Protoplasma (2013) 250:1239–1249
DOI 10.1007/s00709-013-0505-z

ORIGINAL ARTICLE

Expressed sequence tags and molecular cloning and characterization of gene encoding pinoresinol/lariciresinol reductase from *Podophyllum hexandrum*

Dhammaprakash Pandhari Wankhede · Dipul Kumar Biswas · Subramani Rajkumar · Alok Krishna Sinha

Received: 13 February 2013 / Accepted: 18 April 2013 / Published online: 8 May 2013

Abstract Podophyllotoxin, an aryltetralin lignan, is the source of important anticancer drugs etoposide, teniposide, and etopophos. Roots/rhizome of *Podophyllum hexandrum* form one of the most important sources of podophyllotoxin. In order to understand genes involved in podophyllotoxin biosynthesis, two suppression subtractive hybridization libraries were synthesized, one each from root/rhizome and leaves using high and low podophyllotoxin-producing plants of *P. hexandrum*. Sequencing of clones identified a total of 1,141 Expressed Sequence Tags (ESTs) resulting in 354 unique ESTs. Several unique ESTs showed sequence similarity to the genes involved in metabolism, stress/defense responses, and signalling pathways. A few ESTs also showed high sequence similarity with genes which were shown to be involved in podophyllotoxin biosynthesis in other plant species such as pinoresinol/lariciresinol reductase. A full length coding sequence of pinoresinol/lariciresinol reductase (PLR) has been cloned from *P. hexandrum* which was found to encode protein with 311 amino acids and show sequence similarity with PLR from *Forsythia intermedia* and *Linum* spp. Spatial and stress-inducible expression pattern of *PhPLR* and other known genes of podophyllotoxin biosynthesis, secoisolariciresinol dehydrogenase (*PhSDH*), and dirigent protein oxidase (*PhDPO*) have been studied. All the three genes showed wounding and methyl jasmonate-inducible expression pattern. The present work would form a basis for further studies to understand genomics of podophyllotoxin biosynthesis in *P. hexandrum*.

Keywords *Podophyllum hexandrum* · Suppressive subtraction hybridization · cDNA library · Expressed sequenced tags · Pinoresinol/lariciresinol reductase · RT-PCR

Handling Editor: Klaus Harter

Electronic supplementary material The online version of this article (doi:10.1007/s00709-013-0505-z) contains supplementary material, which is available to authorized users.

D. P. Wankhede · D. K. Biswas · A. K. Sinha (✉)
National Institute of Plant Genome Research,
Aruna Asaf Ali Marg,
New Delhi 110067, India
e-mail: alok@nipgr.ac.in

S. Rajkumar
CSIR-Institute of Himalayan Bioresource Technology, Palampur,
Himachal Pradesh, India

S. Rajkumar
National Bureau of Plant Genetic Resource, New Delhi, India

Introduction

Plants have been source of several important compounds for medicinal purposes for mankind. *Podophyllum hexandrum* (Royle), a member of Berberidaceae family, produces an important aryltetralin lignan, a podophyllotoxin (PT) that possesses antitumor and antihyperlipidemic activities (Farkya et al. 2004). Further, PT is source of clinically important anticancer drugs etoposide, etopophos, and teniposide (Canel et al. 2000). Podophyllotoxin is commercially extracted from the roots/rhizomes of *P. hexandrum* and *Podophyllum peltatum*, the former is superior in terms of podophyllotoxin content (Jackson and Dewick 1984; Van Uden et al. 1989). *P. hexandrum* is a rhizomatous herb, about 15–30 cm in height with perennial root system and grows in the Himalayan region at an altitude of 3,500–4,000 m (Choudhary et al. 1998). Rhizomes and roots of *P. hexandrum* are the preferred source of PT. Owing largely to the indiscriminate harvesting in the natural habitat, *P. hexandrum* has been declared as an endangered species (Farkya et al. 2004).

PT biosynthesis in plants occurs by phenylpropanoid pathway. Although there have been efforts to understand PT biosynthesis, knowledge regarding the same came largely from other plants species such as *Forsythia intermedia* and *Linum* spp. than *Podophyllum* per se. The first step towards lignan biosynthesis is marked by formation of pinoresinol with stereocoupling of two coniferyl alcohol molecules mediated by dirigent protein oxidase (DPO; Fig. 1) (Broomhead et al. 1991; Davin et al. 1997; Xia et al. 2000). Pinoresinol by enantiospecific reduction is converted to lariciresinol and then secoisolariciresinol, and both the steps are catalyzed by pinoresinol/lariciresinol reductase (PLR) (Rahman et al. 1990; Katayama et al. 1993, Dinkova-Kostova et al. 1996). However, the gene encoding PLR has not been cloned from any of the species of *Podophyllum*. Secoisolariciresinol is dehydrogenated by secoisolariciresinol dehydrogenase (SDH) to give matairesinol (Xia et al. 2001). Matairesinol is considered to be a precursor of several important secondary metabolites in different species such as podophyllotoxin in *Podophyllum* (Broomhead et al. 1991; Xia et al. 2000) and trachelogenin in *Ipomea carica* (Schroder et al. 1990). Very little is known about the biosynthetic pathway from matairesinol to podophyllotoxin and is believed to occur via yatein, deoxy-podophyllotoxin (Jackson and Dewick 1984; Farkya et al. 2004). In *Linum flavum*, a cytochrome p450 monooxygenase, deoxy-podophyllotoxin 6-hydroxylase is believed to catalyze step from deoxy-podophyllotoxin to podophyllotoxin (Molog et al. 2001); however, the gene encoding this enzyme has not been cloned so far.

P. hexandrum in spite of being the highest podophyllotoxin-producing plant has been poorly explored with respect to elucidation of podophyllotoxin biosynthetic pathway. Only a few genes encoding enzymes have been cloned and characterized. Elucidations of the podophyllotoxin biosynthetic pathway have been performed in different plant species. As a result, it is not clear to which extent the information available from one plant species would be applicable to other podophyllotoxin-producing plants. Therefore, owing to its rich podophyllotoxin content, *P. hexandrum* could be the best system to elucidate the biosynthesis of this important lignan.

