Journal of Experimental Botany, Vol. 66, No. 20 pp. 6259–6271, 2015
doi:10.1093/jxb/erv333 Advance Access publication 10 July 2015

RESEARCH PAPER

Combined transcriptome and metabolite profiling reveals that *li*PLR1 plays an important role in lariciresinol accumulation in *Isatis indigotica*

Ying Xiao[1,2], Qian Ji[3], Shouhong Gao[2], Hexin Tan[3], Ruibing Chen[3], Qing Li[2], Junfeng Chen[2], Yingbo Yang[1], Lei Zhang[3,*], Zhengtao Wang[1,*], Wansheng Chen[2,*] and Zhibi Hu[1]

[1] The MOE Key Laboratory of Standardization of Chinese Medicines, Institute of Chinese Materia Medica, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China
[2] Department of Pharmacy, Changzheng Hospital, Second Military Medical University, Shanghai 200003, China
[3] Department of Pharmaceutical Botany, School of Pharmacy, Second Military Medical University, Shanghai 200433, China

*To whom correspondence should be addressed. E-mail: zhanglei@smmu.edu.cn, ztwang@shutcm.edu.cn, or chenwansheng@smmu.edu.cn

Received 11 March 2015; Revised 24 May 2015; Accepted 11 June 2015

Editor: Qiao Zhao

Abstract

A lignan, lariciresinol, is an important efficacious compound for the antiviral effect of *Isatis indigotica*, a widely used herb for the treatment of colds, fever, and influenza. Although some rate-limiting steps of the lariciresinol biosynthetic pathway are well known, the specific roles of gene family members in *I. indigotica* in regulating lariciresinol production are poorly understood. In the present study, a correlation analysis between the RNA sequencing (RNA-Seq) expression profile and lignan content by using *I. indigotica* hairy roots treated with methyl jamonate (0.5 μM) at different time points as a source implicated that *I. indigotica* pinoresinol/lariciresinol reductase 1 (*li*PLR1), but not *li*PLR2 or *li*PLR3, contributed greatly to lariciresinol accumulation. Gene silencing by RNA interference (RNAi) demonstrated that *li*PLR1 indeed influenced lariciresinol biosynthesis, whereas suppression of *li*PLR2 or *li*PLR3 did not change lariciresinol abundance significantly. *li*PLR1 was thus further characterized; *liPLR1* was constitutively expressed in roots, stems, leaves, and flowers of *I. indigotica*, with the highest expression in roots, and it responds to different stress treatments to various degrees. Recombinant *liPLR1* reduces both (±)-pinoresinol and (±)-lariciresinol efficiently, with comparative K_{cat}/K_m values. Furthermore, overexpression of *liPLR1* significantly enhanced lariciresinol accumulation in *I. indigotica* hairy roots, and the best line (ovx-2) produced 353.9 μg g^{-1} lariciresinol, which was ~6.3-fold more than the wild type. This study sheds light on how to increase desired metabolites effectively by more accurate or appropriate genetic engineering strategies, and also provides an effective approach for the large-scale commercial production of pharmaceutically valuable lariciresinol by using hairy root culture systems as bioreactors.

Key words: Correlation analysis, *Isatis indigotica*, lariciresinol, lignan, pinoresinol/lariciresinol reductase, RNA-Seq expression profile.

Abbreviations: AACT, acetoacetyl CoA thiolase; ABA, abscisic acid; CAD, cinnamyl alcohol dehydrogenase; CCoAOMT, caffeoyl CoA *O*-methyltransferase; CCR, cinnamoyl CoA reductase; C3H, *p*-coumarate 3-hydroxylase; C4H, cinnamate 4-hydroxylase; 4CL, 4-(hydroxy)cinnamoyl CoA ligase; DIR, dirigent protein; dw, dry weight; FPS, farnesyl diphosphate synthase; fw, fresh weight; HMGR, HMG-CoA reductase; ht-*li*PLR1, His-tag fused *li*PLR1; MeJA, methyl jamonate; MRM, multiple reaction monitoring; PAL, phenylalanine ammonia-lyase; PLR, pinoresinol/lariciresinol reductase; PrR, pinoresinol reductase; qRT-PCR, real-time quantitative PCR; SARS, severe acute respiratory syndrome; SIRD, secoisolariciresinol dehydrogenase; SMMU, Second Military Medical University; SQE, squalene epoxidase; UGT, UDP-sugar-dependent glycosyltransferase.

Introduction

Isatis indigotica Fort. (*Isatis tinctoria*), belonging to the family Cruciferae, is a prevalent Chinese medicinal herb (Zhao, 2007). The root of *I. indigotica* (*Radix Isatidis*), named 'Ban–Lan–Gen' as a traditional Chinese medicine, has been used for the treatment of influenza, epidemic hepatitis, and epidemic encephalitis B for >1000 years (Wang *et al.*, 2000). During the epidemic period of severe acute respiratory syndromes (SARS) and H1N1 influenza, Ban–Lan–Gen demonstrated its significant antiviral effect (Lin *et al.*, 2005; Sui, 2010; Wang *et al.*, 2011). However, the antiviral compounds of *I. indigotica* were unclear until Li (2003) determined that lariciresinol isolated from this plant was useful for the treatment of influenza A1 virus. More recently, Yang *et al* (2013) determined that a lariciresinol derivative, clemastanin B [7S,8R,8'R-(–)-lariciresinol-4,4′-bis-*O*-β-D-glucopyranoside], significantly inhibited human and avian influenza A and B viruses.

Lariciresinol is a type of lignan which has been reported to possess a number of biological activities, including antimicrobial, antioxidant, anti-inflammatory, and antioestrogenic properties (Saleem *et al.*, 2005). However, lariciresinol accumulates at a low abundance in plants. For example, its content in *I. indigotica* is only 68 μg g^{-1} of dry weight (dw) (Li, 2003), which not only constrains the yield of lariciresinol for the pharmaceutical industry, but also affects the use of the *I. indigotica* herb itself. Metabolic engineering provides an attractive approach for increasing the production of such health-promoting compounds in plants.

Plant metabolic pathways are complex and often feature multiple levels of regulation, making the consequences of metabolic engineering difficult to predict. The existence of multicopy gene families in the metabolic pathway suggests divergent functional roles for pathway components involved in the biosynthesis of metabolites. For example, in *Arabidopsis thaliana*, six squalene epoxidases (SQEs) were predicted based on a homology search (SQE1–SQE6) (Benveniste, 2002; Rasbery *et al.*, 2007; Posé *et al.*, 2009); however, only the products of SQE1, SQE2, and SQE3 were able to functionally complement the *Saccharomyces cerevisiae* squalene epoxidase mutant *erg1* (Rasbery *et al.*, 2007). Recently, only SQE3, but not SQE2, was found to functionally complement SQE1 in the promotion of plant squalene epoxidase activity (Laranjeira *et al.*, 2015). At the early stages of *Arabidopsis* seed development, embryo viability relies on farnesyl diphosphate (FPP) or an FPP-derived isoprenoid/sterol precursor provided by farnesyl diphosphate synthase 1 (FPS1) gene products, but not those of FPS2 (Keim *et al.*, 2012). In contrast, *Arabidopsis* HMG-CoA reductase 1 (HMGR1) and HMGR2 are both responsible for the biosynthesis of triterpenes (Ohyama *et al.*, 2007). *Arabidopsis* acetoacetyl CoA thiolase 1 (AACT1) and AACT2 can both complement the lethal *aact* deletion allele (DERG10) in yeast despite displaying different spatial and temporal expression patterns (Jin *et al.*, 2012). With respect to lariciresinol biosynthesis, although some enzymatic steps have been proposed to be particularly important for activating the lariciresinol biosynthetic pathway (Fig. 1) (Humphreys and Chapple, 2002; Nakatsubo *et al.*, 2008; Satake *et al.*, 2013), the specific roles of gene family members in regulating this metabolic flux have rarely been investigated. *PLR*, a gene encoding pinoresinol/lariciresinol reductase, represents one of the most important rate-limiting steps of the lariciresinol biosynthetic pathway (Satake *et al.*, 2013). Hemmati *et al.* (2010) reported that both *Lu*PLR1 and *Lu*PLR2 are expressed in flax seed tissues, whereas only *Lu*PLR2 is expressed in stem and leaf tissues. Zhao *et al.* (2014) reported that pinoresinol reductase 1 (PrR1), but not PrR2, impacts lignin distribution during secondary cell wall biosynthesis in *Arabidopsis.* These studies indicated that PLR family members in plants appeared to play divergent functional roles.

Under the umbrella of a transcription profiling of *I. indigotica* (Chen *et al.*, 2013), 51 biosynthetic genes involved in the lariciresinol pathway were first identified in *I. indigotica* in the present study. Correlation analysis based on profiling changes of transcript abundance and accumulation pf metabolites under methyl jamonate (MeJA) stress indicated that *Ii*PLR1 was positively correlated with lariciresinol, whereas *Ii*PLR2 or *Ii*PLR3 were not. This hypothesis was demonstrated by examining lariciresinol production when *Ii*PLR1, *Ii*PLR2, or *Ii*PLR3 was silenced. *Ii*PLR1 was then well characterized, and the effect of *Ii*PLR1 overexpression on lariciresinol accumulation was also investigated.

Materials and methods

Plant material

The *I. indigotica* plant, grown in the gardens of the Second Military Medical University (SMMU), Shanghai, China, was identified by Professor Hanming Zhang, School of Pharmacy, SMMU. The harvested *I. indigotica* seeds were pre-treated with 75% alcohol for 1 min, washed three times with distilled water, followed by treatment with 0.1% $HgCl_2$ for 5 min and by four rinses with sterile distilled water. The sterilized seeds were then incubated between several layers of sterilized wet filter paper and cultured on Murashige and Skoog (MS) basal medium for germination. The seedlings were grown at 25 °C under 12 h light/12 h dark photoperiod cycles.

Transcriptome analysis

The assembly of *I. indigotica* transcriptome sequences (Chen *et al.*, 2013) was searched for putative genes involved in the lariciresinol biosynthetic pathway. Protein sequences, nucleotide sequences, and expressed sequence tag (EST) records of target genes of other plants were downloaded from the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/). All the sequences were used as queries to search the *I. indigotica* transcription profiling database through the basic local alignment search tool (TBLASTN or BLASTN) to determine candidate gene sequences with an e-value <1e-5. The Pfam database (http://pfam.janelia.org/) (Punta *et al.*, 2012) was used to screen the above putative sequences and identify the conserved protein domains with default parameters. Blastx alignment (e-value <1e-5) between the resulting sequences and protein databases such as the NCBI non-redundant protein database, the Swiss-Prot protein database, the Kyoto Encyclopedia of Genes and Genomes pathway database, and the Cluster of Orthologous Groups database was performed, and the best aligning results were used to determine the target genes and their functional annotations. Sequences corresponding to the target genes involved in the lariciresinol biosynthetic pathway are listed in Supplementary Table S1 available at *JXB* online.

