaffold
 82473_pacid_15411314_transcript
 Hornworts_TCBC_Megaceros_vince
 Mosses_ABCD_Niphotrichum_elong
 Mosses_AWOI_Diphyscium_foliosu
 Mosses_CMEQ_Orthotrichum_lyell
 Mosses_EEMJ_Thuidium_delicatu
 Mosses_FFPD_Ceratodon_purpureu
 Mosses_HRWG_Buxbaumia_aphylla
 Mosses_IGUH_Schwetschkeopsis_f
 Mosses_JADL_Rhynchostegium_ser
 Mosses_JADL_Rhynchostegium_ser
 Mosses_KEFD_Encalypta_strepto
 Mosses_LNSF_Hypnum_subimponens
 Mosses_LNSF_Hypnum_subimponens
 Mosses_NGTD_Dicranum_scaripum
 Mosses_QKQO_Pseudotaxiphylum_
 Mosses_QMWB_Anomodon_attenuatu
 Mosses_RGKI_Leucobryum_glaucum
 Mosses_SKQD_Takakia_lepidozoi
 Mosses_SKQD_Takakia_lepidozoi
 Mosses_VMXX_Leucobryum_abidum
 Mosses_WNGH_Aulacomnium_hetero
 Mosses_WSPM_Rhytidadelphus_lo
 Mosses_XWHK_Funaria_proteins_1
 Mosses_YWNF_Hedwigia_ciliata_p
 Mosses_YWNF_Hedwigia_ciliata_p
 Phpat_019G070100_1_p_pacid_282
 Phpat_021G054100_1_p_pacid_282
 Liverworts_BNCU_Radula_lindenb
 Liverworts_IRBN_Scapania_nemor
 Liverworts_LGOW_Schistochila_s
 Liverworts_UUHD_Porella_pinnat
 Liverworts_WZYK_Bazzania_trilo
 Liverworts_YFGP_Pallavicinia_1
 Liverworts_YFGP_Pallavicinia_1
 SpirogyraGBSM01_1_fasta_august
 Green_Algae_BHKB_Cosmarium_tin
 Green_Algae_DRFX_Closterium_lu
 Green_Algae_HKZW_Mesotaenium_c
 Closterium
 Green_Algae_ISHC_Staurastrum_s
 Green_Algae_JOJQ_Cylindrocysti
 Green_Algae_KEYW_Gonatozygon_k
 Green_Algae_MCHJ_Micrasterias_
 Green_Algae_RPQV_Phymatodocis_
 Green_Algae_RPRU_Staurodesmus_
 Green_Algae_SNOX_Planotaenium_
 Green_Algae_GYRP_Euastrum_afi
 Green_Algae_STKJ_Zygnema_sp_A
 Green_Algae_WCQU_Staurodesmus_
 Green_Algae_WDCW_Mesotaenium_e
 Green_Algae_WDGV_Cosmarium_sub
 Green_Algae_FFGR_Netrium_digif
 Green_Algae_WSJO_Mesotaenium_b
 Green_Algae_XRTZ_Roya_obtusa_p
 Green_Algae_GBGT_Xanthidium_an
 Green_Algae_YSQT_Penium_exigui
 Green_Algae_ZRMT_Mougeotia_sp_
 NitellaGBST01_1_fasta_augustus
 NitellaGBST01_1_fasta_augustus
 Nitella_mirabilis_RNAseq_faa_f
 Green_Algae_BFIK_Entransia_fim
 MesostigmaGBSK01_1_fasta_augus
 Green_Algae_KYIO_Mesostigma_vi
 Catm_CCaMK
 Green_Algae_AYPS_Unidentified_
 Green_Algae_HVNO_Tetraselmis_c
 Green_Algae_ISIM_Nephroselmis_
 Green_Algae_NNHQ_Spirotaenia_m
 Green_Algae_RPGL_Cylindrocysti
 Green_Algae_TNAW_Pyramimonas_p
 Green_Algae_VAZE_Cylindrocysti
 Green_Algae_XIVI_Cymbomonas_sp
 Green_Algae_XJGM_coccoid_prasi
 Green_Algae_XMCL_Prasinococcus
 Green_Algae_YLBK_Cylindrocysti