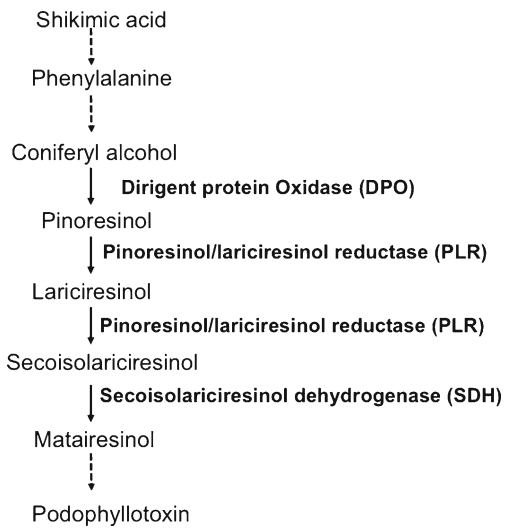


Fig. 1 Podophyllotoxin biosynthesis pathway. *Bold font* indicates enzymes of the respective reactions, *dashed arrows* show involvement of multiple steps

There has been enormous genetic diversity reported in different populations of *P. hexandrum*. Moreover, different populations of *P. hexandrum* have been shown to have different levels of podophyllotoxin. High and low PT-producing plants could form a wonderful system to identify genes involved in PT biosynthesis. Keeping this in mind, we have used suppressive subtractive hybridization (SSH) cDNA library for generation of Expressed Sequence Tags (ESTs) from *P. hexandrum* with the aim of obtaining genes involved in PT biosynthesis.

Here, we report the construction of two subtracted cDNA library from root/rhizome and leaves of *P. hexandrum* and identification of several unique ESTs. Further, from an EST, a full length gene pinoresinol–lariciresinol reductase (*PhPLR*) involved in PT biosynthesis has been cloned. Expression pattern of the cloned gene as well as other known genes of *P. hexandrum* involved in PT biosynthesis is also studied under different stress conditions.

Materials and methods

Plant material

P. hexandrum plants (Yada and Jobrang) were obtained from the CSIR-Institute of Himalayan Bioresource Technology, Palampur, India. Plants were maintained in a green house at National Institute of Plant Genome Research, New Delhi, India. Leaves and root/rhizome of healthy plants were used as plant material for isolation of RNA.

Construction of subtracted cDNA library and sequence analysis

Total RNA was isolated from root/rhizome tissue using RNeasy mini kit (Qiagen). On-column DNAse digestion was performed using "RNase free DNase" (Qiagen). Subtractive hybridization was performed using the PCR select cDNA subtraction kit (Clontech, USA) as per manufacturer's instructions. In brief, first and second strand cDNA was synthesized from 2 μg of mRNA (poly A^+ RNA) from Yada roots/rhizome or leaf (tester population) and Jobrang roots/rhizome or leaf (driver population). Double strand tester and driver cDNA were digested with *RsaI*, and the driver cDNA was ligated with adapters for hybridization reactions. In the first hybridization, an excess of driver cDNA was

hybridized at 68 °C for 10 h with tester cDNA. In the second hybridization, the reaction was incubated at 68 °C overnight for hybridization in the presence of fresh driver cDNA. The subtracted product was then PCR amplified by using oligonucleotide primers complementary to adapters. PCR reaction was carried with initial denaturation (75 °C for 5 min) and 94 °C for 25 s followed by 27 cycles of 94 °C for 10 s, 66 °C for 30 s, and 72 °C for 1.5 min. A nested PCR reaction was carried out with 12 cycles at 94 °C for 10 s, 68 °C for 30 s, and 72 °C for 1.5 min. The final PCR product was checked on a 2.0 % agarose/EtBr gel which showed a clear smear. The subtracted product was purified using PCR Purification kit (QIAGEN, Germany) and ligated in pJET1.2 cloning vector (Fermentas, USA). Ligation reactions were transformed in *Escherichia coli* DH5α cells, and positive clones were selected on Luria Agar plate with ampicillin (100 μg/ml). The clones were screened by colony PCR using vector-specific forward and reverse primers, and clones with sufficiently larger size (>200 bp) were selected for plasmid isolation and sequencing.

Sequencing of plasmids was done using the Big Dye Terminator™ version 3.0 (Applied Biosystems, USA) and analyzed using 3700 ABI Prism 96-capillary sequence analyzer. The Vector and adaptor sequences were removed from the sequences with the help of VECSCREEN program (http://www.ncbi.nlm.nih.gov/VecScreen/VecScreen.html) and further verified by manual inspection. The sequences were screened for homology in the GenBank database using BLASTX (http://www.ncbi.nlm.nih.gov/BLAST/). The high-quality sequences were fed for Blast2GO analysis, a web-based annotation application, following the default parameters (Conesa and Götz 2008; Götz et al. 2008, 2011). Blast2GO analysis comprised three steps, blast to find homologous sequences, mapping to collect GO terms associated to blast hits, and annotation to assign trustworthy information to query sequences. The ESTs were later grouped into ten different categories as per their putative cellular functions.

Stress treatment

Leaves of the healthy plant were subjected to different stress conditions in the following manner. Wounding stress was performed by puncturing the young leaves attached to plants several times across the apical lamina with a surgical blade, which effectively wounded ~40 % of the leaf area.

For methyl jasmonate (MeJA) treatment, cotton ball soaked in methyl jasmonate solution (50 μM) was used to apply MeJA on adaxial as well as abaxial surfaces of the detached leaves. The treated leaves were kept on blotting paper soaked in 1⁄10th Murashige–Skoog basal medium, and the trays containing leaves were sealed with saran wrap. In control experiments, leaves were painted in a similar fashion with water containing the same amount of ethanol present in MeJA solution of the treatment. For UV treatment, plants were exposed under under UV C lamp (253 μmol m^{-2} s^{-1}) for 2 min and harvested at various time periods. For each treatment, leaves were harvested at different time points by snap freezing in liquid nitrogen, and stored at 80 °C for further analyses.

