Network of GRAS Transcription Factors Involved in the Control of Arbuscule Development in *Lotus japonicus*<sup>1</sup>[OPEN]

Li Xue, Haitao Cui, Benjamin Buer, Vinod Vijayakumar<sup>2</sup>, Pierre-Marc Delaux<sup>3</sup>, Stefanie Junkermann, and Marcel Bucher*

Botanical Institute, Cologne Biocenter, Cluster of Excellence on Plant Sciences, University of Cologne, D-50674 Cologne, Germany (L.X., B.B.V.V., S.J., M.B.); Department of Plant-Microbe Interactions, Max Planck Institute for Plant Breeding Research, D-50829 Cologne, Germany (H.C.); and Department of Agronomy, University of Wisconsin, Madison, Wisconsin 53706 (P.-M.D.)

Arbuscular mycorrhizal (AM) fungi, in symbiosis with plants, facilitate acquisition of nutrients from the soil to their host. After penetration, intracellular hyphae form fine-branched structures in cortical cells termed arbuscules, representing the major site where bidirectional nutrient exchange takes place between the host plant and fungus. Transcriptional mechanisms underlying this cellular reprogramming are still poorly understood. GRAS proteins are an important family of transcriptional regulators in plants, named after the first three members: GIBBERELLIC ACID-INSENSITIVE, REPRESSOR of GAI, and SCARECROW. Here, we show that among 45 transcription factors up-regulated in mycorrhizal roots of the legume *Lotus japonicus*, expression of a unique GRAS protein particularly increases in arbuscule-containing cells under low phosphate conditions and displays a phylogenetic pattern characteristic of symbiotic genes. Allelic *rad1* mutants display a strongly reduced number of arbuscules, which undergo accelerated degeneration. In further studies, two RAD1-interacting proteins were identified. One of them is the closest homolog of *Medicago truncatula*, REDUCED ARBUSCULAR MYCORRHIZATION1 (RAM1), which was reported to regulate a glycerol-3-phosphate acyl transferase that promotes cutin biosynthesis to enhance hyphopodia formation. As in *M. truncatula*, the *L. japonicus* *ram1* mutant lines show compromised AM colonization and stunted arbuscules. Our findings provide, to our knowledge, new insight into the transcriptional program underlying the host’s response to AM colonization and propose a function of GRAS transcription factors including RAD1 and RAM1 during arbuscule development.

Plant roots absorb inorganic phosphorus and most other nutrients from the soil to maintain normal growth. Phosphorus depletion in the rhizosphere favored the evolution of adaptive strategies in response to low phosphate (Pi) conditions, such as the formation of symbioses with arbuscular mycorrhizal fungi (AMF), which improves Pi uptake at the cost of photosynthesis-derived carbon. About 80% of terrestrial plants are colonized by AMF.

Under conditions of Pi starvation, plant roots exude strigolactones into the rhizosphere, which induce the germination of arbuscular mycorrhizal (AM) fungal spores and subsequent branching of fungal hyphae (Akiyama et al., 2005; Besserer et al., 2006). Strigolactones are short-lived molecules because of an ester bond that readily hydrolyzes in water. As a consequence, it is presumed that germinated external hyphae grow following the concentration gradient of the strigolactones toward the roots. Growing hyphae then contact the root surface and form hyphopodia, preparing for penetration. AM fungal hyphae secrete short-chain lipochitooligosaccharides (LCOs) and short-chain chitin oligomers to provoke plant roots into an early invasion response (Maillet et al., 2011; Genre et al., 2013). Subsequently, in root epidermal and subjacent cortex cells, nuclei move toward the hyphal penetration sites and direct the formation of the pre-penetration apparatus (PPA), lined with microtubules and endoplasmic reticulum, which guide the hyphae into the root endosphere (Genre et al., 2005, 2008). Once the hyphae grow into the cortical cells, they form highly branched structures, the arbuscules (Latin for little tree). Arbuscules are separated from the cytoplasm by the periarbuscular membrane, which is the continuous plant plasma membrane surrounding the arbuscule. The periarbuscular membrane contains a repertoire of proteins that facilitate its distinct functions, such as mycorrhiza-specific Pi transporters (Rausch et al., 2001; Harrison et al., 2002; Paszkowski et al., 2002; Karandashev et al., 2004; Kobae and Hata, 2010; Pumplin et al., 2012) and
half-size ATP-binding cassette transporters stunted arbuscule (STR)/STR2 (Zhang et al., 2010). Arbuscules have a short life span of approximately 4 to 15 d with mature arbuscules subsequently undergoing collapse concomitant with the appearance of septa before disappearing, leaving the cells for recolonization (Sanders et al., 1977).

During evolution, AM symbiosis (AMS) emerged about 400 million years earlier than the rhizobium-legume symbiosis (RLS; Redecker et al., 2000; Parniske, 2008) and is thought to have played a critical role during land colonization by plants (Humphreys et al., 2010). Genetic evidence shows that both symbioses share a common symbiosis signaling pathway (CSSP), which is composed of Leu-rich repeat receptor-like kinase (SYMBIOSIS RECEPTOR-LIKE KINASE [SYMRK] which is composed of Leu-rich repeat receptor-like kinase (RLK) (Humphreys et al., 2010). SYMRK is required for mycorrhizal factor-induced lateral root growth (Maillet et al., 2011; Gobbato et al., 2012). Moreover, DELLA proteins and GA3 were reported to antagonistically regulate arbuscule development in pea (Pisum sativum), M. truncatula, and rice (Oryza sativa; Floss et al., 2013; Foo et al., 2013; Yu et al., 2014). A GA3−deficient mutant showed enhanced mycorrhization in pea and exogenous GA3 repressed mycorrhization in three species. Meanwhile, severely impaired arbuscules were observed in mutants of DELLA orthologs. DELLA interacting protein (DIP1) was identified in rice, and RAM1 was shown to interact with DIP1 (Yu et al., 2014). This suggests that formation of multicomponent GRAS transcription factor complexes is a prerequisite for elicitation of nodulation or mycorrhization (Oldroyd, 2013).

Arbuscule development occurs in genetically separable stages, including (1) PPA formation, (2) fungal cell entry, (3) birdsfoot formation (lowly branched arbuscule resembling a bird’s foot), (4) highly branched hyphae forming a mature arbuscule, and (5) arbuscule collapse including formation of septa to disconnect old arbuscules from the hyphal network (Gutjahr and Parniske, 2013). Defective CCaMK and CYCLOPS resulted in the arrest of PPA formation. The Vapyrin gene from M. truncatula and its ortholog PENETRATION AND ARBUSCULE MORPHOGENESIS1 (PAM1) from Petunia hybrida are required for both AMS and RLS (Feddermann et al., 2010; Pumplin et al., 2010). Down-regulation of Vapyrin by RNA interference (RNAi) displayed a significant reduction in hyphopodium penetration into the epidermal cell, and hyphae failed to enter the cortical cells (Pumplin et al., 2010). In pam1 mutants, although few arbuscules were observed in the cortical cells, their growth was arrested at the early point of branching (Feddermann et al., 2010; Pumplin et al., 2010). Half-size ATP-binding cassette transporters STR/STR2 were shown to be essential for mature arbuscule formation. In M. truncatula and rice, loss of function mutations in STR genes resulted in stunted arbuscules (Zhang et al., 2010; Gutjahr et al., 2012), suggesting a role of STR proteins in the transport of metabolites that are essential for a functional root-AMF interaction. Mycorrhiza-specific Pi transporters, such as M. truncatula PHOSPHATE TRANSPORTER4 (MtPT4), Rice PHOSPHATE TRANSPORTER11 (OsPT11), and maize (Zea mays) PHOSPHATE TRANSPORTER1,6 (Pht1,6), localize to arbuscule-containing cells and are indispensable for maintenance of arbuscule function (Javot et al., 2007; Yang et al., 2012; Willmann et al., 2013). Loss-of-function mutants and RNAi lines of MtPT4, OsPT11, and maize pht1,6 mutant led to reduced root colonization with premature degeneration of the arbuscules. In addition, RNAi knockdown of VESICLE ASSOCIATED MEMBRANE PROTEIN72 in M. truncatula (Gobbato et al., 2012). GPAT initiates de novo glycerolipid synthesis and participates in the biosynthesis of cutin and suberin, which are the polymer matrices of the cell wall (Beisson et al., 2007; Li et al., 2007). Phenotypically, RAM1, but not NPS1, is required for mycorrhizal factor-induced lateral root growth (Maillet et al., 2011; Gobbato et al., 2012).

Plant-specific GRAS transcription factors play an important regulatory role in root and shoot development, the GA3 and phytochrome A signaling pathways, abiotic stress, and symbiosis (Di Laurenzio et al., 1996; Peng et al., 1997; Pysh et al., 1999; Bolle et al., 2000; Greb et al., 2003; Kaló et al., 2005; Smit et al., 2005; Fode et al., 2008). The GRAS protein family was divided into DELLA, HAIRY MERISTEM, LILium LONGIFLORUM SCARECROW-LIKE, PHYTOCHROME A SIGNAL TRANSDUCTION1, LATERAL SUPPRESSOR, SCR, SHORT-ROOT, and SCARECROW-LIKE3 subfamilies, each with distinct conserved domains and functions (Tian et al., 2004). In symbiosis, the CSSP is divergent from GRAS complexes. NSP2 forms a DNA-binding complex with another GRAS transcription factor, NODULATION SIGNALING PATHWAY1 (NSP1), to elevate early nodulin gene expression during rhizobial infection (Hirsch et al., 2009). NSP2 interacts with GRAS transcription factor REDUCED ARBUSCULAR MYCORRHIZATION1 (RAM1) to induce RAM2, encoding a glycerol-3-phosphate acyl transferase (GPAT), to facilitate hyphopodia formation on the root surface during mycorrhizal colonization (Oldroyd, 2013).
or SOLUBLE N-ETHYLMALEIMIDE-SENSITIVE FACTOR ATTACHMENT PROTEIN RECEPTOR family member L. japonicus VESICLE TRANSPORT THROUGH T-SNARE INTERACTION12 in L. japonicus led to distorted arbuscular development, which provided evidence for a role of these proteins in vesicle trafficking during arbuscule formation (Ivanov et al., 2012; Lota et al., 2013). Moreover, two L. japonicus mutants were reported to be affected in distinct steps of arbuscule development but not during RLS (Groth et al., 2013).

While transcriptional changes in arbuscule-containing cortical cells were recently reported (Hogekamp et al., 2011), the sophisticated regulatory mechanisms underlying these changes are still poorly understood. Here we report that a unique GRAS protein, REQUIRED FOR ARBUSCULE DEVELOPMENT1 (RAD1), regulates arbuscule development. RAD1 interacts with RAM1 (ARBUSCULE DEVELOPMENT1) that are involved in arbuscule development. LjNSP2, respectively, providing support for the existence of a GRAS protein-protein interaction network in arbuscule development.

RESULTS

Gene Expression Profiling Reveals Transcription Factors Related to AM Colonization in L. japonicus

To investigate transcriptional regulation of arbuscule development and to screen for candidate transcription factor genes, roots of L. japonicus Gifu-129, colonized for 8 weeks by Rhizophagus irregularis and resupplied or not with Pi, were subjected to deep sequencing analysis of messenger RNA (RNAseq). Two independent biological replicates exhibited similar mycorrhization rates and induction of arbuscular marker gene expression (LjPT4) were used. Application of a minimum 2-fold induction at a significance level of $P < 0.01$ was used as a cutoff for significantly up-regulated genes. As a result, 45 genes encoding transcription factors or transcriptional regulators were significantly up-regulated by R. irregularis compared with nonmycorrhizal roots grown in a similar low-Pi environment (Fig. 1A; Supplemental Table S1). These genes encoded GRAS (18 members), APETALA2 (AP2)/ETHYLENE RESPONSE FACTOR (ERF; eight), NO APICAL MERISTEM, ARABIDOPSIS TRANSCRIPTION ACTIVATION FACTOR and CUP-SHAPED COTYLEDON (NAC)-domain family members (four), MYELOBLASTOSIS (MYB; three), C2H2 zinc finger (five), MADS DNA-binding domain (MADS [from MINICHROMOSOME MAINTENANCE, AGAMOUS, DEFICIENS and SERUM RESPONSE FACTOR]; two), BASIC LEUCINE ZIPPER (bZIP; one), AUXIN RESPONSE FACTOR (ARF; one), WRKY DNA binding domain (one), NODULE INCEPTION (NIN)-LIKE protein (one), and LATERAL ORGAN BOUNDARIES (LOB) domain (one) transcription factors, indicating the complexity of transcriptional regulation of AMS. These are distinguishable in four general expression patterns: genes that were overall highly expressed (11), genes that were only responsive to AM under low-phosphorus conditions (23), genes with high expression under low Pi regardless of AM status (nine), and the rest of intermediate status (i.e. with no exact classification; two; Supplemental Table S2).

To investigate the distribution of these genes in plant species, a reciprocal blast survey in AM nonhosts (L. angustifolius, C. rubella, T. halophila, B. rapa, A. lyrata, and Arabidopsis) and AM hosts (M. truncatula, P. trichocarpa, C. papaya, G. raimondii, and V. vinifera) was carried out. Except for the partial sequences of chr1.CM1413.480.r2,d and L. japonicus SELECTED GENOME SHOTGUN ASSEMBLY (LjSGA)_055804.0.1, the remaining 43 transcription factor coding sequences were clustered into three groups: conserved genes that occur in host and nonhost species (8), host-specific or correlated (16+4), and potential L. japonicus-specific genes (15), respectively (Fig. 1B). Interestingly, out of the 16 AM host-specific genes, 10 were absent in L. angustifolius (Fig. 1B), which is a known AM nonhost but still retains the ability to associate with nitrogen-fixing rhizobia (Oba et al., 2001; Schulze et al., 2006). This suggested the important role of these 10 transcription factors in AMS. To further determine the absence of the AM host-specific transcription factors in the Brassicaceae, microsynteny analysis was performed. We compared genomic blocks encompassing these genes in M. truncatula and P. trichocarpa and the corresponding blocks in Arabidopsis. Indeed, this syntenic analysis supports the orthology between both host species for most of the putative host-specific genes (15 of 16; Supplemental Text S1). By contrast, none of the syntenic regions in Arabidopsis contains homologs of these genes, supporting their absence in this well-characterized nonhost species (Supplemental Text S1).

To verify AM-dependent expression of the previously described genes identified by RNAseq, quantitative reverse transcription (qRT)-PCR was performed. We tested eight genes, and all of them showed a similar expression pattern compared to the RNAseq results, although the fold induction rates were slightly different (Supplemental Fig. S1). Mycorrhization and AM marker gene expression are generally suppressed under high-Pi conditions, which was recently suggested to be mediated by GA3 signaling and low levels of DELLA proteins (Nagy et al., 2009; Breuillin et al., 2010; Chen et al., 2011; Floss et al., 2013; Lota et al., 2013). Consistently, the expression of all eight tested mycorrhiza up-regulated transcription factors was repressed after a 2-week high-Pi treatment (Supplemental Fig. S1).

LORE1a Insertion Mutant rad1 Displays Impaired Mycorrhization

To identify important transcriptional regulators in AMS, mutants of these candidate transcription factor genes, containing LOTUS RETROTRANSPOSON 1A (LORE1a) transposon insertions (Fukai et al., 2012;
Figure 1. Global analysis of mycorrhiza up-regulated transcription factors. A, Hierarchical clustering of mycorrhizal (+M) and nonmycorrhizal wild-type Gifu-129 expression profiles based on transcript abundance of the genome-wide 45 AM up-regulated transcription factors identified by RNAseq (adjusted $P < 0.01$, fold change $> 2$). Gifu-129 plants were grown for 8 weeks at low Pi (5.5 $\mu$M; low P) or high Pi (7.5 mM; high P) conditions, respectively, or they were grown at low Pi for 6 weeks.
Urbański et al., 2012), were screened for potential defects in AM5 with R. irregularis. Subsequently, mycorrhization rate and arbuscule morphology were surveyed at 8 weeks postinoculation (wpi). Mycorrhizal roots of a mutant with a LORE1a insertion in the 5’ untranslated region (UTR) of a GRAS protein gene were found to contain a high number of stunted arbuscules (Supplemental Fig. S2A). We therefore named the gene RAD1 (chr4. CM1864.540.12.1m) and the corresponding mutant as rad1-1. The full-length RAD1 complementary DNA (cDNA) sequence corresponding to the intronless gene is 1530 bp long, and the predicted protein size is 56.7 kD. To further confirm the function of RAD1 in AM development, two other alleles, rad1-2 and rad1-3, were identified with LORE1a insertions in the RAD1 coding region (Fig. 2A). The RAD1 transcripts were detected in three allelic mutants (Supplemental Fig. S2B). Both the 5’ and 3’ partial transcripts of RAD1 were significantly reduced in rad1-1, which indicated that the insertion in the 5’ UTR leads to a knockdown of RAD1 gene expression. In rad1-2, the 5’ partial transcript of RAD1 was barely detectable, whereas the 3’ transcript could be identified at a very low level. In rad1-3, the 5’ transcript level was comparable to that in Gifu-129, whereas the 3’ transcript was not detectable.

Although the total mycorrhization rates in both rad1-1 and rad1-3 were indistinguishable from that in the wild type at 6 wpi, a significantly reduced degree of mycorrhization was observed in rad1-2, especially concerning the percentage of sectors with arbuscules, vesicles, and hyphae (Fig. 3A). In both rad1-2 and rad1-3, arbuscule number and size were significantly less than in wild-type Gifu-129, accompanied by accelerated arbuscule degeneration (Figs. 2B and 3B). In rad1-2 and rad1-3, 52% and 40%, respectively, of arbuscules ranged from 10 to 20 μm in size, whereas 18% in the wild type were of this diameter (Fig. 3B). Likewise, 36% of arbuscules in rad1-1 ranged from 10 to 20 μm in size, which was an insignificant difference from the control due to relatively large variation between biological replicates. In addition, the percentage of arbuscules ranging from 50 to 60 μm was much less in all three mutant alleles. In particular, arbuscules larger than 40 μm were rarely observed in rad1-2. The arbuscule phenotype was more severe in rad1-2 than that in rad1-1 and rad1-3, which was consistent with the different RAD1 transcript levels in the three allelic mutants. Intriguingly, the arbuscules in the mutants collapsed prematurely with increased numbers of septa compared with Gifu-129 (Fig. 2C). This indicated that the GRAS transcription factor RAD1 is essential for regulating arbuscule morphology and senescence.

To reveal transcriptional changes resulting from defective RAD1, expression of AM marker genes correlating to mature arbuscule formation was studied in rad alleles relative to Ubiquitin. Six weeks postinoculation, expression of marker genes of arbuscule development stage IV, such as STR and RAM2 as well as Lipt4, was significantly reduced in rad1 allelic mutants (Fig. 3, C–E). Primers specific for R. irregularis large subunit ribosomal RNA were used to demonstrate fungal RNA levels (van Tuinen et al., 1998; Fig. 3F). Expression of AM-inducible marker genes was also compromised in rad1-2 and rad1-3 in a time course experiment (Supplemental Fig. S3). These data indicated that impairment of RAD1 inhibits the formation of highly branched arbuscules and the corresponding marker gene expression. Additionally, constitutive overexpression of RAD1 in L. japonicus hairy roots resulted in a significantly increased mycorrhization rate and number of sectors with arbuscules (Supplemental Fig. S4A). Exemplarily, three root samples were analyzed for RAD1 overexpression, exhibiting on average a 4-fold higher RAD1 expression in the Cauliflower mosaic virus (CaMV) 35S promoter 35Sp:RAD1 transgenic roots compared with hairy roots harboring the empty vector (Supplemental Fig. S4B). Moreover, expression levels of Lipt4 were also increased in these roots (Supplemental Fig. S4C). This result further confirmed RAD1 function in arbuscule development.