B

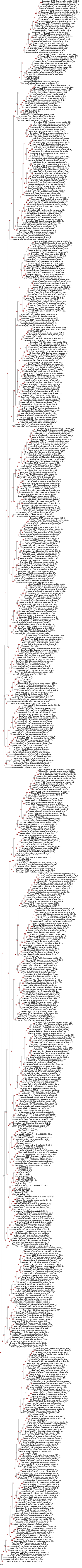
S F Y E K T W K G I S
 S F Y E K T W K N I T
 T F Y E K T W K N I T
 S F Q D H T W K T I S
 S F D E H T W K T I S
 N F E E Y T W K T I S
 D L T D D T W S T V S
 S M E E D T W R G I S
 D M D E P S W R G V S
 T M D E D S W S G I S
 S M D E Q S W G G I S
 S M E E V S W R G I S
 S M E E D S W R G V S
 S M D E Q S W G G I S
 S M D E Q S W G D I S
 S M D E S W R G I S
 S M E E S W R G I S
 S M D E S W R G I S
 S M D E Q S W S G I S
 T M D E D S W R G I S
 S M E E Q S W G G I S
 S M D E S W R G I S
 S M E E L S W R G I S
 S M E E D T W K G V T
 S M E E D T W K G V T
 T M D E S W R E I S
 N M D E E S W R G I S
 N M D E E S W G G I S
 T M E E D S W R G I S
 S M D E V S W R G I S
 S M E E E S W G C I S
 S M E E E S W R G I S
 K F G E D S W K D V S
 N F Q E E S W R T V S
 N F Q E E T W G T V S
 K F E D E T W R T V S
 N F Q E D S W R S V S
 N F E E E T W K T V S
 N F E E E T W K T V S
 T F S D S I W T S V S
 D F D E S T W N S I S
 D F D E Y T W K T I S
 S F D E S T W K S V S
 D F D E S S W K S I S
 D F D E S T W N S I S
 S F D E H S W K T V S
 D F E E A T W A A I S
 D F D E P T W A A I S
 D F D E S T W N T I S
 D F D E D S W R H V S
 D F D E T T W N S I S
 D F D E S T W S T I S
 D F D E A T W S T I S
 D F D E Y T W S N V S
 D F D E G T W N S I S
 D F D E P S W K T V S
 D F D D P C W A T V S
 D F D E S C W K T V S
 D F D E A T W N S I S
 D F D E S T W D S V S
 G F E E A T W K S V S
 N F S Y S T W R S V S
 D F K Y H T W R S V S
 N F S Y S T W R S V S
 D L E D E C W R E I S
 D F D D P V W D T V S
 D F D D P V W D T V S
 T F D D P A W H G V S
 S F V H P V W S K I S
 N M A V F P V C M L S
 D F N D P A W G G V S
 S F D D P V W Q G I S
 D F D D S A W K P V S
 D F D E P I W K N V S
 D F D D S A W K P V S
 D F D D Q Q W D E V S
 D F D D P A W A Q V S
 D F N D P V W A T V T
 D F D D S A W K P V S