Cloning of full length *PhPLR*

The first strand cDNA was synthesized according to the manufacturer's guidelines of the SMART rapid amplification of cDNA ends (RACE) cDNA amplification kit (Clontech, Palo Alto, CA) using PowerScripts reverse transcriptase, SMART II oligonucleotide, and 5′-RACE CDS primer provided with the kit. First PCR was performed with the UPM (Universal primer) provided with the kit as forward primer and gene-specific primer 5′GSP as the reverse primer. The PCR was carried out by first denaturing the cDNA at 94 °C for 3 min followed by 35 cycles of amplification (amplification 94 °C for 30 s, 63 °C for 30 s, 72 °C for 60 s) and a final extension at 72 °C for 5 min. The nested PCR was then performed with the internal primer, nested universal primer provided with the kit as a forward primer and the nested gene-specific primer 5′NGSP as the reverse primer with the annealing temperature of 60 °C. The PCR product was purified and cloned into pJET1.2 cloning vector. Positive clones were screened by colony PCR, and the clones that showed expected size were selected for sequencing.

Phylogenetic analysis

Phylogenetic analysis of PhPLR and PLR from other plant species was performed using "phylogeny.fr" (www.phylogeny.fr) by "One Click" mode. The default pipeline included the following programs: MUSCLE for multiple alignment, Gblocks for automatic alignment curation, PhyML for tree building, and TreeDyn for tree drawing (Dereeper et al. 2008, 2010). For branch support, "PhyML" was run with the aLRT statistical test (Anisimova and Gascuel 2006).

Semiquantitative RT-PCR

Total RNA was extracted using RNeasy mini kit along with the on-column Rnase free DNAse treatment (Qiagen). Aliquots of 2 μg of total RNA from each sample were used for cDNA synthesis using Revert AidTM H minus first strand cDNA synthesis kit (Fermentas, USA). First strand cDNA (1 μl) was used as template in a 25-μl PCR reaction volume using gene-specific primers. The PCR reaction for each target gene was performed at 94 °C for 3 min followed

by 24–30 cycles of amplification (94 °C for 30 s, 55–60 °C for 30 s, 72 °C for 60 s) with final extension for 5 min at 72 °C. PCR amplification of actin was used as control, and amplification was carried out using gene-specific primers under the same conditions as described above to estimate if equal amounts of RNA among samples were being used in RT-PCR reaction. The RT-PCR reactions were repeated at least three times, and representative results are shown in the "Results" section.

Results

Construction of subtractive cDNA library, sequencing, and establishing ESTs of *P. hexandrum*

In order to identify genes involved in podophyllotoxin biosynthesis, subtractive cDNA libraries were synthesized using high and low PT-producing plant populations. Plants from populations Yada and Jobrang have been shown to have high and low PT content, respectively. Forward subtractive cDNA libraries were synthesized each from root/rhizome and leaves using Yada as tester and Jobrang as a driver with the aim of getting genes responsible for podophyllotoxin biosynthesis. From the cDNA libraries constructed from root/rhizome and leaf of *P. hexandrum*, ~3,000 cDNA clones were screened by PCR amplification. The cDNA clones of sufficiently larger size (>200 bp) were subjected for DNA sequencing. The vector sequences and bad quality sequences were removed after screening thorough VecScreen (NCBI) and verified by manual inspection. Thus, a total of 1,141 clones with high-quality sequences were obtained (908 ESTs from root/rhizome cDNA library and 233 ESTs from leaf cDNA library). The sequences were assembled using CAP3 Sequence Assembly Program (Huang and Madan 1999). From the total 1,141 ESTs, 354 unique ESTs were obtained, 265 from root/rhizome and 89 from leaf tissue. The unique EST sequences have been submitted in the GenBank database under GenBank accession nos. JZ142050–JZ142403.

Unique ESTs were annotated following BLASTX and Blast2Go analysis (Altschul et al. 1997, Conesa et al. 2005; Conesa and Götz 2008; Götz et al. 2008, 2011). The ESTs were classified into ten functional categories according to their putative cellular functions (Fig. 2).

The largest category (52 %) contained EST sequences with hypothetical/unknown function or with no similarity to available sequences in the database. The large subset of ESTs with unknown function or no hits indicates presence of putative novel genes in this category. The third largest category of ESTs after miscellaneous (13.6 %) was ESTs related to metabolism (10.2 %). The other putative functional categories were of photosynthesis (1.4 %), defense/stress related (8.5 %), signalling (3.4 %), translation/post-translation (1.4 %), transcription/post-transcription (2.8 %), transport (2.8 %), and cell development/housekeeping (3.9 %). Interestingly, there were ESTs showing similarity with genes encoding enzymes involved in PT biosynthesis such as PLR and DPO. Details of important ESTs primary and secondary metabolism, defense, transport, and signalling have been shown in Table 1.

Cloning of full length pinoresinol–lariciresinol reductase

Among the clones identified in the SSH library of high and low podophyllotoxin-yielding plants, a 584-bp-long EST (PhR208, accession no. JZ142346) was obtained which showed strong homology with pinoresinol-lariciresinol reductase from different plant species including *F. intermedia* and *Arabidopsis thaliana*. Taking PhR208 sequence as a base, 5′-RACE strategy was used to clone the full length gene. Two reverse gene-specific primers (5′ GSP and 5′ NGSP) were designed based on the sequence of the PhR208 for 5′ RACE. After two PCR amplifications, a single band of about 500 bp was obtained. Based on the middle region sequence and the sequence of clone obtained from 5′ RACE, the full-length cDNA sequence was deduced