Expression of RAD1 Is Root Specific and Enhanced in Arbuscule-Containing Cells

The gene expression of RAD1 was detectable in root tissue by reverse transcription (RT)-PCR in colonized and noncolonized plants under low-Pi conditions (Supplemental Fig. S5A). To elucidate the spatial pattern of RAD1 gene expression, a 1.9-kb promoter region of RAD1 was fused to GUS generating the chimeric gene RAD1:pro:GUS for Agrobacterium rhizogenes-mediated transformation of hairy roots. Discosoma spp. RED FLUORESCENT PROTEIN encoded by the pRedroot vector was used for the selection of successful genetic transformation. Consistent with the

---

**Figure 1.** (Continued) with a subsequent shift from low Pi to high Pi for 2 weeks (shift). Significantly up-regulated transcription factors are shown in a heat map drawn by using R. Log10 values of normalized counts per transcript plus one were used. Transcription factor families are labeled above the genes. The closest homologs and RAD1 were labeled below the plot. Asterisks indicate the genes whose expression was confirmed by qRT-PCR in Supplemental Figure S1. B. The 43 AM up-regulated transcription factors were clustered into three groups in AM host and nonhost species. Amino acid sequences of these L. japonicus AM up-regulated transcription factors were utilized for reciprocal blast in AM nonhost Brassicaceae (Capsella rubella, Thellungiella halophila, Brassica rapa, Arabidopsis lyrata, and Arabidopsis [Arabidopsis thaliana]), Lupinus angustifolius, and AM hosts (M. truncatula, Populus trichocarpa, Carica papaya, Gossypium raimondii, and Vitis vinifera). Symbols indicate the presence (+) or absence (−) of the capacity for symbiosis. Detected homologs are labeled with dark blue. Lack of homolog detection was labeled with white.
qRT-PCR results (Supplemental Fig. S1), GUS was not detected under high-Pi conditions (Fig. 4, A and B). RAD1 could be induced by low Pi, and moderate GUS staining was detected in the cells closest to the vascular cylinder (Fig. 4, C and D). After colonization with *R. irregularis*, RAD1pro::GUS expression was observed in the presence of intraradical hyphae in epidermal cells and in the outer cortical cells (Fig. 4, E–G). Interestingly, as the development of arbuscules in the cortical cells progressed, the highest GUS activity was observed in the well-developed arbuscule-containing cells (Fig. 4, H–J). GUS activity progressively decreased in cells harboring aging and senescent arbuscules. The activity of RAD1 promoter was induced as early as internal hyphae appeared in the root tissue, which suggests that RAD1 could influence the growth of intracellular hyphae upon induction of the gene by inter- and/or intracellular hyphal growth. Overall, the temporal pattern of RAD1pro::GUS expression during arbuscule development supports the role of RAD1 in arbuscule formation.

**Symbiotic Function of RAD1 Is Characteristic for AMS**

AMS and RLS share common symbiosis regulatory pathways at the early stages of microbial root colonization. To check whether RAD1 is in line with or divergent from the CSSP, transcription of RAD1 was studied in roots forming nodules. The nodulation marker gene *L. japonicus* NODULE INCEPTION was induced significantly, but the expression levels of AM marker genes LjPT4 and SUBTILASE M1, as well as RAD1 remained unchanged in roots carrying nodules (Supplemental Fig. S6A). Accordingly, visual inspection revealed that normal pink nodules were formed in all three genotypes, and there was no significant difference in the number of nodules between Gifu-129, rad1-1, and rad1-3 plants (Supplemental Fig. S6B). These observations led us to conclude that RAD1 function is essential for normal mycorrhization, rather than a common prerequisite for both root symbioses.

**RAD1 Displays an Evolutionary Pattern Characteristic of Phylogenetically AMS-Related Genes**

To better characterize RAD1 and determine its specificity among GRAS proteins, all 18 mycorrhiza-up-regulated

---

**Figure 2. Mycorrhizal phenotype of rad1 allelic mutants.** A, RAD1 (chr4.CM1864.540.r2.m) gene structure is shown at scale, and LORE1a insertion sites are illustrated in mutants rad1-1 (~71 bp upstream of the translation start codon [ATG]), rad1-2 (359 bp downstream of start ATG), and rad1-3 (1,379 bp downstream of start ATG). B, Confocal microscopy images of *R. irregularis* in the roots of Gifu-129 and the rad1 allelic homozygous mutants. Images at right show the overlay of fluorescence and bright-field images. Scale bars = 100 μm. C, Defective arbuscules in rad1 allelic homozygous mutants. Fungal structures were stained with WGA-Alexa Fluor 488. White arrowheads indicate characteristic mycorrhizal structures: a, arbuscule; ih, intraradical hyphae; eh, extraradical hyphae; da, degenerated arbuscule; and s, septa. Scale bars = 10 μm.
L. japonicus GRAS protein family members, 33 GRAS family members from Arabidopsis, as well as known GRAS transcription factors involved in symbiosis such as LjNSP1, LjNSP2, MtNSP1, MtNSP2, OsDIP1, MtDELLA1/2, and MtRAM1, were subject to phylogenetic tree analysis (Fig. 5). RAD1 and the closely related protein LjSGA_122341.1 clustered separately from the Arabidopsis GRAS members. Such an absence in Arabidopsis is a common pattern of genes involved in AMS (Delaux et al., 2014). We further investigated the phylogenetic distribution of RAD1 in land plants by looking in genomes and transcriptomes of 18 other hosts and nine
Whether the additional copy was lost in species with a single RAD1 sequence, such as *M. truncatula* and *L. japonicus*, or this duplication took place only in specific lineages would require a much deeper sampling within these clades. In addition, putative orthologs of RAD1 were also identified in monocots, lycophytes, and liverworts, suggesting an ancient origin of this transcription factor (clade III). To further test the orthology of these potential homologs, we conducted a synteny analysis on a subset of host species and Arabidopsis as a nonhost representative. This analysis supported the orthology of all the host RAD1, including those from monocots and *Selaginella moellendorffii* (Supplemental Fig. S7).

**GRAS Proteins Interact with RAD1 in Vitro and in Vivo**

As GRAS transcription factors have been reported to form dimers with GRAS family members (Itoh et al., 2002; Hirsch et al., 2009; Gobbato et al., 2012), RAD1 was used as bait and seven GRAS proteins representative of the different clades in the phylogenetic tree were used as prey in yeast two-hybrid (Y2H) assays to identify potential interacting proteins that may facilitate RAD1 function. Our data showed that RAD1 failed to form a homodimer in the yeast system. Two specific interactions were demonstrated in vitro (Supplemental Fig. S5B). One of them was the protein encoded by chr1.cm1852.30.r2.m, which shares 80% sequence identity with MtRAM1, suggesting it is the orthologous RAM1 in *L. japonicus*. Interestingly, RAD1 also directly interacted with LjNSP2 (Supplemental Fig. S5B), a protein with dual regulatory functions in root symbiosis signaling and appropriate establishment of RLS and AMS (Kaló et al., 2005; Hirsch et al., 2009; Liu et al., 2011; Maillet et al., 2011; Laressergues et al., 2012).

GRAS domains are composed of five highly conserved sequential motifs: Leu heptad repeat I (LHRI), VHIID, Leu heptad repeat II, PFYRE, and SAW (Bolle, 2004). The LHRI motif and the C-terminal PFYRE and SAW motifs are required for GRAS transcription factor interaction (Hirsch et al., 2009). To study the specificity of interaction between RAD1 and RAM1, truncated versions of RAD1 were generated. As RAM1 exhibited strong autoreactivation, it was used as prey in our Y2H assay. As shown in Figure 6A, deletion of the N terminus including the LHRI motif and the C-terminal PFYRE and SAW (Bolle, 2004). The GRAS domains are composed of five highly conserved sequential motifs: Leu heptad repeat I (LHRI), VHIID, Leu heptad repeat II, PFYRE, and SAW (Bolle, 2004). The LHRI motif and the C-terminal PFYRE and SAW motifs are required for GRAS transcription factor interaction (Hirsch et al., 2009). To study the specificity of interaction between RAD1 and RAM1, truncated versions of RAD1 were generated. As RAM1 exhibited strong autoreactivation, it was used as prey in our Y2H assay. As shown in Figure 6A, deletion of the N terminus including the LHRI motif and specifically the SAW motif from RAD1 abolished the interaction between RAD1 and RAM1. None of the RAD1 versions interacted with the GRAS control protein, indicating the specificity of the interaction between RAD1 and RAM1 (Fig. 6A).

In vitro glutathione S-transferase (GST) pull-down assay was subsequently performed to verify the direct interaction between RAD1, LjNSP2, and RAM1, respectively. Our results showed that recombinant RAM1 and LjNSP2 proteins purified from *Escherichia coli* have the ability to precipitate RAD1 in vitro, thus confirming the Y2H results (Fig. 6B).

To further support the RAD1-RAM1 interaction, the subcellular localization of both proteins was investigated.

**Figure 4.** Histochemical analysis of proRAD1:GUS expression in mycorrhizal and nonmycorrhizal hairy roots of *L. japonicus*. A and B, Absence of GUS activity in nonmycorrhizal roots and root tips under high-Pi conditions. Scale bars = 100 μm. C and D, GUS activity in nonmycorrhizal roots and root tips under low-Pi conditions. Scale bars = 100 μm. E to G, GUS activity was observed in arbuscule-containing and adjacent cells. E, Fluorescence image of WGA-Alexa Fluor 488-stained intraradical fungal structures. F, Bright-field image of magenta GUS-stained root sector shown in E. G, Alexa Fluor 488 and bright-field image were merged to show colocalization of intraradical fungal structures. F, Bright-field image of magenta GUS-stained intraradical fungal structures. F, Bright-field image of magenta GUS-stained root sector shown in E. G, Alexa Fluor 488 and bright-field image were merged to show colocalization of intraradical fungal structures. F, Bright-field image of magenta GUS-stained root sector shown in E. G, Alexa Fluor 488 and bright-field image were merged to show colocalization of intraradical fungal structures. F, Bright-field image of magenta GUS-stained root sector shown in E. G, Alexa Fluor 488 and bright-field image were merged to show colocalization of intraradical fungal structures. F, Bright-field image of magenta GUS-stained root sector shown in E. G, Alexa Fluor 488 and bright-field image were merged to show colocalization of intraradical fungal structures. F, Bright-field image of magenta GUS-stained root sector shown in E. G, Alexa Fluor 488 and bright-field image were merged to show colocalization of intraradical fungal structures. F, Bright-field image of magenta GUS-stained root sector shown in E. G, Alexa Fluor 488 and bright-field image were merged to show colocalization of intraradical fungal structures.
using GFP-protein fusions. Particle bombardment of onion (Allium cepa) epidermal cells demonstrated GFP-RAD1 localization in the nucleus (Supplemental Fig. S8A). A. rhizogenes-mediated transformation of L. japonicus hairy roots confirmed the nuclear localization of GFP-RAD1 (Supplemental Fig. S8B). The RAD1-RAM1 and RAD1-LjNSP2 interaction was confirmed in the nucleus by bimolecular fluorescence complementation (BiFC) assay (Fig. 6D; Supplemental Fig. S9). Moreover, RAD1 did not form homodimers in vivo, whereas RAM1 homodimerized in the nucleus and cytoplasm. RAM1, RAD1, and LjNSP2 did not interact with the split C terminus of Yellow Fluorescent Protein (YFP) or the N terminus of Yellow Fluorescent Protein alone, which suggested the specific interaction between them. Coimmunoprecipitation (Co-IP) was also used to confirm the interaction between RAD1 and RAM1 in vivo. To this end, hemagglutinin (HA)-tagged RAD1 and GFP-tagged RAM1 were transiently expressed in Nicotiana benthamiana leaves. Indeed, RAM1 was able to capture RAD1 in vivo when both proteins were transiently and simultaneously expressed (Fig. 6C). In sum, these data support our view that RAD1 forms multicomponent GRAS transcription factor complexes including RAM1 and/or NSP2 involved in late fungal colonization of host roots in AMS.

L. japonicus ram1 Mutant Fails to Establish Mycorrhizae

To elucidate the significance of RAM1 in L. japonicus–R. irregularis association, two ram1 insertion alleles were identified. The ram1-I contains a LORE1a insertion in the...
Figure 6. RAD1 interacts with RAM1 in vitro and in vivo. A, Interaction between truncated versions of RAD1 and full-length RAM1. Conserved domains in GRAS proteins are color coded. BD, GAL4 DNA-binding domain; SD, synthetic minimal media based upon yeast (*Saccharomyces cerevisiae*) nitrogen base supplemented with Glc and amino acids (aa) except those indicated (L, W, H); 3AT, 3-amino-1,2,4-triazole. cGRAS encoding GRAS protein fused to the GAL4 activation domain (AD) was used as a negative control. B, In vitro pull-down assay. GST fusion proteins RAM1 (GST-RAM1) and NSP2 (GST-NSP2) or GST alone (control) were used to precipitate His-tag fused RAD1 (His-RAD1). + indicates the presence of the recombinant protein. WB: anti-His, Western blotting with anti-His antibody detecting His-bait fusion proteins. Purified GST fusion protein through glutathione-Sepharose beads was analyzed by SDS-PAGE and Coomassie Brilliant Blue in-gel protein staining (CBB staining). Arrows at left demonstrate presence of GST fusion proteins as designated. C, In vivo Co-IP of HA-RAD1 by GFP-RAM1. In vivo Co-IP was performed using *N. benthamiana* leaves infiltrated with *Agrobacterium rhizogenes* GV3101 strain harboring the desired interacting pair (HA:RAD1/GFP:RAM1). Leaf extracts were subjected to immunoprecipitation (IP) using anti-GFP antibody. Anti-HA and anti-GFP antibodies were used for western blotting of Co-IP sample. Symbols represent the presence (+) or absence (−) of the various proteins. D, Interaction between RAM1 and RAD1 by BiFC assay in *N. benthamiana*. Leaves were infiltrated with
second exon of the RAM1 gene, and the second allele ram1-2 has a LORE1a insertion in the first exon (Supplemental Fig. S10A). Homozygosity of ram1-1 and ram1-2 plants was confirmed by genotyping with sequence-specific primers (Supplemental Fig. S10B). When inoculated with R. irregularis, Gifu-129 established a normal mycorrhiza. In contrast, in ram1-1, not only was the total mycorrhization rate lower compared with the wild type, but intracellular branched hyphae instead of arbuscules appeared in the cortex (Fig. 7, A and B). In ram1-2, although the total mycorrhization rate was similar to the wild type, the arbuscules were stunted and arrested at birdsfoot stage III. This indicates that RAM1 is required for arbuscule development, which was consistent with RAM1 function in M. truncatula (Gobbato et al., 2012, 2013).

Subsequently, we studied the expression pattern of L. japonicus RAM1. Our qRT-PCR result showed that transcript levels of RAM1 were detected in all tissues tested and were strongly enhanced in mycorrhizal roots and root tips (Fig. 8A). Induction of RAM1 gene expression cohered with the colonization of R. irregularis, which was suppressed in rad1 alleles (Fig. 8B). The 1.8-kb promoter region of RAM1 was fused to GUS to test the tissue-specific expression pattern of RAM1. Under low-Pi conditions, GUS activity could be detected weakly in root primordia, where lateral roots emerge from the primary root (Fig. 8C). In the presence of R. irregularis, RAM1 was induced in the cortical cells (Fig. 8, G and H). Interestingly, GUS activity could also be detected in the root tips where intercellular hyphae barely reach (Fig. 8, E and F). Likewise, RAM1 promoter activity could also be detected in lateral root primordia (Fig. 8D). This indicates that RAM1 is strongly expressed in arbuscule-containing cells and root meristems, suggesting a dual role in arbuscule development and in root development.

DISCUSSION

Here we demonstrated that the GRAS protein encoding gene RAD1, the expression of which was increased in arbuscule-containing cells, was required for arbuscule development. RAD1 was shown to be specific for AMS, rather than for RLS (Supplemental Fig. S6). This phenotype is supported by a phylogenetic pattern specific to AM-related genes (Figs. 1B and 5). Premature collapse of arbuscules was observed in three rad1 allelic mutants, and the temporal profile of marker gene expressions was recorded, supporting the role of RAD1 in arbuscule development as well as in overall fungal colonization of the cortex. RAD1-RAM1 interaction could be demonstrated both in vitro and in vivo (Fig. 6; Supplemental Fig. S5B). Our results support a molecular role of RAD1 oligomerization with at least two GRAS-type transcriptional regulators, RAM1 and LjNSP2, in Pi-dependent arbuscule development.

Transcriptional Network Underlying Cellular Reprogramming during Late Colonization in AMS

Microbial reprogramming of root cellular development in the AMS is under intense study and is accompanied by remarkable progress in understanding the molecular mechanisms underlying mainly the precontact and early colonization phase of AM development (Oldroyd, 2013). Only a few cis-regulatory elements were reported to be essential for AM-specific gene expression in arbuscule-containing cells, whereas little is known about components of the corresponding transcriptional machinery (Chen et al., 2011; Lota et al., 2013; Xie et al., 2013). Overall, we found 45 transcriptional regulators that were significantly induced during mycorrhiza formation and the expression of which was under the control of Pi (Fig. 1; Supplemental Fig. S1; Supplemental Table S1). This common expression pattern of multiple regulatory proteins strongly suggested simultaneous and multilevel regulation of developmental and functional processes during late colonization in AMS, which remain to be mechanistically dissected. It has previously been reported that several transcription factor genes were induced after a 6-h treatment by AM fungal Myc-LCO signals in M. truncatula (Czaja et al., 2012). Comparing the encoded proteins of these early induced M. truncatula genes with the 45 mycorrhiza-regulated transcription factors from L. japonicus did not result in the identification of orthologous pairs, which suggested that expression of transcription factors during arbuscule development is likely to be triggered in a Myc-LCO-independent way.

Only a few of these transcription factor genes were identified as mycorrhiza regulated in earlier works (Guether et al., 2009; Hogeckamp et al., 2011; Gaude et al., 2012; Gobbato et al., 2012; Volpe et al., 2013). According to the Plant Transcription Factor Database, there are 46 GRAS family members in L. japonicus (Guo et al., 2008). Eighteen of these were significantly induced in mycorrhizal roots compared with non-mycorrhizal roots, 10 of which appear in the AM host and are not detected by an extensive BLAST survey in the nonhost species (Fig. 1B; Supplemental Text S1). In sum, we have expanded the number of mycorrhiza-regulated GRAS proteins and members of other transcription factor families, providing a global view of transcriptional regulators in mycorrhizal L. japonicus.
roots. It was previously postulated that GRAS protein complex formation is crucial in the activation of either RLS or AMS (Gobbato et al., 2012). Our work will serve as a basis for more detailed protein interaction network and coexpression analysis, respectively, in the context of arbuscule development.
Figure 8. Transcriptional expression pattern of RAM1 in L. japonicus.
A, RT-PCR analysis of RAM1 transcript levels in different tissues of the wild type in the presence or absence of R. irregularis. Rt, Root tip; R, root; Yl, young leaves; Ml, mature leaves. Mean values from three biological replicates are shown (±SD, n = 3). Student’s t test was used. *P < 0.05; **P < 0.01.
B, RAM1 transcript levels in three rad1 alleles in the presence or absence of R. irregularis. Mean values from three biological replicates are shown (±SD, n = 3). Student’s t test was used. *P < 0.05; **P < 0.01.
C to H, Histochemical analysis of proRAM1:GUS expression in the absence of R. irregularis (C) and in the presence of R. irregularis (D–H) in hairy roots of L. japonicus. Arrows indicate lateral root primordia. Scale bars = 200 μm.