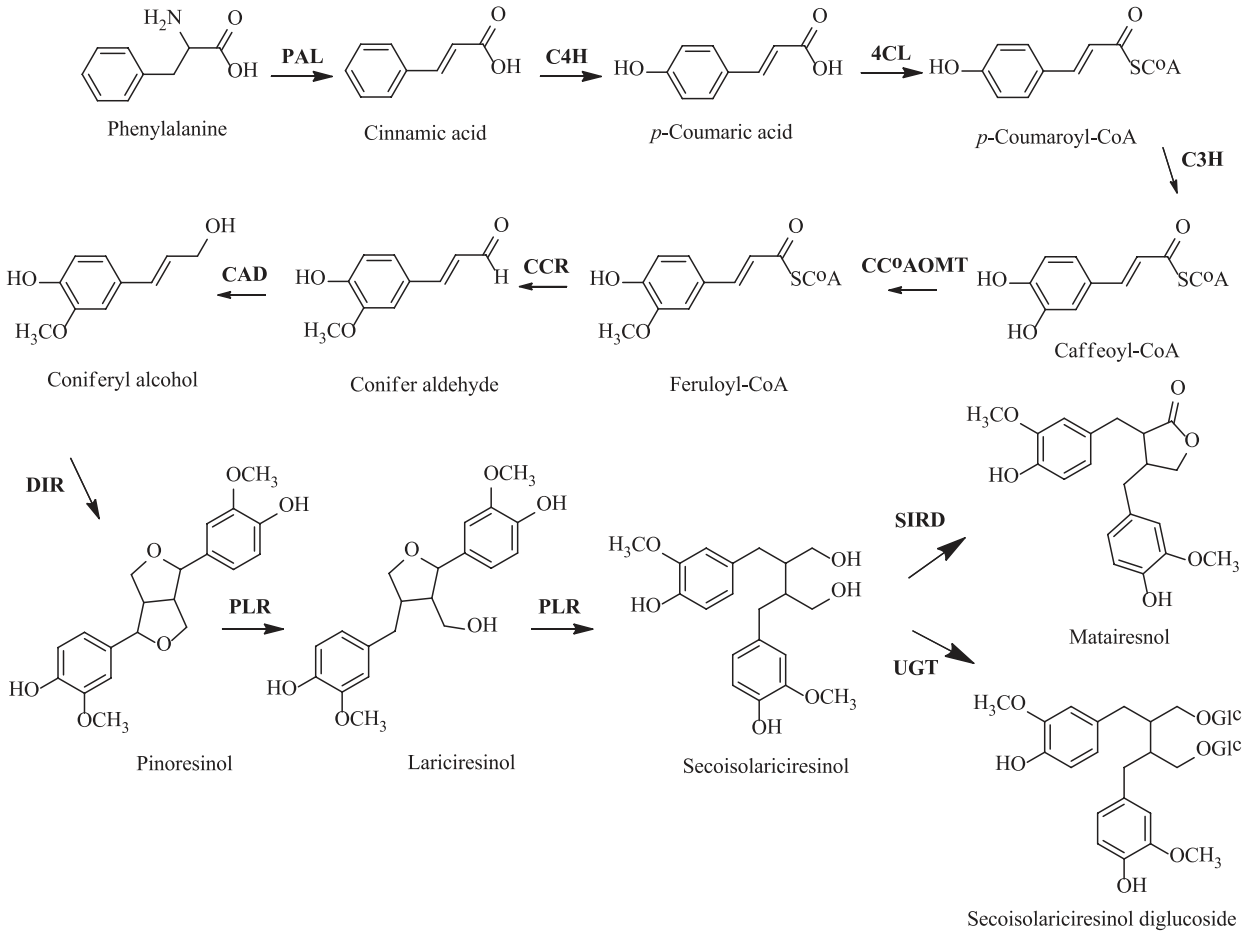

Fig. 1. Lariciresinol biosynthetic pathway and involved genes. PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-(hydroxy) cinnamoyl CoA ligase; C3H, *p*-coumarate 3-hydroxylase; CCoAOMT, caffeoyl CoA O-methyltransferase; CCR, cinnamoyl CoA reductase; CAD, cinnamyl alcohol dehydrogenase; DIR, dirigent protein; PLR, pinoresinol/lariciresinol reductase; SIRD, secoisolariciresinol dehydrogenase; UGT, UDP-sugar-dependent glycosyltransferase.

To gain insight into the gene expression patterns in response to MeJA treatment, the Illumina RNA squencing (RNA-Seq) data provided by Chen *et al.* (2013) were utilized. The RNA-Seq expression profile data were generated using the Illumina HiSeq™ 2000 platform (100 bp, TruSeq SBS kit v3-HS 200 cycles; Illumina), and included the hairy roots of *I. indigotica* treated with MeJA (0.5 μM) at different time points (0, 1, 3, 6, 12, and 24 h). 0 h was used as the control to normalize the expression level of the other time points. A heat map was constructed using the log2-transformed and normalized expression level data in MultiExperiment Viewer (Saeed *et al.*, 2003). Genes with different expression patterns were grouped by hierarchical clustering.

Metabolite analysis

MeJA-treated *I. indigotica* hairy roots were freeze-dried and ground into powder. The dried powder sample (100 mg) was extracted twice with methanol (20 ml) under sonication for 30 min, and then centrifuged at 4500 rpm for 5 min. The supernatant was concentrated to 10 ml total volume, and the extract solution was then filtered through a 0.2 μm membrane filter (Woongki Science, Seoul, Korea) before analysis. Lignan contents were determined using an Agilent 6410 triple quadrupole mass spectrometer (Agilent, USA) coupled to an Agilent 1200 high-pressure liquid chromatograph equipped with a pump (Agilent 1200 G1311A) and an autosampler (Agilent G1329A). Chromatographic separation was performed with a Thermo® C18 column (150 mm×2.1 mm, id, 3.5 μm particle size; Agilent). A mobile phase consisting of acetonitrile (solvent A) and 5 mM ammonium acetate (solvent B) was used, with the flow rate set at 0.3 ml min^{-1}. A gradient run of 0–4 min in 15% solvent A, 4–4.5 min in 50% solvent A, and finally 4.5–8.5 min in 85% solvent A was determined as optimal. Multiple reaction monitoring (MRM) mode was used for the quantification, and the selected transitions of *m/z* were 357→151 for pinoresinol, 359→329 for lariciresinol, 361→164 for secoisolariciresinol, 357→121 for matairesisnol, and 685→523 for secoisolariciresinol diglucoside. All standards were purchased from Sigma-Aldrich (St. Louis, MO, USA).

Correlation analysis of genes and metabolites

Correlations between the identified biosynthetic genes and five lignans were calculated using the Pearson correlation coefficient using R based on the co-occurrence principle between mRNA and metabolite levels (Saito and Matsuda, 2010). The heat maps of lignans and genes were generated by two-way hierarchical clustering using R. The genes were selected by the criterion that they were correlated with at least one lignan ($|r|>0.6$). The correlation network was generated using Cytoscape (Cline *et al.*, 2007).

Generation and analyses of liPLR1-, liPLR2-. *and* liPLR3-*silenced hairy roots*

IiPLR1, *IiPLR2*, and *IiPLR3* were each suppressed by RNA interference (RNAi) technology. First, the nucleotide sequences of *IiPLR1*, *IiPLR2*, and *IiPLR3* retrieved from RNA-Seq (Supplementary Table S1 at *JXB* online) were aligned using CLUSTALX version 2.0.12 (Larkin *et al.*, 2007), and the conserved sites were shaded

using GeneDoc (Nicholas and Nicholas, 1997). Secondly, according to the alignment results, an appropriate 400–500 bp fragment (Supplementary Table S2) of the coding region of each *IiPLR* member was selected and amplified by PCR using primers *IiPLR1*-s*Nco*I-a*Sal*I and *IiPLR1*-s*Kpn*I-a*Bam*HI for *IiPLR1*; *IiPLR2*-s*Nco*I-a*Sal*I and *IiPLR2*-s*Kpn*I-a*Bam*HI for *IiPLR2*; and *IiPLR3*-s*Nco*I-a*Sal*I and *IiPLR3*-s*Kpn*I-a*Bam*HI for *IiPLR3* (Supplementary Table S3) containing two sets of restriction sites at the 5' end. The PCR product was cloned into pGEM-T vector, sequenced, and then subcloned in opposite orientations on either side of the Pdk intron of the pCAMBIA-1300-pHANNIBAL vector (Wesley *et al.*, 2001) to generate plasmids p*IiPLR1*-RNAi, p*IiPLR2*-RNAi, and p*IiPLR3*-RNAi.

After sequencing confirmation, plasmids p*IiPLR1*-RNAi, p*IiPLR2*-RNAi, and p*IiPLR3*-RNAi, together with pCAMBIA-1300-pHANNIBAL as vector control (CK), were introduced separately into *I. indigotica* leaf explants by using the *Agrobacterium tumefaciens* C58C1 strain, and the generated hairy root lines were screened using hygromycin (100 mg l^{-1}). Some transgenic roots turned brown and aged considerably faster than wild-type root cultures (hairy root lines generated through transformation with the blank C58C1 strain, WT). These lines were discarded, and the remaining hairy root lines were subcultured for 45 d in hormone-free, half-strength MS liquid medium, for DNA extraction, RNA extraction, metabolite analysis, and phloroglucinol-HCl staining.

Genomic DNA was isolated from hairy root samples using the cetyltrimethyl ammonium bromide method (Doyle, 1990). The DNA was then used in PCR analysis for detecting the presence of the inserted *IiPLR* fragment by primers JDPDK-1F and JDPDKR. The transformed status of hairy roots was also verified by PCR for the genes *hpt* and *rolb* (Chilton, 1982). The primer sequences are listed in Supplementary Table S3 at *JXB* online. The PCR was performed under the following conditions: initial denaturation at 94 °C for 3 min followed by 35 cycles of denaturation at 94 °C for 10 s, annealing at 58 °C for 30 s, extension at 72 °C for 1 min, with a final extension at 72 °C for 3 min.

