The evolution of pectate lyase-like genes across land plants, and their roles in haustorium formation in parasitic plant, *Triphysaria versicolor* (Orobanchaceae)

Running title:

Authors: Huiting Zhang[1,2], Eric K. Wafula[1], Elizabeth A. Kelly[1,2], Itsuhiro Ko[3], Alan Yocca[1], Pradeepa Gunathilake[4], Zhenzhen Yang[1,2], Daniel B. Steele[5], John Yoder[5], Loren A. Honaas[6], Claude W. dePamphilis[1,2]*

Author affiliations:

1. Department of Biology, the Pennsylvania State University, University Park, PA, USA
2. The Plant Biology Program, Huck Institute of Life Sciences, the Pennsylvania State University, University Park, PA, USA
3. Plant Science, College of Agriculture, the Pennsylvania State University, University Park, PA, USA
4. Agricultural Biotechnology Center, University of Peradeniya, Peradeniya, Sri Lanka
5. Department of Plant Sciences, UC Davis, Davis, CA, USA
6. Physiology and pathology of tree fruit research unit, USDA ARS, Wenatchee, WA, USA

*Corresponding author: Claude dePamphilis

Abstract

Parasitic plants in Orobanchaceae are noxious agricultural pests that severely impact crops worldwide. These plants acquire water and nutrients from their hosts through a specialized organ called the haustorium. A key step in haustorium development involves cell wall modification. In this study, we identified and analyzed the evolutionary relationships of *pectate lyase-like* (*PLL*) genes across parasitic plants and other non-parasitic land plant lineages. To support detailed examination of gene models and paralogous gene family members, we used published parasitic plant genomes, as well as a recently generated draft genome assembly and annotation of *Triphysaria versicolor*. One particular *PLL* gene, denoted as *PLL1* in parasitic Orobanchaceae, emerged as an important candidate gene for parasitism, as highlighted by our previous comparative transcriptomics research; it belongs to the largest sub-clade of the *PLL* gene family and is highly upregulated in haustoria. To explore its function in haustorium development, we manipulated *PLL1* expression in *T. versicolor*, a model parasitic species from Orobanchaceae, using direct transformation of the parasite and host-induced-gene-silencing (HIGS). For HIGS, we generated transgenic *Medicago truncatula* and *Arabidopsis thaliana* hosts expressing hairpin RNAs targeting the *PLL* gene in *T. versicolor*. Both experimental approaches effectively reduced the *PLL1* expression in the haustorium of *T. versicolor*, resulting in an increased frequency of poorly adhered parasites. These findings demonstrate that *PLL1* plays a crucial role in haustorium development and suggest it as a promising target for managing parasitic weeds. Notably, the success of HIGS even before the establishment of a functional haustorium highlights the possibility of early intervention against parasitism.

Introduction

Pectins, a family of complex polysaccharide polymers, are major components of primary cell walls and the middle lamellae of land plants. In higher plants, pectin functions in maintaining the mechanical and structural integrity of plant cell walls, contributes to intercellular adhesion [1], and plays a role in plant defense [2,3]. Several classes of pectin degrading enzymes have been identified in bacteria, fungi, and plants. One of those is pectate lyase (PL), an enzyme that cleaves the α-1,4-linked galacturonic acid in homogalacturonan by beta-elimination. PLs have been well documented to play key roles in invasive growth in plant pathogens such as *Erwinia carotovora* and *Bacillus* spp. as extracellular pathogens in soft rot disease [4–6]. In plants, *pectate lyase-like* genes (*PLLs*) play essential roles in various developmental and reproductive processes. During pollen tube growth, *PLL* functions by loosening pollen cell walls to initiate pollen tube emergence [7–10]. Plant *PLLs* are also involved in the activation of plant defenses

by releasing plant cell wall oligogalacturonides as defense elicitors in response to various plant hormones and stresses [11]. In addition, studies of *PLLs* in horticulture crops revealed their roles in the development of important crop traits [12]. *GhPEL76*, a *PLL* gene in cotton, regulates the organ elongation in Arabidopsis and fiber elongation in cotton [13]. Gene silencing assays in tomato and peach shed light on the roles that *PLLs* play in fruit ripening and softening [14,15].

Haustoria, the feeding structures that are present in all parasitic plants, are evolutionarily novel organs in the plant kingdom. Several studies have shown that parasitic plants have recruited haustorial genes from ancestral gene functions in nonparasitic plant tissues to adapt to a parasitic lifestyle [16–18], including those involved in pectin modification and degradation processes [17,19–25]. One such example is the *PLL* gene family [17]. In a study conducted by Yang *et al*., comparison between *PLL* gene expression patterns in the parasites and their orthologs in nonparasitic plants suggested that ancestral genes in parasitic plants were commonly expressed in floral tissues and that the haustorial expression was gained through regulatory neofunctionalization after a genome duplication event [17]. Haustorial specific *PLL* genes from three parasitic species in Orobanchaceae were proposed to be good candidates for functional assays [17] and could be potential gene silencing targets for parasitic weed management [26]. Dynamic regulations of cell wall modification enzymes, including PLLs, were also observed in another parasitic plant, *Cuscuta*, during haustorium development and parasite-host interaction [23–25].

Unlike functional characterization in model plant species where high-efficiency protocols have been developed and tested exhaustively, functional assays in parasitic plants remain a challenging endeavor. A method was proposed by Zhang 2020 [27], to select the best candidate genes for functional analysis in non-model organisms using a combination of transcriptome data, gene family phylogeny, and molecular evolution tests. To qualify as a good candidate gene, several analyses are needed on the parasitic plant *PLL* genes, such as validation of gene models, analysis of gene expression patterns on a uni-gene level, and identification of paralogous genes that might have functional redundancy; all of which can be facilitated with a gene family analysis. However, a comprehensive view of the evolutionary history of the PLL family is lacking. Sun and van Nocker used the Neighbor-Joining method to construct a *PLL* family phylogeny in Arabidopsis [28], which served as the baseline for more detailed phylogenetic studies of *PLLs* across land plants [29,30] or for a specific lineage [14,31]. In the aforementioned cases, however, only a limited number of species were selected and the phylogenetic relationships between some large groups remains unresolved or lacks strong support. Additionally, phylogenetic relationships from those studies conflict each other. To gain a deeper understanding of the evolutionary history of the *PLL* gene family in plants, especially in parasitic plants, and to verify whether the *PLLs* identified by Yang *et al.*

meet the requirement of good candidate genes for functional studies, a well resolved, robust phylogeny that includes parasitic plants is needed.

In this study, with a goal of investigating the evolution of *pectate lyase-like* genes in parasitic plants and examining the involvement of *PLLs* in haustorium development, we implemented a phylogenetic approach to study the *PLL* gene family across land plants and gene knockdown assays on a selected *PLL* gene in parasitic plants. A haustorially expressed *Triphysaria versicolor PLL* gene, *TrVeBC3_16311.1*, designated as *TrVePLL1*, was selected for functional assays in the model parasitic plant *T. versicolor*. Both transient transformation via hairy root transformation [32,33] and host induced gene silencing (HIGS) assays [34] were used to knockdown the expression of the target gene. Knockdown assays resulted in fewer mature haustoria with weakened haustorial connections, demonstrating the involvement of *TrVePLL1* in plant parasitism and indicating that the HIGS knockdown strategy could contribute to parasite control in agriculturally damaging parasites.

Materials and methods

Pectate-lyase like gene identification and initial phylogenetic inference

The following processes were accomplished using the PlantTribes2 [35] tools in Galaxy (https://usegalaxy.org/) [36], following the method from Zhang et al. [37].

Genes (*Arabidopsis PLL* genes), genomes (*Cuscuta* spp., *Ipomoea* spp., *Lindenbergia philippensis*, and *Striga hermonthica*), and transcriptomes of interest (*T. versicolor* build 4, *S. hermonthica* build 4, *P. aegyptiaca* build 6, and *L. philippensis* build 2 obtained from the Parasitic Plant Genome Project II, PPGP II) were processed to predict protein coding regions using *AssemblyPostProcessor* with TransDecoder [38]. A list of genomes and transcriptomes and data sources can be found in Supplemental Table S1. The post-processed sequences were classified into the pre-computed orthologous gene family clusters (orthogroups) of the 26Gv2.0 othrogroup scaffold from PlantTribes2 with *GeneFamilyClassifier* using the dual BLASTp [39] and HMMER hmmscan approach [40]. The classified sequences were then assigned with the corresponding orthogroups' metadata including functional annotation and super orthogroup assignment. Annotations of orthogroups within the same super orthogroups under all classification stringency were investigated to identify missing *PLL* genes. A total of four orthogroups were identified as *PLLs* from the PlantTribes2 26Gv2.0 scaffold based on both super orthogroup and functional annotation (Supplemental Table S2).

Sequences classified into desired orthogroups were integrated into the scaffold gene models using the *GeneFamilyIntegrator* tool and then all merged into one fasta file

for downstream process. Multiple sequence alignments were generated using *GeneFamilyAligner* using MAFFT's L-INS-i algorithm [41] and the corresponding coding sequences were forced onto the amino acid alignment. Sites present in less than 10% of the aligned DNA sequences, and sequences with less than 50% alignment coverage were removed with trimAL [42]. Maximum likelihood (ML) phylogenetic trees were estimated from the trimmed DNA alignments using two programs. First, a RAxML tree [43] was constructed using the PlantTribes2 tool *GeneFamilyPhylogenyBuilder* with default settings plus 1000 bootstrap replicates to estimate the reliability of the branches on the trees. Second, IQtree2 [44] was used with default settings to infer phylogenetic trees. In addition, 2000 ultrafast bootstrapping and the -bnni flag were used for UFBoot tree optimization [45]. *PLL* gene family trees with sequences from only autotrophic plants were constructed using the same methods after removing parasitic plant sequences for comparison.

Assembly and annotation of the *Triphysaria versicolor* genome

Genome assemblies have been shown to be useful resources to validate and improve gene models [37,46–49]. A draft assembly and annotation of the *Triphysaria versicolor* genome was used to facilitate gene model validation. This genome was first described by Daniel B. Steele, 2021 as a part of a doctoral dissertation [50], where the material and methods were described in detail. A brief summary of the methods is provided below.

High-molecular-weight genomic DNA was isolated from frozen above-ground material of a single field collected *Triphysaria versicolor* individual using a modified CTAB method. The genomic DNA was size-selected for large fragment (>40 kb) enrichment on a BluePippin instrument (Sage Science, Beverly, MA, USA) and barcoded according to the instructions from 10X Genomics (Pleasanton, CA, USA). The resulting products were sheared to a peak size of 350 bp on an E220 Focused Ultrasonicator (Covaris, Woburn, MA, USA). The sequencing library was prepared using the 10X Genomics Chromium Genome Kit V2, and sequenced on an Illumina Hiseq X sequencer in PE150 mode (Illumina, San Diego, CA, USA).

The genomic sequencing generated 713.5 million paired-end reads. After barcode removal, the cleaned reads were input to supernova (v2.1.1) [51] and assembled using default parameters. The haploid assembly was used for annotation using MAKER (v3.01.02-beta) [52]. Transcriptome sequences for *Triphysaria versicolor* generated using RNASeq data from the same dissertation work were provided as EST evidence. Protein sequences from *Mimulus guttatus* were acquired from Mimubase (http://mimubase.org/FTP/Genomes/Mguttatus_256_v2.0/, last accessed: May 27th 2025) and provided as protein homology evidence. A brief summary of genome assembly and annotation can be found in Supplemental Table 3.