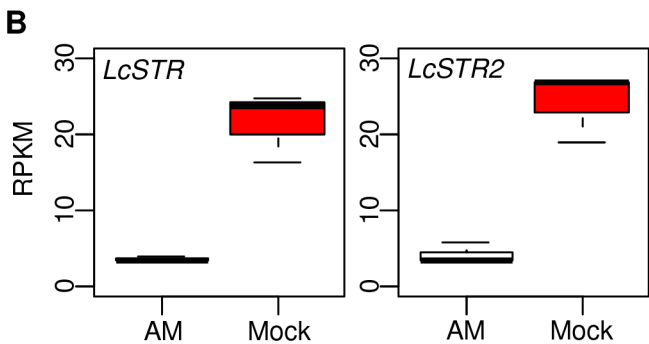
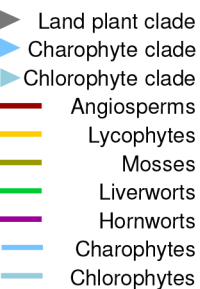
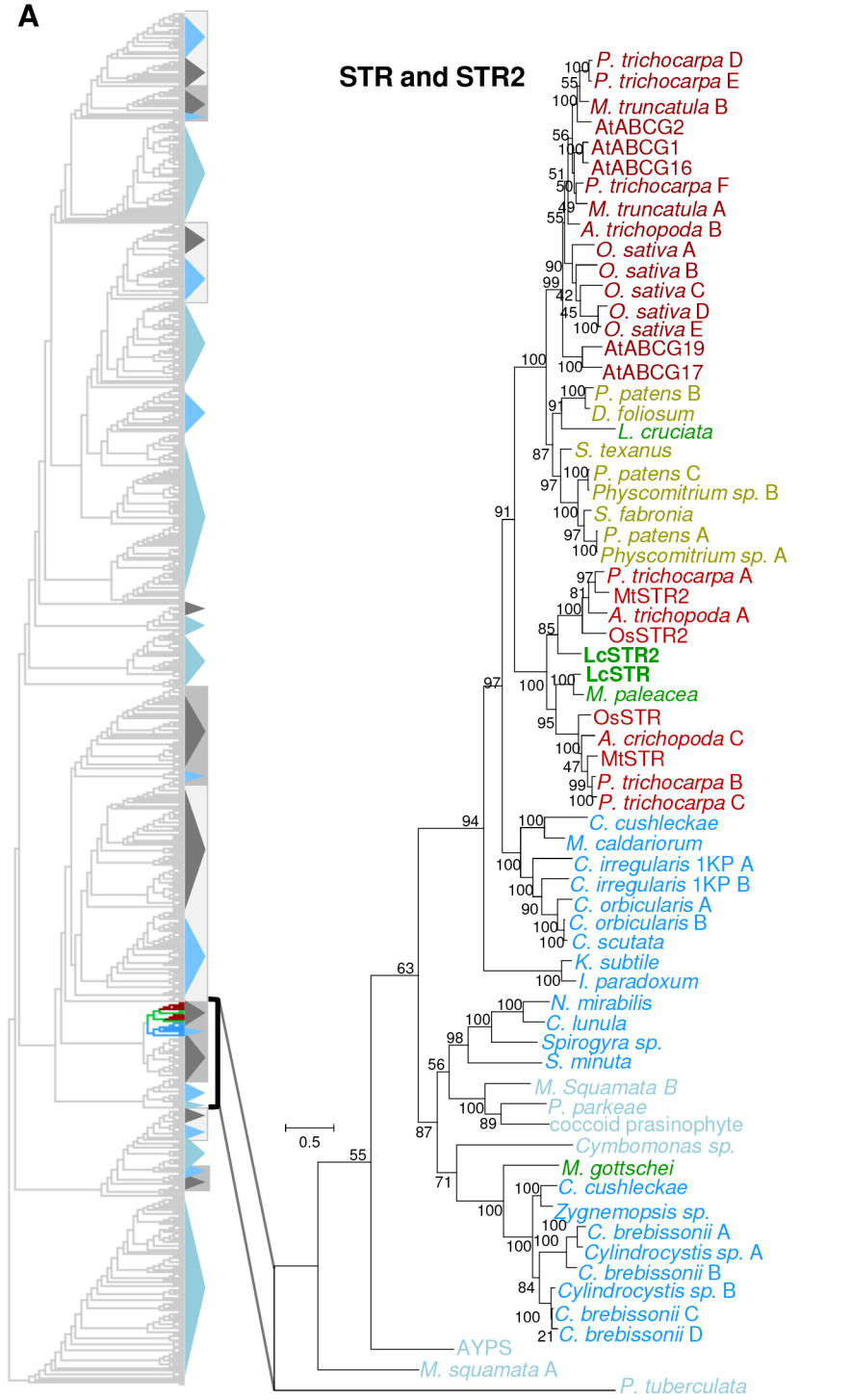
C