Regulatory Network of GRAS Proteins in Arbuscular Mycorrhizae

To date, several GRAS transcription factors involved in mycorrhizal signaling or mycorrhizal colonization have been identified, namely NSP1, NSP2, RAM1, DELLA, and DELLA-interacting proteins (Liu et al., 2011; Maillet et al., 2011; Gobbato et al., 2012; Delaux et al., 2013; Floss et al., 2013; Yu et al., 2014). In M. truncatula, NSP1 and NSP2 are essential for Nod factor-induced nodulation (Kaló et al., 2005; Smit et al., 2005). Their homologs LjNSP1 and LjNSP2 in L. japonicus were also reported to be required for nodulation (Heckmann et al., 2006). On the other hand, NSP1 and NSP2 were reported to regulate strigolactone biosynthesis both in M. truncatula and rice, but nsp1nsp2 double mutants exhibited well-developed and more arbuscules compared with wild-type roots, while the colonization was significantly reduced (Liu et al., 2011). In contrast, nsp2-2 mutant exhibited abolished Myc-LCO-triggered root branching, and quantitative reassessment of the mycorrhizal phenotype revealed a significantly reduced degree of colonization in nsp2-2 and nsp1 mutants when compared with wild-type plants, showing that NSP2 and NSP1 are involved in mycorrhizal signaling (Maillet et al., 2011; Delaux et al., 2013). In M. truncatula, the GRAS transcription factor RAM1 was reported to be required for AMS and to regulate the expression of RAM2 encoding a GPAT of the cutin biosynthesis pathway, which facilitates hyphopodia formation (Gobbato et al., 2012; Wang et al., 2012). Further research showed that RAM1 and RAM2 were also involved in arbuscule development (Gobbato et al., 2013). It remains to be investigated whether cooperative activity between RAD1 and RAM1 directs RAM2 expression in roots undergoing mycorrhization. Gobbato et al. (2012) also reported on protein-protein interaction between RAM1 and NSP2 in support of a role for NSP2 in AMS. Our work showed that RAD1 interacted with LjNSP2 in vitro and in vivo (Figure 6B; Supplemental Figs. S5B and S9). Recent work demonstrated that della1/della2 double mutant exhibited severely impaired development of arbuscules, and dominant DELLA protein lacking the DELLA domain restored the arbuscule formation in the presence of GA3, which indicated the positive regulatory role of DELLA in arbuscule development (Floss et al., 2013). In rice, DIP1 was identified as a protein interacting with DELLA, and the knockdown of DIP1 led to less mycorrhizal colonization (Yu et al., 2014). Interestingly, Yu et al. also found that DIP1 interacted with RAM1, although RAM1 did not interact with DELLA directly. Further genetic studies will likely reveal common functions of RAD1 and its interaction partners in AMS.

Proposed RAD1 Mode of Action

The mycorrhizal colonization in the rad1-2 mutant line was reduced, and all three rad1 alleles showed
accumulation of prematurely stunted arbuscules. Likewise, the ram1-1 mutant exhibited compromised mycorrhization, and both ram1 mutant alleles showed arrested arbuscules (Fig. 7, A and B). Promoter-GUS activity strongly suggested the concurrence of RAD1 and RAM1 expression in the cortex (Figs. 4 and 8). Like RAM1, gene expression analysis revealed a broad effect of RAD1 on AM-related genes, including RAM2 (Figs. 3D and 7E). Like NSP2, RAD1 has no predicted DNA-binding domains, whereas predicted DNA-binding residues are enriched in the N terminus of RAM1 when the BindN+ web-based tool was used for prediction (Hirsch et al., 2009; Wang et al., 2010). Considering the different proteins interacting with RAD1 and their proposed roles throughout AM development, we propose that RAD1 regulates subsets of mycorrhiza-inducible genes in cooperation with RAM1, which is likely to have DNA-binding activity (Oldroyd, 2013). However, to date we cannot exclude that the reduced induction of these genes may be a secondary effect of abnormal arbuscules. Nevertheless, we speculate that the decision to activate either one of these subsets of genes is defined by the formation of specific heterocomplexes of RAD1 with other regulatory proteins, promoting mycorrhiza-specific responses. The mycorrhiza-signaling pathway(s), which promotes RAD1 protein complex formation, and the mechanisms that dictate complex formation remain to be defined.

Many open questions still offer food for thought concerning AM research. The gene Medtr8g109760.1 (M. truncatula genome version 3.0) or Medtr4g104060 (M. truncatula genome version 3.5), the closest homolog of RAD1 in M. truncatula sharing 74% identity at its corresponding amino acid sequence level, was reported to be expressed in arbuscule-containing cells and in adjacent cells as was shown by laser-capture microdissection of mycorrhizal roots followed by microarray hybridization of labeled cDNA (Gaude et al., 2012). This suggested that RAD1 shares similar functions across different plant species. This hypothesis is also supported by the conservation of RAD1 in mycorrhizal host plants, including basal land plants (Fig. 5). Interestingly, microRNA miR5204* has the ability to cleave the messenger RNA derived from Medtr8g109760.1 (Devers et al., 2011). As the expression of miR5204* was correlated with the plant Pi status and up-regulated in mycorrhizae, it was proposed that miR5204* spatially regulates GRAS gene expression during AMS. Future work will likely shed more light on posttranscriptional regulation of the RAD1-dependent regulatory network.

MATERIALS AND METHODS

Plant Materials, Growth Conditions, and Transformation

L0REa insertion lines used in this article were kindly provided by Dr. Stig U. Andersen and Jens Stougaard (http://users-mb.au.dk/pmgrp/) under the following accession numbers: rad1-1 (30001039-2), rad1-2 (30030576), rad1-3 (30052260), ram1-1 (30002740), and ram1-2 (30082472). Lotus japonicus Gifu-129 was used in all experiments, and all L0REa insertion mutants that we used were in the Gifu-129 background. Gifu-129, rad1-1, rad1-2, rad1-3, ram1-1, and ram1-2 were grown in the greenhouse at 22°C and 16-h-day/8-h-night periods. For phenotypic and gene expression analysis, plants were grown on 0.8% (w/v) agar for 5 d in the dark and then transplanted to a soil/sand mixture (1.9 [w/w] with fungal inoculum (Rhizopogon irregularis, BRC75). Plants were grown in one-half-strength Hoagland solution with 5 μM NH4H2PO4, for 6 to 8 weeks. Agrobacterium rhizogenes-mediated hairy root transformation of L. japonicus was conducted as described previously (Lota et al., 2013). Transgenic roots were selected based on Discosoma spp. RED FLUORESCENT PROTEIN (pRedRoot)- or GFP-derived fluorescence (pH7WG2D). Hairy root transformation was mediated by A. rhizogenes strain 15864 as described.

Analysis of Mycorrhization and Measurement of Arbuscule Size

The procedures for mycorrhiza colonization analysis were based on the magnified intersection method from McGonigle et al. (1990) with some modification. For the mycorrhization rate, 15 root fragments, each about 1 cm long, were stained with trypan blue or wheat germ agglutinin (WGA) Alexa Fluor 488, and placed onto microscopic slides. Using 20× objective magnification, over 150 views per slide were surveyed and classified into five groups: no colonization, only hyphae (H), hyphae with arbuscules (H+A), hyphae with vesicles (V+H), and hyphae with arbuscules and vesicles (V+A+H). The percentages of each group were calculated by the number of each sector divided by total views. For arbuscule size measurement, after WGA-488 staining, about 30 images were taken from each slide and viewed using 40× magnification of Leica CLSM500. Each ecotype had three slides from three biological replicates. For the wild type, the images were taken close to the end of the infection units, as described previously (Javot et al., 2011). For rad1 mutants, images were randomly taken in the colonized region. Each arbuscule was measured along the longitudinal direction of the root. Then, the sizes of arbuscules from each slide were classified into seven clusters from 0 to 10 μm to >60 μm. The numbers in different clusters were divided by the number of all measured arbuscules from each slide to get the percentage of arbuscule size classes. Mean values of the respective percentages from three biological replica were used for the diagram shown in Figure 3B.

Constructs and Plasmids

For tissue-specific RAD1 gene expression analysis, a 1.9-kb promoter region was amplified by primers 5′-ACTgtagtcGCTTtCTTACTGCTTGC-3′ and 5′-AGGCGGgAACCtAAGATAGCtAAAC-3′ from transformation competent artificial chromosome clone TM2604, with lower case letters indicating suitable restriction sites. gtccg and cccggg are the Kpn1 and SmI1 enzyme digest sites, respectively. The PCR fragment was digested by Kpn1 and SmI1 and inserted into the binary vector pRedRoot-GUS to generate pRedRoot-RAD1pGUS. The 1.8-kb promoter region of RAM1 was amplified by primers 5′-ACTgtagtcGCTTtCTTACTGCTTGC-3′ and 5′-AGGCGGgAACCtAAGATAGCtAAAC-3′ to generate the pRedRoot-RAM1pGUS. For the overexpression assay of RAD1, the cDNA sequence of RAD1 was amplified by PCR. Then the PCR product with attB sequence ends was used for a Gateway recombination reaction to generate pH7WG2D-SSpro-RAD1. For subcellular localization of RAD1 protein in onion cells, the pGWbE-RAD1 plasmid encoding a nuclear N-terminal GFP fusion to RAD1 was generated by a Gateway recombination reaction and delivered into onion cells by particle bombardment. Cells expressing the 35S CaMV promoter-driven GFP gene were used as a control. Gateway vectors pSP/pACT were used for Y2H and Gateway vectors pIA-sYFPn-pBatTL-B and pIA-sYFPc-pBatTL-B (from Dr. Fransicka Turk, MPI) were used for BiFC assay. The primers are described in Supplemental Table S3.

RNA Transcript Profiling

RNaseq was performed as described previously (Willmann et al., 2013). In brief, >1 μg of molecular grade DNA-free RNA from each sample was used for transcriptome analysis. Library construction was performed according to the manufacturer’s protocol (Illumina, TruSeq). RNaseq was performed on Illumina HiSeq2000 at the Cologne Center for Genomics (http://portal.ccg.uni-koeln.de/cccc/). Paired-end reads of 100 bp each were generated, and six
samples were loaded onto one flow cell lane using multiplexing. Barcodes of sequences were removed, and the quality control of reads was done with the Galaxy FASTQ tool (Giardine et al., 2005; Blankenberg et al., 2010; Goecks et al., 2010). Short reads were mapped to the L. japonicus transcriptome (Satò et al., 2008) using Burrows-Wheeler Alignment tool with default parameters and an insert size of 200 (Li and Durbin, 2010). Hits per gene were counted using HTseq-count (developed by Simon Anders from EMBL Heidelberg). Statistical analysis of differentially expressed genes was performed with DESeq (Anders and Huber, 2010). Genes with fold change ≥ 2 and P value ≤ 0.05 were identified as being differently expressed. Gene annotation was performed using Blast2GO (Conesa et al., 2005).

Quantitative Real-Time PCR Analysis

Total RNA was extracted using the NucleoSpin RNA Plant extraction kit (Machery-Nagel) according to the manufacturer’s instructions. Total RNA of 0.5 to 1 μg was used to synthesize first strand cDNA using Thermo Scientific RevertAid H Minus Reverse Transcriptase. Gene expression was determined by qRT-PCR (ABI 7500) using the SYBR green PCR master mix (Applied Biosystems). Relative changes in gene expression from real-time quantitative PCR experiments were analyzed with the 2−ΔΔCt method, which was normalized to the endogenous reference Ubiqutin gene and relative to the reference sample (Livak and Schmittgen, 2001). Gifu-129 grown under low-Pi conditions was used as a reference sample without annotation. The gene-specific primers used in this article can be found in Supplemental Table S3.

Trypan Blue Staining and WGA-Alexa Fluor 488 Staining

For mycorrhizal structure visualisation, trypan blue staining was performed as described (Brundrett et al., 1984). For WGA-Alexa Fluor 488 staining, roots were first washed with distilled water and then soaked in 30% (v/v) ethanol overnight. Roots were then incubated in a 20% (w/v) KOH solution for 3 d at room temperature. After rinsing with water, the roots were incubated in 0.1 M HCl for 2 h. After rinsing with water and subsequently with 1× phosphate-buffered saline, roots were stained in 1× phosphate-buffered saline buffer containing 0.2 μg mL−1 WGA-Alexa Fluor 488 overnight in the dark. For long-term storage, roots were stored at 4°C in the staining solution. Zeiss confocal microscopy (LSM 700) was used to detect the WGA-Alexa Fluor 488 (excitation/emission maxima at approximately 495/519 nm) signal.

Sequence Collection and Phylogenetic Analysis

To collect sequences corresponding to potential RAD1 orthologs, RAD1 amino acid sequence was searched in publicly available genomes using BLASTp. The list of species and the sequences used are indicated in Supplemental Table S3. Hits with an e value of less than 10−50 were selected for the phylogenetic analysis. In addition, for nonhost species, the best hits were reciprocally blasted on the Medicago truncatula genome to confirm the absence of orthologs. Collected sequences were aligned using MAFFT (http://mafft.cbrc.jp/alignment/server/) and the alignment manually curated with Bioedit. Phylogenetic trees were generated by both the Neighbor-Joining and maximum-likelihood algorithms in MEGA5. Paratranscriptome collection and phylogenetic analysis of RAD1 were aligned using MAFFT (http://mafft.cbrc.jp/alignment/server/) and the alignment manually curated with Bioedit. Phylogenetic trees were generated by both the Neighbor-Joining and maximum-likelihood algorithms in MEGA5. Partial gap deletion (90%) was used together with the FFT substitution model (gamma 2 parameter rate). Bootstrap values were calculated using 100 replicates.

Reciprocal Blast Analysis

Reciprocal blasts were carried out using amino acid sequences of the 43 transcription factors (TFs) (Supplemental Table S1). First, each TF was used as a query in a BLASTp search on genomes of host and nonhost species (Supplemental Table S3). Best equal hits were collected for each species and used as a query in reciprocal BLASTp on the L. japonicus genome (http://www.kazusa.or.jp/lotus/). A reciprocal blast was considered positive if the initial TF was among the equal best hits.

Synteny Analysis

As L. japonicus is not available in Gevo (https://genomesevolution.org/CoGe/Gervo.ppl), we used M. truncatula. Given the extremely high identity between L. japonicus and M. truncatula genes, we assumed that M. truncatula was a good proxy. An approximately 200-kb-sized region in the M. truncatula genome containing transcription factors identified by RNAseq was compared to the syntonic region in Arabidopsis (Arabidopsis thaliana; Columbia-0) and Populus trichocarpa using CoGe/Gervo (https://genomesevolution.org/CoGe/ Gervo.ppl) as described in Delaux et al. (2014). For RAD1, in addition to the previously mentioned genome, we included rice (Oryza sativa), Vitis vinifera, Sorghum bicolor, maize (Zea mays), and Selaginella moellendorffii.

GST Pull-Down Assay

The pull-down assay was processed as previously described, with modifications (Cui et al., 2010). GST-RAM1, pGEX-LjNSP2, pGEX-RAM1, and pET-RAD1 were generated by Gateway recombination reactions, and GST was used as a negative control. In brief, 50 mL of bacterial cell culture expressing GST-tagged proteins was resuspended with 15 mL of lysis buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 3 mM dithiothreitol [DTT], and 1× complete mini-protease inhibitor cocktail [Roche]) after centrifugation at 5,000 rpm for 10 min. The sample was sonicated (2 min twice on ice at 40% power using Sonics vibra-cell ultrasonic processor VCS500) and centrifuged at 15,000 rpm for 20 min; then, the supernatant was incubated with 100 μL of glutathione agarose beads for 30 min with rotation at 4°C and washed three times with wash buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl, 3 mM DTT, 0.2% [v/v] Triton X-100). Then, the beads containing purified GST protein, GST-LjNSP2, GST- and RAM1 were incubated with soluble bacterial lysate (20 mg) containing His-RAD1 for 20 min at 4°C, respectively. After washing four times with wash buffer, the proteins were eluted with 100 μL of 150 mM Tris-HCl glutathione (25 mM Tris-HCl, pH 9.0). Western blotting using approximately 1% of input and 6.7% of eluted protein was performed for protein detection.

Co-IP and BiFC

The Co-IP assay was performed as described previously (Cui et al., 2010). For transient expression of proteins in Nicotiana benthamiana, leaves were infiltrated with mixed Agrobacterium tumfaciens GV3101 pMPP9 strains (final optical density at 600 nm of 0.5) harboring a vector that generates HA-RAD1 or GFP-RAM1 protein under the control of the CaMV 35S promoter, respectively, in combination with GV3101 strain expressing RNA silencing suppressor p19 protein. One gram of leaf tissue was harvested 3 d after infiltration and homogenized in 1.5 mL of IP buffer (50 mM Tris-Cl, pH 7.5, 150 mM NaCl, 5 mM EDTA, 2 mM DTT, 0.1% [v/v] Triton X-100, and 1× complete mini-protease inhibitor cocktail [Roche]). After two times of centrifugation at 20,000g for 10 min, the supernatants were incubated with 50 μL of HA agarose beads for 1 h at 4°C and then washed six times with IP buffer. The beads were eluted with 60 μL of SDS loading buffer (100 mM Tris-Cl, pH 6.8, 4% [v/v] SDS, 0.2% [w/v] bromophenol blue, 20% [v/v] glycerol, and 200 mM DTT), and then 20 μL of each elution was analyzed by western immunoblotting.

For the BiFC assay, leaves of N. benthamiana were infiltrated with A. tumefaciens GV3101 pMP90 strains harboring either pBatTL-RAD1-YFP or pBatTL-RAD1-YFP, in combination with pBatTL-RAM1-YFP, or empty vector, as described by Hu et al. (2014). Cells were kept in infiltration buffer (10 mM MgCl2, 10 mM MES, pH 5.6, 0.5 mM acetylstyringone) for 2 to 3 h and mixed with p19 before infiltration. Images were taken by Leica CLSM D500 S 3 d after infiltration.

Subcellular Protein Localization

pGWB6-RAD1 was used for subcellular localization in onion epidermal cells by particle bombardment, as described by Folkers et al. (2002). In addition, transgenic hairy roots containing pGWB6-RAD1, generated by inoculation of L. japonicus seedling stems with transformed A. rhizogenes, containing pGWB6-RAD1 from 1-month-old wild-type L. japonicus plants, were used for RAD1 subcellular localization. GFP fluorescence was detected via Zeiss LSM710 confocal microscopy.