Manual curation of parasitic plant *PLL* genes and final phylogeny inference

The putative *Triphysaria versicolor PLL* transcripts were subject to manual curation with webAUGUSTUS [53] (https://bioinf.uni-greifswald.de/webaugustus/, last accessed: May 7th 2024). A BLASTn search was performed on each transcript against the *T. versicolor* genome assembly described above to identify the genomic region corresponding to the *PLL* transcript. A genomic fragment containing the target transcript and 3kb upstream and downstream of the targeting region was extracted with Geneious R9 and used as the *genome file* input. Next, a BLASTp search was performed on the protein sequence predicted from each transcript against the protein sequences from the PLLs identified from the previous step. Sequences with the highest bitscore from closely related species (*Striga asiatica*, *Mimulus guttatus*, *Lindenbergia philippensis*, and *Solanum lycopersicum*) and Arabidopsis were selected as the *cDNA evidence file* input. *Solanum lycopersicum* was selected as the model organism for AUGUSTUS' *species* parameters. Genes from both strands were reported and analyzed. Full length genes were identified by aligning the post-processed transcript, gene models predicted by AUGUSTUS, and the reference sequences used as cDNA evidence for AUGUSTUS prediction. For cases where AUGUSTUS default setting failed to predict full length genes, the transcript(s), predicted exons, and *Striga asiatica* gene were mapped to the *T. versicolor* 10X genome assembly and the coding region were manually determined by using *S. asiatica* gene as a reference. A final BLASTn search was performed on the final gene model against the assembly and the results are available in Supplemental Table 4. Surprisingly, the *TrVeBC3_16311.1* gene identified by Yang et al., 2015 was missing in the *T. versicolor* transcriptome build 4. Manual annotation using the method described above was able to recover this gene from the genome. The same method was used to improve *Cuscuta campestris* gene models, with *C. australis* and *Ipomea* orthologs as reference.

Using the curated gene sequences, multiple sequence alignment and phylogenetic tree construction were repeated using the same methods described above. These final phylogenetic trees were used to infer monophyletic groups, which were named following the convention proposed in the Sun and van Nocker, 2010 research. Genes belonging to each clade were compared across the four trees (Supplemental Table S5). Because all the four trees share the same clade topology and almost the same genes were present in each clade, the IQtree constructed phylogenies were used for downstream analysis due to higher bootstrap support values.

Domain prediction and amino acid frequency analysis

Domain structures of proteins in each clade were predicted using the command line version InterProScan (v5.59.91) [54] and searched against all the databases. To predict the presence of transmembrane domains and signal peptides, the protein sequences were submitted to the SignalP-6.0 [55] server (https://dtu.biolib.com/SignalP-6, last accessed: April 2025) and TMHMM server v2.0 [56] (https://services.healthtech.dtu.dk/services/TMHMM-2.0/, last accessed: April 2025). Default setting was used for TMHMM, whereas 'Eukarya' and 'fast' mode was selected for SignalP-6.0.

The Pec_lyase_C domains predicted in each sequence by InterProScan were isolated and a MAFFT alignment was performed on the coding sequences comprising the focal domain. Sequences from the same clade identified from the gene family phylogenetic tree including all parasitic plants investigated in this research were extracted from the alignment and submitted to WebLOGO (https://weblogo.berkeley.edu/logo.cgi, last accessed: May 8th 2024) [57] to create the amino acid frequency plot.

Selective constraint analysis

To infer adaptive or purifying selection, the EasyCodeML software (v1.21) [58] was used to perform hypothesis testing using the CodeML program [59] from the PAML package. Since some gene relationships were not clear in the full phylogenetic tree, genes from each clade and subclade were extracted and clade/ subclade trees were constructed using the same method as described above. Sub-trees are available in Supplemental File 1. Branches housing parasitic plant *PLL* genes and selected non-parasitic asterid genes were identified from the subtrees and subjected to the test (Supplemental File 2). The branch-model in CodeML was first tested in the clades with parasitic plant *PLL* genes using the present parameters (nested models). The parasite genes were selected as foreground and the rest were identified as background. A likelihood ratio test was performed afterwards to determine whether the branch model (M2) fits better than the one ratio model (M0). For the cases where the branch model fits significantly better than the one ratio model, a branch-site test was performed to identify sites under positive selection for the indicated foreground lineages. The preset (nested models) parameters were used to perform the branch-test and the same branches were selected for this test as were used in the branch model. The Bayes empirical Bayes (BEB) method [60] was used to identify amino acids that are evolving under selection.

Plant material

Seeds of *Triphysaria versicolor* were obtained from the Yoder lab at University of California, Davis. *Arabidopsis thaliana* Columbia-0 seeds were obtained from the Anderson lab at the Pennsylvania State University (PSU). *Medicago truncatula* Jemalong A17 seeds were kindly provided by David Stout from USDA Agricultural Research Service, Pullman, WA and propagated in the Buckhout greenhouse at PSU.

Seed germination: *T. versicolor* seeds were surface sterilized with 70% ethanol for 10 min with shaking and then washed twice with distilled water. Next, seeds were scarified with *Triphysaria* sterilization solution (TSS: 50% household bleach, 0.01% Triton X-100) for 30 min. on a platform shaker at 95 rpm and then washed twice with distilled water. The treated *T. versicolor* seeds were transferred to petri dishes with plant growth medium (1/4 x Hoagland's basal salt and nutrient mix, 7.5 g/L plant tissue culture grade agar, pH of 6.1), and sealed with Micropore surgical tapes (3M, Maplewood, MN, USA). Seeds were allowed to stratify for 3 days at 4°C in the dark before moving to a 16°C growth chamber with a 12-hour light regime under 30 µmoles photons/m^2/sec light intensity.

M. truncatula seeds were scarified by soaking in 18M H_2SO_4 for 8 min. with gentle stirring every 2 min.. After the seeds were thoroughly washed with reverse osmosis water five times, they were sterilized with TSS for 5 min. with gentle shaking. The sterilized seeds were then washed with distilled water 10 times. Scarified and sterilized seeds were transferred to a new 50 ml tube with 20 ml distilled water and incubated on a shaker with gentle agitation at room temperature for 6 hours to overnight. Water was decanted after the incubation period and seeds were left in the tubes to grow in the dark for 12-24 hours at room temperature. The germinated seeds were transferred to petri dishes with the same growth medium used for *T. versicolor* and incubated in a growth room at 25°C with a 16-hour light regime under 100 µmoles/m^2/sec light intensity.

A. thaliana seeds were sterilized using 30% bleach with 0.1% SDS for 20 min., and shaken with a vortex shaker every 10 min. The sterilized seeds were washed with distilled water five times and transferred to petri dish containing MS growth medium (1/2 x Murashige and Skoog basal salt mixture, 10g/L sucrose, and 8g/L plant tissue culture grade agar, pH of 5.7). The treated seeds were incubated in a 22°C growth chamber with a 16-hour light period.

Parasite and host co-culture: *M. truncatula* and *A. thaliana* plants with robust root systems were transferred to co-culture medium (1/4 x Hoagland's basal salt and nutrient mix, 7.5 g/L sucrose, 7.5 g/L plant tissue culture grade agar, pH of 6.1). Host roots were carefully arranged approximately along the radius of the 150 mm petri dish. *Triphysaria* root tips were placed 0.5-1 mm from the host roots. Co-culture plates were sealed with Micropore surgical tape (3M, Maplewood, MN, USA) and placed on growth

racks with a 10-degree angle from level. The growth racks were placed in a growth room at 25°C with a 16-hour light regime under 100 μmoles/m^2/sec light intensity until tissue collection. Note *Arabidopsis* was only used for the Triphenyl tetrazolium chloride staining assay, while *M. truncatula* was used as the host for all other assays.

Haustorium development assay and tissue collection

To track changes in *TrVePLL1* expression levels during haustorium development, wild type *Triphysaria versicolor* was co-cultured with *Medicago truncatula* for up to 10 days. Haustorium tissue was collected at 12, 24, 48, 72, 120, and 240 hours post inoculation (HPI). Pre-attachment haustoria (at 12 and 24 HPIs) were collected by making cuts adjacent to the swollen region of the parasite roots. The post-attachment haustoria (48-240 HPI) were harvested by making cuts in both the parasite and host roots within 1mm adjacent to the haustoria. Harvested tissue was immediately frozen in liquid nitrogen and stored at -80C until processing.

Plasmid construction

To construct the hairpin RNAi vector, a 503-nucleotide region near the 3' end of the *TrVePLL1* open reading frame was amplified with primers containing four nucleotides, CACC, at the 5' end of the forward primer. To prevent off-targeting in both the parasitic plant and the host plant, this fragment was searched against the *T. versicolor* transcriptome assembly and the host genomes (i.e. *Medicago truncatula* and *Arabidopsis thaliana*) using BLAST. No perfect matches longer than 20 continuous nucleotides were found in any of the three targets. RNA isolation and cDNA synthesis was described below in the 'RNA isolation and transcript level analyses' section. The PCR product was recombined into the pENTR™/D-TOPO® entry vector using the pENTR™ Directional TOPO® Cloning kit (Invitrogen, Waltham, MA, USA) following the manufacturer's protocol. The entry construct was confirmed by Sanger sequencing and sequence comparison with the reference. Next, a Gateway® LR recombination reaction was performed to transfer the target amplicon into the pHG8-YFP parent vector (cite QR1 paper) facilitated by Gateway® LR Clonase® II enzyme mix (Invitrogen, Waltham, MA, USA). A double enzymatic digestion with restriction endonucleases XbaI and XhoI was performed to screen for positive hairpin vectors. The selected vector is named

pHpTvPLL1 and was transformed into *Agrobacterium rhizogenes* strain MSU440 via electroporation. Primer sequences can be found in Supplemental Table S6.

Hairy root transformation

Triphysaria versicolor transformation process was performed following Bandaranayake and Yoder, 2018 [61] with modifications. *Agrobactrium rhizogenes* was grown on mannitol-glutamic acid: Luria-Bertani (MG/L) medium containing 400 uM acetosyringone (Sigma-Aldrich, St. Louis, MO, USA) overnight or until the bacteria covered the petri dish and formed a thick lawn. Ten to fifteen-day old *T. versicolor* seedlings were used for transformation. A clear cut was made at the joint between the hypocotyl and the root. The wounded hypocotyl end of the seedling was dipped into the *A. rhizogenes* culture several times. Petri dish containing Murashige and Skoog (MS) medium with 3% sucrose was treated with freshly prepared acetosyringone at the concentration of 400 uM. The *A. rhizogenes* infiltrated plantlets were transferred onto the aforementioned plates and then incubated vertically in a 16°C growth chamber with a 12-hour light regime under 30 μmoles/m^2/sec light intensity for 7 to 10 days. To initiate root growth, plantlets were transferred to new growth medium (1/4 x Hoagland's basal salt and nutrient mix, 7.5g/L sucrose, 7.5 g/L plant tissue culture grade agar, pH of 6.1) with 300 uM Timentin and incubated for another 14 days. Transgenic calli and roots were identified by the visualization of YFP fluorescent signal using Zeiss SteREO Discovery.V12 with X-Cite® series 120 Q lamp under a YFP filter. Non-transgenic calli were discarded and non-transgenic roots were removed. Plants with positive signals were moved to a new growth medium and the tips of the transgenic roots were trimmed off to allow proliferation. Plates containing the transgenic plants were incubated vertically in a growth room at 25°C with a 16-hour light regime under 100 μmoles /m2/sec light intensity. Plants were screened and transferred to a new medium every 2 to 3 weeks until used for co-culture experiments. Because the transgenic *T. versicolor* roots have been through several rounds of root proliferation, an initial co-culture experiment was conducted with *T. versicolor* transformed with the empty vector and wild type *M. truncatula* to test their haustorium formation capability. Three to five plates were collected for each of the three biological replicates when using seedlings (control), and each plate contains 10 to 14 roots. Six to eight plates were collected for each replicate when using the transgenic line, and each plate contains three to eight roots with YFP signal. Haustorium formation rates were assayed five and seven days post co-culture. Statistical analysis is performed for this dataset and is described below.

Host plant, *i.e. Medicago truncatula*, transformation was performed following Biosson-Dernier et al., 2001 [62] with modifications. *A. rhizogenes* were prepared as described above. *M. truncatula* seeds were germinated as described above and seedlings grown in the dark for 36-48 hours were used for transformation. To make a

wound in the host plant seedling, the terminal 2-3 mm of the root tip was excised. The wounded plant was dipped in the *A. rhizogenes* culture and then transferred to growth medium (1/4 x Hoagland's basal salt and nutrient mix, 7.5 g/L plant tissue culture grade agar, pH of 6.1) treated with freshly made acetosyringone at the concentration of 400 uM. The plates were placed vertically in the 16°C growth chamber for 5 days and transferred to fresh plates and incubated in the 25°C growth room. Transformed plants were subject to YFP screening 2 weeks post infiltration. Plants with no transgenic signal were discarded and roots with no YFP signal were removed. Plants were screened and transferred to a new medium every 2 to 3 weeks until used for co-culture experiments. To test the presence of transformed DNA fragments in the host roots, *M. truncatula* roots (transformed roots with or without YFP signal and wild type roots) were collected by cutting the whole root system from the plant. Three biological replicates each containing 5-8 *M. truncatula* plants were collected. The collected tissue was ground in liquid nitrogen and DNA isolated using the Macherey-Nagel (Düren, Germany) NucleoSpin® Plant II kit following instructions in section 5.1 from the user manual. Concentrations of the extracted DNA samples were assessed on the Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) using dsDNA BR (broad range) Assay Kit. Same primers used to construct the RNAi construct were used to amplify a fragment of *TrVePLL1*. Gel electrophoresis was used to visualize the presence or absence of the amplicons.