The expression profiles of *IiPLR1*, *IiPLR2*, and *IiPLR3* were investigated by using real-time quantitative PCR (qRT-PCR) analysis. Total RNAs from hairy root samples were extracted using TRIzol Reagent (GIBCO BRL) (Jaakola *et al.*, 2001) and then reverse transcribed to generate cDNA as a template. Gene-specific primers (*IiPLR1*-qRT-F and *IiPLR1*-qRT-R; *IiPLR2*-qRT-F and *IiPLR2*-qRT-R; *IiPLR3*-qRT-F and *IiPLR3*-qRT-R) were designed using Primer Premier 5.0 and optimized using oligo 7; the *I. indigotica* actin gene reported by Li *et al.* (2014) was used as the internal reference gene (Supplementary Table S3 at *JXB* online). The qRT-PCR was performed according to the manufacturer's instructions (Takara, Japan) under the following condition: 30 s pre-denaturation at 95 °C, one cycle; 10 s denaturation at 95 °C, 20 s annealing at 58 °C, 20 s collection fluorescence at 72 °C, 40 cycles. Quantification of gene expression was done with the comparative CT method.

Lariciresinol accumulations were analysed by liquid chromatography–mass spectrometry as described above.

The hairy root samples were stained with phloroglucinol-HCl for 20 min for detection of lignans, lignins, or phenolic derivatives according to Hano *et al.* (2006).

liPLR1 *gene isolation and sequence analysis*

Total RNAs of *I. indigotica* seedlings were extracted using TRIzol Reagent (GIBCO BRL) (Jaakola *et al.*, 2001) and then reverse transcribed to generate cDNA. Gene-specific primers for full-length *IiPLR1* were designed based on the corresponding sequence retrieved from RNA-Seq (Supplementary Table S1 at *JXB* online). *IiPLR1* was amplified by PCR using Pfu Ultra DNA polymerase (TransGen Biotech), the gene-specific primers *IiPLR1*-ORF-F and *IiPLR1*-ORF-R (Supplementary Table S3), and the cDNA as a template. The amplified PCR product was purified and cloned into the PMD18-T vector and then sequenced.

Translation of the open reading frame (ORF) and molecular mass calculation of the predicted protein were carried out on Vector NTI Suite 8. The amino acid sequence alignments of *Ii*PLR1 with other known PLRs (Supplementary Table S4 at *JXB* online) were performed using CLUSTALX version 2.0.12 (Larkin *et al.*, 2007), and the conserved sites were shaded using GeneDoc (Nicholas and Nicholas, 1997). Phylogenetic relationships were analysed using the Neighbor–Joining method with the pairwise deletion option in MEGA 5.05. Tree reliability was estimated using a bootstrap analysis of 1000 replicates (Tamura *et al.*, 2011).

Expression profile of liPLR1 *in different tissues and under various treatments*

Leaves of 2-month-old *I. indigotica* seedlings were sprayed with 100 μM MeJA or 100 μM abscisic acid (ABA), and then sampled at 0, 2, 4, 6, 8, 12, and 24 h post-treatment. For UV-B treatment, the seedlings were exposed to 1500 J m^{-2} UV-B light for 30 min, and then sampled at 0, 5, 10, and 30 min during treatment, and at 30, 60, and 120 min post-treatment.

The expression profile of *IiPLR1* in different tissues (roots, stems, leaves, and flowers) and under various treatments was investigated by using qRT-PCR analysis as described above.

Enzyme assay

Full-length *IiPLR1* cDNA was cloned into plasmid pET32a(+) (Novagen, USA) using *Nco*I/*Sac*I restriction sites to generate the *IiPLR1*-pET construct. The sequences of primers *IiPLR1*-pET-F and *IiPLR1*-pET-R are given in Supplementary Table S3 at *JXB* online. After sequencing confirmation, the *IiPLR1*-pET construct was transformed into *Escherichia coli* BL21(DE3) cells for fusion protein expression. *Escherichia coli* BL21(DE3) cells harbouring *IiPLR1*-pET were grown in LB medium at 30 °C. When the A_{600} of the cultures reached 0.6, expression of the gene was induced for 12 h by adding 1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) (Merck, Germany). Protein purification was performed using a His Spin Trap column following the manufacturer's instructions (GE Healthcare). The purity of the His-tag-fused *Ii*PLR1 (ht-*Ii*PLR1) was examined by 12% SDS–PAGE, and the protein concentration was determined by the Bradford method (Bradford, 1976) with bovine serum albumin (BSA) as the standard.

*Ii*PLR1 activity was assayed according to Fukuhara *et al.* (2013) with slight modifications. The assay mixtures (1 ml) consisted of TG buffer (50 mM TRIS-HCl, 10% glycerol; pH 7.0), 150 μM NADPH, 200 μM (±)-pinoresinol, or 3.5 μM (±)-lariciresinol and 5 μg of purified ht-*Ii*PLR1. Protein, buffer, and substrate were pre-incubated for 5 min at 30 °C. The enzyme reaction was initiated by addition of NADPH and terminated by addition of 300 μl of ethyl acetate after 30 min. The assays were extracted with ethyl acetate (3 × 300 μl in total). The combined ethyl acetate phases were dried under vacuum. The residue was dissolved in 1000 μl of methanol and subjected to LC-MS analysis (see 'Metabolite analysis' above). For determination of V_{max} and K_m values, 10 different concentrations of (±)-pinoresinol (between 10 μM and 500 μM) or (±)-lariciresinol (between 0.5 μM and 20 μM) as substrates and 1 μg of purified ht-*Ii*PLR1 were used. Incubations were carried out at 30 °C for 5 min (within the linear kinetic range). The assay mixtures without ht-*Ii*PLR1 were used as controls. The consumption of substrates was calculated for kinetic analysis. V_{max} and K_m values were determined from Lineweaver–Burk plots, and k_{cat} was determined by dividing V_{max} by the enzyme concentration.

Generation and analyses of liPLR1*-overexpressing hairy roots*

To construct the *IiPLR1* overexpression vector, full-length *IiPLR1* cDNA was cloned into plasmid PHB-flag using *Bgl*II/*Xba*I restriction sites to generate the PHB-*IiPLR1*-flag construct. The sequences of primers *IiPLR1*-ovx-F and *IiPLR1*-ovx-R are given in Supplementary Table S3 at *JXB* online.

After sequencing confirmation, plasmid PHB-*IiPLR1*-flag, together with PHB-flag as vector control (CK), were separately introduced into *I. indigotica* leaf explants by using the *A. tumefaciens* C58C1 strain and the hairy root lines generated were screened using hygromycin (100 mg l^{-1}). The vigorous hairy roots (~100 mg) were inoculated into 150 ml conical flasks containing 40 ml of hormone-free, half-strength MS liquid medium, cultured on an orbital shaker (120 rpm) at 25 °C in the dark, and the culture medium was routinely refreshed every 9 d. The fresh weight (fw) of root tissues from flasks of cultures (determined as the difference between the whole flask with and without the harvested root tissues) was recorded at days 9, 18, 27, 36, and 45 after inoculation. At day 45, the hairy roots were harvested for DNA extraction, RNA extraction, protein extraction, metabolite analysis, and phloroglucinol-HCl staining.

Genomic DNA was isolated and then used in PCR analysis for detection of the presence of the inserted *IiPLR1* fragment. Primers *IiPLR1*-ovx-1F and *IiPLR1*-ovx-1R (Supplementary Table S3 at *JXB* online) were designed specifically to cover the gene sequence and the vector sequence for detecting exogenous *IiPLR1* transformations. The transformed status of hairy roots was also verified by PCR for the genes *hpt* and *rolb* (Chilton, 1982).

PCR-positive hairy roots harbouring the *IiPLR1* gene were also analysed for *IiPLR1* expression by western blot assay. Protein was extracted from hairy roots by grinding in ice-cold lysis buffer (50 mM HEPES/pH 7.5, 150 mM NaCl, 1 mM EDTA/pH 8.0, 1% Triton X-100, 1% SDS) supplemented with a cocktail of protease inhibitors (Roche). Aliquots of 30 μg of total soluble proteins per sample, determined by the Bradford method (Bradford, 1976), were fractionated on a 12% SDS–PAGE mini-gel and blotted to a nitrocellulose membrane. The blots were probed with the DYKDDDDK Tag antibody (Mouse) (1:1000, Santa Cruz Biotechnology), washed in PBST three times, reacted with goat anti-mouse IgG conjugated with horseradish peroxidase (1:5000, Amersham), washed, and exposed to X-ray film using the enhanced chemiluminescence method according to the manufacturer's instructions (Amersham). For actin detection, anti-actin goat polyclonal IgG (Santa Cruz) and anti-goat IgG, were used as a primary and a secondary antibody at a final dilution of 1:1000 and 1:5000, respectively.

IiPLR1 transcript and lariciresinol content determination, as well as phloroglucinol-HCl staining, were performed following the methods described above.

Statistical analysis

Experiments were performed in triplicate, and all results are expressed as the mean ±SEM.

Results

Gene–lignan pathway network analysis

Fifty-one ESTs corresponding to lariciresinol biosynthetic genes (Supplementary Table S1 at *JXB* online) were identified from *I. indigotica* through analysis of transcriptome data. To assess the impact of the expression of these genes on lignan production, variation in gene expression and lignan accumulation was exploited by using MeJA-elicited *I. indigotica* hairy roots harvested at different time points (0, 1, 3, 6, 12, and 24 h after treatment). The RNA-Seq expression profile showed that almost all tested genes except *IiDIR15* responded to MeJA treatment, indicating their important role in MeJA-mediated transcriptional reprogramming. These genes responded to MeJA with different patterns; some (e.g. *IiDIR1* and *Ii4CL12*) were dramatically up-regulated, some (e.g. *Ii4CL11* and *IiCCoAOMT2*) were down-regulated, and others (e.g. *IiDIR6* and *IiDIR7*) showed 'up and down' changes but without clear trends (Fig. 2A). Since jasmonates orchestrate a diverse set of physiological and developmental processes in plants (Pauwels *et al.*, 2008), it is suggested that the various response patterns might be caused by several different regulatory mechanisms. The lignan content determination by LC-MS showed, as expected, that the signalling molecule MeJA greatly triggered lignan production but with different patterns. Each lignan showed wide ranges of concentrations in these hairy roots, indicating that the MeJA-elicited *I. indigotica* hairy roots contain significant lignan variation. For example, lariciresinol accumulation first gradually decreased, with a minimum amount of 36.4 μg g^{-1} dw at 3 h after treatment, and then gradually increased and peaked at 24 h with an amount of 77.0 μg g^{-1} dw (Fig. 2B).

A correlation matrix of all pairwise comparisons between gene expression and lignan concentrations was created, and the full set of correlation coefficients (Supplementary Table S5 at *JXB* online) is presented as a heat map in Fig. 3A, which makes the genes positively correlated (red), negatively (blue) correlated, or uncorrelated (yellow), with each lignan easily distinguishable. For example, *Ii*PLR1 is found to be positively correlated with the pharmaceutically important lariciresinol, with a high correlation coefficient value (r=0.98), whereas *Ii*PLR2 and *Ii*PLR3 seem uncorrelated with any of the measured lignans.