Histochemical GUS Analysis

pRedRoot-RAD1proGUS was introduced into hairy roots of Gifu-129 using A. rhizogenes-mediated transformation. Transformed hairy roots were harvested 6 and 8 weeks after colonization by R. irregularis. Magneta GUS staining was used to detect GUS activity, and WGA-Alexa Fluor 488 staining was used to detect the presence of R. irregularis in root tissues. Fluorescence imaging was performed with a Zeiss fluorescent microscope (Microscope Axio Imager.D2).
Y2H Assay

Using pAS-RAD1 as bait, cDNA sequences of eight GRAS family members were amplified by PCR and subcloned into the pACT2 vector to be used as prey. Ten micrograms of each 1, 0.1, 0.01, and 0.001 optical density at 600 nm dilutions of yeast (Saccharomyces cerevisiae) cell suspension was dropped on synthetic minimal media plates based upon yeast nitrogen base supplemented with Cc and amino acids without Leu and Trp or without Leu, Trp, and His. An AP2 transcription factor encoded by the chr2.CM0608.1100.r2.m gene was used as a negative control. To determine the domains required for RAD1 and RAM1 interaction, the DNA fragments encoding full-length and truncations of RAD1 (121–509 amino acids), ANL (200–509 amino acids), AS (1–420 amino acids), and ΔPS (1-340 amino acids) were cloned into pAS bait vector by using the Gateway system. cGRAS, encoding mycorrhiza-induced GRAS gene chr3.CM0106.470.r2.d, was used as a negative control. One millimolar 3-amino-1,2,4-triazole was applied to suppress autoactivation.

Nodulation Assay

To generate nodulated roots, plants were inoculated with Mesorhizobium loti strain as described in Liu et al. (2003). One week after germination, seedlings were transferred to the sand and soil mixture as described previously. Ten days later, seedlings were inoculated with an M. loti cell suspension of 0.1 OD. Three weeks postinoculation, plants were harvested, and nodules were counted and examined by microscopy.

Sequence data from this article can be found in the Lotus database (http://www.kazusa.or.jp/lotus/) or GenBank under the following accession numbers: RAD1, chr4.CM1864.540.r2.m; RAM1, chr1.CM1852.30.r2.m; LjNSP2, chr1.CM1976.90.r2.m; STK, chr4.CM0042.2570.r2.d; RAM2, CM0905.50.r2.d and CM0905.160.r2.d; SbtM1, chr2.CM0221.2780.r2.m; LjPT4, chr1.CM2121.10.r2.a; cGRAS, chr5.CM0106.470.r2.d; and Ubiquitin, DQ49171.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. AM-inducible transcription factor genes repressed by high-Pi conditions.

Supplemental Figure S2. Mycorrhiza morphology in rad1-1 and RAD1 transcript analysis in rad1 LORE1 insertion lines.

Supplemental Figure S3. Mycorrhization rate and AM-inducible marker gene expression in rad1-2 and rad1-3 at 5 and 7 wpi.

Supplemental Figure S4. Mycorrhizal structures and RAD1 gene expression in transgenic hairy roots overexpressing chimeric 35Spro:RAD1.

Supplemental Figure S5. RAD1 gene expression profile in mycorrhizal and nonmycorrhizal L. japonicus and GRAS proteins that interact with RAD1.

Supplemental Figure S6. RAD1 gene expression in nodulated roots and nodulation in rad1-1 and rad1-3.

Supplemental Figure S7. Synteny analysis of RAD1 on AM host species and nonhost Arabidopsis.

Supplemental Figure S8. Nuclear localization of RAD1.

Supplemental Figure S9. Interaction among RAM1, RAD1, and LjNSP2 by BiFC assay in N. benthamiana.

Supplemental Figure S10. Gene structure of RAM1 in wild-type L. japonicus and ram1 mutant.

Supplemental Table S1: List of mycorrhiza up-regulated transcription factors generated by DESeq.

Supplemental Table S2: Forty-five mycorrhiza up-regulated transcription factors under different treatments.

Supplemental Table S3: Sequences of oligonucleotide primer pairs used.

Supplemental Text S1: Microsynteny analysis of potential symbiotic transcription factors.

Supplemental Text S2: Sequences used for phylogenetic tree.
signals from symbiotic microbes reveal MiMNPV and MiDOMD-dependent reprogramming of host gene expression by arbuscular mycorrhizal fungal lipochitooligosaccharides. Plant Physiol 159: 1671–1685.


nucleoporin is required for induction of Ca2+ spiking in legume nodule development and essential for rhizobial and fungal symbiosis. Proc Natl Acad Sci USA 103: 359–364


Supplemental Fig. S1. AM-inducible transcription factor genes repressed by high Pi condition. Gifu-129 plants grown at low Pi (5 μM) conditions were used as controls. After inoculation with *R. irregularis*, plants were grown at low Pi conditions for 8 weeks or 6 weeks at low Pi and re-supplied for two weeks with high Pi (7.5mM). Roots of four plants in one pot were harvested as one sample. Three technical replicates were used to calculate SD.
Supplemental Fig. S2. Mycorrhiza morphology in rad1-1 and RAD1 transcript analysis in rad1 LORE1a insertion lines.

(A) Arbuscules in Gifu-129 and rad1-1 in the presence of *R. irregularis* at 8wpi. eh, external hypha. a, arbuscule. da, degenerated arbuscule. Scale bars are 100 µm.

(B) RAD1 transcripts were detected in Gifu-129, rad1-1, rad1-2 and rad1-3 in the presence or absence of *R. irregularis* at 6wpi. 32 cycles were performed. Representative results from one out of three independent experiments are shown.
Supplemental Fig. S3. Mycorrhization rate and AM-inducible marker gene expression in rad1-2 and rad1-3 at 5 and 7 wpi. Three biological replicates were utilized. Error bars represent SD. Student’s t-test was used to calculate confidence level. * indicates $P<0.05$, ** indicates $P<0.01$. Similar results was obtained from another independent experiment.
Supplemental Fig. S4. Mycorrhizal structures and RAD1 gene expression in transgenic hairy roots overexpressing chimeric 35Spro:RAD1.

(A) Relative root length colonized (RLC) in wild type hairy roots transformed with the empty vector (EV) or with a construct containing 35Spro:RAD1 (oxRAD1). Percentage of RLC from transgenic hairy roots with EV (n=3) or oxRAD1 (n=8), respectively, was quantified. Relative RLC was calculated comparing the RLC from the transgenic red fluorescent hairy roots with the RLC from non-transgenic roots on the same composite plants. Significant differences of means between EV and oxRAD1 were determined by Student’s t-test (* indicates P<0.05).

(B) RAD1 gene expression is shown in transgenic red fluorescent hairy roots (+F) and non-transgenic non-fluorescent roots (-F) from three composite oxRAD1 plants and one control plant with EV in presence of R. irregularis. Error bars are from technical replicates.

(C) LjPT4 gene expression in transgenic fluorescent hairy roots (+F) and non-transgenic roots (-F) from three composite oxRAD1 plants and one control plant with EV in presence of R. irregularis. Error bars are from technical replicates.
Supplemental Fig. S5. RAD1 gene expression profile in mycorrhizal and non-mycorrhizal *Lotus* and GRAS proteins that interact with RAD1.

(A) RT-PCR analysis of RAD1 transcript levels in different tissues of wild type in the presence or absence of *R. irregularis*. Rt, root tip; R, root; Yl, young leaves; Mi, mature leaves. For Ubiquitin transcript amplification 28 PCR cycles, for RAD1 32 PCR cycles were used.

(B) Y2H analysis of interaction of RAD1 with mycorrhiza-regulated GRAS proteins. RAD1 was fused to the GAL4 DNA binding domain (BD). An AP2 transcription factor (chr2.CM0608.1100.r2.m) fused to the GAL4 activation domain (AD) was used as negative control. SD/-LW and SD/-LWH indicate a synthetic dropout medium lacking Leu and Trp and lacking Leu, Trp, and His, respectively.
Supplemental Fig. S6. *RAD1* gene expression in nodulated roots and nodulation in *rad1-1* and *rad1-3*. 

(A) Relative expression levels of *LjPT4, SbtM1, LjNIN* and *RAD1* genes detected by qRT-PCR in nodulated plant roots.

(B) The number of nodules per plant in *rad1-1, rad1-3* and Gifu-129 is given (n=6). Two experiments were performed independently with similar results. Error bars represent SD.
**Supplemental Fig. S7.** Synteny analysis of \textit{RAD1} on AM host species and non-host \textit{Arabidopsis thaliana}. Synteny analysis of a ~200 kb region encompassing the \textit{RAD1} locus in \textit{M. truncatula}. Green and blue boxes above and below the dashed lines represent \textit{Medicago} genes. Orthologous genes in other species are indicated above and below: \textit{Arabidopsis thaliana} (red), \textit{Populus trichocarpa} (orange), \textit{Vitis vinifera} (dark blue – \textit{RAD1}-like; dark-green \textit{RAD1}), \textit{Sorghum bicolor} (green), \textit{Oryza sativa} (pink), \textit{Zea mays} (light green) and \textit{Selaginella moellendorffii} (violet). The \textit{RAD1} locus is red-boxed.
Supplemental Fig. S8. Nuclear localization of RAD1.

(A) Nuclear localization of RAD1 in onion epidermal cells after biolistic bombardment with 35Spro:GFP or 35Spro:GFP:RAD1, respectively. Images with green fluorescence (GFP) alone and merged with bright-field image (merge) are shown as indicated. Scale bar is 100 µm.

(B) Nuclear localization of RAD1 in Lotus hairy roots expressing 35Spro:GFP or 35Spro:GFP:RAD1, respectively. Scale bar is 20 µm.
**Supplemental Fig. S9.** Interaction between RAM1, RAD1 and LjNSP2 by BiFC assay in *N. benthamiana*. Leaves were infiltrated with mixtures of *A. tumefaciens* strain GV3101 to coexpress RAD1 and RAM1 or LjNSP2 fused to half of split YFP, respectively as indicated at left. Three days after infiltration, images were captured from five different leaves for each plasmid combination. Scale bars are 50 µm.
Supplemental Fig. S10. Gene structure of RAM1 in wild type Lotus and ram1 mutant.

(A) Gene structure of RAM1 in Lotus and the positions of the LORE1a insertions in ram1-1 and ram1-2. Gray arrows indicate the primers matching with the genome sequence used for genotyping. Black arrows indicate P2R primer on the left border of the LORE1a transposon.

(B) Identification of homozygous ram1-1 and ram1-2 through genomic DNA amplification. Primer pairs as described in (A) were used for PCR and are indicated at left. Ubiquitin gene was used as a control.
**Supplemental Table S1:** List of mycorrhiza up-regulated transcription factors generated by DESeq.

<table>
<thead>
<tr>
<th>Seq. Name (LotusCDSv2.5)</th>
<th>Putative annotation of transcription factor</th>
<th>RNA-seq data</th>
<th>fold change (+R.i./-R.i.)</th>
<th>Padj</th>
<th>resVarA</th>
<th>resVarB</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr6.CM0679.550.r2.d</td>
<td>MADS-box</td>
<td></td>
<td>7.0e-04</td>
<td>6.09</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>chr6.CM1613.300.r2.m</td>
<td>C2H2 Zinc finger</td>
<td>M specific</td>
<td>5.1e-03</td>
<td>0.24</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>chr1.LjB18K24.70.r2.a</td>
<td>MYB</td>
<td>M specific</td>
<td>1.42e-102</td>
<td>17.63</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>chr3.CM0111.130.r2.d</td>
<td>NAC</td>
<td>M specific</td>
<td>5.10e-17</td>
<td>0.82</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>chr3.CM2163.240.r2.m</td>
<td>GRAS</td>
<td></td>
<td>970.83</td>
<td>6.73e-89</td>
<td>8.54</td>
<td>5.1e-04</td>
</tr>
<tr>
<td>chr2.CM0608.1100.r2.m</td>
<td>AP2</td>
<td></td>
<td>586.44</td>
<td>1.18e-06</td>
<td>12.14</td>
<td>1.3e-03</td>
</tr>
<tr>
<td>chr3.CM2163.230.r2.m</td>
<td>GRAS</td>
<td></td>
<td>462.42</td>
<td>3.11e-111</td>
<td>12.07</td>
<td>8.84e-06</td>
</tr>
<tr>
<td>chr1.CM1852.30.r2.m</td>
<td>GRAS</td>
<td></td>
<td>438.56</td>
<td>8.73e-109</td>
<td>2.33</td>
<td>9.84e-06</td>
</tr>
<tr>
<td>chr5.CM0239.240.r2.m</td>
<td>GRAS</td>
<td></td>
<td>170.03</td>
<td>1.89e-40</td>
<td>0.64</td>
<td>0.01</td>
</tr>
<tr>
<td>chr2.CM0021.530.r2.m</td>
<td>C2H2 Zinc finger</td>
<td></td>
<td>165.30</td>
<td>5.66e-65</td>
<td>0.18</td>
<td>6.40e-05</td>
</tr>
<tr>
<td>chr4.CM0075.50.r2.m</td>
<td>ERF</td>
<td></td>
<td>164.31</td>
<td>3.43e-23</td>
<td>4.27e-04</td>
<td>0.01</td>
</tr>
<tr>
<td>chr2.CM0081.1990.r2.m</td>
<td>AP2</td>
<td></td>
<td>150.70</td>
<td>5.7e-69</td>
<td>10.70</td>
<td>3.88e-04</td>
</tr>
<tr>
<td>chr1.LjB18K24.100.r2.a</td>
<td>MYB</td>
<td></td>
<td>120.88</td>
<td>1.56e-80</td>
<td>15.0</td>
<td>2.46e-03</td>
</tr>
<tr>
<td>chr6.CM0041.30.r2.a</td>
<td>AP2</td>
<td></td>
<td>78.55</td>
<td>4.72e-62</td>
<td>0.08</td>
<td>1.72e-04</td>
</tr>
<tr>
<td>LjB17L21.50.r2.d</td>
<td>MADS-box</td>
<td></td>
<td>44.10</td>
<td>7.51e-07</td>
<td>2.13e-03</td>
<td>0.09</td>
</tr>
<tr>
<td>chr1.CM0029.1590.r2.a</td>
<td>ARF</td>
<td></td>
<td>43.89</td>
<td>5.02e-48</td>
<td>0.1</td>
<td>8.24e-04</td>
</tr>
<tr>
<td>chr1.CM1413.480.r2.d</td>
<td>WRKY</td>
<td></td>
<td>38.12</td>
<td>5.42e-26</td>
<td>0.49</td>
<td>0.08</td>
</tr>
<tr>
<td>LjSGA_032180.1</td>
<td>C2H2 Zinc finger</td>
<td></td>
<td>24.65</td>
<td>2.11e-30</td>
<td>0.76</td>
<td>0.02</td>
</tr>
<tr>
<td>chr5.CM1667.220.r2.a</td>
<td>GRAS</td>
<td></td>
<td>19.68</td>
<td>6.69e-15</td>
<td>0.63</td>
<td>0.28</td>
</tr>
<tr>
<td>LjSGA_122341.1</td>
<td>GRAS</td>
<td></td>
<td>13.04</td>
<td>9.07e-19</td>
<td>1.46</td>
<td>0.14</td>
</tr>
<tr>
<td>chr6.CM0139.1440.r2.d</td>
<td>GRAS</td>
<td></td>
<td>11.74</td>
<td>4.35e-35</td>
<td>0.74</td>
<td>4.56e-05</td>
</tr>
<tr>
<td>chr2.CM1835.10.r2.m</td>
<td>NAC</td>
<td></td>
<td>8.96</td>
<td>3.65e-09</td>
<td>1.05</td>
<td>1.5e-03</td>
</tr>
<tr>
<td>LjSGA_020219.1</td>
<td>GRAS</td>
<td></td>
<td>8.93</td>
<td>5.03e-03</td>
<td>0.60</td>
<td>0.10</td>
</tr>
<tr>
<td>LjSGA_012098.1</td>
<td>AP2</td>
<td></td>
<td>6.80</td>
<td>3.26e-09</td>
<td>0.04</td>
<td>0.25</td>
</tr>
<tr>
<td>chr4.CM0680.320.r2.m</td>
<td>GRAS</td>
<td></td>
<td>6.02</td>
<td>9.55e-17</td>
<td>11.47</td>
<td>0.10</td>
</tr>
<tr>
<td>LjSGA_029723.1</td>
<td>AP2</td>
<td></td>
<td>6.00</td>
<td>8.98e-14</td>
<td>1.73</td>
<td>0.10</td>
</tr>
<tr>
<td>chr1.CM0012.840.r2.m</td>
<td>NAC</td>
<td></td>
<td>6.04</td>
<td>2.71e-23</td>
<td>4.12</td>
<td>0.70</td>
</tr>
<tr>
<td>chr3.CM0127.880.r2.m</td>
<td>GRAS</td>
<td></td>
<td>5.77</td>
<td>6.12e-13</td>
<td>6.14</td>
<td>0.01</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Low P+M</td>
<td>High P+M</td>
<td>shift</td>
<td>Low P</td>
<td>High P</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
<td>----------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>chr1.CM0122.2130.r2.d</td>
<td>4.81</td>
<td>2.83e-13</td>
<td>1.06</td>
<td>0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr3.CM0106.470.r2.d</td>
<td>3.57</td>
<td>1.50e-13</td>
<td>2.51</td>
<td>0.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr1.CM0375.530.r2.a</td>
<td>3.54</td>
<td>9.86e-12</td>
<td>1.44e-03</td>
<td>1.05e-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr2.CM0018.140.r2.m</td>
<td>3.20</td>
<td>8.07e-10</td>
<td>0.01</td>
<td>0.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr3.CM0106.760.r2.m</td>
<td>3.17</td>
<td>9.21e-04</td>
<td>0.01</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr4.CM1864.540.r2.m</td>
<td>2.93</td>
<td>2.01e-10</td>
<td>40.77</td>
<td>1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr5.T15C06.70.r2.m</td>
<td>2.51</td>
<td>4.23e-04</td>
<td>11.13</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr6.CM0018.140.r2.m</td>
<td>3.54</td>
<td>6.97e-06</td>
<td>0.21</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr1.CM0284.80.r2.d</td>
<td>2.31</td>
<td>1.60e-03</td>
<td>1.21</td>
<td>2.98e-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr6.CM0314.840.r2.d</td>
<td>2.30</td>
<td>3.35e-03</td>
<td>0.46</td>
<td>2.86e-03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr3.CM0176.150.r2.m</td>
<td>2.27</td>
<td>3.43e-03</td>
<td>0.40</td>
<td>5.15e-05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>chr3.CM0460.10.r2.d</td>
<td>2.11</td>
<td>2.85e-04</td>
<td>1.50</td>
<td>0.19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

"resVar" values indicate the degree of the variance among the biological samples. If resVarA or resVarB are <15, the fold changes are robust. “M specific” indicates mycorrhiza specific induction of gene expression.