Haustorium categorization and tissue collection

To test the effect of *PLL* knockdown on haustorium development, co-culture experiments were performed on the following parasite-host pairs: 1) *T. versicolor* transformed with *pHpTvPLL1* and co-cultured with wild type host; 2) *T. versicolor* with *pHG8-YFP* (empty vector control) co-cultured with wild type host; 3) Wild type *T. versicolor* co-cultured with host plant transformed with *pHpTvPLL1*; 4) Wild type *T. versicolor* co-cultured with host plant transformed with *pHG8-YFP*. Parasite-host co-cultures were monitored daily. Parasite root tips that started haustorial initiation later than 24 hours post inoculation were marked and not taken into consideration in the final analyses. Transgenic *T. versicolor* roots were allowed to grow in the presence of a wild type host for 7 days before harvesting. Two-week-old *T. versicolor* seedlings were allowed to grow in the presence of a transgenic host root for 5 days before collection.

At the end of the co-culture period, to test the quality of the haustorial connection, a gentle force was applied to the haustorium: host root was secured with one pair of tweezers and parasitic root adjacent to the haustorium was pulled gently. Haustoria that breaks easily from the host root with the gentle force were categorized as 'failed' haustoria, while the rest were categorized as 'mature' haustoria. When counting root tips forming mature or failed haustorium, in order to avoid potential observer bias, a random

number was assigned to each co-culture plate and the researcher recorded the haustorium type without knowing the treatment of each plate.

Five or six biological replicates were created for each co-culture pair. For transformed *T. versicolor* and wild type host co-culture, 50-60 transgenic roots from 8-10 independent *T. versicolor* root cultures were used in each biological replicate. For wild type *T. versicolor* and transgenic host co-culture, 80-100 *T. versicolor* seedlings were inoculated onto 8-10 transgenic host roots in each biological replicates. Statistical tests were performed on the results and are described below. The co-culture and haustorium categorization assays were performed two more times at a smaller scale. Results from the three experiments are summarized in Supplemental Table S7.

To assay transcript levels of *TrVePLL1* in haustoria from the co-culture experiment, failed and mature haustoria were collected. The failed haustoria were collected using the same methods for collecting pre-attachment haustoria. The mature haustoria were harvested using the same method for harvesting post-attachment haustoria. Three biological replicates were transferred immediately to liquid nitrogen for nucleic acid isolation. The rest were divided into two sets: one set for ink-vinegar staining and the other set for Toluidine Blue O staining (methods described below). Those samples were transferred immediately to the corresponding staining/ fixation solutions.

RNA isolation and transcript level analyses

RNA was isolated for cDNA synthesis in the plasmid construction process and for qPCR assays to estimate gene expression levels. Haustorial tissues collected for RNA isolation were ground in 350 ul RA1 buffer (Macherey-Nagel, Düren, Germany) mixed with 1% beta-mercaptoethanol (Qiagen, Hilden, Germany) using Kimble Chase glass tissue grinders (part# KT885450-0020). Total RNA was isolated from the macerated tissue using the Macherey-Nagel NucleoSpin® RNA Plant kit following the instructions from the user manual (section 5.1). For each sample, RNA concentration was assessed on the Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) using the RNA HS Assay Kit, while RNA purity was assessed with a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

cDNA was synthesized from RNA used for transcriptome sequencing in the PPGP project [17] or RNA collected for this project. 200 nanograms of RNA from each sample were converted to cDNA using the AzuraQuant™ cDNA Synthesis Kit (Azura Genomics, Raynham, MA, USA).

Primers for transcript level measurement were designed using Primer3 (v. 0.4.0, http://bioinfo.ut.ee/primer3-0.4.0/) [63,64]. To test primer specificity, melting curve

analysis was performed on each PCR reaction and primer pairs yielding more than one peak in the melting curve were not used.

The reverse transcription products were diluted 10-fold, and 1ul was used for each the qRT-PCR reaction with PerfeCTa SYBR Green FastMix (QuantaBio, Beverly, MA, USA). The reaction was performed with StepOnePlus Real-Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA) with the following program: Step 1. 95°C for 5 min.; Step 2. 95°C for 0.5 min.; Step 3. 60°C for 0.5 min.; Repeat Step 2 and 3 for 40 cycles; Melt curve: 0.5 increments from 95°C to 25°C.

The average crossing point (Ct) values from three technical replicates were used to represent each biological replicate and three biological replicates were measured for each sample. The 2^(-ddCt) method [65] was used to calculate the fold change in expression of target genes relative to the constitutively expressed gene *TrVePP2A* [33]. Statistical tests were performed on the results and are described below.

Haustorium staining assays

Three staining assays were deployed to examine morphological changes within the haustoria formed from the transgenic root and the control roots. Triphenyl tetrazolium chloride (TTC) staining was performed to observe metabolic activities in the haustoria [66]; Toluidine blue O (TBO) staining was performed to investigate changes in cell wall components in the haustoria [67], while modified ink-vinegar staining [below] was performed to visualize xylem bridges and other internal features of the haustorium. Samples for TBO and ink-vinegar staining assays were collected from the same experiment that categorized haustoria (see *Haustorium categorization and tissue collection*, above). A separate co-culture experiment was performed to collect tissue for the TCC staining assay. Three biological replicates were collected for each staining assay.

TTC staining was performed following Porter et al., 1947 [68] with modifications. *M. truncatula* root accumulates red or green pigments in co-culture conditions, therefore, *Arabidopsis* was used in this experiment instead. 1% tetrazolium solution was prepared and stored at 4°C in the dark. Haustoria connected to the same host root were collected together and immediately transferred to a 50ml falcon tube containing 30 ml of 1% tetrazolium solution. The plant tissue was immersed in tetrazolium solution and then incubated at room temperature in the dark for 24 hours before screening. Zeiss SteREO Discovery.V12 was used to visualize the staining result and capture images.

TBO staining was performed on haustoria infiltrated with 4% paraformaldehyde. The Leica TP1020 - automatic tissue processor at PSU Core Facility was used to dehydrate and infiltrate the tissue with paraffin (process described in Supplemental table S8). One to several haustoria were embedded in a paraffin mold with the host root oriented on the vertical axis using the Leica EG 1150C tissue embedding workstation.

Sections with the thickness of 8um were made with a rotary microtome (Thermo Fisher Scientific, Waltham, MA, USA) and de-waxed with a Leica autostainer XL with program 3 (Supplemental table S8). TBO solution was added to regions on the slides where tissue was present and incubated at room temperature of 30 seconds. Slides were cleaned with distilled water. Stained slides were visualized, and images were captured with an Olympus BX61 microscope.

Ink-vinegar staining of the haustoria was performed following Vierheilig et al., 1998 [69] with modifications. Host roots harboring one to several haustoria were collected in tissue embedding cassettes and placed in glass staining dishes with clearing solution (10% KOH solution). Samples were incubated at room temperature for 4 to 5 days and examined under dissecting scope daily to determine whether the clearing process was completed. Once the tissue is clear, cassettes were moved to a 1 L stainer and rinsed in cold running water for 5 min. The cassettes were then submerged in acidified water (5 ml white table vinegar in 95 ml water) for 5 min and rinsed briefly under cold tap water. The cleared and rinsed samples were placed in a preheated 5% ink solution (5 ml india ink in 95 ml water) in a water bath at 90°C for 30 min. Cassettes were removed to allow staining solution to drain, and then cassettes were rinsed under cold running water until water runs clear. Stained samples were moved to sterilized acidified water overnight to remove excessive ink staining. Samples were visualized and images were captured with Zeiss SteREO Discovery.V12.

Statistical analyses

Due to the overall low sample numbers, both parametric and non-parametric statistical analyses were performed. For pairwise comparison, a Mann-Whitney U test (non-parametric) and a Student's t-test (parametric) were performed. For multiple comparisons, a Kruskal–Wallis test (non-parametric) and an ANOVA test (parametric) were performed. Tests for normality and equal variance were performed prior to the parametric test to validate the assumptions. Shapiro-Wilk test was performed to test for normality, whereas a F-test (pairwise) or Bartlett's test (multiple) were performed to test for equal variance. For the qPCR data, a one-tailed p-value was used to test whether the gene expression level decreased in the knockdown lines. Two-tailed p-values were used for all the other tests. Raw data and statistics are available in Supplemental Table S9.

Results

Phylogenetic analysis of the pectate lyase-like gene family across land plants

To investigate the evolutionary relationships of *PLL* genes, Maximum Likelihood (ML) trees were constructed using two datasets. The first dataset included PLL genes from autotrophic land plants represented in the PlantTribes2 26Gv2.0 orthogroup scaffold, along with sequences from *Ipomoea nil [70]* and *I. trifida [71], and Lindenbergia philippensis;* representing the nonparasitic sister group of the parasitic Convolvulaceae and Orobanchaceae, respectively. The second dataset expanded on the first by incorporating *PLL* genes from parasitic plants (species listed in Supplemental Figure S1 and data source listed in Supplemental Table S1).

Each dataset was used to construct a separate ML tree. The autotroph-only tree served as a reference to establish a baseline topology, as parasitic plant sequences can be highly divergent due to their unique lifestyle or poorly annotated due to methodological errors, potentially distorting branching patterns and reducing node support. The second tree, based on the combined dataset of autotrophs and parasitic species, was used to place parasitic *PLL* genes in a broader phylogenetic context and to compare them with homologs in their closest nonparasitic relatives.

Most of the five major clades identified in previous studies [28–31] were strongly supported (bootstrap > 70) in both the autotroph-only and the combined autotroph-plus-parasite phylogenetic trees (Figure 1A, Supplemental Figure S2). Genes from all sampled angiosperm lineages were present in most clades. Clade 5, the earliest diverging monophyletic group (Figure 1A), contains genes from all land plants and forms the sister group to the remaining clades. Clades 4 and 3 form the second main monophyletic group, corresponding to seed plant-wide and eudicot-wide clades, respectively. These two clades are highly similar in sequence and form a well-supported monophyletic group along with two *PLL* genes from *Selaginella moellendorffii*.

A notable expansion of Clade 4 *PLLs* was observed in most rosid species, but not in other major lineages, including the parasitic plants (Supplemental Figure S2). Furthermore, Clade 4 included *PLL* genes from *Striga* and *Triphysaria*, but none from *Cuscuta* or *Phelipanche*, in contrast to the other clades where all four parasitic lineages were represented. The grouping of Clade 3 and 4 has previously been observed in McCarthy et al., [29], but not in the other three publications.

Clades 1 and 2 comprise the third major monophyletic group. Each can be further divided into two subclades, 1-1 and 1-2, and 2-1 and 2-2. Subclades 1-1 and Clade 2 include *PLL* genes from all major angiosperm taxa, while Subclade 1-2 additionally contains genes from loblolly pine (*Pinus taeda*), indicating a broader seed plant-wide

distribution. In contrast to previous phylogenies inferred using Neighbor-Joining methods [28,30,31], where Subclade 2-2 appears as a sister to Clades 1 and 2, and Subclade 2-1 was classified within Clade 1 (Supplemental Figure S3), the present analysis provides strong support for the monophyly of Clade 2, a relationship originally suggested by McCarthy et al., [29].

The haustorial *PLL* genes previously proposed for functional studies from the Yang *et al.,* study were classified into Subclade 1-1 (Supplemental Figure S2 and Supplemental File 1), the largest clade in the *PLL* family. Additionally, these haustorial genes reside on a branch with a longer branch length than most other genes, indicating higher sequence divergence in the coding region. Expression analysis from the Parasitic Plant Genome Project II (PPGP II) transcriptome data (Supplemental Figure S4) revealed that among the four *Triphysaria PLL* genes in Subclade 1-1, *TrVePLL1* is specifically expressed in haustorial stages, which provides additional strong support for this gene as a good candidate for functional analysis.