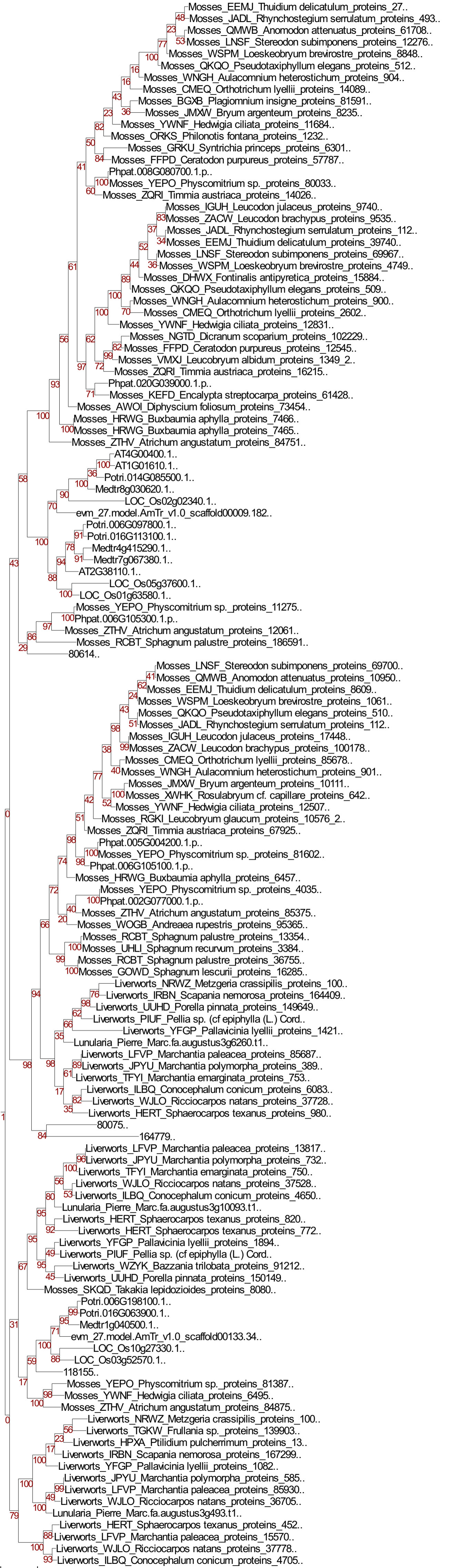
D

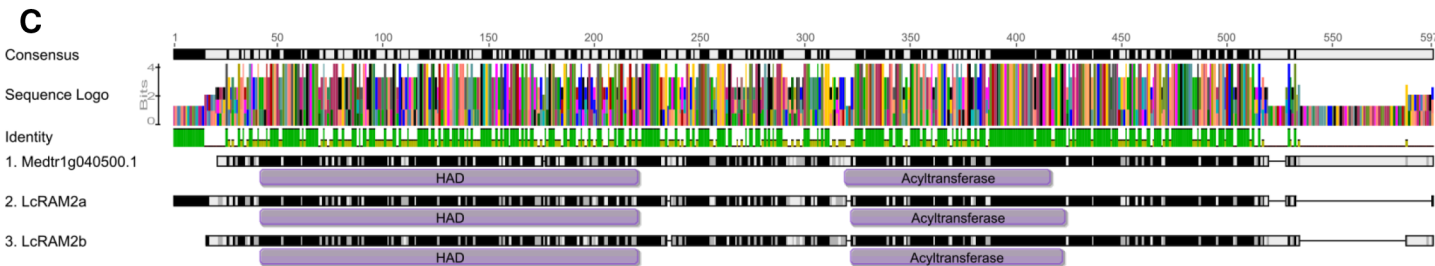
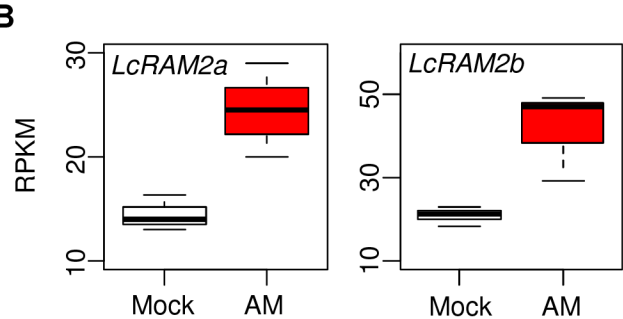