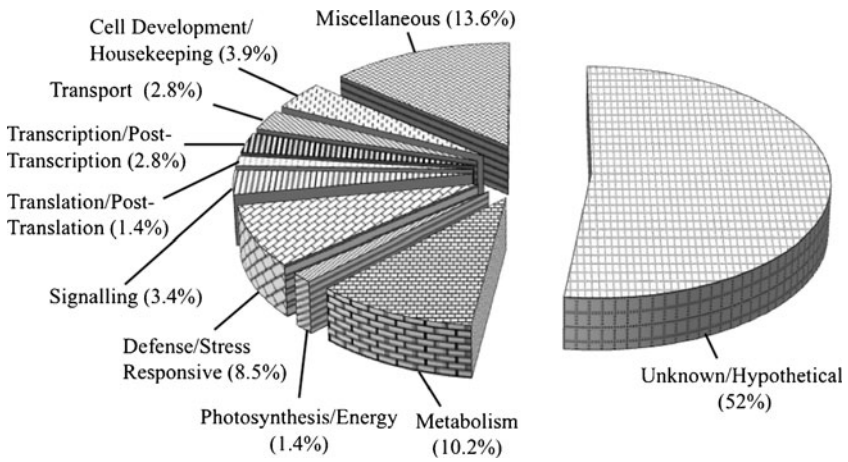


Fig. 2 Functional classification of unique ESTs of *P. hexandrum*

Table 1 Important ESTs from roots/rhizome and leaf of *P. hexandrum* subtracted cDNA libraries

EST name	Accession no.	Size (bp)	Putative function based on similarity with sequences available in NCBI database	*E* value
PhR9	JZ142147	509	Cytochrome p450 like tbp	1.20E−21
PhR18	JZ142156	485	Phenylalanine ammonia-lyase	2.64E−84
PhR23	JZ142161	525	Ring finger and chy zinc finger domain-containing protein 1	8.37E−09
PhR67	JZ142205	268	Copper amine oxidase family protein	7.57E−40
PhR73	JZ142211	176	Vacuolar sorting protein 35	9.64E−26
PhR84	JZ142222	530	Diacylglycerol kinase	8.19E−46
PhR86	JZ142224	628	Succinate dehydrogenase subunit 4	1.07E−60
PhR96	JZ142234	240	S-norcoclaurine synthase 1-like	2.85E−30
PhR106	JZ142244	748	Heavy metal associated domain-containing protein	7.57E−19
PhR108	JZ142246	782	Defensin-like protein 19	7.69E−04
PhR159	JZ142297	578	Ascorbate peroxidase	4.97E−20
PhR176	JZ142314	408	Farnesylcysteine lyase-like protein	1.08E−37
PhR178	JZ142316	606	Pyrophosphate-energized vacuolar membrane proton pump	6.00E−107
PhR184	JZ142322	513	Pathogenesis-related protein	3.31E−48
PhR187	JZ142325	303	Disease resistance response protein	3.53E−13
PhR188	JZ142326	606	Cold induced protein	8.52E−15
PhR196	JZ142334	427	Antifungal protein	3.33E−12
PhR197	JZ142335	335	Protein kinase family protein	3.05E−22
PhR216	JZ142354	217	Casein kinase ii subunit beta-like	7.63E−16
PhR224	JZ142362	141	G-type lectin s-receptor-like serine threonine-protein kinase At2g19130-like protein	2.52E−12
PhR241	JZ142379	646	Senescence-associated protein	1.87E−54
PhR263	JZ142401	221	Mago nashi-like protein	1.90E−44
PhR264	JZ142402	121	hin1-like protein	4.02E−12
PhL4	JZ142053	256	Coronitine insensitive 1like protein	1.07E−04
PhL11	JZ142060	297	Importin subunit beta-1-like	3.57E−52
PhL33	JZ142082	401	Myb domain protein 28	9.27E−41
PhL41	JZ142090	416	Heat shock protein 83-like	1.58E−72
PhL51	JZ142100	130	Nematode-resistance protein	5.84E−11
PhL55	JZ142104	490	Maternal effect embryo arrest 9 protein	1.71E−55
PhL59	JZ142108	927	K(+) efflux antiporter chloroplastic-like	5.23E−99
PhL66	JZ142115	437	Potassium channel	8.78E−74
PhL77	JZ142126	523	Mitogen-activated protein kinase 4	1.18E−93
PhL87	JZ142136	289	Protein kinase	3.52E−42

in which a 47-bp 5′ untranslated region was found upstream of first ATG codon. A complete coding region for the cDNA clone was then amplified using a primer pair complementary to 5′ UTR and 3′ region of the gene and reconfirmed by sequencing. Owing to its high homology with PLR of other plants, we named the gene as *PhPLR*. *PhPLR* sequence has been submitted to GeneBank (accession no. EU855792). Full-length *PhPLR* cDNA sequence and deduced amino acids have been shown in Fig 3a.

Sequence analysis of *PhPLR*

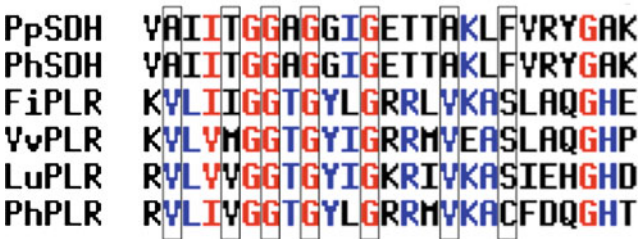
PhPLR shows a 936-bp-long open reading frame encoding protein of 311 amino acids with a calculated molecular weight of about 34.8 kDa and pI of 6.64. The predicted molecular weight is in agreement with GST fused PhPLR as observed on SDS–PAGE (Supplementary Fig. 1). The sequence analysis of PhPLR protein using SIGNAL pv3.0 (Bendtsen et al. 2004) revealed a possible *N*-glycosylation site at residues 214–217 (NKTI) and 299–302 (NYTT), although no secretary signal was found. Further 13 possible protein phosphorylation sites were observed at residues 50, 203, and 227 (protein kinase C-type), residues 50, 62, 68, 134, 206, 227, 247, 249, 281, and 302 (casein kinase II-type). Among the invariant amino acids known among PLRs, NAD(P)H binding region with the conserved glycine residues with the sequence GXGXXG (where X is any residue) and six conserved hydrophobic residues were