Conserved domain structure among PLL proteins

Domain search results (Figure 1B) showed that all sequences used in the phylogenetic inference contain a Pec_lyase_C domain, which is responsible for the primary catalytic function in pectate modification. In addition, proteins in Subclade 2-2 contain a unique Pec_lyase_N terminal domain. This domain likely originated in a common ancestor of angiosperms, as Subclade 2-2 includes representatives from all sampled angiosperms. Secretory signal peptides (SSP) were frequently detected at the N-terminus of proteins in Clades 1 through 4, but only a single Clade 5 gene, *Selmo_v1.0_412910*, harbors this domain (Supplemental File 2). This pattern suggests two possible scenarios: 1) this SSP was present in the common ancestor of all clades but lost in Clade 5, or 2) it was gained in the common ancestor of Clade 1-4 after divergence from Clade 5. Additionally, several proteins in Subclade 2-1 contain C-terminus transmembrane domains, suggesting a novel domain acquisition event by the ancestral protein of this clade.

An amino acid frequency plot (LOGO) was generated for the Pec_lyase_C domain from each clade to visualize sequence divergence (Figure 1C). Clade 5, the sister group to all the other *PLL* clades, exhibits an 18-amino acid deletion as well as additional sequence divergence relative to the others. Despite this divergence, key functional residues corresponding to putative Ca^{2+} binding sites, substrate binding sites, and the catalytic site, as predicted from previously characterized PL protein structures [72–75], were highly conserved in all clades.

High sequence divergence and adaptive evolution in parasitic plant PLLs

Long branches in the phylogenetic tree, indicative of high sequence divergence, may reflect underlying biological processes such as adaptive evolution. To investigate potential shifts in selective pressure, a two-step analysis was performed within branches housing parasitic plant genes in the *PLL* gene family. First, a branch test was conducted with parasitic plant lineages designated as the foreground and asterid nonparasitic genes as the background. This analysis identified five branches exhibiting relaxed purifying selection and one branch under positive selection (Supplemental Table S10). The branch with positive selection signal contains *TrVePLL1*, a gene proposed to be an ideal candidate for functional assays by Yang et al., 2015 [17]. Second, branch-site tests were applied to these six branches to detect specific amino acid sites under adaptive evolution. Three of the five branches with relaxed purifying selection, as well as the branch containing *TrVePLL1*, exhibits amino acid sites with a signature of adaptive evolution in the foreground sequences (Supplemental Table S10).

TrVePLL1 is upregulated during haustorium development

TrVeBC3_16311.1, a single *Triphysaria versicolor* gene, was chosen for functional assay and designated as *TrVePLL1.* This gene contains seven exons and six introns, encoding a 1338bp mRNA that translates into 446-amino acid protein (Figure 2A). Similar to proteins in Subclade 1-1, TrVePLL1 contains 3 functional motifs: an N-terminal signal peptide, a pectate lyase C domain, and a C-terminus low complexity motif (Figure 2A). Previous research demonstrated that *TrVePLL1* underwent neofunctionalization via an expression shift [17], and results from this study suggest additional evidence of adaptive protein evolution. These features make *TrVePLL1* a strong candidate for downstream characterization using RNAi-based molecular approaches.

Transcriptome data from PPGP [17] and PPGP II (unpublished, accessible at http://ppgp.huck.psu.edu/) indicated that *TrVePLL1* is upregulated during early haustoria development (stage 3) (Supplemental Figure S4D). To validate this pattern and examine its expression profile at higher temporal resolution, transcript levels of *TrVePLL1* were measured from *T. versicolor* roots or haustorial tissues collected at 0, 12, 24, 48, 72 hours, as well as 5 and 10 days after incubation with *Medicago truncatula* roots (Figure 2B). Between 0 and 24 hours post inoculation (HPI), swelling occurred just behind the root tip of the parasite and the epidermal cells produced long haustorial hairs that adhered to the host root. Between 24 and 72 HPI, the haustorium apex extended toward the host and initiated penetration; during this phase, *TrVePLL1* expression increased sharply, reaching a plateau at ~200-fold above baseline by 72 HPI. Elevated transcript levels persisted through 5 days post inoculation, coinciding with the formation of vascular connections, and declined by 10 days, when the haustorium had matured [76].

Expression levels of *TrVeBC3_3448.1* (annotated as *TrVeBC4_p_c2216_g69* in PPGP II assembly, and designated as *TrVePLL2)*, was also measured in the same plant material as a control. Consistent with the transcriptome data, no significant expression changes were observed post inoculation. Two additional *TrVePLL* genes in Clade 1-1 were not tested due to very low transcript abundance in the PPGP II transctiptome datasets.

Reduction of mature haustorium formation was achieved by knocking down *TrVePLL1* via direct transformation and HIGS

To investigate whether *TrVePLL1* contributes to haustorium formation, RNA interference (RNAi) was used to knock down gene expression via two methods: 1) transient hairy root transformation of *Triphysaria versicolor*, and 2) host induced gene silencing (HIGS), in which the silencing construct is transformed into the host plant.

A 503-base pair fragment of the *TrVePLL1* gene was cloned into the *pHG8-YFP* RNA silencing vector [33] to generate the *pHpTvPLL1* hairpin construct. On average, 73% and 68% *T. versicolor* root transformed with the empty vector (*pHG8-YFP*) and *pHpTvPLL1*, respectively, showed YFP signals (Supplemental Figure S5), higher than previously reported efficiencies [33]. Approximately 80% of the *M. truncatula* plants developed transgenic roots in both lines, consistent with previous publications [34]. The presence of the RNAi construct in *M. truncatula* roots was confirmed via PCR (Supplemental Figure S6). No morphological differences were observed between roots transformed with pHG8-YFP and pHpTvPLL1.

To generate sufficient root material roots for haustorium assays, transformed *T. versicolor* roots were maintained in tissue culture for up to three months with periodic root trimming to stimulate proliferation. These roots remained competent for haustorium formation, similar to non-transformed seedlings as shown in Table 1.

Table 1. Haustorium development in transgenic and non-transgenic *T. versicolor* roots over time.

Days post inoculation	% of root with haustoria[a]		
	Seedlings	pHG8-YFP (empty vector)	p-value (MWU test)[b]
5	83 ± 8 (n=125)	62 ± 11 (n=76)	0.121
7	88 ± 4 (n= 117)	83 ± 10 (n=95)	0.7

a Value are means ± standard deviation from three biological replicates. n = the total number of roots in all replicates.

b Statistical tests were performed to assess whether the number of roots forming haustorium are different in the seedling group *vs.* the plants transformed with an empty vector, pHG8-YFP. Both parametric and non-parametric statistical tests were performed, and results can be found in Supplemental Table S9A.

In the first experiment, haustoria were assayed from wild type *M. truncatula* inoculated with transgenic *T. versicolor*. Seven days post inoculation, both control and knockdown lines developed haustoria that appeared swollen, hairy, and attached to the host. In the control lines, most haustoria are firmly connected to the host root, namely mature haustoria, while a small percentage (~28%, Figure 3A) are loosely connected and detect readily from the host upon the application of a small mechanical force with forceps. Those loosely connected haustoria were deemed as 'failed haustoria'. In contrast, the *TrVePLL1* knockdown lines exhibited a higher proportion (~45%) of failed haustoria. This phenotype was reproducible across three independent experiments (Supplemental Table S5).

The same haustorium assay was repeated using a host-induced gene silencing (HIGS) system, where wild type *T. versicolor* seedlings were co-cultured with transgenic *M. truncatula*. In the control group (empty vector), ~69% of roots formed mature haustoria, and ~20% formed failed haustoria (Figure 3C). When co-cultured with *M. truncatula* expressing pHpTvPLL1, only ~49% of *T. versicolor* roots developed mature haustoria, while ~44% produced failed haustoria. These results were highly consistent with those from the direct transformation assay. Statistical analyses using both parametric and non-parametric methods were performed for both experiments, and results are provided in Supplemental Table S9B.

To determine whether the increase in failed haustoria correlated with the expected decrease in transcript level of *TrVePLL1*, a qRT-PCR assay was conducted. In the direct transformation system, *TrVePLL1* transcript levels were higher in haustorium of the same type (mature or failed) from the control compared to the knockdown line (Figure 3B). In the HIGS system, a similar pattern was observed, though failed haustoria from the control group were not sampled due to their low number. Both mature and failed haustoria from *T. versicolor* co-cultured with pHpTvPLL1-transformed hosts showed reduced *TrVePLL1* transcript levels compared to the control (Figure 3D). Interestingly, although transcript reduction was more pronounced in the direct transformation assay, the proportion of failed haustoria was comparable, if not higher, under HIGS. Statistical analyses were performed with both parametric and non-parametric methods and the results can be found in Supplemental Table S9C.

Staining of haustoria reveals morphological similarities and differences between control and knock down lines

It is difficult to distinguish, through visual inspection of haustoria, the failed from the mature haustoria or a haustorium formed in the knock down lines versus the control lines. To examine potential differences in internal haustorium structure, three staining assays were performed on haustoria collected from transgenic *Triphysaria versicolor* co-cultured with wild type hosts.

The ink-vinegar staining method revealed the presence of one or more xylem connections (Figure 4A, B) in most haustoria. Over 70% of mature haustoria developed single or multiple xylem bridges, while only about 50%of failed haustoria developed xylem connections (Figure 4C, and Table 2). Statistical analysis was conducted to assess differences in xylem bridge formation among four haustorium groups: mature and failed haustoria from both control and *TrVePLL1* knockdown lines. For the first comparison, which evaluated the percentage of haustoria that formed at least one xylem bridge, the failed haustoria from the control group were excluded due to low sample size. A Kruskal–Wallis test yielded a p-value of 0.063, indicating a marginal difference among the remaining groups. To further examine this trend, a one-way ANOVA was performed and revealed a significant difference ($p = 0.013$). Post hoc Tukey's HSD test showed that the frequency of xylem bridge formation in failed haustoria from the knockdown line was significantly lower than in both mature haustoria from the same line and mature haustoria from the control line (Supplemental Table S9D). In the second comparison, which evaluated the average number of xylem bridges per haustorium, all four groups were included. No significant differences were observed among them (Table 2; Supplemental Table S9E).

In the Triphenyl tetrazolium chloride (TTC) staining assay, haustoria were stained to a deep red, especially at the host-parasite interface (Figure 4E, F), consistent with high metabolic activity. In some instances, intense staining was also observed at root tips (Figure 4G). No clear differences were observed between failed and mature haustoria, or between the knockdown and the control lines in terms of TTC staining patterns.

Toluidine blue O (TBO) staining, distinguishes cell wall components based on acidity and chemical composition [67], was used to examine histological features in sectioned haustoria. However, haustorium sections processed from the *TrVePLL1* knock down line and the control line were not visibly different (Figure 4H, I). It's worth mentioning that only mature haustoria were used for sectioning in the TBO study, as failed haustoria frequently detached from the host root during the dehydration and embedding process.

Table 2. Xylem connection observations in transgenic haustoria. Statistical analysis results can be found in Supplemental Table 9D and E.

Type of haustoria	% roots with xylem bridge[a]	Average number of xylem bridges in a haustorium[b]
pHG8-YFP mature	67 ± 6 (n= 62)	1.32 ± 0.85 (n=41)
pHG8-YFP failed	53 (n= 17)[c]	1.30 ± 0.67 (n=10)
pHpTvPL1 mature	72 ± 8 (n= 46)	1.28 ± 0.79 (n=39)
pHpTvPL1 failed	50 ± 4 (n= 40)	1.15 ± 0.50 (n=20)

a Value are means ± standard deviation from three biological replicates (except for *pHG8-YFP* failed). Three to five plates were collected for each replicate. n = the total number of roots in all replicates.
b only haustorium formed xylem bridge were included in the calculation

Discussion

A robust phylogenetic analysis reveals the evolutionary history of the *PLL* genes.