**Supplemental Table S2:** Forty-five mycorrhiza up-regulated transcription factors under different treatments.

<table>
<thead>
<tr>
<th>Overall highly expressed genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene ID</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>chr1.CM0375.530.r2.a</td>
</tr>
<tr>
<td>chr3.CM0176.150.r2.m</td>
</tr>
<tr>
<td>chr1.CM0012.840.r2.m</td>
</tr>
<tr>
<td>chr5.CM0200.2670.r2.d</td>
</tr>
<tr>
<td>chr3.CM0460.10.r2.d</td>
</tr>
<tr>
<td>chr2.CM0018.140.r2.m</td>
</tr>
<tr>
<td>chr3.CM0106.470.r2.d</td>
</tr>
<tr>
<td>chr1.CM0122.2130.r2.d</td>
</tr>
<tr>
<td>chr6.CM0139.1440.r2.d</td>
</tr>
<tr>
<td>chr6.CM0314.840.r2.d</td>
</tr>
<tr>
<td>chr3.CM0106.760.r2.m</td>
</tr>
</tbody>
</table>
### Genes which were only responsive to AM under low P conditions

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Low P+M</th>
<th>High P+M</th>
<th>shift</th>
<th>Low P</th>
<th>High P</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr2.CM0021.530.r2.m</td>
<td>301.577336</td>
<td>0.3172006</td>
<td>3.7890705</td>
<td>1.8244533</td>
<td>1.2694607</td>
</tr>
<tr>
<td>LjSGA_032180.1</td>
<td>144.082462</td>
<td>6.8876778</td>
<td>9.0619715</td>
<td>5.8439403</td>
<td>7.6067857</td>
</tr>
<tr>
<td>chr2.CM1835.10.r2.m</td>
<td>44.693256</td>
<td>8.3700899</td>
<td>15.0244244</td>
<td>4.9887356</td>
<td>5.4743866</td>
</tr>
<tr>
<td>chr3.CM0111.130.r2.d</td>
<td>43.491953</td>
<td>0.5308103</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LjSGA_026563.1</td>
<td>144.082462</td>
<td>6.8876778</td>
<td>9.0619715</td>
<td>5.8439403</td>
<td>7.6067857</td>
</tr>
<tr>
<td>chr1.LjB18K24.100.r2.a</td>
<td>485.867106</td>
<td>3.812835</td>
<td>5.8745779</td>
<td>4.019487</td>
<td>2.9554226</td>
</tr>
<tr>
<td>chr2.CM1835.10.r2.m</td>
<td>44.693256</td>
<td>8.3700899</td>
<td>15.0244244</td>
<td>4.9887356</td>
<td>5.4743866</td>
</tr>
<tr>
<td>chr3.CM0111.130.r2.d</td>
<td>43.491953</td>
<td>0.5308103</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>chr1.CM1852.30.r2.m</td>
<td>145.41238</td>
<td>2.226832</td>
<td>0.8517816</td>
<td>0.8552047</td>
<td>0</td>
</tr>
<tr>
<td>chr6.CM1867.220.r2.a</td>
<td>56.110791</td>
<td>1.4824121</td>
<td>1.2337258</td>
<td>2.8507238</td>
<td>1.2694607</td>
</tr>
<tr>
<td>chr5.CM0239.240.r2.m</td>
<td>44.693256</td>
<td>8.3700899</td>
<td>15.0244244</td>
<td>4.9887356</td>
<td>5.4743866</td>
</tr>
<tr>
<td>chr3.CM1812.300.r2.m</td>
<td>6.53636</td>
<td>1.6960217</td>
<td>0.4258908</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>chr2.CM0127.880.r2.m</td>
<td>117.540714</td>
<td>12.0600506</td>
<td>19.4895084</td>
<td>20.3825447</td>
<td>12.183845</td>
</tr>
<tr>
<td>chr5.CM0239.240.r2.m</td>
<td>44.693256</td>
<td>8.3700899</td>
<td>15.0244244</td>
<td>4.9887356</td>
<td>5.4743866</td>
</tr>
<tr>
<td>chr3.CM0111.130.r2.d</td>
<td>43.491953</td>
<td>0.5308103</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>chr1.CM1852.30.r2.m</td>
<td>145.41238</td>
<td>2.226832</td>
<td>0.8517816</td>
<td>0.8552047</td>
<td>0</td>
</tr>
<tr>
<td>chr6.CM1677.220.r2.a</td>
<td>56.110791</td>
<td>1.4824121</td>
<td>1.2337258</td>
<td>2.8507238</td>
<td>1.2694607</td>
</tr>
<tr>
<td>chr5.CM0239.240.r2.m</td>
<td>44.693256</td>
<td>8.3700899</td>
<td>15.0244244</td>
<td>4.9887356</td>
<td>5.4743866</td>
</tr>
<tr>
<td>chr3.CM1812.300.r2.m</td>
<td>6.53636</td>
<td>1.6960217</td>
<td>0.4258908</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>chr2.CM0127.880.r2.m</td>
<td>117.540714</td>
<td>12.0600506</td>
<td>19.4895084</td>
<td>20.3825447</td>
<td>12.183845</td>
</tr>
<tr>
<td>chr5.CM0239.240.r2.m</td>
<td>44.693256</td>
<td>8.3700899</td>
<td>15.0244244</td>
<td>4.9887356</td>
<td>5.4743866</td>
</tr>
<tr>
<td>chr3.CM1812.300.r2.m</td>
<td>6.53636</td>
<td>1.6960217</td>
<td>0.4258908</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>chr2.CM0127.880.r2.m</td>
<td>117.540714</td>
<td>12.0600506</td>
<td>19.4895084</td>
<td>20.3825447</td>
<td>12.183845</td>
</tr>
</tbody>
</table>

### Low Pi induced genes regardless of AM status

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Low P+M</th>
<th>High P+M</th>
<th>shift</th>
<th>Low P</th>
<th>High P</th>
</tr>
</thead>
<tbody>
<tr>
<td>LjT15C06.70.r2.m</td>
<td>103.902922</td>
<td>2.9712519</td>
<td>5.1106894</td>
<td>41.3637576</td>
<td>6.3173675</td>
</tr>
<tr>
<td>chr4.CM1864.540.r2.m</td>
<td>1502.546805</td>
<td>18.216164</td>
<td>41.9738268</td>
<td>513.4417307</td>
<td>32.7066181</td>
</tr>
<tr>
<td>LjSGA_073109.1</td>
<td>475.682844</td>
<td>23.3235125</td>
<td>25.1443536</td>
<td>88.3434921</td>
<td>16.8496397</td>
</tr>
<tr>
<td>chr4.CM0680.320.r2.m</td>
<td>176.496151</td>
<td>5.295247</td>
<td>0.807835</td>
<td>29.3052966</td>
<td>3.3719237</td>
</tr>
<tr>
<td>LjSGA_029723.1</td>
<td>123.780115</td>
<td>11.8592965</td>
<td>14.5985516</td>
<td>20.6106325</td>
<td>10.0958136</td>
</tr>
<tr>
<td>chr6.CM0401.30.r2.a</td>
<td>320.224498</td>
<td>2.7512145</td>
<td>3.1923322</td>
<td>4.0765089</td>
<td>6.3173675</td>
</tr>
<tr>
<td>chr3.CM0075.50.r2.m</td>
<td>396.305693</td>
<td>3.5992253</td>
<td>4.6408521</td>
<td>2.2502557</td>
<td>2.9554226</td>
</tr>
<tr>
<td>chr4.CM0608.1100.r2.m</td>
<td>284.202937</td>
<td>0</td>
<td>0</td>
<td>0.4846234</td>
<td>0</td>
</tr>
<tr>
<td>chr4.CM0075.50.r2.m</td>
<td>70.258661</td>
<td>0</td>
<td>0</td>
<td>0.4267024</td>
<td>0.4264798</td>
</tr>
</tbody>
</table>

### Intermediates

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Low P+M</th>
<th>High P+M</th>
<th>shift</th>
<th>Low P</th>
<th>High P</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr6.CM1613.300.r2.m</td>
<td>6.53636</td>
<td>1.6960217</td>
<td>0.4258908</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LjSGA_055804.0.1</td>
<td>44.863102</td>
<td>16.0993508</td>
<td>16.0823821</td>
<td>14.168024</td>
<td>15.1736565</td>
</tr>
</tbody>
</table>
Normalized counts per transcript from different treatments are shown above.

**Supplemental Table S3:** Sequences of oligonucleotide primer pairs used.

Primers for RT-PCR:

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Primer name and primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lj<em>NIN</em></td>
<td>qLjNIN-F: TGGATCAGCTAGCATGGAAT</td>
</tr>
<tr>
<td></td>
<td>qLjNIN-R: TCTGCTTCTGCTGTGTCAC</td>
</tr>
<tr>
<td>SbtM1</td>
<td>qSbtM1-F: TGTATGCTGCTGCTGAAAAAAACAAT</td>
</tr>
<tr>
<td></td>
<td>qSbtM1-R: CTTCTTGACCTTTTGCAATAATGGAATT</td>
</tr>
<tr>
<td>STR</td>
<td>qSTR-F: CTGGACAGATCGACCGTCTCT</td>
</tr>
<tr>
<td></td>
<td>qSTR-R: GTGGCCATCAAGCTGGTATT</td>
</tr>
<tr>
<td>RAD1</td>
<td>q-F1: CCGAGGCTCATGCTAGGTCCACT</td>
</tr>
<tr>
<td></td>
<td>q-R1: CCCCAATGGGTTTCATGCTATCC</td>
</tr>
<tr>
<td></td>
<td>q-F2: ATGGTGAGCAAGGATCAGAG</td>
</tr>
<tr>
<td></td>
<td>q-R2: TTAACATTTCCAGCAAGAAG</td>
</tr>
<tr>
<td>RAM1</td>
<td>qRAM1-F: GGAGGTTTCTTGGAGCACTG</td>
</tr>
<tr>
<td></td>
<td>qRAM1-R: CTTCCATGATCTTCCTCCA</td>
</tr>
<tr>
<td>RAM2</td>
<td>qRAM2-F: GGGATGGACCCGTTTTACTT</td>
</tr>
<tr>
<td></td>
<td>qRAM2-R: GACTTTGGTTGGCATGCGCTTCA</td>
</tr>
<tr>
<td>Lj<em>PT4</em></td>
<td>qLjPT4-F: TCCAAGCGGAGCAAGACAGA</td>
</tr>
<tr>
<td></td>
<td>qLjPT4-R: TCTGCTGAGGTTCTGCTGAG</td>
</tr>
<tr>
<td>Ubiquitin</td>
<td>qUbiquitin-F: TACCGCTTGCTCCGTCCTT</td>
</tr>
<tr>
<td></td>
<td>qUbiquitin-R: AACAACAGCAGCACACGACAATCC</td>
</tr>
<tr>
<td>R.<em>irregularis</em> large subunit rRNA (RiLSU*)</td>
<td>LR1: GCTATCAATAAGCGGAGGA</td>
</tr>
<tr>
<td></td>
<td>8.22: AACTCCCTCAGCCTCCAGA</td>
</tr>
<tr>
<td>chr3.CM0106.760.r2.m</td>
<td>q30-F: AGGATGCTGCAAGAGCATT</td>
</tr>
<tr>
<td></td>
<td>q30-R: CTCCCTGACAGAGTCAGGC</td>
</tr>
<tr>
<td>chr3.CM0127.880.r2.m</td>
<td>q21-F: GAAGGCTAGGGCAAGAAAAA</td>
</tr>
<tr>
<td></td>
<td>q21-R: TCCAAAGTGCTCACACAGC</td>
</tr>
<tr>
<td>Chromosome</td>
<td>Gene</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
</tr>
<tr>
<td>chr2.CM0608.1100.r2.m</td>
<td>q3-F: AGAGGAGTAGCAAGGCACCA</td>
</tr>
<tr>
<td>chr3.CM2163.230.r2.m</td>
<td>q-8-F: TGCTGTGAAGACAGGTGAA</td>
</tr>
<tr>
<td>chr3.CM2163.240.r2.m</td>
<td>q-5-F: TCTCTACCAAAGTGGAAGG</td>
</tr>
<tr>
<td>chr2.CM1835.10.r2.m</td>
<td>q17F: GCAACTGATTCTTGGGCAAT</td>
</tr>
<tr>
<td>chr3.CM0106.470.r2.d</td>
<td>q26-F: TCCTCACCAGAATCTAC</td>
</tr>
</tbody>
</table>

Primers for genotyping:

<table>
<thead>
<tr>
<th>LORE1a mutant</th>
<th>primer name and primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>rad1-1</td>
<td>1039F: GGAATGCTATTCAACTCAACATC</td>
</tr>
<tr>
<td></td>
<td>1039R: GAACGGAGAATCAACACATC</td>
</tr>
<tr>
<td>rad1-2</td>
<td>576F: CCGAGGCTCATGCTAGTCCACT</td>
</tr>
<tr>
<td></td>
<td>576R: CCCAAATGGGTTCATGCTATCC</td>
</tr>
<tr>
<td>rad1-3</td>
<td>2260F: AGTCGTTGGCCTTTTGAATTCCTG</td>
</tr>
<tr>
<td></td>
<td>2260R: TGGATGGTTTGAATCACCAGGGC</td>
</tr>
<tr>
<td>ram1-1</td>
<td>2740F: GCTTGTTGCTGAAGCATGGAACAA</td>
</tr>
<tr>
<td></td>
<td>2740R: TCCGGAGGAGTGTGCTAATTCG</td>
</tr>
<tr>
<td>ram1-2</td>
<td>82472F: TCAATTGGCTGGCAGGGAAGGTTC</td>
</tr>
<tr>
<td></td>
<td>82472R: TGCAGTCCAAAGAAAAAGCCTCA</td>
</tr>
<tr>
<td>Ubiquitin</td>
<td>GS-UB-F: CGTGAAGGCTCAAGGATAAG</td>
</tr>
<tr>
<td></td>
<td>GS-UB-R: CGATATCTACTTTTCAAGAGG</td>
</tr>
<tr>
<td></td>
<td>P2R: CCATAGGGCTTCGGTAATCTTAGG</td>
</tr>
</tbody>
</table>

Note: Endogenous fragment: F+R, LORE1a insertion fragment: F+P2R

Primers for Y2H:

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>primer name and primer sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>pAS-RAD1 /</td>
<td>RAD1-attB1:</td>
</tr>
<tr>
<td>pACT-RAD1</td>
<td>GGGGACAAGTTTGTACACAAAAAGCAGGCTCTATGTCCCTCTCTTTATAGTG</td>
</tr>
<tr>
<td></td>
<td>RAD1-attB2:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCCTTTGTCAAGAAAGCAGGTGGGTCTAATCATTTCCAGGAAGCTG</td>
</tr>
<tr>
<td>pACT-RAM1</td>
<td>RAM1-attB1:</td>
</tr>
<tr>
<td></td>
<td>GGGGACAAGTTTGTACAAAAAAAGCAGGCTCTATGATCGATGGGTCAATCATTTCCAGGAAGCTG</td>
</tr>
<tr>
<td></td>
<td>RAM1-attB2:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCCTTTGTCAAGAAAGCAGGTGGGTCTAATCATTTCCAGGAAGCTG</td>
</tr>
<tr>
<td>pACT-AP2</td>
<td>AP2-attB1:</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>pACT-</td>
<td>Chr3.CM2163.230.r2.m-attB1:</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>GGGGACCAAGTTTGTACAAAAAAGCAGGCTCTATGGAGGAGCTGAGATATGAC</td>
</tr>
<tr>
<td></td>
<td>Chr3.CM2163.230.r2.m-attB2:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCACCTTTGTACAAGAAAGCTGGGTCTCATTTCCTGAAACTCCATGCTG</td>
</tr>
<tr>
<td>pACT-</td>
<td>Chr3.CM0127.880.r2.m-attB1:</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>GGGGACCAAGTTTGTACAAAAAAGCAGGCTCTATGGAAAGACAAGGGCTTAAAAC</td>
</tr>
<tr>
<td></td>
<td>Chr3.CM0127.880.r2.m-attB2:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCACCTTTGTACAAGAAAGCTGGGTCTTAAAACCTCCAGGCTGATATAG</td>
</tr>
<tr>
<td>pACT-</td>
<td>LjT15C06.70.r2.m-attB1:</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>GGGGACCAAGTTTGTACAAAAAAGCAGGCTCTATGGACATTACTCCTTATAC</td>
</tr>
<tr>
<td></td>
<td>LjT15C06.70.r2.m-attB2:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCACCTTTGTACAAGAAAGCTGGGTCTCAAGTACGCTAGCTGGGCT</td>
</tr>
<tr>
<td>pACT-LjNSP2</td>
<td>LjNSP2-attB1:</td>
</tr>
<tr>
<td>----------</td>
<td>------------------</td>
</tr>
<tr>
<td></td>
<td>GGGGACCAAGTTTGTACAAAAAAGCAGGCTCTATGGAAATGGATATAGATTTG</td>
</tr>
<tr>
<td></td>
<td>LjNSP2-attB2:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCACCTTTGTACAAGAAAGCTGGGTCTATGCACTTCGGAAATCAGTATTCTG</td>
</tr>
<tr>
<td>pAS- RAD1ΔN</td>
<td>RAD1-121-attB1:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCAAGTTTGTACAAAAAAGCAGGCTCTATGGAGGAGCTGAGATATGAC</td>
</tr>
<tr>
<td></td>
<td>RAD1-attB2:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCACCTTTGTACAAGAAAGCTGGGTCTCATTACATTTCAAGAAAGCTG</td>
</tr>
<tr>
<td>pAS- RAD1ΔPS</td>
<td>RAD1-attB1:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCAAGTTTGTACAAAAAAGCAGGCTCTATGGACATGGATATAGATTG</td>
</tr>
<tr>
<td></td>
<td>LjNSP2-attB2:</td>
</tr>
<tr>
<td></td>
<td>GGGGACCACCTTTGTACAAGAAAGCTGGGTCTATGCACTTCGGAAATCAGTATTCTG</td>
</tr>
</tbody>
</table>
| pAS- RAD1 ΔS | RAD1-attB1:  
GGGGACACAGTTTGTACAAAAAAGCAGGCTCTCTATGTG 
CCTCCTCCTTATAGTG  
RAD1-420-attB2:  
GGGGACACACCTTTGTACAAAAAAGCAGGCTCTCATACCTTC 
TTTGCCCTCCTTATAGTG |
|-------------|-----------------------------------------------|
| RAD1-200-attB1:  
GGGGACACAGTTTGTACAAAAAAGCAGGCTCTatgTCCA 
TGATGAACATCATG  
RAD1-attB2:  
GGGGACACACTTCTTGTTCAAGAAAGGCTGCTTAAACAT 
TTCCAGCAAGAAAGCTG |

**Primers for BiFC:**

| pBatTL-RAD1-YFPn/  
pBatTL-RAD1-YFPc | RAD1-attB1:  
GGGGACACAGTTTGTACAAAAAAGCAGGCTCTCTATG 
CCTCCTCCTTATAGTG  
RAD1withoutTGA-attB2:  
GGGGACACACTTCTTGTTCAAGAAAGGCTGCTGacatttc 
cagcaagaagctg |
|-------------------|-----------------------------------------------|
| pBatTL-RAM1-YFPn/  
pBatTL-RAM1-YFPc | RAM1-attB1:  
GGGGACACAGTTTGTACAAAAAAGCAGGCTCTAT 
GATCAATTCAATGTGCTGG  
RAM1withoutTGA-attB2:  
GGGGACACACTTCTTGTTCAAGAAAGGCTGCTGCGCA 
TCTCCATGCAGAGG |
| pBatTL-LjNSP2-YFPn | LjNSP2-attB1:  
GGGGACACAGTTTGTACAAAAAAGCAGGCTCTAT 
GAAATGGATATAGATTG  
LjNSP2withoutTGA-attB2:  
GGGGACACACTTCTTGTTCAAGAAAGGCTGCTGCA 
ACAATCTGATTCTG |
**Supplemental Text S1.** Synteny analysis of potential symbiotic transcription factors. A ~200kb sized region in the *Medicago truncatula* genome containing transcription factors identified by RNAseq was compared to the syntenic region in *Arabidopsis thaliana* (col-0, pink) and *Populus trichocarpa* (green) using CoGe:GEvo ([https://genomewolution.org/CoGe/GEvo.pl](https://genomewolution.org/CoGe/GEvo.pl)) as described in Delaux et al. 2014. *Medicago* was used as a proxy for *Lotus* given that this tool is not available for *Lotus*. The target gene in *Medicago* is red-boxed together with orthologs from *Populus* and *Arabidopsis* when present. The vertical orange stripes are regions that contain unknown nucleotides that were inserted during the scaffolding of assembled contigs. However, the position and relative size of the gap (orange stripe) is known due to the size of the genomic library and/or a genomic map.