a

```
 1 gaggactactatactcaagagagaaaagaaagaactaacaagttaag

48 atggctaagagcagagttctcattgttgggggtacagggtaccta
    M  A  K  S  R  V  L  I  V  G  G  T  G  Y  L   15
93 ggaaggagaatggtaaaggcttgctttgaccaaggtcacactaca
    G  R  R  M  V  K  A  C  F  D  Q  G  H  T  T   30
138 tatgttctacatcgtcaagagattggcgtcgacatcgataagatc
    Y  V  L  H  R  Q  E  I  G  V  D  I  D  K  I   45
183 cagatgctactttcgttcaaggagcaaggggcacaccttgttgag
    Q  M  L  L  S  F  K  E  Q  G  A  H  L  V  E   60
228 ggctcattcaatgatcatcgcagccttgttgaggccgtgaaattg
    G  S  F  N  D  H  R  S  L  V  E  A  V  K  L   75
273 gttgatgttgttatatgtactatttctggagtacatataaggagc
    V  D  V  V  I  C  T  I  S  G  V  H  I  R  S   90
318 catcagatattgttacaacttaagcttgttgaagcaatcaaggaa
    H  Q  I  L  L  Q  L  K  L  V  E  A  I  K  E  105
363 gctggaaatgtaaagcgattcttgccatcggagtttggtatggat
    A  G  N  V  K  R  F  L  P  S  E  F  G  M  D  120
408 ccagcacggatggcacatgcaatggaacctggaagggcaacattt
    P  A  R  M  A  H  A  M  E  P  G  R  A  T  F  135
453 gacgaaaagatggtggtgagaaaggcgatcgaagatgctaaaatt
    D  E  K  M  V  V  R  K  A  I  E  D  A  K  I  150
498 cctcatacttatgcttcagcaaattgttttgctggatacttttg
    P  H  T  Y  A  S  A  N  C  F  A  G  Y  F  L  165
543 gggggtctgtgtcaattcggtaagatcataccctcgaaggaaagt
    G  G  L  C  Q  F  G  K  I  I  P  S  K  E  S  180
588 gtgattttaagcggagatggcaatgtaaaaggtatttatgtggat
    V  I  L  S  G  D  G  N  V  K  G  I  Y  V  D  195
633 gaatacgatattgcgacttacaccatcaagaccatggatgatcct
    E  Y  D  I  A  T  Y  T  I  K  T  M  D  D  P  210
678 cgtacgctgaacaaaacgatctatattaggcctccagcgaacatt
    R  T  L  N  K  T  I  Y  I  R  P  P  A  N  I  225
723 ctatctcaaagagaggtggtagaaatttgggagaaactcattggg
    L  S  Q  R  E  V  V  E  I  W  E  K  L  I  G  240
768 aaagtgttagacaagtcttccttatcagaggaggacttcctggct
    K  V  L  D  K  S  S  L  S  E  E  D  F  L  A  255
813 ctcatgaaaggtctgagtcatggacatcaggcaggattgacacat
    L  M  K  G  L  S  H  G  H  Q  A  G  L  T  H  270
858 tattatcatgtttcctatgaggggtgccttacaaattttgaagta
    Y  Y  H  V  S  Y  E  G  C  L  T  N  F  E  V  285
903 gaggatggagtagatgcttcaaagctttatccacaagtgaattac
    E  D  G  V  D  A  S  K  L  Y  P  Q  V  N  Y  300
948 actacagtgtccgaatatctcaaacgatatttgtag 983
    T  T  V  S  E  Y  L  K  R  Y  L  *          311
```

b



◀ **Fig. 3** Full-length cDNA sequence and deduced amino acid sequence of *P. hexandrum* pinoresinol lariciresinol reductase (*PhPLR*) (**a**). The start and the stop codon are indicated by a *rectangle* and an *asterisk* (*), respectively. The NADPH binding domain conserved residues are *circled*. Amino acids residues (214–217, 299–302) at possible *N*-glycosylation site are indicated by *italic* and *boldface type.* The possible protein phosphorylation sites are indicated by *squares*. The complementary regions of gene-specific primers used for 5′ RACE are *underlined.* The nucleotide sequences complementary to forward and reverse gene-specific primers used for amplification of full-length cDNA clone are shown in *boldface letters. Numbers to the left and right* of the sequence indicate the nucleotide and the amino acid residue, respectively. **b** Alignment of the amino acids sequences of the regions of polypeptide involved in coenzyme binding of selected reductases and dehydrogenses. Amino acid sequences included are secoisolariciresinol dehydrogenases from *P. peltatum* (PpSDH-AAK38664), *P. hexandrum* (PhSDH-ABN14311), pinoresinol lariciresinol reductase from *F. intermedia* (FiPLR-AAC49608), *P. hexandrum* (PhPLR-ACF71492), and *L. usitatissimum* (LuPLR-CAH60858) and isoflavone reductase-like protein from *V. vinifera* (VvIR-CAI56335)

observed in PhPLR and also other dehydrogenases *like* alcohol dehydrogenase from *Drosophila melanogaster* and lactate dehydrogenase from dogfish (Dinkova-Kostova et al. 1996). In the similar line, a comparison was made between PhPLR, SDH from *P. peltatum* and *P. hexandrum*, PLR from *F. intermedia* and *Linum usitatissimum* and isoflavone reductase (IR) from *Vitis vinifera*. It has been observed that the NAD(P)H binding region was highly conserved not only in PLRs but also in SDH and IR. The glycine residues and adjoining four hydrophobic amino acids of PLRs were aligned perfectly with SDH and IR (Fig. 3b).

A BLAST search (Altschul et al. 1997) was conducted with the translated amino acids of PhPLR against the non-redundant protein database at National Centre of Biological Information. Significant homology was noted for PhPLR with the PLRs from other plants, *F. intermedia* (GenBank accession number U81158; 75 % identity and 84 % similarity) *Linum strictum* (EU107358; 75 % identity and 85 % similarity), *Linum album* (AJ849358; 74 % identity and 84 % similarity), Isoflavone Reductase of *Ricinus communis* (CAI56335; 74 % identity and 86 % similarity), phenyl caumaran benzyl ether reductase like protein of *Populus trichocarpa* (XM002326541; 75 % identity and 85 % similarity), and *A. thaliana* PLR (R20416; 61 % identity and 78 % similarity). The homology between PhPLR and PLR from different plant species can be viewed in multiple sequence alignment (Fig. 4).