Compared to the published phylogenetic estimation of the *PLL* gene family [11, 28-31], the results from this research are more thorough and robust. The Bayesian tree constructed by Palusa et al. divided 62 *PLL* proteins from *Arabidopsis thaliana*, *Populus trichocarpa*, and *Oryza sativa* into 3 distinct groups [11]. With rice as the only monocot representative in that study, group II and a few other subgroups were annotated as dicot-specific. However, with increased sampling from diverse monocot species, the phylogeny presented in this research identified monocot genes in these groups. The only true dicot specific clade is Clade 3, which was identified as the *PLL4* subgroup in group I from the Palusa et al. results. McCarthy *et al.* expanded the species list from three to ten and built a high-quality phylogeny to infer the ancestry of the *PLL* gene family [29]. The main limitations in the McCarthy *et al*. phylogeny were the lack of resolution between some major clades and the absence of asterid representatives due to the then available genomes. By sampling a wider range of plant species, and including genes from eight asterid genomes and transcriptomes, all the clades in this study were resolved into monophyletic groups with strong support, with little ambiguity remaining in terms of the relationship between major clades. There have been other attempts to resolve the phylogeny of PLLs across land plants with additional plant genomes including asterids [30,31]. However, both publications failed to provide supporting evidence for the clade-to-clade relationship as we do in this study.

How does *TrVePLL1* facilitate parasitism? Two hypotheses

In the process of pollen tube growth, pectin deposition and the action of several pectin processing enzymes have been proposed to play essential roles. As a fast-growing tissue that penetrates through other plant tissues, the apical cell walls of the pollen tube must remine flexible enough to allow incorporate new cell wall material, while also maintaining sufficient strength to withstand turgor pressure [77]. As major cell wall components, pectic molecules have been reported to be actively delivered to the cell surface via vesicles [8,78–80]. These pectic molecules may be newly synthesized and secreted from the Golgi apparatus, or recycled from older cell walls that have been degraded by pectic enzymes [9]. In parallel, pectate lyase-like proteins have also been proposed to loosen the cell wall of the stylar cells, facilitating pollen tube penetration [9].

Haustorium penetration, which is morphologically similar to pollen tube growth, may have adopted analogous molecular strategies. In fact, some *PLLs* in the parasitic Orobanchaceae are derived from ancestral genes formerly expressed in floral tissues of their non-parasitic relatives [17]. In addition, Losner-Goshen *et al.* used immunolabeling with pectin-specific antibodies and observed the disappearance of pectins from the cell walls of a sunflower root at the parasite-host interface, while the host cells themselves stayed intact [81]. More recently, *Wakatake et al.* (2023) provided additional support for the idea that parasitic plants penetrate host tissues intracellularly, rather than disrupting host cells directly [82].

Taken together, we propose two possible mechanisms by which *TrVePLL1* may function during haustorium development. *Hypothesis 1*: *TrVePLL1* facilitates the mobilization of parasite-derived pectin molecules by degrading demethylated pectin from older parasite cell walls, generating monomers that can be recycled as building material for new walls to grow at the apical tip of the haustorium. *Hypothesis 2*: *TrVePLL1* promotes haustorial penetration in the host tissue by targeting and degrading pectin within the host cell wall or middle lamella. To further evaluate these hypotheses, further experiment, such as high-resolution tracking of TrVePLL1 protein localization or ultrastructural analysis of pectin content at the host–parasite interface, will be required.

Working towards durable resistance in a changing environment

Despite the high heterozygosity level of gene sequences in the obligate outcrossing species *Triphysaria versicolor*, the success of the RNAi-based knockdown experiments to affect haustorial development is encouraging. Notably, the success of using host induced gene silencing (HIGS) in reducing gene expression and altering haustorial outcomes highlights the potential of this approach as a strategy for managing weedy parasitic plants such as *Striga*. The effectiveness of HIGS prior to the establishment of a mature haustorium suggests that parasitism can potentially be

disrupted before significant damage to the host occurs. Functional studies of *TrVePLL1* orthologs in *Striga hermonthica* and *S. asiatica* (designated as *StHePLL1* and *StAsPLL1*) should be prioritized if genetic manipulation becomes feasible. *StHePLL1* is of particular interest, not only because of its high sequence divergence (reflected by the long branch length in Supplemental Figure S2) and high expression level in haustorial stages [17], but also because it can be a candidate gene for parasitic weed control. However, targeting a single gene may be insufficient for durable suppression, as genetically diverse parasitic populations may overcome silencing rapidly. A multi-targeted HIGS approach, as proposed by Dr. Jonathan Gressel, in which RNAi constructs targeting many parasitism related genes could be "stacked" into the host, offering a more robust and sustainable control strategy [83]. To fulfill this goal, more functional characterization studies like the experiments conducted in this research needs to be executed to generate a good set of target genes.

Gene redundancy also presents a challenge when targeting a single gene for functional disruption. In *Triphysaria versicolor*, 27 *PLL* genes were identified, including five that are significantly differentially expressed and another seven that are most highly expressed (although not differentially expressed) in haustorial stages (Supplemental Figure S7). Although *TrVePLL1* is the only haustorially expressed *PLL* gene in Subclade 1-1 and possesses several unique features that qualify it as a strong candidate for functional assay, it is possible that partial functional redundancy with other PLLs contributed to the moderate knockdown phenotype, in which over 50% of haustoria still formed mature, stable host connections. To tackle this limitation, future gene-silencing efforts may require the design of constructs targeting multiple candidate genes or adoption of the "stacked" gene-silencing strategy described above.

Conclusions

Our results provide novel insights into the evolutionary history of the pectate lyase-like genes in land plants and offer evidence for the involvement of *TrVePLL1* in haustorium development in parasitic Orobanchaceae. The *PLL* gene family tree, constructed using sequences from 34 plant species representing major groups of land plants, including six parasitic plants and the carnivorous plant *Utricularia*, revealed new relationships among the five major clades, with strong supporting evidence. *TrVePLL1*, a gene from *Triphysaria versicolor* predicted to be involved in parasitism [17], was identified as the only haustorial-specific gene in Subclade 1-1. Additionally, parasitic plant genes in the branch housing *TrVePLL1* showed signatures of positive selection. Together, these findings highlights *TrVePLL1* as a strong candidate gene for functional assays. Knockdown of *TrVePLL1* expression through transforming *T. versicolor* root with an antisense hairpin construct, or via host-induced gene silencing (HIGS), reduced the formation of mature haustoria. This study is among the first to provide experimental

evidence that a cell wall processing enzyme plays a crucial role in haustorium development [19,21]. With the rapid advancements in genome and transcriptomic sequencing technologies, along with improved data processing methodologies, scientists are now well-equipped to identify additional candidate genes potentially involved in plant parasitism. However, the functional characterization of those genes remains a challenge. Continued research, such as the study presented here, is needed to fully understand their roles and explore new strategies to disrupt parasitic mechanisms.

References

1. Ridley BL, O'Neill MA, Mohnen D. Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. Phytochemistry. 2001;57: 929–967. doi:10.1016/S0031-9422(01)00113-3 PMID - 11423142

2. Eticha D, Stass A, Horst WJ. Cell-wall pectin and its degree of methylation in the maize root-apex: significance for genotypic differences in aluminium resistance. Plant Cell Environ. 2005;28: 1410–1420. doi:10.1111/j.1365-3040.2005.01375.x

3. Starr MP, Moran F. Eliminative split of pectic substances by phytopathogenic soft-rot bacteria. Science. 1962;135: 920–921. doi:10.1126/science.135.3507.920

4. Collmer A, Keen NT. The role of pectic enzymes in plant pathogenesis. Annual Review of Phytopathology. 1986. doi:10.1146/annurev.py.24.090186.002123

5. Nakajima N, Ishihara K, Tanabe M, Matsubara K, Matsuura Y. Degradation of pectic substances by two pectate lyases from a human intestinal bacterium, *Clostridium butyricum-beijerinckii* group. J Biosci Bioeng. 1999;88: 331–333. doi:10.1016/s1389-1723(00)80020-1

6. Kotoujansky A. Molecular Genetics of Pathogenesis by Soft-Rot *Erwinias*. Annu Rev Phytopathol. 1987;25: 405–430. doi:10.1146/annurev.py.25.090187.002201

7. Ohtsuki T, Taniguchi Y, Kohno K, Fukuda S, Usui M, Kurimoto M. Cry j 2, a major allergen of Japanese cedar pollen, shows polymethylgalacturonase activity. Allergy. 1995;50: 483–488. doi:10.1111/j.1398-9995.1995.tb01183.x

8. Ischebeck T, Stenzel I, Heilmann I. Type B phosphatidylinositol-4-phosphate 5-kinases mediate *Arabidopsis* and *Nicotiana tabacum* pollen tube growth by regulating apical pectin secretion. Plant Cell. 2008;20: 3312–3330. doi:10.1105/tpc.108.059568 PMID - 19060112

9. Wu Y, Qiu X, Du S, Erickson L. PO149, a new member of pollen pectate lyase-like gene family from alfalfa. Plant Mol Biol. 1996;32: 1037–1042. doi:10.1007/BF00041387 PMID - 9002602

10. Jiang J, Yao L, Yu Y, Liang Y, Jiang J, Ye N, et al. *PECTATE LYASE-LIKE 9* from *Brassica campestris* is associated with intine formation. Plant Sci. 2014;229: 66–75. doi:10.1016/j.plantsci.2014.08.008

11. Palusa S, Golovkin M, Shin S, Richardson DN, Reddy ASN. Organ-specific, developmental, hormonal and stress regulation of expression of putative pectate lyase genes in *Arabidopsis*. New Phytol. 2007;174: 537–550. doi:10.1111/j.1469-8137.2007.02033.x PMID - 17447910

12. Uluisik S, Seymour GB. Pectate lyases: Their role in plants and importance in fruit ripening. Food Chem. 2020;309: 125559. doi:10.1016/j.foodchem.2019.125559

13. Sun H, Hao P, Gu L, Cheng S, Wang H, Wu A, et al. *Pectate lyase-like* Gene *GhPEL76* regulates organ elongation in *Arabidopsis* and fiber elongation in cotton. Plant Sci. 2020;293: 110395. doi:10.1016/j.plantsci.2019.110395

14. Xu Z, Dai J, Kang T, Shah K, Li Q, Liu K, et al. *PpePL1* and *PpePL15* Are the Core Members of the Pectate Lyase Gene Family Involved in Peach Fruit Ripening and Softening. Front Plant Sci. 2022;13: 844055. doi:10.3389/fpls.2022.844055

15. Yang L, Huang W, Xiong F, Xian Z, Su D, Ren M, et al. Silencing of *SlPL*, which encodes a pectate lyase in tomato, confers enhanced fruit firmness, prolonged shelf-life and reduced susceptibility to grey mould. Plant Biotechnol J. 2017;15: 1544–1555. doi:10.1111/pbi.12737

16. Cui S, Wakatake T, Hashimoto K, Saucet SB, Toyooka K, Yoshida S, et al. Haustorial Hairs Are Specialized Root Hairs That Support Parasitism in the Facultative Parasitic Plant *Phtheirospermum japonicum*. Plant Physiol. 2016;170: 1492–1503. doi:10.1104/pp.15.01786

17. Yang Z, Wafula EK, Honaas LA, Zhang H, Das M, Fernandez-Aparicio M, et al. Comparative transcriptome analyses reveal core parasitism genes and suggest gene duplication and repurposing as sources of structural novelty. Mol Biol Evol. 2015;32: 767–790. doi:10.1093/molbev/msu343 PMID - 25534030

18. Yoshida S, Cui S, Ichihashi Y, Shirasu K. The Haustorium, a Specialized Invasive Organ in Parasitic Plants. Annu Rev Plant Biol. 2016;67: 643–667. doi:10.1146/annurev-arplant-043015-111702 PMID - 27128469

19. Leso M, Kokla A, Feng M, Melnyk CW. Pectin modifications promote haustoria development in the parasitic plant *Phtheirospermum japonicum*. Plant Physiol. 2023;194: 229–242. doi:10.1093/plphys/kiad343

20. Dixit S, Upadhyay SK. Melting the wall: plant parasitism entails pectin modification. Plant Signal Behav. 2023;18: 2252219. doi:10.1080/15592324.2023.2252219

21. Jhu M-Y, Ichihashi Y, Farhi M, Wong C, Sinha NR. *LATERAL ORGAN BOUNDARIES DOMAIN 25* functions as a key regulator of haustorium development in dodders. Plant Physiol. 2021;186: 2093–2110. doi:10.1093/plphys/kiab231

22. Kurotani K-I, Wakatake T, Ichihashi Y, Okayasu K, Sawai Y, Ogawa S, et al. Host-parasite tissue adhesion by a secreted type of β-1,4-glucanase in the parasitic plant *Phtheirospermum japonicum*. Commun Biol. 2020;3: 407. doi:10.1038/s42003-020-01143-5