Correlation coefficient cut-off values were applied to construct gene–metabolite correlation networks. Figure 3B presents one example with a cut-off $|r|>0.6$: a total of 27 unigenes could be correlated with one or more of the five measured lignans, and among them seven (*Ii*PAL1, *Ii*4CL7, *Ii*C3H, *Ii*DIR1, *Ii*DIR5, *Ii*DIR9, and *Ii*PLR1) were positively correlated with lariciresinol (Table 1).

RNAi silencing of IiPLR1, IiPLR2, *and* IiPLR3 *in hairy roots*

*Ii*PLR family members (*IiPLR1*, *IiPLR2* and *IiPLR3*) were selected to validate the predicted gene–metabolite correlation. A 400–500 bp fragment (Supplementary Table S2 at *JXB* online) selected from the unconserved region of each member (Supplementary Fig. S1) was inserted upstream and downstream of the Pdk intron in both sense and antisense orientations to generate the RNAi vector (Fig. 4A). The transformants were identified by PCR analysis: all of the hairy roots contained the *rolb* gene, which was evidence of transformation by pRiA4 (Chilton, 1982). The hygromycin resistance gene *hpt* was detected in both RNAi (*IiPLR1*-RNAi, *IiPLR2*-RNAi, and *IiPLR3*-RNAi) and CK lines. Compared with CK, RNAi lines contained an additional 400–500 bp fragment when amplified by using primers JDPDK-1F and JDPDKR (data not shown).

qRT-PCR analysis showed that the transcript level of the target gene (*IiPLR1*, *IiPLR2*, or *IiPLR3*) was significantly suppressed through corresponding RNAi manipulation, with the transcript level of the other two *IiPLR* members displaying a slight fluctuation. There was no significant difference in *IiPLR1*, *IiPLR2*, and *IiPLR3* transcript abundance between WT and

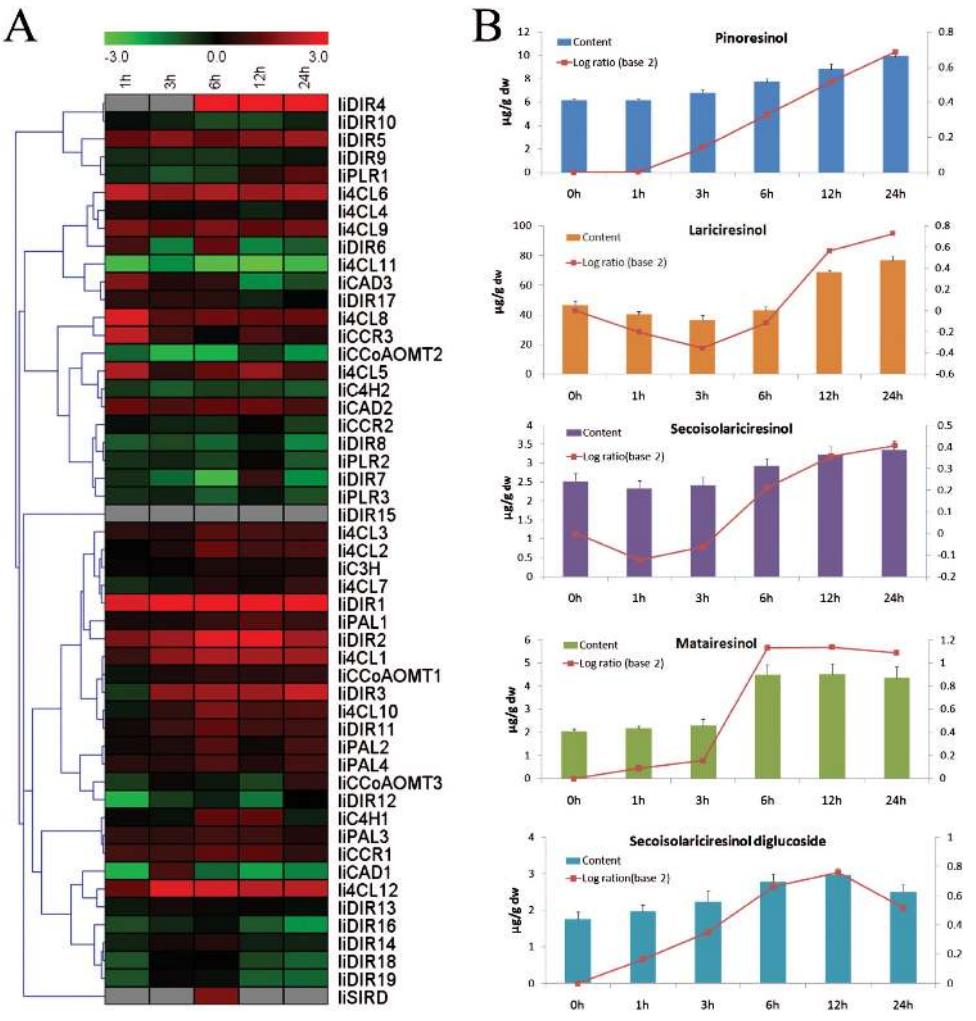

Fig. 2. The transcript abundance of lariciresinol biosynthetic genes and lignan accumulation in response to MeJA treatment. Hairy roots of *I. indigotica* are treated with 0.5 μM MeJA for 0, 1, 3, 6, 12, and 24 h. 0 h is designated as the control. (A) Heat map showing gene transcript levels. The colour bar indicates fold change transcript differences on a natural log scale (treatment/control). The rows in the heat map indicate the genes clustered by their expression patterns. (B) Lignan accumulation in MeJA-treated hairy roots. The different colour bars represent the content of different lignans (μg g^{-1} dw); the red curves represent the relative abundance, which is calculated as log ratio (base 2) of content in treatment to control.

CK lines (Fig. 4B). Metabolite analysis showed that the lariciresinol content in the *IiPLR1*-RNAi lines was significantly reduced compared with that in WT lines, and the reduced lariciresinol accumulation in different *IiPLR1*-RNAi lines (1-1, 1–3, 1–4, 1–6, and 1–7) was found to correlate well with the corresponding decrease of *IiPLR1* mRNA levels. Line 1–4, with the lowest *IiPLR1* mRNA level (~5% of that of the WT, Fig. 4B), accumulated the least lariciresinol (2.0 μg g^{-1} dw), which was ~3.6% of that of the WT control (56.0 μg g^{-1} dw). However, there was no significant difference in lariciresinol content between the WT and CK, *IiPLR2*-RNAi, or *IiPLR3*-RNAi lines (Fig. 4C). These results indicated that *IiPLR1* is the only one of the *IiPLR* members that influences lariciresinol accumulation.

The *IiPLR1*-RNAi hairy roots presented a weaker browning after staining with phloroglucinol-HCl compared with their CK and WT counterparts (Fig. 4D), indicating a decrease of lignans, lignins, and/or wall-bound or secreted phenolic derivatives (Hano *et al.*, 2006). Since *IiPLR1* is involved in the lignan biosynthetic pathway, and lignins share the same precursor monolignols with lignans (Davin and Lewis, 2000), it is speculated that not only lignans (e.g. lariciresinol) but also lignins are strongly impacted by *IiPLR1* suppression. However, there was no phenotype difference observed in WT and CK, *IiPLR2*-RNAi, and *IiPLR3*-RNAi hairy roots after phloroglucinol-HCl staining (Fig. 4D).

Isolation and expression characteristics of liPLR1

A full-length cDNA sequence of *IiPLR1* (GenBank accession no. JF264893) was isolated from *I. indigotica* (Supplementary Fig. S2 at *JXB* online). This gene codes for a predicted polypeptide of 317 amino acids with a calculated molecular mass of 35.6 kDa and an isoelectric point of 5.64. The protein–protein BLAST analysis showed that *Ii*PLR1 has 85% and 82% amino acid identity to *A. thaliana* pinoresinol reductase *At*PrR2 and *At*PrR1, respectively. Similarly to other known PLRs and PrRs, *Ii*PLR1 contains an NADPH-binding motif at the N-terminus (Supplementary Fig. S3). Phylogenetic analysis indicated that *Ii*PLR1 has the closest relationship with *At*PrR2 from *A. thaliana*, which also blongs to the family Cruciferae (Supplementary Fig. S4).

The expression profiles of *IiPLR1* in different tissues and under different stresses were investigated by qRT-PCR. The results showed the *IiPLR1* gene was

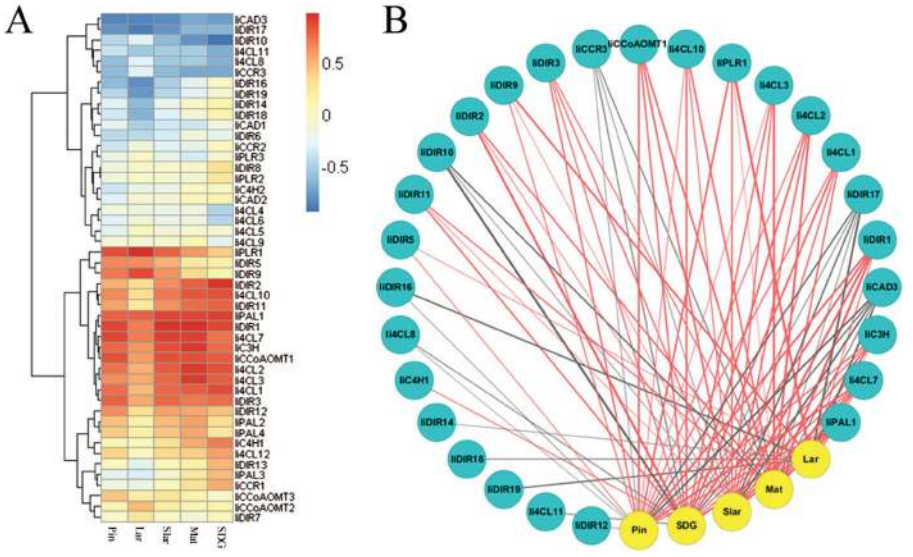

Fig. 3. Correlation between lariciresinol biosynthetic gene expression and lignan accumulation. (A) Heat map showing gene–lignan correlation coefficients. The rows in the heat map are the genes clustered by their expression patterns, and the columns are lignan metabolites. (B) Gene–lignan correlation network with a cut-off $|r|>0.6$. Genes and lignans are drawn as blue and yellow circles, respectively. Red and grey lines indicate positive and negative correlations, respectively. The thickness of lines represents the level of correlation. Pin, pinoresinol; Lar, lariciresinol; Slar, secoisolariciresinol; Mat, matairesinol; SDG, secoisolariciresinol diglucoside.