**Examples of genes conserved in host and non-host species**

**chr2.CM0021.530.r2.m**

![Image of chr2.CM0021.530.r2.m](image)

Strong syntenic support for both *Arabidopsis* and *Populus*.

**chr3.CM0106.760.r2.m**

![Image of chr3.CM0106.760.r2.m](image)

Strong syntenic support for both *Arabidopsis* and *Populus*.

**chr3.CM0460.10.r2.d**
Strong syntenic support for *Populus* but low support for *Arabidopsis*.

**Genes missing in in non-host species**

*Present in Lupinus (potential common symbioses transcription factors)*

LjSGA_045015.1

chr5.CM1667.220.r2.a

chr3.CM0106.470.r2.d

chr4.CM0075.50.r2.m
Absent in Lupinus (potential AMS specific transcription factors)

chr1.CM0012.840.r2.m

chr1.CM0284.80.r2.d

chr1.CM1852.30.r2.m
Low syntenic support even for *Populus*.
Supplemental Text S2: Sequences used for phylogenetic tree.

*Aquilegia coerulea*

> Aco Aqua_123_00013

MISDTFHDEKIEGVRGLDLSLSAMAFYPFPYLSTLENSVSTWFLPSTDTRN
KRLKPTLSDELIGNSPYSGSNNSLFSXGSCITANSTNLSLPLKLQR
HIWTTQRFLLAEAVEEAAAAAMIGEDGERVDEGSIDMGRLVQLLVSCAEAV
ACRDRSHASTLLSRLANLVFGTSFQRVASCFFVQGLDRLALVQPLGAVGV
VAPTLMFAMCSKKEELRLVYDVCQPHIQFHVFVANTSILEAFEGS/LVHV
VDLGMTPMLPGHQPWRRIIHLANRAGQPPRRLRITGVGNSGDQLQTIGDEL
EAYAQLGINFDILQLHCVVKESRGALNISLQPQHELSPKVLVELQDSSHNGP
FFLGRFMEALHYYSAIFDSLDAMLPKYDTRRAKIEQFYFAEIEIKNIVSCEGPA
**Arabidopsis thaliana**

>AtSCL1

MVEQTQVREHIIKARVMSDLRSAPSSYRNPKLTYLNLNENGNNGVSSAQIFDP
DRSKNPCLTFDSYSQSYEKFYLDSPTDEFQVQHIPGSGAVSFFSGLDSEFYPQ
RPVLCQSMEFQFLPLDSTSTSSTRLLGDYQAVSYSVSMVDVEFQDEQMRSKIQ
ELERALLLDGDDK MGIDNMLEDSEWSYQNESEQHDSPEK SSSADSNSHV
SSKEVVSQATPKQILSCARALESEGKLEEALESVMNLRQIVS IQGDPSQRIAAY
MVEGLAARMAASGKFIYRALKCPEPDSDERLAAMQVLFEVCPCFKFGLFAA
NGAILEAIGGEVEHIIDFNDNQYNQLMTLSRAELPGKPRLLRTIGEDPSV
QRSIGGRLRIQLRLEQLAEDNGVSDKFKKAMPSKTISVIPSTLGCKPGETLIVNFA
FQLHMPDESVTQNRDELLHMVKSLPQKLTVQVQDVQNTNTSPFPRFIE
AYEYYSAFESLDMTPLPESQERMNVERQLCQARDIVNIAEGERIYEEA
AGKWRARMMMAFPFKPMASAKVTNNIQNLKKYKLEEMGEHLFCH
WEEKSLIASWR

>AtGAI

MKRDHIIIHIIHIHQDKKTMMMNEDDNGMDELLAVLGYKRVRESEMAVADQA
KLEQLEVMSNVQEDDLSQLATEVHYHPAELYTDLDSMLTDNPPSSNAE
YDLKAIPGDAILNQFAIDSASSSSNQGGGDDTDYNKRLKCSNGVVEETTATAE
STRHVLVDSQENGVRLHALLACAEAVQKENLTVALVKQIGFLAVSQQIG
AMRKVATYFAEALRSLPSQPIDHLSLSDTLQMHHYETCPYLFKAHFTA
NQAILEAFQGKKRNVHVIDFMSQGQLQPALMQALARPGGGPFLRTGIGPP
APNDFYLVHEVGLHAMLAEAHVEFEGFVANLADLDASMLELPSEIES
VAVNSVFLHKLKLLGRPGAIDKLVGLYVMQIKPEIFTVVEQESNHISPIFDFRFE
SLHYYSTLDSLEGPSVQGAVLQGICNCSNVACDGPDVFETHLSEQ
WRNRFGSAGFAAHIGSNQDILMKLQGFGYRVEESDGLMLWGHT
TRPLIASTAWKLNSTN

>AtSCL3

MVAMFQEDNSSVASSPLQVFSTMSLNRPTTALLSSSFHCLKDLKPEERGLY
LIHLLLTCANHVISGLSQNANAAEEQLSHLAPDGDTMQRIAAYFTEALANRI
LKSWPGYLYLNAQTFRTNVSEEEHVRRLFEMIPFLKVSYTTNRAILEAM
EGEKMVHVIDLDASEPAQWLLALLQAFSNRPEGPHPLRTGVJHQQEVEQOM
AHRLEIIEAEKLDPFQNPVVSRLDCLNEQLVRKVGTGALVSVSLQLHFTLA
SDDDLMRKNCALRFQQNSGVGLQRVLMSHSGAAARENDMSNNNGYSP
SGDASSSLPLPSSGRDTSFSNLAIWGLPSKVMVTVQDSDHNGLMERLES
LYTAYAALFDICLTVQRTSQDRIKVEKMLFEEIKNIIGSFERRHEKLEK
WSQFLDAGNGNVPILSYAMLQARLQGCGFQGFYRIKESGGCAVCWQDR
PLYSVSAWCRK

>AtSCL4

MAYMCTDSGNLMAIAQQVQKQKQEQQEQQQQQHQHDQINOIFGINPLSLNWPWP
TSGLFGSLGASFDPDFQVTVGTGDSNDPGFPPNLDHHAATTGGGFRLSDFFG
GTGGGEFEDFDEMELTISGGDSVADGPDCTHWDNDPVYVPPDFDFTYPNSR
LSVQPSDNRVIDTSSPLPPTLWPPSSPLSIPPLTHESPTKEDPETNDSDEDDDF
>AtSCL8
MESGFSGGGGSDFYGGGGRSIPPPGPTVINVGNNPQTTYRNQIPG1FFDQ
IGNRVAHGFSGKRTLADFQAQQHQQQQQQQQQFYSQAALNAFLSRSVKP
RNYQNFQSPSPIMDLTSVNDMLFGGGSGQRYGLPVPRSRQTTQQQSDYGLF
GGIRMIGGSGINQYPTLTVPCIFPVQRVHSENEMLSNL'RELQKLLDDDDE
SGGDDDVSTINNSNDWIQLNVLTPNPNNPPLVSFPSSSSSSSSPSTTSTTSVC
SRQTVMEIATIAEGKTEATEILARVSQPNEKEVLDFVMAALRSGRA
SPVTELYGKEHLISTQQLYELSPCFKLMGEAAALNAFLAADNNDDGMMIPHI
DFDINEGGQYVNLNLRTLLLRRNGKSSQNSQPVVKITAVANVYGLVDDGGE
ERTKAVGDLSLQGDRLGISVSNFVTVSRLGLDNRELSGCPDTELAVNLAF
KLYRVPDESVCTENPRDELRVKKRPGTVTLVEQEMSNTAPFLGRVSES
CACYGALLESVESTPVSTNSDRAKVEEGIRKLGVNAVACEGDIERCEVFGK
WRRMRMMSAGFELMPLSEKIAEMKSRGNVRHGFVTKEDEENNGVCFGMG
ALTVASAWR

>AtSHR
MDTLFLRLVSLQQQSDSIIITNQLSSRSRTTTTGPQATAYHYNFPQONDVVEE
CFNFIMDDEDLSSSSSSHNNHHHHNNPNTYYPSTPTTPYHPATSTSPSTAAM
AALASPYSSSGIGHNDPASFIPPPTPSFSANAKWASVLEAAARAFSDKDT
ARAPQILDNLNESSPYGDEEQKLYLSFLAQLFMRNTSGLGRCYRTMTAAM
ATEKTCSTESTRTKVLKQEVSPWATFGHVAANGAILEAVDGEAKIHIHVDISS
TFCTQWPTTLEALTRSDDTPHRLTTVTVANKFNDQTASHRMKIDGEIGNR
MEKFARLMGVPKFNNISSHVGDLSEFIDNLNSDLPKVEVLAINCGAMGHIS
RGSPRDAMISFSRGHRPRIVTVVEEADLVGEEEGGFDEFLRGFECRLWFR
VCFSWEESPFRTSNERLMLERAGRAIVDLVACEPSDTERKWSRR
MRNSGFGAEGYSDEVADDVRAALRRYKEGVWSMVQCPDGAIGIFLCWDRQ
PPVWASAWRPT

>AtRGA
MKRDHHFQFQGRLSNHGTSSSSSSSISKDMMMVKEEDGGGNNMDDELLAVL
GYKVRSSMAEVALKLEQLETMMSVQEDGLSLHATDVTYHNPSELYSLWD
NMLSLNPPPLPASSNGLDPLVPSIECGFAPSDYDLKWIPGNAYQFPADDSS
SSNNQNKRKSCSSPDMVSMTSTSTGTQIGGIVTTTVTTTTTTTTAAGESTRSVI
LVDSQENGVRVLHALMACAEIQQNNLTLEALVIIKQICLAVSQAAMGRK
VATYFAEALARRYLSQPNQIDHCLSDLTMHYETCPYLKFAHTANQAI
LEEFEGGKRVHIIDFSMNQQLQWPALMQALAREGGPPTFRTLTTGIAPPDN
SDHLHEVGCQLAQLEAIHFEVEYRGFVANSLADLSMELRPSDTEAVAV
NSVFELHKLLRGPIEGEIKLGVQVKKQIKPVIFTVVEQSNHNGPVFLDRTESL
HYYSTLFDSELVPNSQDKVMSSEVLKQICNVLACEPGDRVERHELTQSW
GNRFHSSGALPAHLGNSNAFKQASMLLSVFNSQGYRVEESNGCLMLGWHTR
PLITTSAWKLSTAAY

>AtSCL11
MDALLQVSDGFRFENGSCCKPRNNLESGNNLFDPDFHESQNNQSSPNDSPPT
VCLDNSPVLYRDIMLDMDEDFVGISRDALQAAERSFYIIEIIQQQSPESDQN
TSSSSDQNSGDQDFCPSSTTDSSALVSSGESQKYHRNDEEDELNRNRNK
QPAIFVSEEMELAVKLEHVLVVCNTQNEEEEERTVITKQSTPNRAGRAKGSSN
EFYTMVLESLDAAPPGLVKIKKEIAEFVLRPKISAIAVETAADRRHTGMETWR
EAFCAAGMRPIQSLQFADFQAECLLEKAQVRGFHVAKRQGELVLCHGRAL
VATSARWF
>AtSCL16
MQIPTLDSMANKLHKPPPLKLTVIASDAEFHPPLPLGISEDYELGKLVNFA
TTTNVAMEFRISSSYDGLSSLIEQLRIPDI PFNEALVNNCHMLHYIPDEILT
SNLRSVFKELRDLNPITVTILEDSDFTSTNFISRRLSLYNYMWIPYTAEMF
LTRGSEQRWQYEADISWKIDNVVAKEGAERVERLEPKSR
>AtRGL1
MKREHNBRESSAGEGGSSSMTTIVKEEAAAGVDELVLVGYKVRSSDMADVA
HKLEQLEMLVGLDGNLSDETVHYNPSDLGWSVEMLSDLDDPRIQEKPSDE
YDLRAIPGSAYPRDEHVTRSSKRTRIESELSSSTRSVVVLDSQETGVRLVHAL
LACAEAVQQNLKLDALVKHVGLLASSQAGAMKVFAYFAEGLARRYRI
YPRDDDVALSSFSDTLQIHFYESCYPYLKFAHFTANQAILEFATAEKVHVIDGL
NGHLQWPALIQALARPNPGPPFRTLGIGSYTDIEQVVGWKLQQLASTIGVF
EFKSIALNLNSDLKPEMLDDGRSEVAVSNVFELHRLLAHPSIDKFLSTIKSI
RPDIMTVVEQEANHNGTFLDRFTELSHYYSSLFDSEGPPSQRVMSELFLG
RQILNYACEGEDRVERHELNLQNWRNFLGKFPSIGSNAYKQASMLLLAR
YAGADGYNEEGLGWWQTRPLIATSARWNRV
>AtSCL18
MLTSFKSSSSSEDATAATTENPPPLCIASSSAATSASHHRLRLFFTAANFVSQS
NFTAAQNLLSLSNLSPPHDGSTERLVLHFTKALSVRIRNNQQDQTATENWAT
TTNEMTMSNSTVTSSVCKEQFLFRTKNNNSDFECCYYLWNKLQTPFIRFGLH
TANQAILDATEETNDGALHILDLDISQGLQWPLMQALAERSNSPSSPPSLRI
TGCGRDVTGLNRTGDRLTRFADSLQLQFQFHHTLIVEEDLAGLLLQRLLALS
AVQGETIAVNCVHFLHKIFNDDGDGMIHGFLSAIKSLSRIVTMAEREANHGD
HSFLNRFSEAVDHMAIFDSLEATLPNSRELTLEQRWFGEKILDVAAEET
ERKQRHRREFIWWMKRFGFVNVPSGFSALSQAQLLRLHLHYPSEYNLQFL
NNSLFGLGWQRNLPSW
>AtRGL2
MKRGYGETWDDPPPPLASPASRSSEGPGSMDKKKADDNNSNMDDELLAV
GYKVRSSMAEVAQKLEQLEMVLSNDVSTVLDVSNVHYNPSDLNSNWES
MLSELNNPASSDSLDTTRSCYDREYDLRAIPGLSAFPKEEEVFDEEASSKRIRL
GSWCESSEDESTRSVVLDSQETGVRLVHALVACAEIHQENWNLDALVVRK
VGLAGSQAGAMKVFAYFAQLARRIRYDYTDAETDVCAAVNSFEEVLEM
HYESCYPYLKFAHFTANQAILEAVTARRHVIDGLQNQMGWQPALMQALA
LRPGPPSFRLTGIPQTPENSDSLQLQQLWLQAQFAQNMGEFEKFGLAEL
SDLPEMFETREPESELVSNVFELHRLLARSGSLDKNLNTVKAIPSIVTVVE
QEANHNGIFVLDFNEALHYYSSLFDSEGPSQVRMSVYLGQQLNV
VAAESGDRVERHETAAQWRIRMKSAQGDFPIHLSAGFQASMLLSLTYATGD
GYRVEENDGCLMIGWQTRPZITSSAWKLA
>AtSCR
MAESGDFAQGQPPPHSLRTTSSGSSSSNRGPPPPPPLLVMVRFKRLASEMSS
NPDDNNSSRPPRVSHELDSNYNTVTPQQPSLTTAATVSSQPNPPSLVCGFS
AtSCL27

MPLSFERFQGEGVFGLSSSSSFYSQDIWSQDKTEAKQEDLGYVGGFLPEP
TSVLDAFLRSPSLASYSTTTTLSHSSHGGGTTVNTTVTAGDNNNNKCSQ
MLDLDLGDGVLSASSPGQEQSLRLIMDPSGAFVFDFPGFSGSGPVPASPVS
DNSLNLCCNFQFEITNPACLEINPSNHCLFYNPLSPPAPKRFNSGLHQPVFFPS
DPDGPHERPVRRHQFQFPFYHNNQQQQQFPPSSSSTAVAMPVPSPMAGDDD
QSVIIQFLFNSAEELIGTTGNNGHDHTVLAQGILARLNNHHLNTSSNKHSPFQRA
ASHIAEALLLSLHNSPSSHPLETTPNLIRIAAYRFSSTESPFLQFVNFNTANQILES
CNESGFDRIIIIDFDDVYGGQWSSLMQELASGVMGRRRRNASLKLTLTVFAPPP
STVSEDEFELFTEENKTFAGEVKFPEIELLSVELENNPAYWPLSLRSSEKEAI
AVNLPSVMVASGYLPLRLFKQQLSPINIVCSDRGCDRNDAPFPNANVHSLOY
HTSLLLELDANQNQDSSIERFVWVQPSIEKLMMKHRWIERSPWIRLFTQCG
FSPASLSQMAEAQAECLLQRNPRVGFGHEKQRQSSLMVCMQVRKELTVSAWK
C

AtSCL28

MLAGCSSSSLLSPTRRLEASEVAATASATVSAHFMQTQLDLPCCSSSSFSPKETP
SSRPLGRSISSLNSNNSNNPKIERRKTTSGCGLKQNIKLPLLATTTRNGEFSW
NNDNRRGKSLRKLALAEEDCSLRAKRTCKCNEEGGFWFHEFTGQDSSSPALP
FSLTCSGDDEEKFVEQISPLPNVDSVITELAGIGKDVSLQPPAAYKE
AGSGBSTASSESSHRLQHEPNTNSMRTYSHRGATEERTGNNPNNNNND
LQRFELVNLTLGCLDAIRNSIAINHFIARTGDLASRGPRGRTPMLRATYIE
ALALRVARMPWPHIFHIAPPREFDRTVEDESIGNARFLNQVTPPIKFIHFTANE
MLLRAFEGKERVHIIIDFIDKQGLQWPSFQLASRINPPHHVIRTIGESKLEL
ETGDRLHFANEAMLQFEHPVPVDRLEDVRLWMLHKEGESAANNCMQ
MHKTYLDGTGAAIRDFLGLIRSTNPILALGMLAEQAEHNSQLETRVCNSLKY
SAMFDAIHNTLADSLMRKVkiemfqlFREINVAEGSQRERHGFRHRW
RMLEQLGFRLSGLVSEELVQKSMLRRMYGSDNEGFFNVERSDEDNGEGGR
GGGVTLRWSEQindyTISAWTTGNN

AtSCL29

MLLEETPNQTLHDHVLSWLEDVSLSPLPGFDDSYYLHEDGSGTQDQTWEDQT
QDPEHGFQFSQDSLAAAYVGCEATNLEVTEAPSIDDLPLPQIPPNDQSRK
RSHDGFLEAQVKKASSRKRKAISKSEKSKDGKEWAEKLNNPCALAIT
ASNSSRQVHYLCVLSELASSSGDANRLLAALFLAFLQLHHSLSSSVSSHWFPSF
TFASEAEMFQKTLLKFLFYEVSFWALPPMNANSAILQLAQDPKDLDDLHIIID
IGVSHGMQWPWPTLLEASCRLFGGPPVRIRTIVSLTDATIPSVGPPNYGSQQL
LGFARSLKINLQISVLDKQLIDTSPHENVLIVCAQFRILHHLKHSINDERGETL
KAVSRLEPKGVLCCENNGECSSSADFAAGFSSKLEYWVKFLSTSSGKEENS
EERKLMGEATKVLVMNAGDMNEGKEKWYRMEAGFFVEAFEDAVDGA
KLRLRYNNWEIRMDGDTFAGLMWKGEAVSFCSLWK