23. Jhu M-Y, Farhi M, Wang L, Zumstein K, Sinha NR. Investigating Host and Parasitic Plant Interaction by Tissue-Specific Gene Analyses on Tomato and *Cuscuta campestris* Interface at Three Haustorial Developmental Stages. Front Plant Sci. 2021;12: 764843. doi:10.3389/fpls.2021.764843

24. Bradley JM, Butlin RK, Scholes JD. Comparative secretome analysis of *Striga* and *Cuscuta* species identifies candidate virulence factors for two evolutionarily independent parasitic plant lineages. BMC Plant Biol. 2024;24: 251. doi:10.1186/s12870-024-04935-7

25. Bawin T, Didriksen A, Faehn C, Olsen S, Sørensen I, Rose JKC, et al. *Cuscuta campestris* fine-tunes gene expression during haustoriogenesis as an adaptation to different hosts. Plant Physiol. 2023;194: 258–273. doi:10.1093/plphys/kiad505

26. Fernández-Aparicio M, Delavault P, Timko MP. Management of Infection by Parasitic Weeds: A Review. Plants. 2020;9. doi:10.3390/plants9091184

27. Zhang H. Phylogenetic and functional studies on parasitism genes in Orobanchaceae and Convolvulaceae. dePamphilis C, editor. Ph.D, The Pennsylvania State University. 2020. Available: https://search.proquest.com/openview/be0f185db348e0573dc527fc9fb57397/1?pq-origsite=gscholar&cbl=18750&diss=y

28. Sun L, van Nocker S. Analysis of promoter activity of members of the *PECTATE LYASE-LIKE* (*PLL*) gene family in cell separation in *Arabidopsis*. BMC Plant Biol. 2010;10: 152. doi:10.1186/1471-2229-10-152 PMID - 20649977

29. McCarthy TW, Der JP, Honaas LA, dePamphilis CW, Anderson CT. Phylogenetic analysis of pectin-related gene families in *Physcomitrella patens* and nine other plant species yields evolutionary insights into cell walls. BMC Plant Biol. 2014;14: 1–14. doi:10.1186/1471-2229-14-79 PMID - 24666997

30. Zheng Y, Yan J, Wang S, Xu M, Huang K, Chen G, et al. Genome-wide identification of the *pectate lyase-like* (*PLL*) gene family and functional analysis of two *PLL* genes in rice. Mol Genet Genomics. 2018;293: 1317–1331. doi:10.1007/s00438-018-1466-x

31. Sun H, Hao P, Ma Q, Zhang M, Qin Y, Wei H, et al. Genome-wide identification and expression analyses of the *pectate lyase* (*PEL*) gene family in cotton (*Gossypium hirsutum* L.). BMC Genomics. 2018;19: 661. doi:10.1186/s12864-018-5047-5

32. Tomilov A, Tomilova N, Yoder JI. *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes* transformed roots of the parasitic plant *Triphysaria versicolor* retain parasitic competence. Planta. 2006;225: 1059–1071. doi:10.1007/s00425-006-0415-9 PMID - 17053892

33. Bandaranayake PCG, Filappova T, Tomilov A, Tomilova NB, Jamison-McClung D, Ngo Q, et al. A single-electron reducing quinone oxidoreductase is necessary to induce haustorium development in the root parasitic plant *Triphysaria*. The Plant Cell Online. 2010;22: 1404–1419. doi:10.1105/tpc.110.074831 PMID - 20424175

34. Bandaranayake PC, Yoder JI. Trans-specific gene silencing of acetyl-CoA carboxylase in a root-parasitic plant. Mol Plant Microbe Interact. 2013;26: 575–584. doi:10.1094/MPMI-12-12-0297-R PMID - 23383721

35. Wafula EK, Zhang H, Von Kuster G, Leebens-Mack JH, Honaas LA, dePamphilis CW. PlantTribes2: Tools for comparative gene family analysis in plant genomics. Front Plant Sci. 2022;13: 1011199. doi:10.3389/fpls.2022.1011199

36. The Galaxy Community. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res. 2022;50: W345–W351. doi:10.1093/nar/gkac247

37. Zhang H, Wafula EK, Eilers J, Harkess AE, Ralph PE, Timilsena PR, et al. Building a foundation for gene family analysis in Rosaceae genomes with a novel workflow: A case study in *Pyrus* architecture genes. Front Plant Sci. 2022;13: 975942. doi:10.3389/fpls.2022.975942

38. Haas BJ. Transdecoder. In: GitHub [Internet]. 2024 [cited 2024]. Available: https://github.com/TransDecoder/TransDecoder

39. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10: 421. doi:10.1186/1471-2105-10-421 PMID - 20003500

40. Eddy SR. Accelerated Profile HMM Searches. PLoS Comput Biol. 2011;7: e1002195. doi:10.1371/journal.pcbi.1002195 PMID - 22039361

41. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30: 3059–3066. doi:10.1093/nar/gkf436 PMID - 12136088

42. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated

alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25: 1972–1973. doi:10.1093/bioinformatics/btp348

43. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30: 1312–1313. doi:10.1093/bioinformatics/btu033

44. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol. 2020;37: 1530–1534. doi:10.1093/molbev/msaa015

45. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol. 2013;30: 1188–1195. doi:10.1093/molbev/mst024

46. Pilkington SM, Crowhurst R, Hilario E, Nardozza S, Fraser L, Peng Y, et al. A manually annotated *Actinidia chinensis* var. chinensis (kiwifruit) genome highlights the challenges associated with draft genomes and gene prediction in plants. BMC Genomics. 2018;19: 257. doi:10.1186/s12864-018-4656-3

47. Marx H, Minogue CE, Jayaraman D, Richards AL, Kwiecien NW, Siahpirani AF, et al. A proteomic atlas of the legume *Medicago truncatula* and its nitrogen-fixing endosymbiont Sinorhizobium meliloti. Nat Biotechnol. 2016;34: 1198–1205. doi:10.1038/nbt.3681

48. Li Y, Wei W, Feng J, Luo H, Pi M, Liu Z, et al. Genome re-annotation of the wild strawberry *Fragaria vesca* using extensive Illumina- and SMRT-based RNA-seq datasets. DNA Res. 2018;25: 61–70. doi:10.1093/dnares/dsx038

49. Yocca A, Akinyuwa M, Bailey N, Cliver B, Estes H, Guillemette A, et al. A chromosome-scale assembly for “d”Anjou’ pear. G3 (Bethesda). 2024;14. doi:10.1093/g3journal/jkae003

50. Steele DB. The Transcriptional Regulation of Host Recognition and Prehaustorium Development in *Triphysaria versicolor*. PhD, UC Davis. 2021.

51. Weisenfeld NI, Kumar V, Shah P, Church DM, Jaffe DB. Direct determination of diploid genome sequences. Genome Res. 2017;27: 757–767. doi:10.1101/gr.214874.116 PMID - 28381613

52. Holt C, Yandell M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. BMC Bioinformatics. 2011;12: 491. doi:10.1186/1471-2105-12-491

53. Stanke M, Steinkamp R, Waack S, Morgenstern B. AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic Acids Res. 2004;32: W309–12. doi:10.1093/nar/gkh379

54. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30: 1236–1240. doi:10.1093/bioinformatics/btu031

55. Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol. 2022;40: 1023–1025. doi:10.1038/s41587-021-01156-3

56. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305: 567–580. doi:10.1006/jmbi.2000.4315

57. Crooks GE, Hon G, Chandonia J-M, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14: 1188–1190. doi:10.1101/gr.849004

58. Gao F, Chen C, Arab DA, Du Z, He Y, Ho SYW. EasyCodeML: A visual tool for analysis of selection using CodeML. Ecol Evol. 2019;9: 3891–3898. doi:10.1002/ece3.5015 PMID - 31015974

59. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24: 1586–1591. doi:10.1093/molbev/msm088

60. Yang Z, Wong WSW, Nielsen R. Bayes empirical bayes inference of amino acid sites under positive selection. Mol Biol Evol. 2005;22: 1107–1118. doi:10.1093/molbev/msi097

61. Bandaranayake PCG, Yoder JI. Factors affecting the efficiency of *Rhizobium rhizogenes* root transformation of the root parasitic plant *Triphysaria versicolor* and its host *Arabidopsis thaliana*. Plant Methods. 2018;14: 61. doi:10.1186/s13007-018-0327-2 PMID - 30026789

62. Boisson-Dernier A, Chabaud M, Garcia F, Bécard G, Rosenberg C, Barker DG. *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. Mol Plant Microbe Interact. 2001;14: 695–700. doi:10.1094/MPMI.2001.14.6.695

63. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. Nucleic Acids Res. 2012;40: e115–e115. doi:10.1093/nar/gks596

64. Koressaar T, Remm M. Enhancements and modifications of primer design program Primer3. Bioinformatics. 2007;23: 1289–1291. doi:10.1093/bioinformatics/btm091

65. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25: 402–408. doi:10.1006/meth.2001.1262

66. Stūrīte I, Henriksen TM, Breland TA. Distinguishing between metabolically active and inactive roots by combined staining with 2,3,5-triphenyltetrazolium chloride and image colour analysis. Plant Soil. 2005;271: 75–82. doi:10.1007/s11104-004-2027-0

67. Pradhan Mitra P, Loqué D. Histochemical staining of *Arabidopsis thaliana* secondary cell wall elements. J Vis Exp. 2014. doi:10.3791/51381

68. Porter RH, Durrell M, Romm HJ. The Use of 2,3,5-Triphenyl-Tetrazoliumchloride as a Measure of Seed Germinability. Plant Physiol. 1947;22: 149–159. doi:10.1104/pp.22.2.149 PMID - 16654086

69. Vierheilig H, Coughlan AP, Wyss U, Piche Y. Ink and vinegar, a simple staining technique for arbuscular-mycorrhizal fungi. Appl Environ Microbiol. 1998;64: 5004–5007. doi:10.1128/AEM.64.12.5004-5007.1998

70. Hoshino A, Jayakumar V, Nitasaka E, Toyoda A, Noguchi H, Itoh T, et al. Genome sequence and analysis of the Japanese morning glory *Ipomoea nil.* Nat Commun. 2016;7: 13295. doi:10.1038/ncomms13295

71. Wu S, Lau KH, Cao Q, Hamilton JP, Sun H, Zhou C, et al. Genome sequences of two diploid wild relatives of cultivated sweetpotato reveal targets for genetic improvement. Nat Commun. 2018;9: 4580. doi:10.1038/s41467-018-06983-8

72. John P. Vogel , Theodore K. Raab , Celine Schiff , Shauna C. Somerville. PMR6, a *Pectate Lyase–Like* Gene Required for Powdery Mildew Susceptibility in *Arabidopsis*. The Plant Cell. 2002;14: 2095–2106. doi:10.1105/tpc.003509

73. Yujia Leng , Yaolong Yang , Deyong Ren , Lichao Huang , Liping Dai , Yuqiong Wang , Long Chen , Zhengjun Tu , Yihong Gao , Xueyong Li , Li Zhu , Jiang Hu , Guangheng Zhang , Zhenyu Gao , Longbiao Guo , Zhaosheng Kong , Yongjun Lin , Qian Qian , Dali Zeng. A rice *PECTATE LYASE-LIKE* gene is required for plant growth and leaf senescence. Plant Physiology. 2017;174: 1151–1166. doi:10.1104/pp.16.01625

74. Scavetta RD, Herron SR, Hotchkiss AT, Kita N, Keen NT, Benen JAE, et al. Structure of a Plant Cell Wall Fragment Complexed to Pectate Lyase C. The Plant Cell. 1999;11: 1081–1092. doi:10.1105/tpc.11.6.1081

75. Yoder MD, Jurnak F. The Refined Three-Dimensional Structure of Pectate Lyase C from *Erwinia chrysanthemi* at 2.2 Angstrom Resolution (Implications for an Enzymatic Mechanism). Plant Physiol. 1995;107: 349–364. doi:10.1104/pp.107.2.349

76. Westwood JH, dePamphilis CW, Das M, Fernández-Aparicio M, Honaas LA, Timko MP, et al. The Parasitic Plant Genome Project: New Tools for Understanding the

Biology of Orobanche and Striga. Weed Sci. 2012;60: 295306. doi:10.1614/ws-d-11-00113.1