Table 1. *Lariciresinol biosynthetic genes positively correlated with the five measured lignans with a cut-off $|r|>0.6$.*

Lignan	Lignan-correlated genes
Pinoresinol	*Ii*PAL1, *Ii*4CL1, *Ii*4CL2, *Ii*4CL3, *Ii*4CL7, *Ii*4CL10, *Ii*C3H, *Ii*CCoAOMT1, *Ii*DIR1, *Ii*DIR3, *Ii*DIR5, *Ii*DIR9, *Ii*DIR12, *Ii*PLR1
Lariciresinol	*Ii*PAL1, *Ii*4CL7, *Ii*C3H, *Ii*DIR1, *Ii*DIR5, *Ii*DIR9, *Ii*PLR1
Secoisolariciresinol	*Ii*PAL1, *Ii*4CL1, *Ii*4CL2, *Ii*4CL3, *Ii*4CL7, *Ii*4CL10, *Ii*C3H, *Ii*CCoAOMT1, *Ii*DIR1, *Ii*DIR2, *Ii*DIR3, *Ii*DIR9, *Ii*DIR11, *Ii*PLR1
Matairesinol	*Ii*PAL1, *Ii*4CL1, *Ii*4CL2, *Ii*4CL3, *Ii*4CL7, *Ii*4CL10, *Ii*C3H, *Ii*CCoAOMT1, *Ii*DIR2, *Ii*DIR3, *Ii*DIR11, *Ii*PLR1
Secoisolariciresinol diglucoside	*Ii*PAL1, *Ii*C4H1, *Ii*4CL1, *Ii*4CL2, *Ii*4CL3, *Ii*4CL7, *Ii*4CL10, *Ii*C3H, *Ii*CCoAOMT1, *Ii*DIR1, *Ii*DIR2, *Ii*DIR3, *Ii*DIR11

constitutively expressed in roots, stems, leaves, and flowers of *I. indigotica*, with the strongest expression in roots and the lowest in flowers (Fig. 5A). *IiPLR1* responded to MeJA, ABA, and UV-B treatments, but with different patterns of variation. After treatment with 100 µM MeJA, the expression level of *IiPLR1* increased sharply and peaked at 4 h, with an increase of ~3.8-fold, and then decreased until 24 h (Fig. 5B). When *I. indigotica* was treated with 100 µM ABA, *IiPLR1* expression increased gradually and reached a maximum at 12 h after treatment, which was almost 14-fold higher than the expression before treatment (Fig. 5C). For UV-B treatment, the response was very fast; *IiPLR1* expression increased at 5 min and decreased at 10 min under UV-B, but the highest expression appeared at 30 min after UV-B was turned off (~4.6-fold of that before treatment) (Fig. 5D).

Functional expression of liPLR1 *in* E. coli

To carry out the biochemical characterization of *Ii*PLR1, the recombinant enzyme prepared using *E. coli* was tested *in vitro* for its activity toward racemic pinoresinol and lariciresinol. The *IiPLR1* gene was cloned in pET32a(+) to obtain *IiPLR1*-pET, and expressed in *E. coli* BL21(DE3). SDS–PAGE showed the production of a 53 kDa protein. This size is close to the predicted molecular mass of ht-*Ii*PLR1 (52 922.4 Da). ht-*Ii*PLR1 was purified to near homogeneity by Ni affinity chromatography (Fig. 6A).

When purified ht-*Ii*PLR1 (5 µg ml^{-1} of protein) was incubated with 200 µM (±)-pinoresinol for 30 min, (±)-pinoresinol decreased, whereas (±)-lariciresinol and (±)-secoilariciresinol appeared. No activity was measured in extracts form cells containing the expression vector lacking an insert (CK). It was presumed that *Ii*PLR1 catalysed the sequential steps: first it reduced (±)-pinoresinol to form (±)-lariciresinol, and then it reduced (±)-lariciresinol to give rise to (±)-secoilariciresinol. This assumption was demonstrated by incubating (±)-lariciresinol (3.5 µM) individually with ht-*Ii*PLR1 (5 µg ml^{-1}). After 30 min of incubation, ht-*Ii*PLR1 reduced (±)-lariciresinol efficiently to (±)-secoilariciresinol, whereas no activity was measured in the control extract (Fig. 6B). These results indicate that *Ii*PLR1 is able to catalyse both (±)-pinoresinol and (±)-lariciresinol.

Kinetic analysis of *Ii*PLR1 using both (±)-pinoresinol and (±)-lariciresinol as substrates was conducted. As shown in Table 2, the K_m value of *Ii*PLR1 for (±)-pinoresinol (65.4 ± 2.76 µM) was 26 times higher than that for (±)-lariciresinol (2.50 ± 0.14 µM), but there was no significant difference in activities between (±)-pinoresinol and (±)-lariciresinol with regard to k_{cat}/K_m values [0.91 ± 0.11 µM^{-1} min^{-1} for (±)-pinoresinol, 1.59 ± 0.05 µM^{-1} min^{-1} for (±)-lariciresinol].

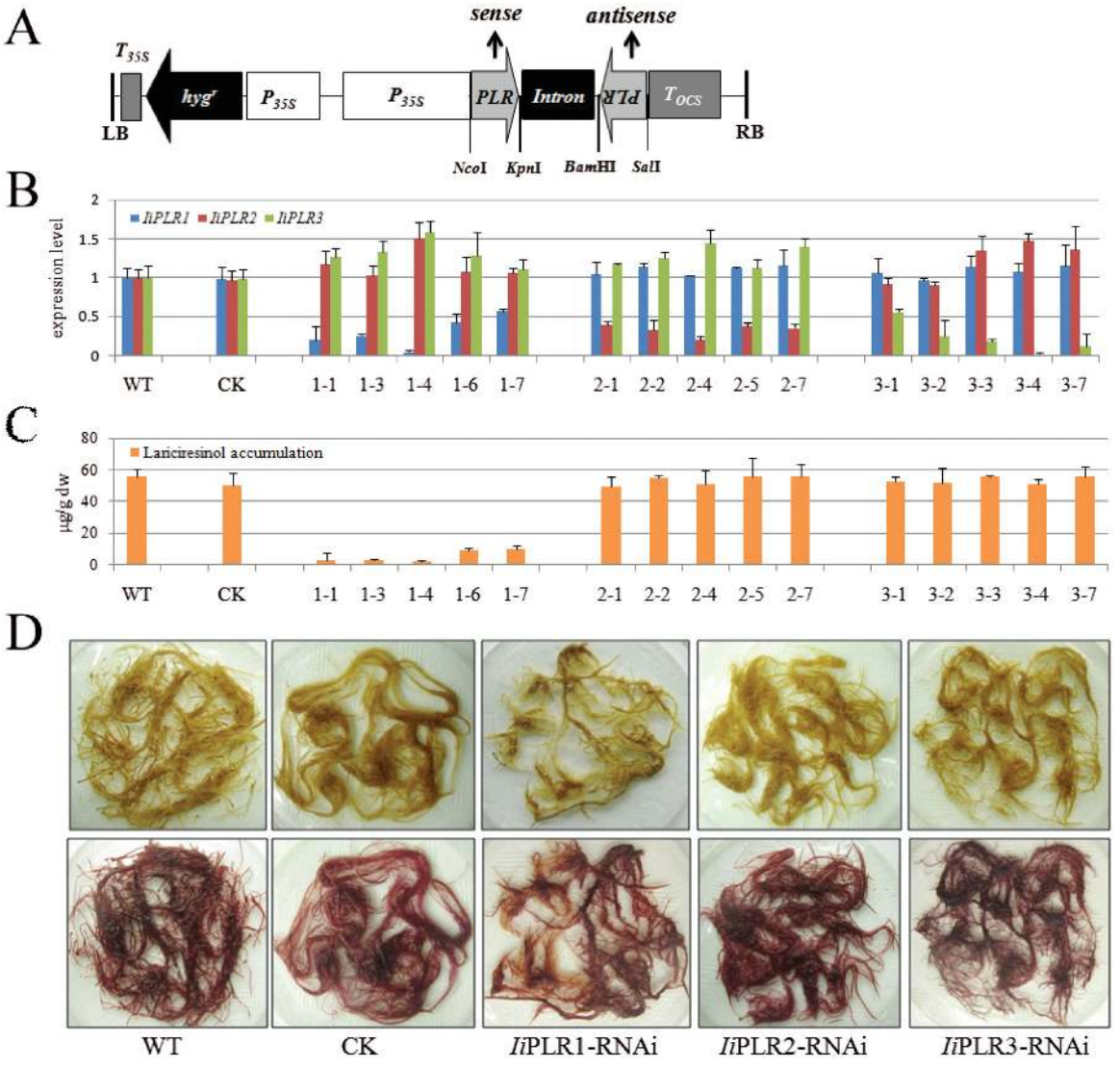

Fig. 4. RNAi silencing of *li*PLR1, *li*PLR2, and *li*PLR3 in *I. indigotica* hairy roots. (A) Schematic representation of the RNAi vector. P_{35S}, CaMV 35S promoter; *hyg*^{r}, hygromycin resistance gene; *Intron*, the Pdk intron; *PLR*, the 400–500 bp fragment of the coding region of *liPLRs*; T_{35S}, CaMV 35S terminator; T_{ocs}, terminator of the octopine synthase gene; LB, T-DNA left border; RB, T-DNA right border. Restriction sites are marked. (B) *li*PLR1, *li*PLR2, and *li*PLR3 transcript abundance and (C) laricresinol content in *li*PLR1-RNAi, *li*PLR2-RNAi, and *li*PLR3-RNAi lines. (D) Phenotype of *li*PLR-RNAi hairy roots before (upper) and after (lower) phloroglucinol-HCl staining.

Overexpression of liPLR1 *in* I. indigotica *hairy roots*

The *Ii*PLR1 overexpression vector PHB-*IiPLR1*-flag (Fig. 7A) and its vector control PHB-flag were separately introduced into *I. indigotica* to generate *Ii*PLR1-OVX and CK lines. PCR analyses showed that all these lines contained both *rolc* and *hpt* genes, and the exogenous *IiPLR1* gene was detected in *Ii*PLR1-OVX lines (data not shown).

qRT-PCR analysis showed that in comparison with levels of the WT control, *IiPLR1* transcript levels in all the seven independent *Ii*PLR1-OVX lines (ovx-1, ovx-2, ovx-4, ovx-5, ovx-7, ovx-8, and ovx-9) were significantly enhanced (~5.4-, 10.5-, 2.1-, 3.2-, 3.2-, 8.8-, and 7.3-fold). No significant difference was observed in the *IiPLR1* transcript between WT and CK lines (Fig. 7B). Moreover, the accumulation of the flag-tagged *Ii*PLR1 (*Ii*PLR1-flag fusion protein) in *Ii*PLR1-OVX lines was examined by western blot analysis using an antibody against flag. A strong cross-reaction signal corresponding to a polypeptide with an expected mol. wt of 52 kDa for the *Ii*PLR1-flag fusion protein was present in all the seven tested root lines, and the strengths of *Ii*PLR1-flag signal in these independent lines was found to match well with the levels of *IiPLR1* transcript. As expected, no signal was observed using total protein extracted from WT and CK controls (Fig. 7C). These results indicated that the exogenous *IiPLR1* gene had successfully integrated and was expressed in *I. indigotica*.