AtSCL30
GIGTVQYotalSSKKTSTSDMLKAYQTISVCPFKKIAIFANHSIMRLASSANA
KTIHHIIGFDGDSDFQWPSLIIHLRAWRGSSCLKLRITGLRPQRFPAEGVIECR
RLAKYCYQKFNFEYNAIAQKVESIKLEDKLKEGEFVAVNSLRFRNLLDET
VAVHSPRDTVLKLRKIKPDKFVPIGILGSYNAPFVTFRFREVFHYSSLDMD
DTNLTDPRMVMPFKEFREIMVNVACEGTERVERPESYKQWARAMR
AFQRQIPKELEKLVQLLMSGYKPKEFVDQDDCHWLLQGWWKGRIVYGGSI
WVPFFFFVGRATRVLIMDPNFALSELNGFEYFDGNPNLLTDPMDQYPPMTD
LKVYSEIMENQDYQSMFDYDLAKRTMEMLQQVITDSQNSFSPASLSI
TNSWDAESIDESAYSADPQPVNEMVKMSFDAESALQFKKGVEEASKFLP
NSDQWVINDISERSRSDSKEMLQDQRVKKHERFEEVRSSKQFASNV
EDSKVTDMPDKVLLIDCQDTPQDTIDSEIQAIRRSSKINEGKGGGKKSQV
VDFRITLTHCAQAISTGDKATTEFLLLLQIQSQSPLGDAQRLHACFANALEG
RLQGSTGPMIQTYNATLSSLKDTAADTIRAYRVLSSSPFVTLMYFSSWIMI
LDVAKADVLHIVDFGILYGQWPMTQISIDRDKPVRKLRTGIELPCQGRP
ERIEETGRRLAYECKRFNFVPFSEKIAASQNWETIEREDLRPENVLAVNLG
RLKNLQDTEGSENCPRLRRLRNMNPVDVHAIVNGSNAPFFRKEAV
YHYSALEDFMFDSTPLPRDKERIFERFARYGREAMNVIAACEADRVPERETY
RQWQVMRAGFQKQKTIKPELVFRGKLKKWRYHKDFVDENSKWLLQG
WKGRITYASSCWVPA
>Al 493792
MTTKRIDFSPSSDDPSAAAKRRVNDIFPEETLENDGAAAIIKLSSLQLQCA
EYVATNHLREASTLLEISEIEICSPFGSSPERVAYFAQALQTRVISSYLSGACTP
LSEKPLTVVSQRQLSLAQTFNSVSLIKFSSHFTANQAIFQALDGEDSVHIIIDLD
VMQGLQWPALFHILASRPKRKLRSRITGFQGSDDLALSTGRRLADFASSLNPFP
EFPHEPVGKIGNLIDPSQLGTRQGEAVVVHWMQHRLYDVTDGDLETLEIRRRK
PNLITVVEQELSYYDGGFLGFFVEALHYYSALFDALDGLGEGSFERTFVEQ
IVLATEIRNIVAHGRRRRRRMKWKEELNRVGRFQVSLRGNPAMQAGLGLLM
LPWNGYTLVEENGLTRLGWKDSLTLTASAWKQPF

Brachypodium distachyon
>Bd Bradi2g57940
MGMSPTFTTSCYCTEQADICTETQELDYHHSYYGIEDVSLDEVELELAAS
SGRHHAKTFVVDHSSHPYHISWPPPPQASADVESRVRKQQFRDVLVSECKQK
VEAMEAMQHSPPLSGGGVGFQDQQQGDSGHGVAVGEGSSNGGTDGMRL
QLVLVCAEAECRDRQAASSLRELQAGAPVHTAFQRVASCVRQGLAD
RLALAHPPALGPAFICPQFFFFSSASCAGRGEALAVYECPYLRFAHFFVA
NASILEAEFGESKVHVVDLGMTLGDLRAHQWRALLDGLAARGVVARPARVR
VTGVGARVDAMRAVGELEAYAEELMCEVRAIDRTLESLHVDDLGVEAD
EAVAINSVELCHVCRESGRALNSVTLQTRKLAPKAFVVLQVDAGHNGPFFL
GRFMALYHYAALFDAALDPRYDARRAVEQFHGAEIFANVVGCEGAA
VERHERAQWRRMSRAGFQSMPIKMAKAREWLEENAGGTGYTVAEEKG
CLVLGKWKGPVIAASCWKC*

Carica papaya
>Cp evm.TU.supercontig_34.30
Citrus cinensis

>Cs orange1.1g040125m.g

MVYDFLFNTHKLNYNDQRSNQAIQLGLQLSIVDCYPYLPMMSDNAAASSMML QPRDQKRLKRTISVADSIAGDGSTSTLSRSSSSSSLNPLQLFRDHIWYTYQT RYLAATVEEAAAMATKVDGCGGDQDGQTADGMRVLQLLIAEAVAC RDKSHASALLSELRAALVFQGSSQVRVASCFVQLADRLASVQPLGAVGSFA PSMNMDIAGSREEEAFRLVYEICPHIQFGFHVANSSILEAFEGSFHVVDL GMTLGLPRQQWURLLESANQAGQQPPRLRITAVGLCQFQSGIDELCKY AKTYGINLEFSVESNLENLQTDKIKVENELVNVSNLQLHCVCERSGALN SVLQIHLSKPVVLVEQDSSHNGPFLFGRFMEALHYSAIFDSLAMDLPK YDTKRAKIEQFYFAEEIKINVSCEPGARVERHERVDQWRRRMSRAGFQAAPMK MINQAQKWLNKVKCNYGTVEYVEEKGLVGLWKSPIATSCWKC*

Citrus clementina

>Clc Ciclev10031305m.g

MEWGMYWRHCLLDKKNEVMGLDLNYSATACYRCFNSTLANHSCTWS TSDTRNHKRIKKSIDISSGSSSSLRCLNSLNSLPLQFRDHIWYTYTQLAY EAMEEAAALAMMAKGDTEVEKEEGNGDGKMLQVQQLIAEAVACRDKAHAS ALLLEEVRNALVFQGTSQVRASCFVQLARLAVQPLGAVGVGPAMA AITQSRDESLSFLYEICPQIQFGHFVANASILEAFEGSFHVVDL GMNLGLPRQWURLLESANRAQPPFLRITAVGLCQFQSGIDELCKY AKTYGINLEFSVESNLENLQTDKIKVENELVNVSNLQLHCVCERSGALN SVLQIHLSKPVVLVEQDSSHNGPFLFGRFMEALHYSAIFDSLAMDLPK YDTKRAKIEQFYFAEEIKINVSCEPGARVERHERVDQWRRRMSRAGFQAAPMK MINQAQKWLNKVKCNYGTVEYVEEKGLVGLWKSPIATSCWKC*

>Clc Ciclev10019777m.g

MVYDFLFNTHKLNYNDQRSNQAIQLGLQLSIVDCYPYLPMMSDNAAASSMML QPRDQKRLKRTISVADSIAGDGSTSTLSRSSSSSSLNPLQLFRDHIWYTYQT RYLAATVEEAAAMATKVDGCGGDQDGQTADGMRVLQLLIAEAVAC RDKSHASALLSELRAALVFQGSSQVRVASCFVQLADRLASVQPLGAVGSFA PSMNMDIAGSREEEAFRLVYEICPHIQFGFHVANSSILEAFEGSFHVVDL GMTLGLPRQQWURLLESANQAGQQPPRLRITAVGLCQFQSGIDELCKY AKTYGINLEFSVESNLENLQTDKIKVENELVNVSNLQLHCVCERSGALN SVLQIHLSKPVVLVEQDSSHNGPFLFGRFMEALHYSAIFDSLAMDLPK YDTKRAKIEQFYFAEEIKINVSCEPGARVERHERVDQWRRRMSRAGFQAAPMK MINQAQKWLNKVKCNYGTVEYVEEKGLVGLWKSPIATSCWKC*
FYFAEEIKIVSCEGPARVERHERHVDQWRRRMSRAGFQAAPMKMINQAKW
LNKVKVEGTYVVEEKGLVGWKSPIATSCWK*C*
>20816852-peptide|Clementina|Ciclev10018153m.g|Ciclev10018153m
MINSVCGTVGLKSENSSSSIKHPTPNESSISETKRTPQSSDLQPPSLTPPSL
NPFPPLKEIDGVEVQSPDCPSTETFSDHLDGFMISSPVRLPSTSAYNN
NYNNYNAQSMGQSSLGSCSPFRFQSAVGFASSSLRAGQSPHLKVFNSPGNQ
YMQPENFLALTTIADFLEDYQKGDYGMSPMKICGEGGGGGSSQLIDMPTPL
VPATLECMPTNNPSFSGVPESSSSTAGAQHNDIYQMSGAAPLSQQL
QQEQLQERQQTQQLQAQQQQONLNSLMPVLPSSEQEDSGGLQVLHL
LLACAEAVAKEDFMLARRYHLHNLNVSPLGDSMQRVAACFEALSARLAA
TLLTKPSTTPTFSPFPNSLEVLLKLYIQIVYQACPYYVFHAFANTQAIAFEA
EERVHVIDLDILQGYQWPAMQALARAAPPQFLRITVGATIESAKETGRCL
TELASHLHVPFEFHPVEQLEDKPMFNRVGEALAVNAVNRHLVPSNCL
GNLLAMIRDQAPNITIVEQEOASINGPYFLGRFLEALHYSAIFDSDATPPD
SAQRAKVEQYFAPEIRNIVACEGGERTARHERLEKWRKIMEQGFRGVLPSA
NAVTSQKILLGLYSCDYRLTEDNGCLLLGGWQDRALLAASAWRC*

**Citrus sinensis**

>Cs orange1.1g037028m.g
MAPDFQWRHGLCLDDKNEAMGLDLNYSATACYRFCNFLSTLANHSCTWST
SEDTRNKHRIKKTSMIDESSGSRLCNLSNSLPRLQFREDHVWVYTQRYLAIE
AMEEAAALAMIDKDETVKEEGNGDMKVLQVLICAAECVACRDKAHASAL
LESLRVNALVFHTSFQRVASCFLQGSLDRLALVQGPAVGVGSAAKSMAIT
SERDESLSLYEICPQPQFGHFWANISIAEFEGSLVHVVDLMTLGLPGR
QWHSMLQVNSRGKPKRLITGVNCERLGEIGDELKRYADGLKNFFLE
AVEKESLETQAKDINVEDGEVLVMNSILELHCVVKESRGLANSVLQRLHQL
SPKVVMLVEQDSSHPPFFLGRFMEALHYSAIFDSDLAMLKPYTDKRAKIE
QFVFAEEIKIVSCEGPARVERHERHVDQWRRRMSRAGFQASPMQVAKQ
WLNVQVTCRTYEEIEEKGLVCLVWKSPIAASCWK*C*

**Cucumis sativus**

>Csa Cucsa.196250
MVLDGDGSSFSTDFSTSVKTEDDMGSGAAHWSLDDTTASSRWRVISFS
DEFHKKLKETTESTPETDESGSNSSNSNSLSLSSCSDSLSTGFRAHWYNYQR
YLAAEAVEEAAAIAINAAESAESADAGMRLHLLLVCACAEAVACDRSH
ASILLSLRANALVGFSFQRVASCFVQLADRLALVQGPLYVGFLPIMSVR
DHSSDRKKKDEALNLAYEIPHIQFHFHAVSSILEVFEGENSHVLDLGMAF
GLPYGHQWHSILRLEASSNNRLRLRTVIGLSVNRVMEGKLAHAEGV
QVQEFVAVEGNLENLRPQDIKLHGDALVITSFOMCHVCVKESRGALTSVLR
MIYDLSPKLVLVQEDSNHAVFPFLGRFMEALHYSAIFDSDLAMLKPYDTR
RAKIEQFYFAEEIKIVSCEGMARVERHERHVDQWRRRMSRAGFQASPIKVMA
QAKQWIGKFKANEYTVIEEEKGLVCLVWKSPIAASCWK*C*

**Eucalyptus grandii**

>Et Euec.194688
MSLYQPLKLYSEDQYDESHVLDLNLSTLECYVSSTSSLQPISIHCFCFLDET
NCKRLVEPSIGIESIESKEIVHRGCSGMRTNSSLSLRLQFREDHIWAYTQRYLA
VEAMEEEAAAIMQDAMLKVQRILIAEAVACRDKTHASALLSELSSKALVF
GTSFQRVASCFCVQGLVDRLLAVQPSGAVGIVGQSAREVMTMNEKKEARLTV
YEICPYIQSHFIAESILEAEFEGESSIHHVDLGMTOQLPHVQRWNLISSLAKR
PGSSSRLCKISGVGNCCEGLTQIDKLIKHYAKGLGNFELMVGQSNLENLQA
EDSFILEGEVILSINQLHCVKESSRAGLNSVQLIHKLPSKLLVLVEQDSSH
NGPFPLGRFMEALHYYSAIFDALAMLKPYDTRAKIEQFFFAEEIKNIVSCE
GPARVERHERIDQWRRRMSRAGFQAPIKMLTKAQQWLAKFKVCYGETIVE
EKGCLVLGWNPSKPIAAASCWKC*

> Eg Eucgr.K02752.1
MLVYCDPDEEEERAGKKRPKRTYSFAESIGNSPRAYTGGFGSGGISRSSSMSS
LDTFPRLHFLRDHIWTSVYAVEAVEAAAAYSVGGGEEEDGTADGMRLL
QLLVACAEAVACRDKLQASSLLSELRSRALVFQTSFQRVASCFCVQLADRLS
LLQPGLTGIFISPAAAPVAAAAAAGADEKKEELAYLICPIYIQFHFVANTSIL
EAFEGENFFVHVDLGMGLPGHHQWRRLLIQSLAGHQQPPRLRITAIGLS
AEKFCSDIGDLSAYALNFSIGINLQILHCWKESSRAGLNSVQMIHNLSPKV
VLVEQDSSHGPFHFCVMEAYLSAIFDSLVMPLKYDTRAKIEQFFYFA
EEIKNIVSCEGPARVERHERIDQWRRRMSRAGFQAPIKMMMAQAKQWLGBK
NKVCYGETVVEEKGCLVLGWNPSKPIAAASCWKC*

Fragaria vesca

>Gve gene22702-v1.0-hybrid
MAPGHGRRFMTEHETDHEIRSNRMSDLELHGLNHSSMACYSSYPYMASL
DQEGANSDDQSSPLNPVSHESRDRLKRTISIDDSLSSITSINGGMMSRSSSTNS
LTTTLPRLHFLRDHIIWTYQYRLAEAVEEAAAAAMINAENGAEEQEDGTADGM
RLVQLLIACAEAVACRDKSHASALLSELARNLAVFSGSQRVASCFCVQLAS
RLALVQPLGAVGFVSPKLNAKDYALMDKKEELRYEICPHQGFHVANS
SILEAEFEGESFYHVVDLGMTHLPQGDQWRGLIHSLITRAGQQPVSRITGV
GLCADRLQIGELEAYADSLGLSVETCVESNLENLKPEDIKLYEGEVLVNS
ILQHLHCWKESSRAGLNSVQMIHELSPKVLVEQDSSHGPFHFCVMEALH
YYSAIFDSLVMPLKYDTRAKIQFFYFAEEIKNIVSCEGPARVERHERIDQW
RRRMSRAGFQAPIKMMMAQAKQWLGBKVCYGETVVEEKGCLVLGWNPSKPI
AASCWKC*

Glycine max

>Gm Glyma17g13680
MAPLYSAFNGEIDGENLPLAFDIWATNWAYGYYPQPAISENSSSKLVDVPF
CDGTIIRDNKRKVRCTCFPYPNVSCHSFNTNSSRSRNPKLHFLRDHYTQQR
YLAAEPVEASEDNTSNSSSESGGEEDGACDVGRLVQLLLIAEAVACRDKSHA
SILLSELKANLVFGSFSORVASCFCVGLTERLNLIQPISAGPMAPAMMIN
DAASDEEMEAYLRLPCIFPHQFYLANSTVLEAEFEGESFHVVDLGSMLG
LRHGHQWRLQSLANRASGVRRLRTIGVGLCVRLQTIGEELSVYANNLGI
NLEFSVNVNKLLENKPEDIEVREEEVLVNSILQHLHCWKESSRAGLNSVQMI
HGLGPKVLVMVEQDSSHGPFHFCVMEALHYYSAIFDSLVMPLKYDTRKA
KMEQFYFAEEIKNIVSCEGPLRMRERHERIDQWRRRMSRAGFQAAPIKMMV
AQSKQWLLKNKVCYGETVVEEKGCLVLGWNPSKPIAAASCWKC*

>Gm Glyma05g03020
Gossypium raimondii

Lotus japonicus

Lotus japonicus
Lupinus angustifolius

Malus domestica

Malus domestica
YSAIFDSDLAMLKYDTRRAKMEQFYFAEEIKNIISCEGRQGGEAREVDQWR
RRMRSQAGFQAAPIKMLAQAKQWLGKIKVCERGYTILEEKGCLVLGWKSKPIV
AASCWK*
> Mdo MD P0000464809
MPPLSLKLHSLFFKYHLRQHQLNHTQAQLQNTDPKFGITSRPEEPVPANPT
FQNGVATKNIHIDPNSSSLRFLPPTVLPKLAPNPSTRVGALLSPSXCSNSDD
GVVYRGYSPDQLVGRHRKRVPFIQLFHGGGFVGSNDSWNDAFCRRMAKL
CDAIYVAVGYLAPSAFEDGVTVLKWVAKQANALVQKGSRIFDSF
GSSMVPWLAAHGDPSRCVLLGVSANCNADYVARKAVEAGDLDPKGVVA
QVLMYPFIFGSTPTRSEIKLANSYFLDKATCMALWAFLQTEEFDLHPAGNP
LMPAPGRPLKTMPPTLTVAQHDWMDRGIAYSEELRKANVDAPLLDDYKD
TVHEFATLDVILETPQAKACAEDITIYKKEASTYTSHISHIVYISPQgDSigQ
ASNTLEYSMTADLYVQPEFCGEITSNLVGFDLNLSSMLCHPSLPLSDDS
ANWVSPFDKTRNHKRLQEHADANYGFKVGTNYYSSFYTGSDFSMCTGSFTS
LPTIQFREDHHASAHKRLYAEAMEEAFALTLMRKEDESEEGGSEDEIKLVOQ
LIAEAEAVACRDKHASALLATRLANAFGSTSFQRVSSCFVQGLAXRLALV
QLPGAGGGVPNVKSKPSAEKPGPSAEKDEALHLYVEICPIQIFQGHFVANASI
LEAFEGGSVHIVDGLMTGLQHGYQWRDLIDRLAREPVGQPLXLHTIGVGS
SSERLQAIGDLLKLQHSMKINXDFSEVSNLENLPKQDFHLVGEIINSILQ
LHCLVKEESRGTLNASALQTLHELPSKLMLVLVEQDTSNGPFPFLGRFMEALHY
SAIFDSDLAMLKYDTRVRKMEQFYFGEIKNIISCEGPARVERHERVDQWR
RRMRRAFQGQAPLRFTMAQAQVKWLETNACEGTVEVDKCLVLGWKSPIIA
TSCWK*