77. Gaffe J, Tieman DM, Handa AK. Pectin Methylesterase Isoforms in Tomato (*Lycopersicon esculentum*) Tissues (Effects of Expression of a Pectin Methylesterase Antisense Gene). Plant Physiol. 1994;105: 199–203. doi:10.1104/pp.105.1.199

78. Heslop-Harrison J, Heslop-Harrison Y. The growth of the grass pollen tube: 1. Characteristics of the polysaccharide particles ("P-particles") associated with apical growth. Protoplasma. 1982;112: 71–80. doi:10.1007/BF01280217

79. Rosen WG, Gawlik SR. Fine structure of lily pollen tubes following various fixation and staining procedures. Protoplasma. 1966;61: 181–191. doi:10.1007/BF01247918

80. Dashek WV, Rosen WG. Electron microscopical localization of chemical components in the growth zone of lily pollen tubes. Protoplasma. 1966;61: 192–204. doi:10.1007/BF01247919

81. Losner-Goshen D, Portnoy VH, Mayer AM, Joel DM. Pectolytic Activity by the Haustorium of the Parasitic Plant *Orobanche* L. (Orobanchaceae) in Host Roots. Ann Bot. 1998;81: 319–326. doi:10.1006/anbo.1997.0563

82. Auriac M-C, Griffiths C, Robin-Soriano A, Legendre A, Boniface M-C, Muños S, et al. The penetration of sunflower root tissues by the parasitic plant *Orobanche cumana* is intracellular. New Phytol. 2024;241: 2326–2332. doi:10.1111/nph.19495

83. Gressel J, Gassmann AJ, Owen MD. How well will stacked transgenic pest/herbicide resistances delay pests from evolving resistance? Pest Manag Sci. 2017;73: 22–34. doi:10.1002/ps.4425

Figures

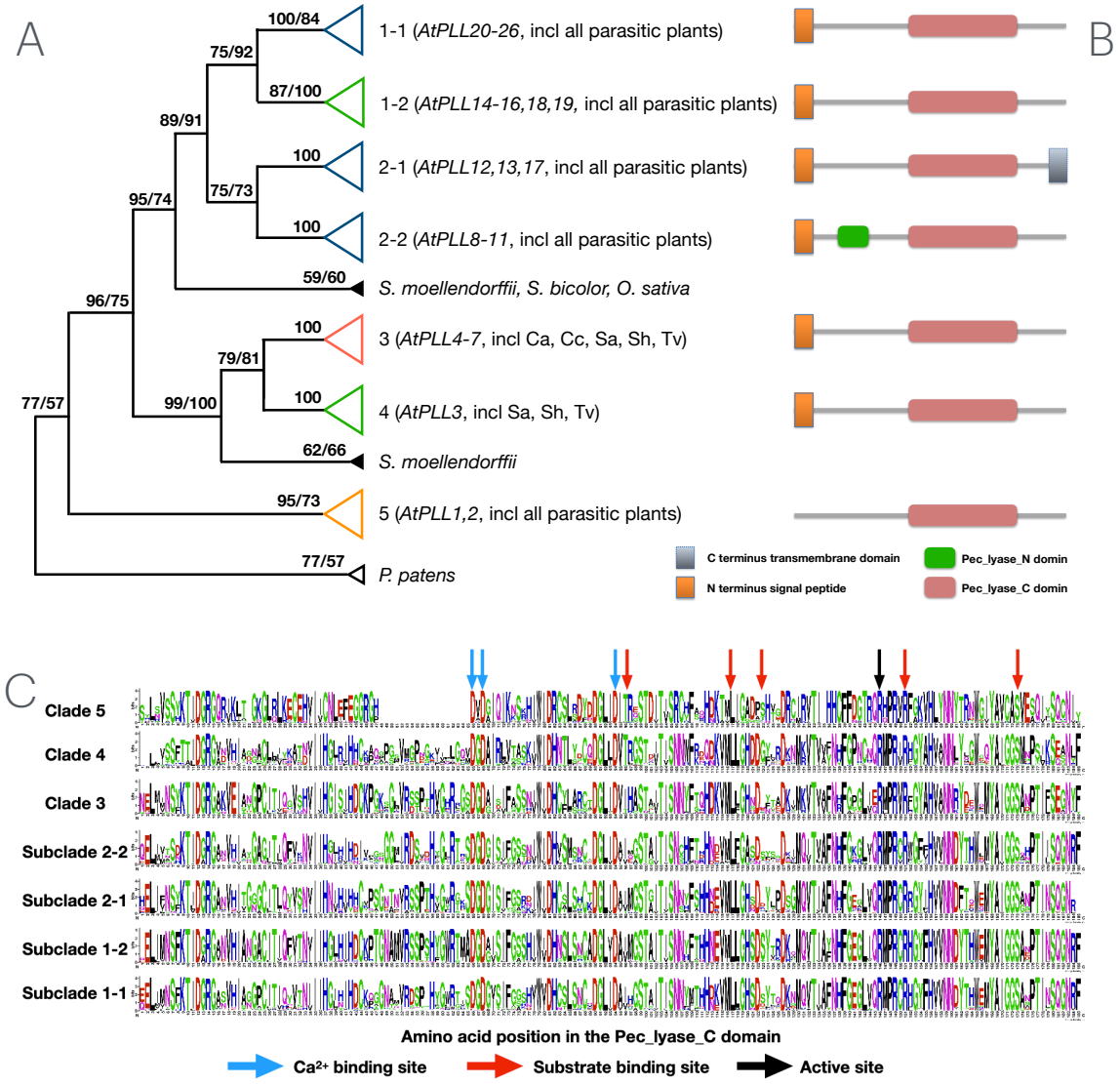

Figure 1. Phylogeny and domain structures of pectate lyase like (PLL) genes across land plants. (A). Simplified phylogenetic tree of the *PLL* gene family constructed with IQtree. Monophyletic groups were represented with triangles and colored according to the range of species present in the clade. Blue: Angiosperm only; Green: Seed plant (Angiosperm + Pine); Pink: Eudicot only; Yellow: Land plant wide; Black: others. Numbers at nodes are bootstrap values. Arabidopsis genes housed in each clade are listed in the parentheses following the clade name along with abbreviations of parasites (Ca: *Cuscuta australis*; Cc: *Cuscuta campestris*; Pa: *Phelipanche aegyptiaca*; Sa: *Striga*

asiatica; Sh: *Striga hermonthica*; Tv: *Triphysaria versicolor*). Numbers at the nodes are bootstrap values, where the number before '/' is value from tree constructed with only autotrophic plants and the number after is value from tree constructed with both autotrophic and parasitic plants. The list of plants used for the analysis is available in Supplemental Figure S1 and data sources are listed in Supplemental Table S1. (B). Domain structure of proteins in each clade. (C). LOGO plot showing a comparison of amino acid frequency at equivalent positions of the Pec_lyase_C domain among the seven major clades and subclades. Arrows above the image indicate functional residues predicted based on published PLL protein structure from previous studies.

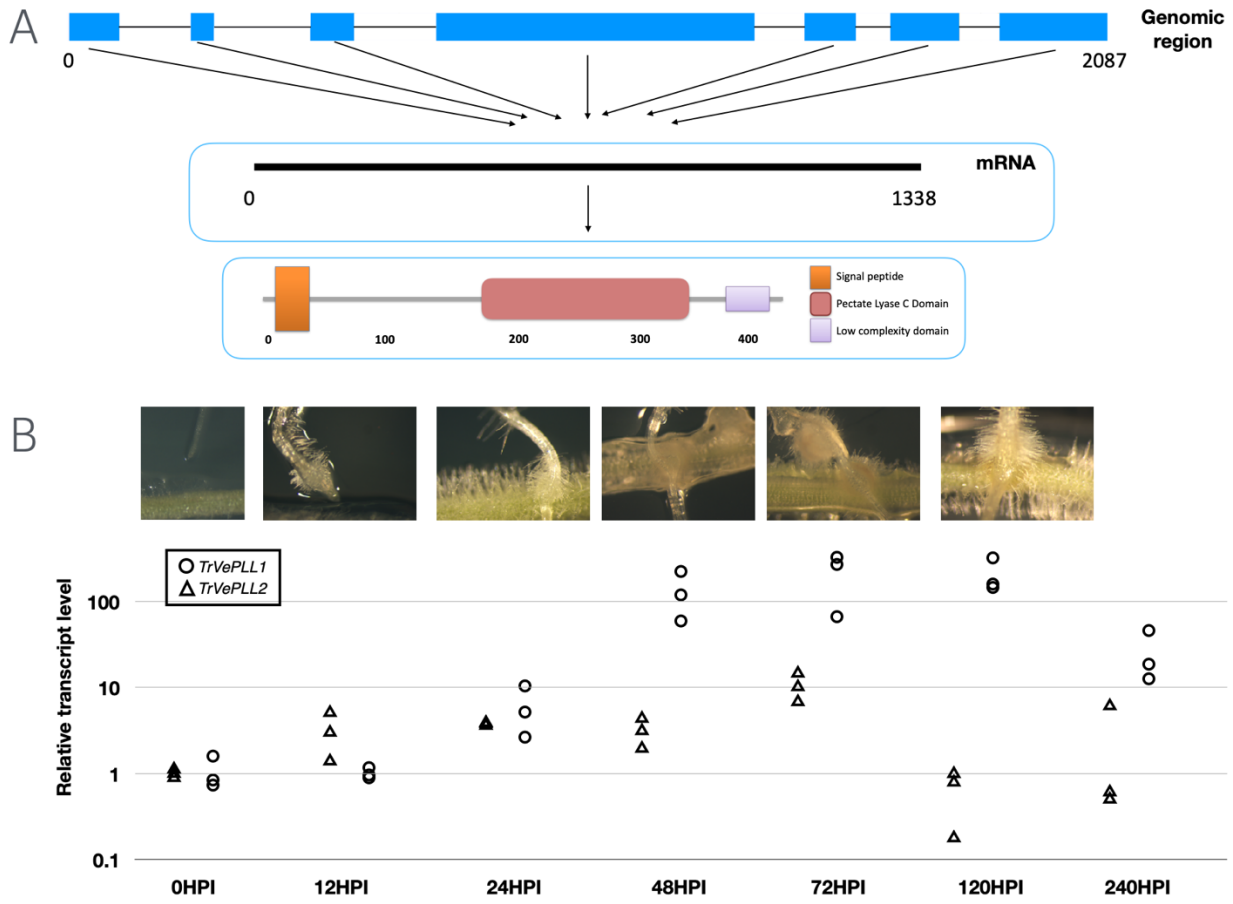

Figure 2. Gene model of *TrVePLL1* and its expression levels during haustorium development in *Triphysaria versicolor*. (A). The seven exons (Blue box) transcribe into a 1341nt mRNA, which translates into the TrVePLL1 protein that contains three predicted functional domains: a signal peptide (Orange), a Pec_lyase_C domain (Pink), and a low complexity domain (Purple). (B). Haustorium development from 0 hour post inoculation till maturation (120 HPI, hours post inoculation) and relative transcript level of *TrVePLL1* and *TrVePLL2*. Total RNA of three biological replicates each consisting of haustoria from 6-8 plants were used for qRT-PCR. Note the log scale y axis.