The relationship between PhPLR with PLRs from other plants was investigated using a parsimonious phylogenetic analysis. Phylogenetic tree was constructed with the help of "phylogeny.fr" (Dereeper et al. 2008, 2010) using amino acid sequence of PLRs from different plant species. As seen in Fig. 5, PhPLR was found to be close to PLR from *F. intermedia* followed by *L. album* as they are grouped together in the same cluster and was most distant from PLR of *Thuja plicata*.

Expression pattern of *PhPLR* and *PhSDH* and *PhDPO* in different tissues and under abiotic stress

The expression pattern of *PhPLR* was investigated at mRNA transcript level in different tissues and stress condition as well as with methyl jasmonate (MeJA) treatment.

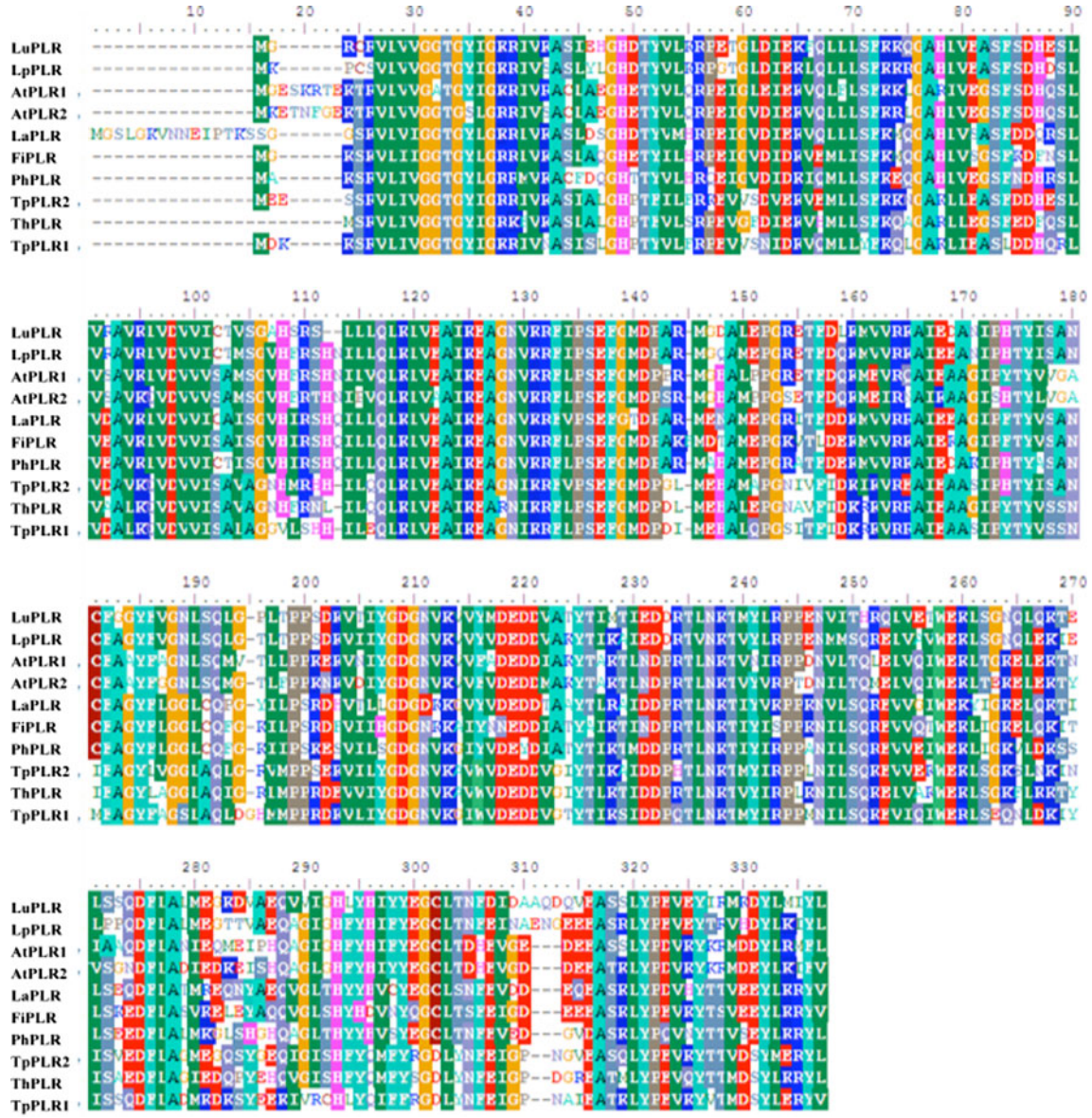


Fig. 4 CLUSTALW 1.82 multiple alignment of translated amino acid sequence of PhPLR with PLRs from other plant species. Amino acids sequences are retrieved from the NCBI database, i.e., *L. usitatissimum* (CAH60858), *L. perenne* (ABM68630), *A. thaliana* (AtPLR1-AAL38690, AtPLR2-AAO22699), *L. album* (CAH60857), *Forsythia x intermedia* (AAC49608), *T. plicata* (TpPLR1- AAF63507), *T. plicata* (TpPLR2-AAF63508), *Tsuga heterophylla* (AAF64184)

The other known genes encoding enzymes involved in podophyllotoxin biosynthetic pathway (Fig. 1) both above and below the action of PLR, such as *DPO* and *SDH*, were also included anticipating their similar expression pattern. Abundant expression of *PhPLR* was observed in leaf, stem, and root/rhizome tissues of *P. hexandrum* with highest expression in leaf and least in stem. *PhDPO* on other hand showed highest expression in leaf and least in root, while *PhSDH* had least expression in leaf compared to stem and root (Fig. 6a).