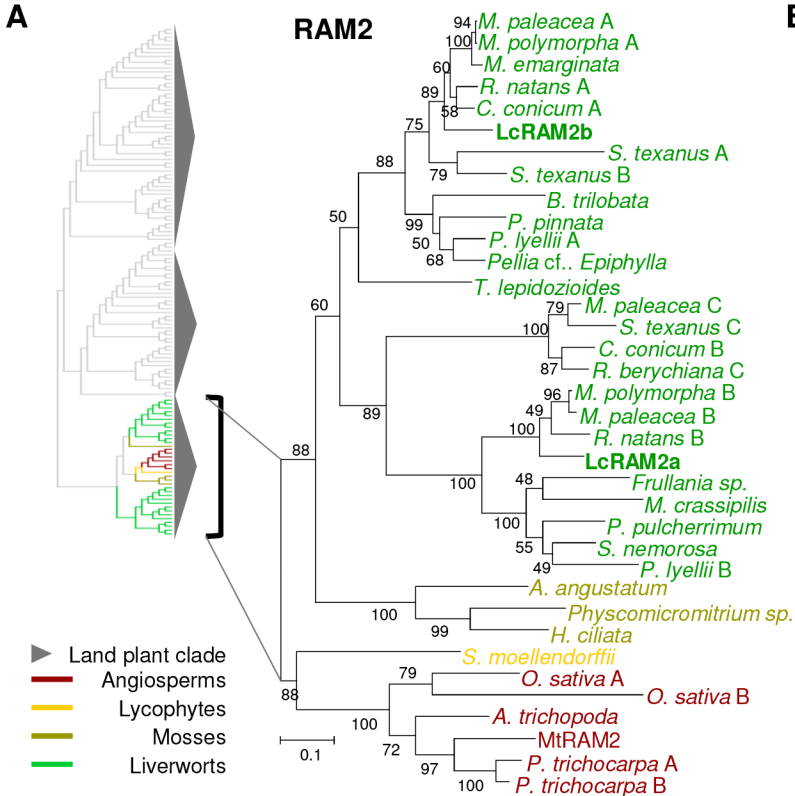


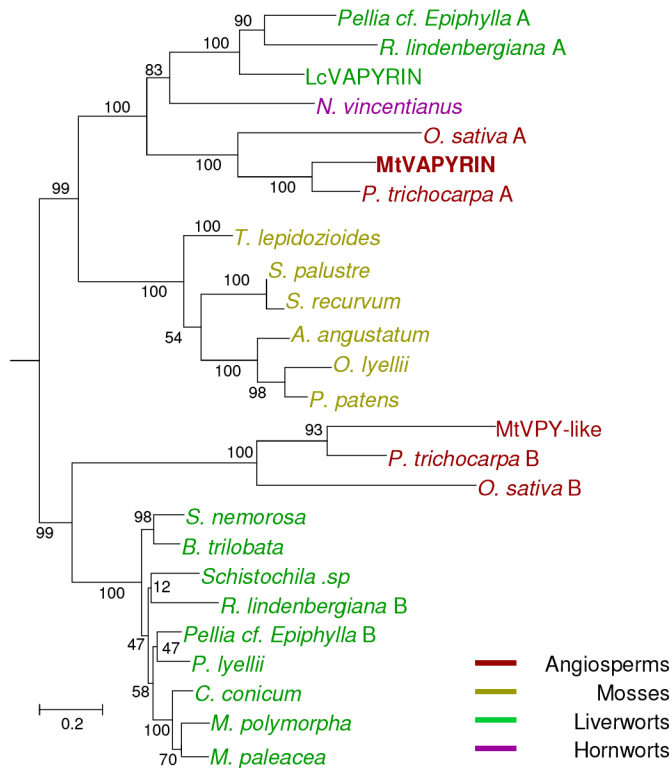
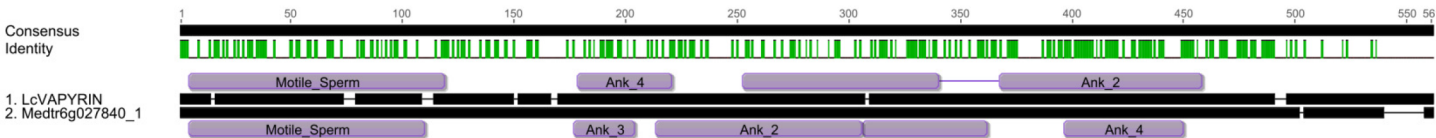