The biomass growth rates of transgenic lines were recorded during the hairy root culture period. All hairy root lines reached the highest growth rate at day 18 and achieved maximum fresh weight at day 45 after inoculation (Supplementary Table S6 at *JXB* online). Particular attention was paid to lines ovx-2, ovx-8, and ovx-9 because of their relatively higher *Ii*PLR1 expression. As shown in Fig. 7D, lines ovx-2 and

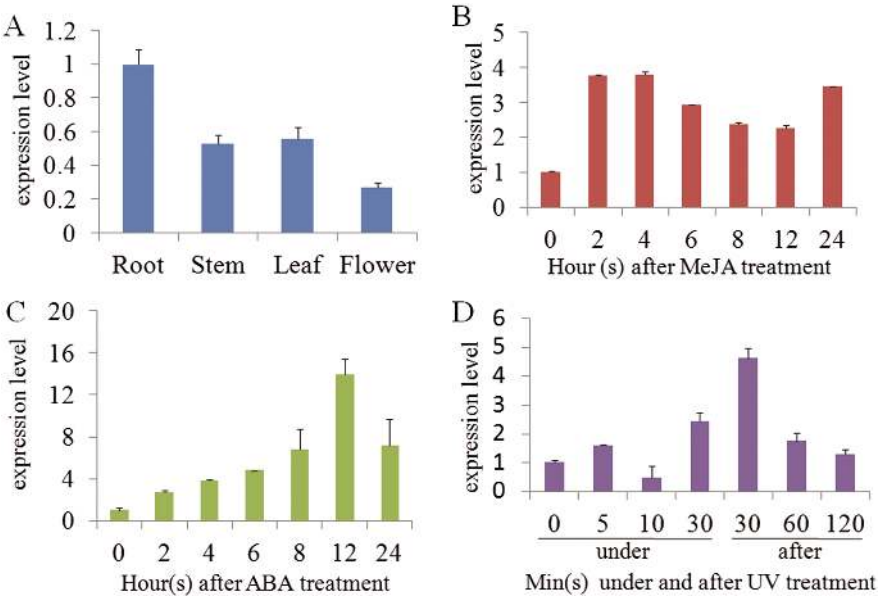

Fig. 5. *Ii*PLR1 expression in different tissues (A) and under MeJA (B), ABA (C), and UV (D) treatments.

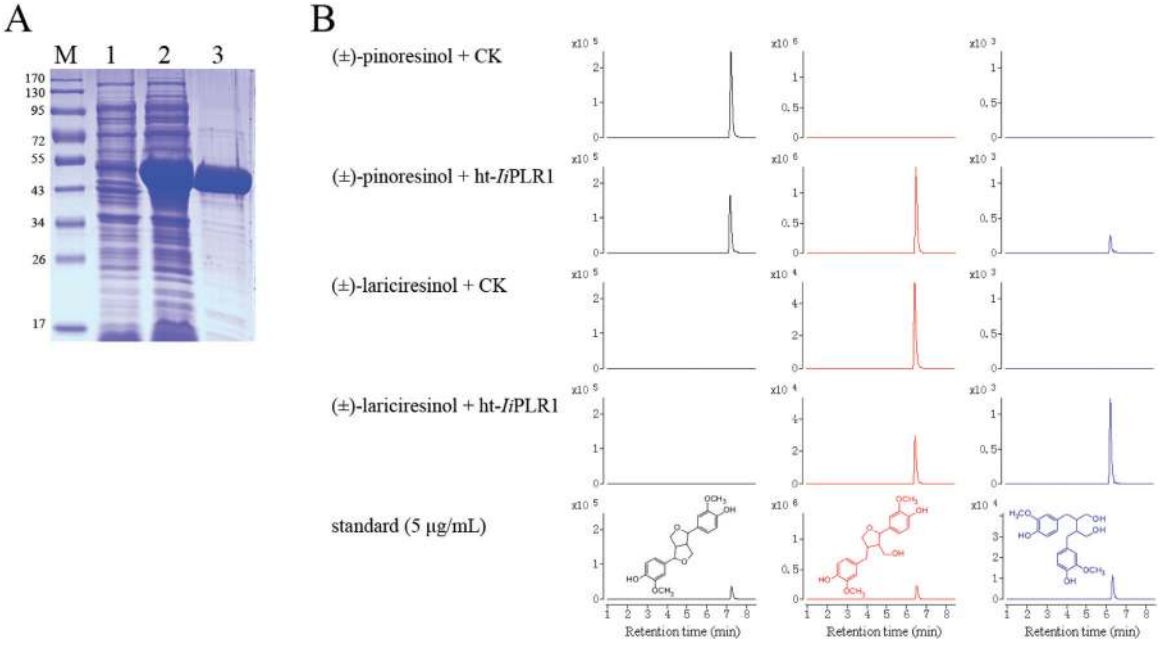

Fig. 6. Characterization of ht-*Ii*PLR1 produced in *E. coli*. (A) SDS–PAGE analysis of expression and purification of ht-*Ii*PLR1. Lanes: M, protein marker; 1, uninduced cell harbouring ht-*Ii*PLR1; 2, IPTG-induced cell harbouring ht-*Ii*PLR1; 3, purified ht-*Ii*PLR1. (B) Conversion of (±)-pinoresinol and (±)-lariciresinol by purified ht-*Ii*PLR1. (±)-Pinoresinol (200 μM) and (±)-lariciresinol (3.5 μM) were separately incubated with *E. coli* cells harbouring the pET32a(+) vector control (CK) and purified ht-*Ii*PLR1 (5 μg ml^{-1} protein). After 30 min of incubation, the reaction products were analysed by LC-MS. (±)-Pinoresinol, (±)-lariciresinol and (±)-secoisolariciresinol were detected by MRM mode with *m/z* 357→151, 359→329, and 361→164 as the monitoring ion pair, respectively. Chromatograms of (±)-pinoresinol, (±)-lariciresinol, and (±)-secoilariciresinol are denoted with black, red, and blue colour, respectively.

ovx-8 presented growth rates comparable with the WT, while the growth rates of ovx-9, along with CK, were significantly lower than that of the WT.

The phloroglucinol-HCl-stained *Ii*PLR1-OVX hairy roots are shown in Fig. 7E. It was apparent that *Ii*PLR1-OVX hairy roots presented a violet-red colour, whereas CK control presented a strong browning, which was similar to the WT (Fig. 4D). This result indicated that *Ii*PLR1-OVX hairy roots accumulated more lignans, lignins, and/or wall-bound or secreted phenolic derivatives than CK and the WT. In agreement with the speculation derived from phloroglucinol-HCl staining of *IiPLR1*-RNAi hairy roots, the result also suggests *Ii*PLR1 expression probably also has a marked effect on lignin accumulation. Metabolite analysis showed that the lariciresinol content in the *Ii*PLR1-OVX lines was significantly higher than that in the WT line, and the lariciresinol accumulation in different independent lines was closely correlated with the corresponding *Ii*PLR1 expression levels. Line ovx-2, with the highest *Ii*PLR1 expression (Fig. 7B, C), produced the most abundant lariciresinol (353.9 μg g^{-1} dw), which was ~6.3-fold

Table 2. *Kinetic properties of purified recombinant IiPLR1*

Substrate	K_m (μM)	V_{max} (μM min^{-1})	K_{cat} (min^{-1})	k_{cat}/K_m (μM^{-1} min^{-1})
(±)-Pinoresinol	65.4±2.76	1.67±0.13	59.6±4.52	0.91±0.11
(±)-Lariciresinol	2.50±0.14	0.11±0.003	3.98±0.09	1.59±0.05

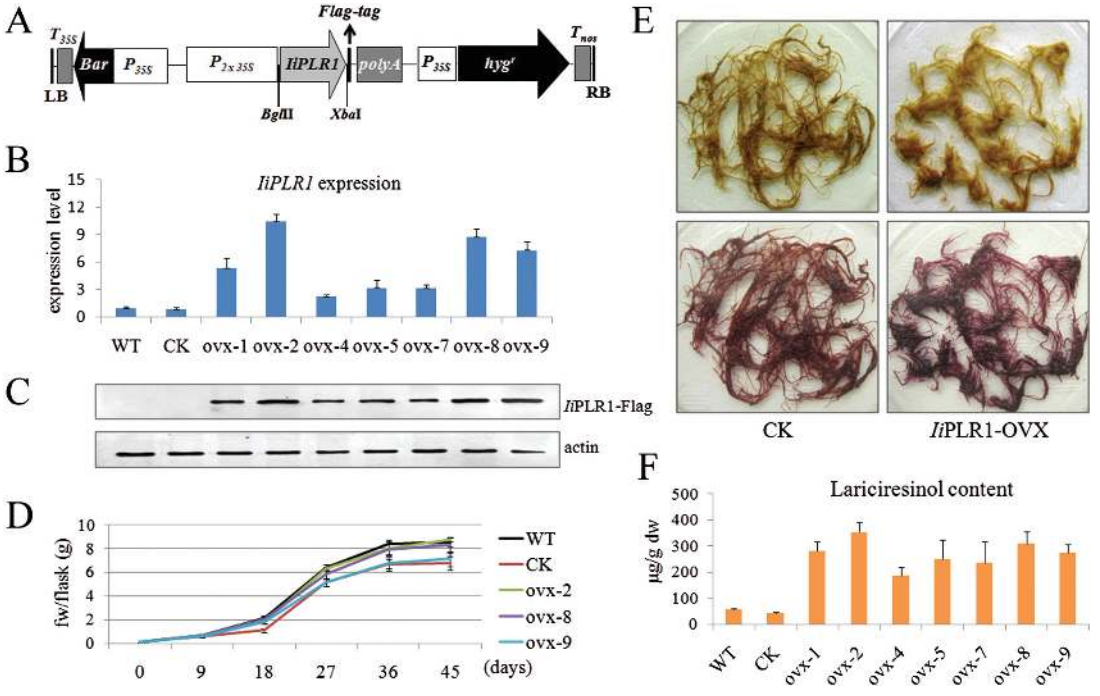

Fig. 7. Overexpression of *Ii*PLR1 in *I. indigotica* hairy roots. (A) Schematic representation of the overexpression vector. P_{35S}, CaMV 35S promoter; $P_{2\times35S}$, CaMV 35S double promoter; *hyg*^{r}, hygromycin resistance gene; *Bar*, herbicide resistance gene; *IiPLR1*, the *IiPLR1* coding sequence; *polyA*, poly(A) tail; T_{35S}, CaMV 35S terminator; T_{nos}, nopaline synthase terminator; LB, T-DNA left border; RB, T-DNA right border. The flag-tag and restriction sites are marked. (B) *Ii*PLR1 transcript abundance, (C) *Ii*PLR1 transcript expression, and (F) lariciresinol content in transgenic hairy root lines. (D) Time courses of growth of hairy root lines ovx-2, ovx-8, ovx-9, WT, and CK. (E) Phenotype of CK control (left) and *Ii*PLR1-OVX hairy roots (right) before (upper) and after (lower) phloroglucinol-HCl staining.

more than in its WT counterpart (56.0 μg g^{-1} dw); line ovx-8, with a relative higher *Ii*PLR1 expression (Fig. 7B, C), accumulated 310.4 μg g^{-1} lariciresinol, which was 5.5-fold more that in the WT. There was no significant difference in lariciresinol content between CK (42.9 μg g^{-1} dw) and the WT (Fig. 7F).