It has been reported that physical damage to the leaves of *P. hexandrum* results in accumulation of podophyllotoxin glucosides and aglucon at the wound site (Oliva et al. 2002). Further, a study in cell culture of *F. intermedia* reports three- to sevenfold increase in accumulation of lignan pinoresinol and matairesinol (both predominant as glucoside) by addition of methyl jasmonate (Schmitt and Petersen 2002). In order to investigate whether *PhPLR* and other genes involved in podophyllotoxin biosynthesis are stress inducible, expression of *PhPLR*, *PhDPO*, and *PhSDH* was studied by semiquantitative RT-PCR upon wounding, UV stress, and application of MeJA. Leaves of *P. hexandrum* were subjected to wounding and UV stress as well as

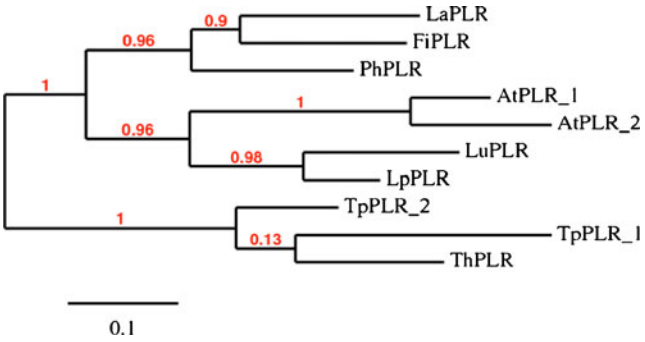


Fig. 5 Phylogenetic relationship between PhPLR and other related PLRs. The alignment consists of amino acid sequence of PLRs from different plants which are used for multiple alignments. Phylogenetic tree was built following maximum likelihood method using PhyML. Values at branch points indicate confidence index of approximate likelihood-ratio test (aLRT)

treatment with methyl jasmonate and analyzed for transcripts accumulation over a time course of 12 h. An increase in the level of *PhPLR* expression was noted with wounding stress and methyl jasmonate treatment (Fig. 6b). In wounding stress, the transcripts accumulation of *PhPLR* was highest at 3 h followed by at 6 h after wounding and declined at 12 h. In methyl jasmonate treatment, an increase in the level of *PhPLR* mRNA transcripts was observed with increasing time up to 6 h and diminished thereafter. Interestingly, the expression pattern of both *PhDPO and PhSDH* followed the near similar pattern to that of *PhPLR* both in response to wounding stress and methyl jasmonate treatment. In response to UV, PhPLR showed higher expression at 0.5 and 1 h after UV treatment to that of control; however, the expression declines thereafter (Fig. 6b). However, *PhDPO* and *PhSDH* also showed different trend to that of *PhPLR* in UV stress.

Discussion

Podophyllotoxin produced from *P. hexandrum* is an important source of clinically important anticancer drugs. Although there have been efforts for chemical synthesis of the compounds, it met with limited success. Elucidation of complete biosynthesis of PT and identification of genes encoding enzymes catalyzing steps in PT biosynthesis and regulators may form the basis for further studies utilizing genetic engineering approach for enhanced and regular production of PT. Such alternative would not only help to provide a continuous supply to industry but also help to maintain the natural *P. hexandrum* populations.

Availability of *P. hexandrum* genotypes with varying PT biosynthetic ability could be the valuable resource for identification of genes involved in biosynthesis and regulation of PT using the genomics approaches. We have used SSH and EST approach to identify genes of *P. hexandrum* that could be involved in regulation and biosynthesis of PT. The EST approach has been successfully utilized in identification of disease-responsive genes in several plant species such as soybean (Iqbal et al. 2005), rice (Jantasuriyarat et al. 2005), ginseng (Goswami and Punja 2008), and abiotic stress such as dehydration and salinity-responsive genes (Lata et al. 2010; Puranik et al. 2011). Sequencing of clones from two SSH libraries, each from root/rhizome and leaf of *P. hexandrum*, resulted in identification of 354 unique ESTs. The majority of ESTs (52 %) showed sequence similarity with genes with unknown function or "no hit" which could help in identification of new genes with novel functions.

In the present work, the identified ESTs also show presence of ESTs with high sequence similarity with pinoresinol/lariciresinol reductase (*PLR*) and dirigent protein oxidase (*DPO*) from different plant species which are

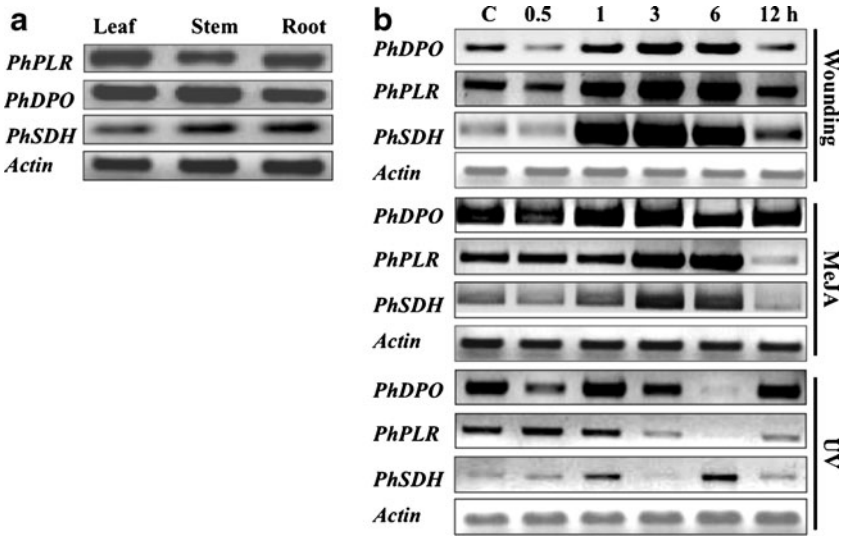


Fig. 6 Expression patterns of genes involved in podophyllotoxin biosynthesis under different tissues and stress conditions in *P. hexandrum*. RT-PCR analysis of expression patterns of *PhPLR*, *PhDPO*, and *PhSDH* in different tissues of *P. hexandrum* (**a**) and upon wounding, UV and methyl jasmonate (MeJA) treatment (**b**). Amplification of actin was used to ensure equal amount of RNA used for cDNA preparation and equal amount of cDNA in PCR reactions

known to be involved in the PT biosynthesis. This suggests efficiency of SSH library in identification of genes of PT biosynthesis.