*Manihot esculenta*

> Me cassava4.1_032725m.g
MSPVPYMSEQSEFSNDSKSYETTFDLSPFCSVPNSADEVENSDPWFDFLDETRSH
KRIKQPECSRTRANSLSNPVPKLFHRDVSYAYERLAEAMEEAAMVTEGE
NEVKDEEGSDGMKVLQQLIAEAEAVACRDKKHASSALLAEAREALVFGTS
FQRVASCZVQQLSRLALLQPLGTAVLQAPQANSVTFFAERKDEALRVLVYDI
CPQIFQGFHVFVANASILEAFEGESSVHIVDGLMTGFLPHQLWNLHSLAKHP
GQPPHRLRTIYGVSNLVAEQLDCLYARSGLNFEFLWVEILENLKPEH
FKLAGEVVINLHQCXKESRGALNSFLQILHELPSKLLILEQDSNHGFP
FLLGRFMEALHYSAIFDSDLAMLKYDTRRAKIEQFEEAIKIISCEGPAR
VERHERLQDQWRMRSQAGFQGQPSTKTMIAKQWLEKANFCEGYAVTEEKSC
LVGLWKSPIIAASCWK*

> Me cassava4.1_031685m.g
MAPSLFSDFSENTEMGSSISSLDSLSAAYYPYLYLPTLENNASICILPFSED1
REHKIRKTRLTAEASLGNNEFDCAGSNSDCSSIRSNSSLNPRLHFR
DHIYRTQRYLAEAEIEATEMDEGDEDEGSTDGMRLVQLLIAACA
VACRDKSHASTLSELRRLNAFLGSSFQRVASCZVFQLGLASLQLPVLGTVG
FTASMNNIDISSDKKEALSLVYIECPIQGFHVFVANSSLEAFEGESFHV
DLGMLTLPHPGHQRWQLQLSRLAAGPRLLRTIGVLCVGRFQITGDELV
EYAKELGINLEFSVSNLENLRDDIKVFDEVLVNVSNQLHCVKESRGA
LNSVLQIIHALSPKVLVLEQDSSNHGFPFLGRFMEALHYSAIFDSDLTMLPR
Medicago truncatula

> MtDELL1
MKREHQESFGGVISNNKNKTNHNLNSKKNINFGECCSMQNTNTKQNMMRE
EKTENGGGMDELLAAALGYKVRSSDMADVQAQKLEQLEMVMGSAQEEGINH
LDSTTVHYDPTDLYSWVTQMTLTELNPDSQIINPDLASLSGSEILNNFNDSE
YDLAIPGMAAYAPQENNATARKMTWSEPSEPEAVSMPAVENTRPVVL
VDQTQETGVRLVHTLMAEAAIQQNKLXAEALVKHISLLASLQTMARVKA
SYFQALARRIYYGNPEETIDSSFEILHMHFYPESYPLFAHTANQAILEAFA
AGARVHVIDFLKQGMQPWALKLALRPGGPPFTPRLTGIGPQPADDNTDAL
QQVQWKLALQAQAQTIGVQFEFRGFCVNSIAALDPNMLEIRPGEAAVAVNSVEL
HTMLARPGSVEKVLNTVKKINPKVTIEVQANHNPVFVFEDFRAHYYSSL
FDSLEGSSNSNSNSNSNSTGLGPSQDDLMSEIYLKQICNVVAYEVGRVDER
HETLTVQWRSMGSAGFEPVHLGSNFKQASTLALFAGGDDGYRVEENGCL
MLGWHTRSIATSAWKLPQNESK

> MtDELL2
MKREHKLEHDMSSGSGKSGVCWEDDDGGGMDELLAVGVYKVXSDMAEV
AQKLEQLEQAMGMGNNHDHDESTIAQHLISNDTHTYVHNPSIDNWLQMTLSNF
DPQPNPNSVSNDDNDLNAIPGKAIYAAYADEFTSRKRKRNESVTVTTESTTTRPI
MVTETQEKIGRLHSMACAEAVEQNNKMAEALVKQIGYLVASEQAGAMRK
VATYFAEGLARRIYDQEHSQHSVDSQHIFQHETCPNLKFAHTANQAILEAQ
GKSSVHVIDFSIQMQMQWPALMQLALALRPGGPPAFRLTGIGPASDSDLHQ
QWVRLAQFQTQTHQFQYRFVANSNLDDLASMEILRSPETESAVANSSLE
LHKLNRPGALKFEVSFRQIRPEIVTVVEQEANHNGPAFLDRFESTHLHYSTL
FDSLEGSSVEPDQKDSMVLQKICNVVACEGTDRVERRHETLNOQWRNFN
SAGFSPVHLGSNFAFKQASMLLALFAGGDDGYKVEENGCCMLGWHTRPLIAT
SAWKLAAANSVSVVSH

> MtNSP1
MTMEPNPTSDHMLWLEGVSXSFPSFSLDPDPYNNGYIHEYIEJWQONQDISNQYQ
IDANTNSNNATNNIVAAANTTTSTTSLEPSNFSNNIPFSDLPPKRNAEDELSLK
KQPQNPKNKRLKSRPMNESDNGAALGETGTVRKSGNGKGAANKGSNN
NGNNKDRGWAELQNLNPCAVAITGGGNLRVQHLVLHELASTTDGANHRL
AAHGLRALTHHLSSSSSTSPGTTITFASTEPRFFQKSLKGYEFSPWSFPNIA
NASILQVLAEEPNLRTLHLDIGVSHGVCWPTFLELARSPGPPFLRLTVV
NASSSTENDQNMETPFSIGPCGDTFSSGLITYAQSLNVLSQKIKLKDHNPLQTL
NASKSDTSSDLVCAQFRKHLHNHNPDSEXFLKVRGMEPKGVILSENN
MECCCSSCGDGFATGYSLWRLDLSTSSAFKRNDSRKEMEGAEAKA
LTNQREMERNERRKECMKEAGFAGEVFGEADIDDGRALLRKYDNNWEM
KVEENSTSVELWWKQPVFSFSLWKLDKQPE

> MtRAM1
MINSLCGSSNSLKEKCLQPNSNNQNTNTHSKKNATNSCGDLEQINVLTQPSLNL
PSLKFDDLGDVEVQSPDSWMEAFFNDLDNDMFMISSPRINNPNSQPASTYN
NCNYNAYAQGMQIQIQLSCSPPRFASQIGSLNSNNQKKGKGLSplhrVFNPN
**Oryza sativa**
> OsDIP1
METMSYPCLSLLIPFSTQFEEISSSSSSLLWSPQAEENPHENANMYEFDADHS HDQIHQDHQFDLMPFQESANEDGDNHSHDQIHQDHEFLETMVQESANEDG DSHQIHQDHEFLEMMAISQESANDLQLQDFDVPANPASASFEFDERLAV AGHENGNNVATQESAGDLLLAGAMADGAVHASAIMSRLDLDLADIA GRRSCEATSPVHDLAYFARGLKLRLISGAATPASSPPPAANWSPAYRMLQE LTPFVKAHFTEANQLALETADDLDVHVDFNVGQVWLMLLKLLLGTI TILQPKLVSLERMSKPPSPSLLAAPPPEFFSDAFAVHTAVMESTASCLV SYDDEAWLSLRVGEEMGPRVEDAVGGRYSLAGGAQMMEGREREVSGF SVAQGKMLAGLGGGFVVHVQEGKGRACLWKRPLISVSLWCPK
> Os LOC_Os01g67650
MGMSPECPNSISDCSQQQSHHHTLTQDSTIIICTNQELDYYRFYDVEAAAF DGNEVLSRFSDKVTRMDHMISSPYQTPWSPAQAADDVVGSSETSRRVKKRF WDVLESCKQVEAMEADMTPATATFRVAGDGDDGGGGGGAGGDADG MRIVLQLVCAEAACAAALLRELQAGPHGTAFQRVASCFVQG LADRLPLAHPPALGPAHASICPCCAGRDRGARGEALVAYECYPYLRAHIF VANACMLEAFGESNVHVDLGMSLGLDRHQQWRLDGLAARASVGKPAR VRTVGARMTDRARAEYAYECLGMYFIRGIRGLESHELHDDLGVD ADEAVAINSLEDHHSVKESRGALNSVLQRTIKLSRPQVLFVEQDAHGHPF LGRFMELHYAALFDDLADALPRYDARRARVEQHGFGEIRNVVCGCEAA RVERHERADQPQRRMSRAGFQSVPKMAKAREWLDEHAGGGGYTVAEK GCLVLGWKGPVIAASCWKC*

**Phaseolus vulgaris**
> Pv Phvul.003G208600
MAPTHCSPFNGEIDGENLHSVDFIWDATDHYYPFPHQPSKSSSTTLVLPFCDGT IRDNKRVKRTVGIPFWDIENLGNHSSLNTSAANNISNRRDPSKHLHRDHIY YTQRYLAAEPEEEAPPETDNSSEGGVGEEDCVDGVRLVQLIACAEAVACRD KSHASILLSELANAVGGEFQYASCFVQGLTERLNLQPTGPVSPMPAM NIMDTASEEMEEAFRLVYELCPIHQFQHFTVNSLVEAFVGEFHVGDMLGTM LGRHGHQWRLMQSRLANKADERVRRRLRAVGLCVDLQSIGDELSVYAT NLGINFEVSVKKNLEPELEDVREEEELAVNSILQHCVKESRGALNSV LQMVHGLGPKVLMMEQDDSSHNGPFLGRFMELESLYSAIFDSLDVMPLKY DTKRAKMEQFYFAEAEIKNIVRPRPLMRHERHDVARQWRMSRAGFQAAPIK MVAQAKOWLQNSKCEYTVEEEKGLVLGWKSPIAVASCWKC*

**Populus trichocarpa**
> Pt Potri.015G091200.1
MAHFSLSGEFNDEVENTSGRLDLSLAMACYPRRPYLPTFDSNAVSWILPFSD DTRDAKRMRRSSLSVESIRSNSSLSYGSSSICRSTNLSNIPKHLHRDHWYT YTQRYLAAEVEAAEASEMINADEGGNEEGNADGMRVLQQLLICAEAVAC RDKSHASALLSELRENALVGSFQVRASCFVQGGLTDRSLVQPQAGVGFP TMINMIDIAKKEEALRLYCEIPHRFGHGFVANAILEAFFESEGFSVFHVVLGM TLGLSHGHQWRRLIESLERAAGKAPSRLRITGVLCDRFRGIDELKEYAKD MGINLEFSAVSNLENLRPEDIKINEGEVLVVSILQHCVKESRGALNSV L
QIVHELSPKVLVLVEQDSSHNGPFFLGRFMEALHYSAIFSDLAMLKPYDTR
RAKMEQFYFAEIEIKIVSCEGPARVERHERVYQWRRMRSAFGQAAPIKMM
AQAKQWLVKNKVCDBGYTVEEVEKGLVLGWSKPIIAASCWCKLINCSSQ*

\textbf{Prunus persica}

> Ppe ppa016418m.g

MSAFDLISHSAMSCYSPYMPLLLEEGNASSALTFFYLDTPNGNKHRLKRTISIAES
MSSSHLYGSGNSSNTNGSIRGTSNSTNLPLHFRDHIWYTQRYLAA
EAVEEEAAAAMINAEGNAGTFADGMRVLQLLIAACAEAVACRDKSHASAL
LSELRAANLVFSSQFRVASCFLQALNLVFLPLAVGFDGGSPMNACKDA
LDKKEALRLVYEICPIHQFQGHFVANSSILEAEGESYHVVDLGMTLGLPHG
DQWRGLIESLATRAGQPSPRLTIGVLYGDRMQIGIDEAYADLGINLEF
SVVSENLELRPEDIKDLDEGEVLVNSILQLHCVVKESRGALNSVLQMHELS
PKILVLEQDSSHNPFFLGRFMEALHYSAIFSDLAMLKPYDTRRAKMEQ
FYFAEIEIKIVSCEGPARVERHERVQWRRMRSAFGQAAPIKMLNKAKQWL
GKINVCETYTLEEGKCLVLGWSKPIAAASCWKC*

\textbf{Ricinus communis}

> Rc 29929.t000249

MPSALYMQSFSFRDYTRYETTGDLNIGYSIPHPVSKNRPTCFDLLHETRSHK
RVKPQGVESIGNIGCYAVSSNGSSINANCLNLSPLHHRDYIRAYTERLYLAIE
AMEEAAAGLMISKKEKEIDGMKLVQLIACAEAVACRDTHASALLSE
LRANALVGTSFQVRVASCFLQSLDRLTLIQPLGAVGLPGAKGTISFTEAKD
EALRLVYEICPQIQGFYFVANATILEAEGESASSHVVDLGMTLGLPHGEQWRN
Selaginella moellendorffii

Setaria italica
LCPYLRFAHFVANASILEAFEGESNVHVVDLGMTLGLDRAHQWRGLLDGLAARAGAKPKRVRVTGVGAPLDTMVAVGRELEAYAEGINLGMRFRAVDRSLES
LHADDDLVADEAVAISSVLEHCVKESRGALNSVLQTVRKLSPRAFVVLVEQDAGHNPPFPLGRFMEALHYYAAVFDALDAALPRYDARRARVEQFHFGEI
RNVVGCEGARVERHERADQWRRRMSRAGFQSVPIRMARARAREWLEENAGGGYVTAVAEKGCVLGKWGBKPIAASCWKC*

*Solanum lycopersicum*
> Sly Solyc03g110950.1

MIHDMANVFLSLEPCNGDQIQYGDPMENSLYLYTHHKLSEYSLNPYTSVLRKNA
PTNMINISSLSNSAFKRLRRTPSLGESFGSTFSTYEETSSSTQGSLPRLIGSSNS
VNSLSLQGPHHRDFHDWALNQRHYLAEAFEEAADIQNEEENGEQMKLQLTLSMADYARSLGINLEFSAVESNLENLKPDKIVYDEGLVNVSLLQH
CVKESRGALNSVLQHELSPKLVLVEQDSSHNGPFLGRFMEALHYYSIAF
DSDLVMLPKYDTRRAKIEQFYFAEEIKINVSCEGPVERHERVDQWRRRMS
RAGFQAAPIMVSAQQAKWVLAKVNHGEFTITEEKCVLGKWKSPIAASCWKC*

*Solanum tuberosum*
> Stu PGSC0003DMP400026755

MIHNMANVFLSLEPCNGDQIQYGDPMENGLYLYSHHKLSEYSLNPYTSVLRKNA
PTNSMIISSLNSDSKFRLRRTPSLGESFGSTFSTYESDSSSGSCLPRIGS
TNSVNSLQGPHHRDFHDWALNQRHYLAEAEAIQNEEENGEQMKLQLTLSMADYARSLGINLEFSAVESNLENLKPDKIVYDEGLVNVSLLQH
CVKESRGALNSVLQHELSPKLVLVEQDSSHNGPFLGRFMEALHYYSIAF
DSDLVMLPKYDTRRAKIEQFYFAEEIKINVSCEGPVERHERVDQWRRRMS
RAGFQAAPIMVSAQQAKWVLAKVNHGEFTITEEKCVLGKWKSPIAASCWKC*

*Sorghum bicolor*
> Sb Sb03g043030

MGTEQPCSTNLFTASSYGTSQQIIHLLLPPQHDSVICETPGMFGPYYYGTQDQD
AAFDGEVELGQASKATRVDYYSSPYQPSWPLRAYAATESSVRKQRFRD
VLESCKQKVEMAMEAEPSVAQEEGEDGLAVGDGGGAGAAAGGGAGAGGG
NGGAGDMRQLVCLAAECVARDRAAAALLRELQAGAPVHTAQFRV
ASCVQGLNRLALAHQPPALGPAASMAFCIPPSCTGRDGAR3ALALAYELCP
YLRFASFFVANASILEAFEGESNVHVDLGMTLGLDRAHQRGLLDQARAG
GAKPARVVRTAVGAPAETMRAGRELEAYAEGGLCLESFRAIDRSLSEHLM
DDLGIAADEAVAISSILEHCVKESRGALNSVLQIRKLSKAFVLLVEQDAG
HNPFFLGRFMEALHYYAAVFDALDAALPRYDARRARVEQFHFGEI
Theobroma cacao

* Theobroma cacao
> Tc Thecc1EG014574
MGHGFLANEFQKTDRIDEVIGDLDELSAMACFYQPFMPIMGDNCWSSLPS
GEIRDTKRLRRTISIPESIGSSGLSSGGNSDSLRSGSTSSLNSFSRHFRDHV
LYTNQRYLAAEVEAAAAMISSESGEDEETADGMRLOQLLIAACAEAVA
CRDKSHASALLSELARALNFQGSSFRQVASCVFQGLADRLLVQPLGTGVLV
APVMNIMDISSDKKKEALRLVYEICPHIQFGHVANSSILEAFEQEGESVHVVDL
GMLTLGLPHGQWHRLIQSANREGKAPSSLRTAVGLSDHRFHIIQGELEYAL
AKDLGMNLEFSVKNLENLRRPEDIKVFGEVLVNVSLQLHCVKESRGAL
NSVLQMIHELSPKVLVLEQDSSSHNGPFFLGRFMEALHYSAIFDSLMAKLKP
YDTRRAMEQFYFAEEIKNIVSCGEPGRVERHVDDQWRRRMSRAGFFQAAP
LMRMTQAKQWLGKNNCVEGYTVVEDEKGLVGLWKSKPIVAASCWKEC*

Vitis vinifera

* Vitis vinifera
> Vv GS1VG01007532001
MASDLLSEERSDEVGLDTSLSAQAYYSYRLPIFQNGSATNWFHYSDEARN
HKRLKRTGSIASIGSSLSSYSGKGSTSSSNSSFINRSSSTNLSLNPRLHRFDHI
WTYTTQRYLAEEAVEEAAAMISAEVEGEVEEGDGMRLVQLLIAACAEAVA
CRDKTHASLLSRLANVFGSFRQVASCVFQGLADRLLVQPLGAVGFIA
PSNPLDTAWEKEEALRLVYEICPHIKFGHVANASILEAFEQENFAHVVDL
GMLTLGLAHQGQWRLHSLANRAGRPPRRLTGVLCVDRFKIEEELAY
AQDLDINDLILQHCVKESRGALNSVLQKINELSPKVLVLEQDSSSHNGPFF
LGRFMEALHYSAIFDSLMAKLKPYDTRRAMEQFYFAEEIKNIVSCGEPARV
ERHERVDQWRRRMSRAGFFQAAPIKMMAQAKQWLGKVKACEGYNIMEEGK
CLVGLWKSKPIVAASCWKEC*

Zea mays

* Zea mays
> Zm GRMZM2G013016
MGTSEQCPCTSTPSFTTTSSSYVTSQIHHLPQHDSVCTEPGLGFPYYYVTD
RQQDAAFDVEGQFQASKATRVDYVSSPYQLSWPLAARRAASESRVRK
RFWDVLESCKQKVEAMEAMESPLVAFQAEADGGGAVGDGGGGGGRGSG
GGADGMRLVQLLLACAEAVACRDRAQAALLRELQAGAPVHGTAFQRVAS
CFVQGLADRLALAHPPALGPAFSCAGRDGGAARALALAYDLCPY
LRFAHVANASILEAFEQGETNVHVLDLGMMLGDLRAHQWRAALDGLARAG
AAARPVVRATVAGPAMADMRAGRELWYAEGLGMCLEFRAVDRSSL
HIDDLGIAADEAVAINSVLELHCVKESRGALNSVLQQTIRLSPKAFVLVEQD
AGHNGPFFLGRFMEALHYYAAVFDALDAALPRYDARRARVEQFHFGAEIRN
VVGCEGAARVERHERADQWRRRMRAGFQSMPIRMAARAREWLEENAGG
GGYTVAEEKGCLVLGWKGKPVIAASCWKC*