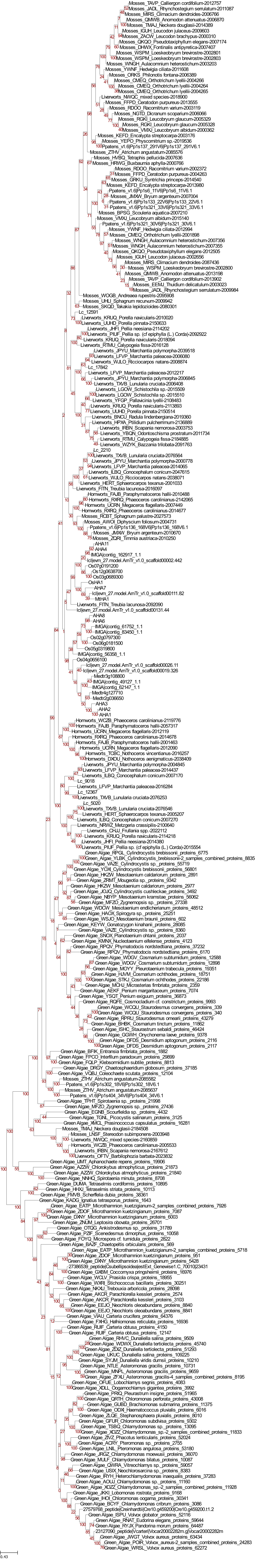


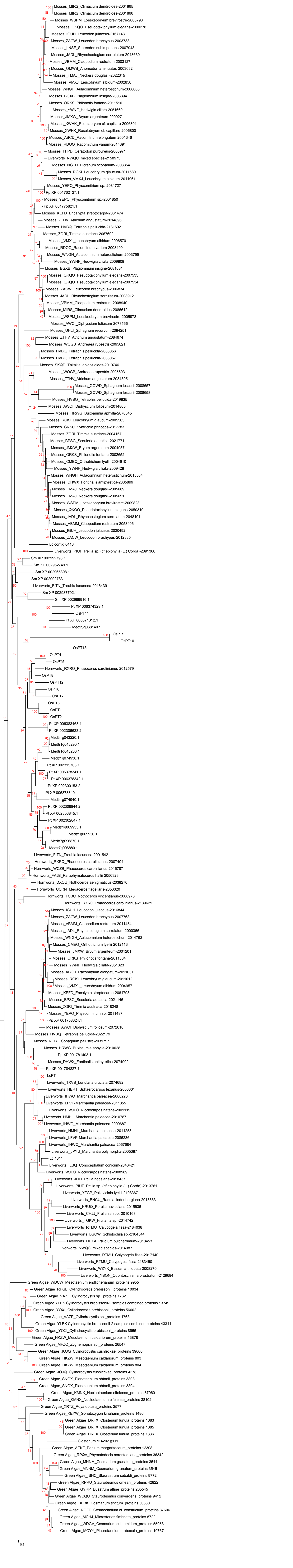




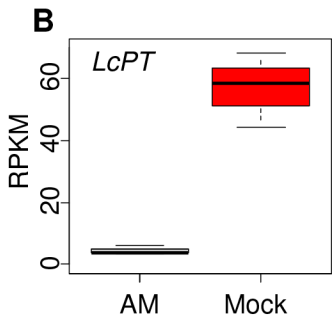
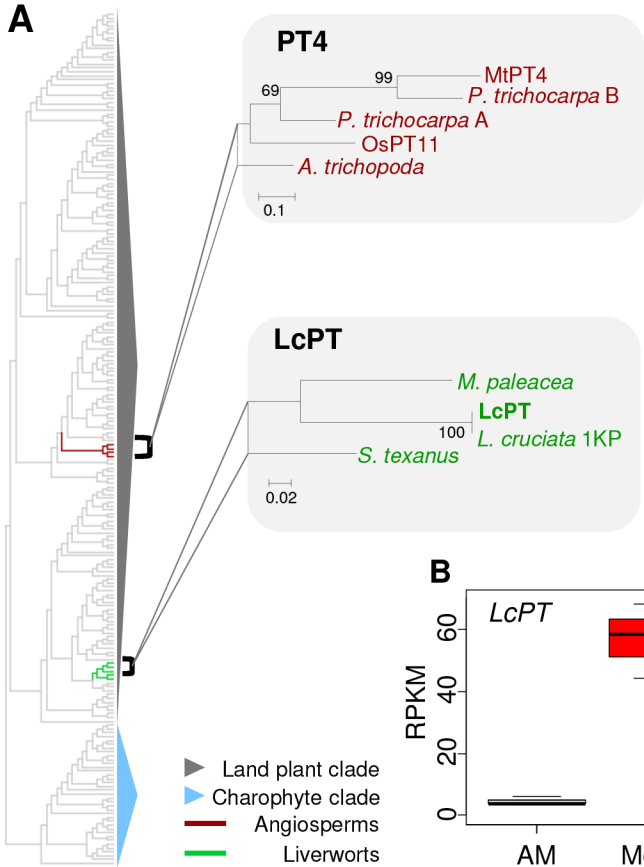


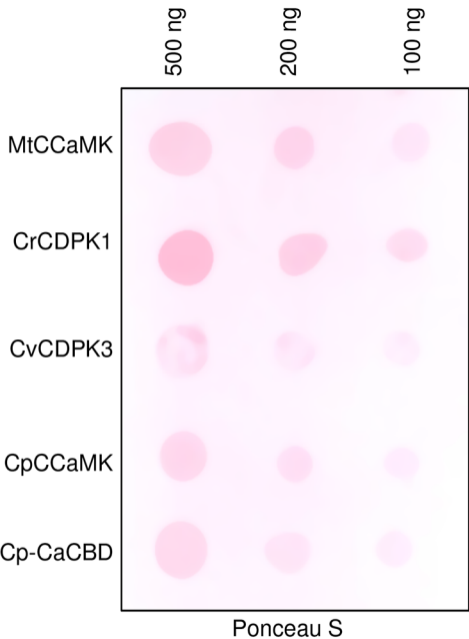
A**B**

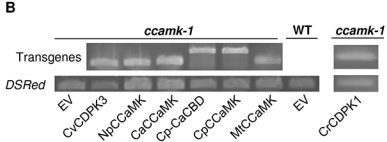
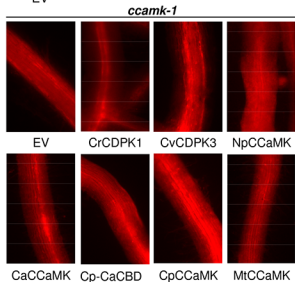
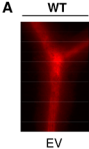




0.1







C

CrCDPK1-F	5'- AGGACAAGAGCGGCTTCATA
CrCDPK1-R	5'- ACGACCACCTTGGTCTTGAG
CvaCDPK3-F	5'- CAGAGTATGAGCGGGACACA
CvaCDPK3-R	5'- TTCGGGCTTGAGATCTCTGT
CpCCaMK-F	5'- ATGGGCGACGTGGAGGAG
CpCCaMK-R	5'- CACCTCAGGCTCCATCG
CaCCaMK-K-F	5'- AACCTCATCCGTCTGCTCAA
CaCCaMK-K-R	5'- CAACGAGCCAAACAGACCAA
NpCCaMK-F	5'- ATGTGCGGCTCATCAGAAAC
NpCCaMK-R	5'- CGGGGTTTCAAACAGCATGT
MtDMI3-F	5'- GACTGAAGCTGCAACTGTGG
MtDMI3-R	5'- TTTGGGCAATGAAAGGTGGA
DSRed-F	5'- ACGAAGGCCACAATACCGTA
DSRed-R	5'- CCACCGTCTTTCAGCTTCAG