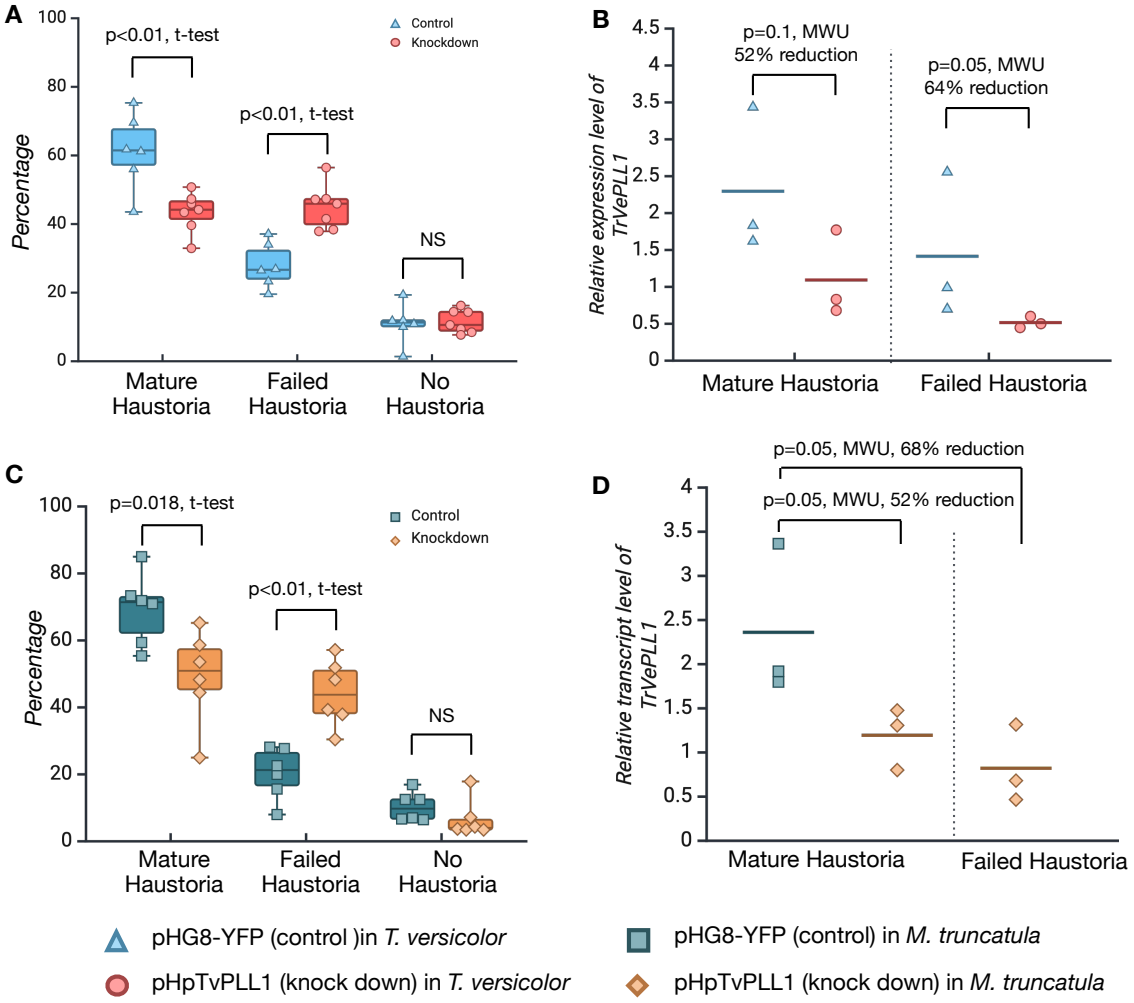

Figure 3. Haustorium formation rate and haustorium categorization in different co-culture experiments and quantification of *TrVePLL1* transcript level. *pHpTvPLL1* and the control vector *pHG8-YFP* were transformed to *Triphysaria* root (A) or *Medicago* root (C) using a hairy root transformation system. Six or seven biological replicates were used in (A) and (C), respectively. 50 to 60 transformed roots from 8-10 independent *Triphysaria* root cultures were analyzed in each biological replicate in (A). 80-100 wild type *Triphysaria* seedlings were co-cultured with 8-10 transformed *Medicago* roots in each replicate in (C). (B) and (D) show steady state transcript levels of *TrVePLL1* determined by qRT-PCR and normalized to the constitutively expressed gene *TrVePP2A*. 8 to10 haustoria were collected for each biological replicate and three biological replicates were generated for each treatment. Expression levels of *TrVePLL1* in *pHpTvPLL1* mature were set as 1. Horizontal lines among the data points represent the means. Statistical tests were performed and results are detailed in Supplemental Table S9. “t-test” stands for Student’s t-test, whereas “MWU” stands for Mann-Whitney U test. Note that no qRT-

PCR was performed on failed haustorium from wild type *Triphasaria* co-cultured with *Medicago* roots transformed with *pHpTvPLL1* due to the small amount of failed haustoria collected from this co-culture experiment.

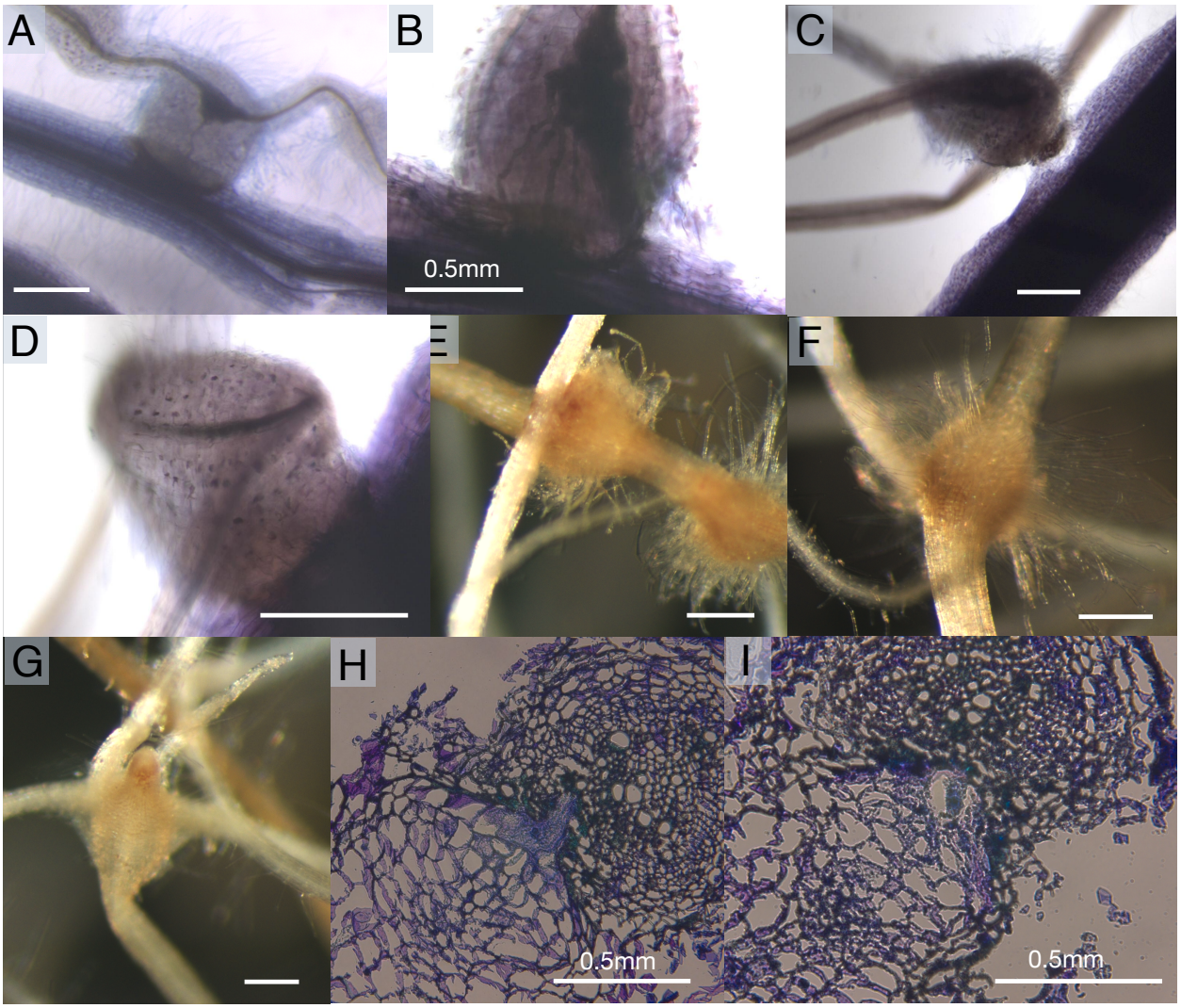

Figure 4. Morphological differences and similarities among haustoria formed by transgenic *Triphysaria* roots. (A)-(D): Ink and vinegar stained haustoria at 7 days post inoculation (DPI), showing difference in xylem connection. (E)-(G): Tetrazolium stained haustoria at 7 DPI. (E) and (F) showing a darker color in the parasite root, especially near the haustorial apex. (G) showed color accumulation at the parasite root tip.
(H) and (I): 7 DPI haustorium sections stained with Toluidine blue O (TBO).
(A), (C), (D), (F), (I) were collected from plants transformed with the hairpin construct, pHpTvPL1, whereas (B), (E), (G) and (H) were collected from plants transformed with the control vector, pHG8-YFP.

Supplemental Figure Legend

Supplemental Figure S1. Species tree of sampled plant taxa in the phylogenetic study. Star (*) indicates parasitic plant species used in the research. Green box presented an alternative species relationship in the Orobanchaceae family, the common ancestor of which was marked with a yellow star. Different colors represent different major land plant groups. Green: asterid; Red: Rosid; Purple: Basal eudicot; Yellow: Monocots; Blue: Basal angiosperm and gymnosperm; Black: Lycophyte and Moss.

Supplemental Figure S2. Phylogenetic trees of the PLL gene family constructed with only autotrophic plant PLL genes using IQtree (A) or RAxML (B) and with PLL genes from both autotrophic and parasitic plants using IQtree (C) or RAxML (D). Different colors represent different major land plant groups. Green: asterid; Red: Rosid; Purple: Basal eudicot; Yellow: Monocots; Blue: Basal angiosperm and gymnosperm; Black: Lycophyte and Moss.

Supplemental Figure S3. Comparison of PLL gene family trees and gene family analysis methods from various publications and this work. (A) Phylogeny from Sun and van Nocker, 2010; (B) Simplified phylogeny from McCarthy et al., 2014; (C) Simplified phylogeny from Sun et al., 2018; (D) Simplified phylogeny from Zheng et al., 2018; and (E) Simplified phylogeny from this work. (F) Summary and comparison of gene family analysis methods from the selected publications. Phylogenies in (B) to (E) are simplified to only show the overall structure with *Arabidopsis* genes for comparison purposes. Bootstrap (bs) values at the nodes represent the bs value of the largest monophyletic group housing the *Arabidopsis* genes shown at the tip. (C) and (D) lack bs values from the original publications. Colored backgrounds indicate clades identified by the Sun and van Nocker, 2010 work, whereas clades on the right side of each phylogeny shows the corresponding clades identified by this study.

Supplemental Figure S4. Normalized transcript level (Transcript per Million, TPM) of *T. versicolor* genes from *PLL* Subclass 1-1. Triplicated PPGP II transcript data was used for (A) to (C), and PPGP I transcript data was used for (D). X axis showing developmental stages (Stage abbreviated as S in the x axis legend). Blue boxes highlighting haustorial stages. *TrVePLL1* is the only *PLL* with principally haustorial expression.

Supplemental Figure S5. *Triphysaria versicolor* (A-F) and *Medicago truncatula* (G-J) transformed with pHG8-YFP. (A), (C), (D), (E), (G), and (I) are epifluorescence photos showing YFP expression. (B), (F), (H), and (J) are bright field photos. (A) and (B) showing a callus expressing YFP signal 2 weeks post *Agrobacterium* inoculation. (C) showing a single root developed from the transgenic callus 3 weeks post *Agrobacterium*

inoculation. (D) showing transgenic root proliferation after removing root tip. (E) and (F) showing transgenic *Triphysaria* root forming haustorium on *Medicago* root 5 days post co-culture. (G) and (F) showing 6 week old transgenic *Medicago* root without parasitization by *Triphysaria*. (I) and (J) are *Triphysaria* roots making haustorium on transgenic *Medicago* roots. C: callus; T: *Triphysaria versicolor* root; M: *Medicago truncatula* root; H: haustorium. Bar: 1mm, unless otherwise indicated.

Supplemental Figure S6. Polymerase chain reaction of *TrVePLL1* using DNA from *Medicago truncatula* roots. TY: transgenic root showing YFP signal; TnY: transgenic root not showing YFP signal; P: positive control, in which the *pHpTvPLL1* construct was used as template; NTC: no template control; nT: non transgenic root. Ladder: from the top to bottom band size: 650bp, 500bp, 400bp, 300bp, 200bp, 100bp.

Supplemental Figure S7. A heat map showing the expression level of all *PLL* genes from PPGP II *T. versicolor* BC4. Note that *TrVePLL1* is missing in BC4, thus not included in this figure. Blue star on the right indicates this gene is differentially expressed in haustorial tissues. Yellow star on the right indicates the highest expression level of this gene is found in haustorial tissues.