Discussion

Plant metabolic pathways are complex and often involve multicopy gene families. It is therefore difficult to target the best intervention points and accurately predict the outcome. Until recently, metabolic engineering in plants relied on the trial and error approach to achieve desirable changes in the metabolic profile. However, technological advances in data mining and modelling are now taking away much of the guesswork by allowing the impact of modifications to be predicted more accurately (Farré *et al.*, 2015). This knowledge-driven approach for engineering requires a well-characterized population as a source of genetic variation. The potency of jasmonates, such as MeJA, to elicit secondary metabolism in cell cultures has made them powerful tools to cause the genetic diversity and help to unravel the complex cellular process (Pauwels *et al.*, 2008). Combined transcript and metabolite profiling approaches to determine the MeJA-mediated regulation of secondary metabolism resulted in the establishment of gene–metabolite networks involved in the biosynthesis of the pharmaceutically valuable terpenoid indole alkaloids in *Catharanthus roseus* (Rischer *et al.*, 2006), in the characterization of enzymatic steps in the isoflavone biosynthetic pathways in *Medicago truncatula* (Naoumkina *et al.*, 2007), and in the rosmarinic acid biosynthetic pathways in *Salvia miltiorrhiza* (Xiao *et al.*, 2009), as well as in the isolation of novel regulators of aliphatic glucosinolate biosynthesis in *Arabidopsis* (Hirai *et al.*, 2007).

Here, MeJA-elicited *I. indigotica* hairy roots were employed as a resource for the identification of candidate lignan-correlated genes. These hairy roots obtained from different time points after MeJA treatment harbour genetic diversity, impacting on the chemical composition of lignans, as shown in Fig. 2. Correlation analysis of gene transcripts and metabolite contents resulted in multiple new hypotheses regarding regulatory relationships between the identified biosynthetic genes and lignans (Fig. 3; Table 1). Seven biosynthetic genes were identified to be significantly correlated with lariciresinol, a pharmaceutically valuable compound of *I. indigotica*. To validate this network analysis, one of the best-supported

nodes, *Ii*PLR1 (predicted to be positively correlated with lariciresinol), together with its family members (*Ii*PLR2 and *Ii*PLR3, predicted to be uncorrelated with lariciresinol) were investigated. RNAi suppression of *Ii*PLR1 significantly inhibited lariciresinol accumulation in *I. indigotica* hairy roots compared with the WT, whereas *Ii*PLR2 or *Ii*PLR3 suppression did not influence lariciresinol accumulation significantly. These results demonstrated that *Ii*PLR1, but not *Ii*PLR2 or *Ii*PLR3, plays an important role in lariciresinol accumulation in *I. indigotica*, consistent with the hypothesized regulatory roles.

IiPLR1 was thus further characterized. Its expression in roots is much higher than that in stems, leaves, and flowers, suggesting that *IiPLR1* is mainly participating in the lariciresinol biosynthesis in roots. This is consistent with the fact that the root of *I. indigotica* 'Ban–Lan–Gen' is the main organ for conventional clinical treatment.

As a major group of secondary metabolites, lignan compounds contribute greatly to plant defence (Touré and Xueming, 2010). It is thus speculated that as a key intervention point regulating lignan production, *IiPLR1* expression should be significantly induced when subjected to environment stresses. In agreement with this, it was observed that *IiPLR1* was inducible by MeJA, ABA, and UV-B, although the extent of the induction varied dramatically. *IiPLR1* was highly induced by ABA like *LuPLR1* from *Linum usitatissimum*, which was reported to play a key role in ABA-mediated regulation of secoisolariciresinol diglucoside biosynthesis in flax (Renouard *et al.*, 2012; Corbin *et al.*, 2013). Compared with MeJA elicitor, ABA was shown to be more effective in up-regulating *IiPLR1* (peaking at 12 h with a 14-fold increase), suggesting that ABA elicitation was a better potential strategy to enhance the accumulation of desired lignans of *I. indigotica* via the transcriptional regulation of the *IiPLR1* gene. *IiPLR1* responded to UV-B very rapidly, but the highest *IiPLR1* transcript appeared at 30 min after UV-B was turned off. It is suggested that the plant defence and repair process in response to UV-B happened sequentially; thus *IiPLR1* needed to be be highly expressed at different times, and the fact that the highest *IiPLR1* transcript was found after UV-B was turned off was probably due to the repair rather than the protection mechanism (Di *et al.*, 2012). However, this remains to be investigated.

The precise catalytic role of *Ii*PLR1 was confirmed by expression in *E. coli* and the conversion of defined substrates. Although *Ii*PLR1 has a high level of amino acid identity (85%) with *At*PrR2, which shows strict substrate specificity toward pinoresinol but no activity toward lariciresinol (Nakatsubo *et al.*, 2008), *Ii*PLR1 reduces both pinoresinol and lariciresinol efficiently like *Fi*PLR1 from *Forsythia intermedia* (Dinkova-Kostova *et al.*, 1996), *Tp*PLR1 and *Tp*PLR2 from *Thuja plicata* (Fujita *et al.*, 1999), as well as *Lp*PLR1 from *Linum perenne* (Hemmati *et al.*, 2007) (Fig. 6B). Moreover, the activity of *Ii*PLR1 toward (±)-lariciresinol is comparable with that toward (±)-pinoresinol with regard to k_{cat}/K_m values (Table 2). The underlying molecular basis of the difference in substrate selectivity between *Ii*PLR1 and *At*PrR2 is worth investigating in the future.

In order to produce a high concentration of lariciresinol, gene constructs containing cDNA clones of *IiPLR1*, driven by a double *Cauliflower mosaic virus* (CaMV) 35S promoter ($P_{2\times35S}$), was introduced into *I. indigotica* hairy roots. As expected, the lariciresinol content significantly increased following introduction of the *IiPLR1* gene. Generally, a high content of secondary metabolites in the tissues is associated with poor growth, and the actual total productivity of secondary metabolites therefore remains low (Jouhikainen *et al.*, 1999). It was found in this study, however, that the growth rate in the high lariciresinol-producing lines ovx-2 (353.9 μg g^{-1} dw, 6.3-fold that of the WT) and ovx-8 (310.4 μg g^{-1} dw, 5.5-fold that of the WT) was not reduced as compared with those with low lariciresinol production including WT and CK lines. The current study provides an effective approach for commercial large-scale production of lariciresinol by using the *I. indigotica* hairy root systems as bioreactors.

In conclusion, the current study sheds light on how to engineer the production of target metabolites effectively using more suitable intervention points or by more refined strategies. In the case of PLR, which is undoubtedly a key rate-limiting step of the lariciresinol biosynthetic pathway, three *Ii*PLRs were predicted from *I. indigotica* based on homology search (*Ii*PLR1–*Ii*PLR3); however, only *Ii*PLR1 influenced lariciresinol biosynthesis. The comprehensive profiling approach described here provides a good example to locate the most potential member(s) rapidlly from a multicopy gene family. Moreover, the present approach of integrated transcriptome and metabolite profiling using MeJA-elicited *I. indigotica* hairy roots as a source of variation of gene expression and metabolites provided insight into multiple previously uncharacterized candidate lignan-correlated genes. One candidate, *IiPLR1*, was demonstrated to indeed be important in lariciresinol biosynthesis, indicating that the analysis is robust and genes identified via this process are worthy of validation.

Supplementary data

Supplementary data are available at *JXB* online.

Figure S1. Nucleotide sequence alignment of *Ii*PLR1, *Ii*PLR2, and *Ii*PLR3.

Figure S2. Nucleotide sequence and the deduced amino acid sequence of *IiPLR1*.

Figure S3. Amino acid sequence alignment of *Ii*PLR1 with other plant PLRs.

Figure S4. Phylogenic analysis of *Ii*PLR1 with other plant PLRs.

Table S1. Sequences corresponding to the lariciresinol biosynthetic genes of *I. indigotica*.

Table S2. The inserted fragments for RNAi silencing of *Ii*PLR1, *Ii*PLR2, and *Ii*PLR3.

Table S3. PCR primers used in this study.

Table S4. Pinoresinol/lariciresinol reductases used in this study.

Table S5. The correlation coefficients between lariciresinol biosynthetic gene expression and lignan concentrations.

Table S6. Biomass accumulation of *Ii*PLR1 transgenic lines.

Acknowledgements

This work was supported by the Natural Science Foundation of China (grant nos. 31100221, 81325024, and 81303160).

References

Benveniste P. 2002. Sterol metabolism. *The Arabidopsis Book* **1,** e0004.

Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. *Analytical Biochemistry* **72,** 248–254.

Chen J, Dong X, Li Q, Zhou X, Gao S, Chen R, Sun L, Zhang L, Chen W. 2013. Biosynthesis of the active compounds of *Isatis indigotica* based on transcriptome sequencing and metabolites profiling. *BMC Genomics* **14,** 857.

Chilton MD, Tepfer DA, Petit A, David C, Casse-Delbart F, Tempe J. 1982. Agrobacterium rhizogenes inserts T-DNA into the genomes of the host plant root cells. *Nature* **295,** 432–434.

Cline MS, Smoot M, Cerami E, *et al*. 2007. Integration of biological networks and gene expression data using Cytoscape. *Nature Protocols* **2,** 2366–2382.

Corbin C, Decourtil C, Marosevic D, Bailly M, Lopez T, Renouard S, Doussot J, Dutilleul C, Auguin D, Giglioli-Guivarc'h N. 2013. Role of protein farnesylation events in the ABA-mediated regulation of the Pinoresinol-Lariciresinol Reductase 1 (*Lu*PLR1) gene expression and lignan biosynthesis in flax (*Linum usitatissimum* L.). *Plant Physiology and Biochemistry* **72,** 96–111.

Davin LB, Lewis NG. 2000. Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis. *Plant Physiology* **123,** 453–462.

Di P, Hu Y, Xuan H, Xiao Y, Chen J, Zhang L, Chen W. 2012. Characterization and the expression profile of 4-coumarate: CoA ligase (*Ii*4CL) from hairy roots of *Isatis indigotica*. *African Journal of Pharmacy and Pharmacology* **6,** 2166–2175.

Dinkova-Kostova AT, Gang DR, Davin LB, Bedgar DL, Chu A, Lewis NG. 1996. (+)-Pinoresinol/(+)-lariciresinol reductase from Forsythia intermedia. Protein purification, cDNA cloning, heterologous expression and comparison to isoflavone reductase. *Journal of Biological Chemistry* **271,** 29473–29482.

Doyle JJ. 1990. Isolation of plant DNA from fresh tissue. *Focus* **12,** 13–15.

Farré G, Twyman RM, Christou P, Capell T, Zhu C. 2015. Knowledge-driven approaches for engineering complex metabolic pathways in plants. *Current Opinion in Biotechnology* **32,** 54–60.

Fujita M, Gang DR, Davin LB, Lewis NG. 1999. Recombinant pinoresinol-lariciresinol reductases from western red cedar (Thuja plicata) catalyze opposite enantiospecific conversions. *Journal of Biological Chemistry* **274,** 618–627.

Fukuhara Y, Kamimura N, Nakajima M, Hishiyama S, Hara H, Kasai D, Tsuji Y, Narita-Yamada S, Nakamura S, Katano Y. 2013. Discovery of pinoresinol reductase genes in sphingomonads. *Enzyme and Microbial Technology* **52,** 38–43.

Hano C, Addi M, Bensaddek L, Crônier D, Baltora-Rosset S, Doussot J, Maury S, Mesnard F, Chabbert B, Hawkins S. 2006. Differential accumulation of monolignol-derived compounds in elicited flax (*Linum usitatissimum*) cell suspension cultures. *Planta* **223,** 975–989.

Hemmati S, Schmidt TJ, Fuss E. 2007. (+)-Pinoresinol/(–)-lariciresinol reductase from *Linum perenne* Himmelszelt involved in the biosynthesis of justicidin B. *FEBS Letters* **581,** 603–610.

Hemmati S, von Heimendahl CB, Klaes M, Alfermann AW, Schmidt TJ, Fuss E. 2010. Pinoresinol-lariciresinol reductases with opposite enantiospecificity determine the enantiomeric composition of lignans in the different organs of *Linum usitatissimum* L. *Planta Medica* **76,** 928–934.

Hirai MY, Sugiyama K, Sawada Y, Tohge T, Obayashi T, Suzuki A, Araki R, Sakurai N, Suzuki H, Aoki K. 2007. Omics-based identification of Arabidopsis Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proceedings of the National Academy of Sciences, USA* **104,** 6478–6483.

Humphreys JM, Chapple C. 2002. Rewriting the lignin roadmap. *Current Opinion in Plant Biology* **5,** 224–229.

Jaakola L, Pirttilä AM, Halonen M, Hohtola A. 2001. Isolation of high quality RNA from bilberry (*Vaccinium myrtillus* L.) fruit. *Molecular Biotechnology* **19,** 201–203.

Jin H, Song Z, Nikolau BJ. 2012. Reverse genetic characterization of two paralogous acetoacetyl CoA thiolase genes in *Arabidopsis* reveals their importance in plant growth and development. *The Plant Journal* **70,** 1015–1032.

Jouhikainen K, Lindgren L, Jokelainen T, Hiltunen R, Teeri TH, Oksman-Caldentey K-M. 1999. Enhancement of scopolamine production in *Hyoscyamus muticus* L. hairy root cultures by genetic engineering. *Planta* **208,** 545–551.

Keim V, Manzano D, Fernández FJ, Closa M, Andrade P, Caudepón D, Bortolotti C, Vega MC, Arró M, Ferrer A. 2012. Characterization of Arabidopsis fps isozymes and *fps* gene expression analysis provide insight into the biosynthesis of isoprenoid precursors in seeds. *PLoS One* **7,** e49109.

Laranjeira S, Amorim-Silva V, Esteban A, Arró M, Ferrer A, Tavares RM, Botella MA, Rosado A, Azevedo H. 2015. Arabidopsis squalene epoxidase 3 (sqe3) complements sqe1 and is important for embryo development and bulk squalene epoxidase activity. *Molecular Plant* (in press).

Larkin MA, Blackshields G, Brown N, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948.

Li B. 2003. Studies on active constituents and quality evaluation of Banlangen. PhD thesis, School of Pharmacy, Second Military University, Shanghai.

Li Q, Chen J, Xiao Y, Di P, Zhang L, Chen W. 2014. The dirigent multigene family in *Isatis indigotica*: gene discovery and differential transcript abundance. *BMC Genomics* **15,** 388.

Lin C-W, Tsai F-J, Tsai C-H, Lai C-C, Wan L, Ho T-Y, Hsieh C-C, Chao P-DL. 2005. Anti-SARS coronavirus 3C-like protease effects of *Isatis indigotica* root and plant-derived phenolic compounds. *Antiviral Research* **68,** 36–42.

Nakatsubo T, Mizutani M, Suzuki S, Hattori T, Umezawa T. 2008. Characterization of *Arabidopsis thaliana* pinoresinol reductase, a new type of enzyme involved in lignan biosynthesis. *Journal of Biological Chemistry* **283,** 15550–15557.

Naoumkina M, Farag MA, Sumner LW, Tang Y, Liu C-J, Dixon RA. 2007. Different mechanisms for phytoalexin induction by pathogen and wound signals in *Medicago truncatula*. *Proceedings of the National Academy of Sciences, USA* **104,** 17909–17915.

Nicholas KB, Nicholas HB Jr. 1997. GeneDoc: a tool for editing and annotating multiple sequence alignments. Distributed by the authors.

Ohyama K, Suzuki M, Masuda K, Yoshida S, Muranaka T. 2007. Chemical phenotypes of the *hmg1* and *hmg2* mutants of Arabidopsis demonstrate the in-planta role of HMG-CoA reductase in triterpene biosynthesis. *Chemical and Pharmaceutical Bulletin* **55,** 1518–1521.

Pauwels L, Morreel K, De Witte E, Lammertyn F, Van Montagu M, Boerjan W, Inzé D, Goossens A. 2008. Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured Arabidopsis cells. *Proceedings of the National Academy of Sciences, USA* **105,** 1380–1385.

Posé D, Castanedo I, Borsani O, Nieto B, Rosado A, Taconnat L, Ferrer A, Dolan L, Valpuesta V, Botella MA. 2009. Identification of the Arabidopsis dry2/sqe1-5 mutant reveals a central role for sterols in drought tolerance and regulation of reactive oxygen species. *The Plant Journal* **59,** 63–76.

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J. 2012. The Pfam protein families database. *Nucleic Acids Research* **40,** D290–D301.

Rasbery JM, Shan H, LeClair RJ, Norman M, Matsuda SPT, Bartel B. 2007. *Arabidopsis thaliana* squalene epoxidase is essential for root and seed development. *Journal of Biological Chemistry* **282,** 17002–17013.

Renouard S, Corbin C, Lopez T, Montguillon J, Gutierrez L, Lamblin F, Lainé E, Hano C. 2012. Abscisic acid regulates pinoresinol-lariciresinol reductase gene expression and secoisolariciresinol accumulation in developing flax (*Linum usitatissimum* L.) seeds. *Planta* **235,** 85–98.

Rischer H, Orešič M, Seppänen-Laakso T, Katajamaa M, Lammertyn F, Ardiles-Diaz W, Van Montagu MC, Inzé D, Oksman-Caldentey K-M, Goossens A. 2006. Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells. *Proceedings of the National Academy of Sciences, USA* **103,** 5614–5619.

Saeed A, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M. 2003. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **34,** 374.

Saito K, Matsuda F. 2010. Metabolomics for functional genomics, systems biology, and biotechnology. *Annual Review of Plant Biology* **61,** 463–489.

Saleem M, Kim HJ, Ali MS, Lee YS. 2005. An update on bioactive plant lignans. *Natural Product Reports* **22,** 696–716.

Satake H, Ono E, Murata J. 2013. Recent advances in the metabolic engineering of lignan biosynthesis pathways for the production of transgenic plant-based foods and supplements. *Journal of Agricultural and Food Chemistry* **61,** 11721–11729.

Sui C. 2010. Establishment of a mouse model for pandemic H1N1 influenza virus and study on effect of Banlangen granules on mice challenged with pandemic H1N1 influenza virus. Masters thesis, Peking Union Medical College, Beijing.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28,** 2731–2739.

Touré A, Xueming X. 2010. Flaxseed lignans: source, biosynthesis, metabolism, antioxidant activity, bio-active components, and health benefits. *Comprehensive Reviews in Food Science and Food Safety* **9,** 261–269.

Wang Y, Qiao CZ, Liu S, Hang HM. 2000. Evaluation on anti-endotoxic action and antiviral action *in vitro* of tetraploid *Isatis indigotica*. *Chinese Traditional Herb* **25,** 327–329.

Wang YT, Yang ZF, Zhan HS, Qin S, Guan WD. 2011. Screening of anti-H1N1 active constituents from *Radix Isatidis*. *International Journal of Molecular Sciences* **28,** 419–422.

Wesley SV, Helliwell CA, Smith NA, Wang M, Rouse DT, Liu Q, Gooding PS, Singh SP, Abbott D, Stoutjesdijk PA. 2001. Construct design for efficient, effective and high-throughput gene silencing in plants. *The Plant Journal* **27,** 581–590.

Xiao Y, Gao S, Di P, Chen J, Chen W, Zhang L. 2009. Methyl jasmonate dramatically enhances the accumulation of phenolic acids in *Salvia miltiorrhiza* hairy root cultures. *Physiologia Plantarum* **137,** 1–9.

Yang Z, Wang Y, Zheng Z, Zhao S, Zhao J, Lin Q, Li C, Zhu Q, Zhong N. 2013. Antiviral activity of *Isatis indigotica* root-derived clemastanin B against human and avian influenza A and B viruses *in vitro*. *International Journal of Molecular Medicine* **31,** 867–873.

Zhao LM. 2007. Pharmacological research and clinical application in *Isatidis indigotica* Fort. *Journal of Chinese Medicine Research* **7,** 141–143.

Zhao Q, Zeng Y, Yin Y, *et al*. 2014. Pinoresinol reductase 1 impacts lignin distribution during secondary cell wall biosynthesis in *Arabidopsis*. *Phytochemistry* **112,** 170–178.