In the PT biosynthetic pathway, deoxypodophyllotoxin is converted to podophyllotoxin, and the reaction is considered to be catalyzed by cytochrome P450 monooxygenase (Molog et al. 2001). However, the gene encoding the same has not been identified and cloned so far. Interestingly, in the present EST collection, there were ESTs with high redundancy showing high sequence similarity with CYP450 monooxygenase from maize and tobacco. Full-length cloning and elaborate characterization of the CYP450 EST would be needed to speculate whether it has any role in PT biosynthesis. Pinoresinol is considered to be the first committed and common entry step towards lignan biosynthesis which later separate into several directions in different plants species. In the PT biosynthesis, the first few steps until matairesinol is formed have been well established in other plants species (Farkya et al. 2004). Pinoresinol is enantiospecifically reduced in two steps leading to secoisolariciresinol by pinoresinol lariciresinol reductase. The same steps have been confirmed in *L. flavum*, a member of Linaceae family (Xia et al. 2000). We here report the cloning and characterization of pinoresinol–lariciresinol reductase (*PhPLR*) gene from the highest podophyllotoxin-yielding plant, *P. hexandrum*. The *PhPLR* gene is 936 bp long and encodes 311 amino acids. Computational analysis revealed the NAD(P)H binding site as seen with the presence of three conserved glycine residues and four conserved hydrophobic residues. The glycine-rich region is considered to play a central role in positioning the NAD(P)H in its correct conformation. It is interesting to note that this NAD(P)H binding site is found conserved not only in SDH (which acts in the same biosynthetic pathway to that of PLR) and IR (Fig. 3b) but also in alcohol dehydrogenase from *D. melanogaster*, cinnamyl alcohol dehydrogenase from *Pinus taeda*, lactate dehydrogenase from the dogfish muscle, and glutathione reductase from human erythrocyte (Dinkova-Kostova et al. 1996). High sequence identity by BLAST search and phylogenetic analysis showed that PhPLR is close to PLR from *F. intermedia* and *L. album*. So it could be speculated that PhPLR may have a role in conversion of pinoresinol to lariciresinol and lariciresinol to secoisolariciresinol as that of PLR of *L. album* and *F. intermedia*. However, the purification and biochemical characterization of GST-fused heterologously expressed recombinant PhPLR would more profoundly reveal its function in catalysis of these important reactions in lignan biosyntheis and its substrate preference. Nonetheless, the similar regulation pattern of mRNA transcripts levels of *PhPLR* to that of *DP* and *SDH* indicates the involvement of *PhPLR* in the podophyllotoxin biosynthetic pathway.

High sequence identity of PhPLR with isoflavone reductase and phenyl caumaran benzyl ether reductase like protein is of considerable interest since isoflavonoids are formed via the related branch of phenylpropanoid pathway metabolism (Gang et al. 1999). For example, phenyl caumaran benzyl ether reductase catalyses reduction of dehydrodiconiferyl alcohol and dihydrodehydrodiconiferyl alcohol to isodihydrodehydroconyferyl alcohol and tetrahydrodehydrodiconiferyl alcohol, respectively (Gang et al., 1999).

The commercial production of podophyllotoxin depends mainly on roots/rhizome of Podophyllum, and the other plants parts are not commonly used for extraction. Moraes et al. (2002) showed that podophyllotoxin from the leaves of *P. peltatum* can be obtained to the amount as high as roots/rhizome of *P. hexandrum* with the use of specific buffer for the extraction of this lignan suggesting that podophyllotoxin is also present in the leaves of these plants. The abundant expression of *PhPLR*, *DP*, and *SDH* observed in the present study, in all the tissue types studied (leaf, stem and root/rhizome), indicates that podophyllotoxin is indeed synthesized in the leaves but could be available in its glycosidic form.

In nature, plants produce secondary metabolites to defend themselves against various biotic and abiotic stress conditions. Lignans are very well known to play an important role in plant defense (Fukuda et al. 1985, Figgitt et al. 1989). The wound inducible expression pattern of *PhPLR*, *DP*, and *SDH* observed in this study is in agreement with the accumulation of podophyllotoxin glucoside in leaves in response to physical damage in *P. hexandrum* (Oliva et al. 2002). Further, increase in the accumulation of pinoresinol and mairesinol lignans in response to methyl jasmonate in *F. intermedia* cell cultures (Schmitt and Petersen 2002) and the observation of methyl jasmonate-inducible expression pattern of podophyllotoxin biosynthetic genes in the present work suggest that the genes involved in podophyllotoxin biosynthtic pathway are methyl jasmonate responsive. It is to note that PhPLR contains 13 possible phosphorylation sites (casein kinase type II protein phosphorylation site), some of which are conserved (Thr-302) in its orthologs as well. Presence of ESTs showing similarity to casein kinase and mitogen activated protein kinase (MAPK) from subtraction library indicates possible higher expression of these genes in high PT-producing plant. Interestingly, MAPKs are also known to be activated by MeJA and wounding in addition to several other stress conditions (Sinha et al. 2011; Raina et al. 2012). It will be interesting to know whether PhPLR activity per se and/or higher accumulation of PT/lignan in response to wounding and MeJA is regulated by these kinases. In this perspective, in *Catharanthus roseus* higher accumulation of transcripts of genes involved in biosynthesis and regulation of monoterpenoid indole

alkaloid (MIA) in response to MeJA and wounding have been observed (Kumar et al. 2007; Peebles et al. 2009; Jaggi et al. 2011; Raina et al. 2012). Recently, involvement of MAPK in wounding, MeJA, and UV has been proposed (Raina et al. 2012). It is also plausible that the genes of podophyllotoxin biosynthetic pathway are regulated by a common regulator which is methyl jasmonate inducible as known in the well-studied MIA pathway in *C. roseus* where methyl jasmonate-responsive AP2 domain transcription factor, ORCA-3, regulates the expression of most of the MIA pathway genes (van der Fits and Memelink 2000). An elaborative work in *P. hexandrum* in identifying more number of genes especially by EST-based investigations and their functional characterization would help to elucidate podophyllotoxin biosynthesis pathway and understand its regulation.

Acknowledgments DPW thanks University Grant Commission, India for fellowships. This work is supported by funding by Department of Biotechnology, India.

Conflicts of interest The authors declare that they have no conflict of interest.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search program. Nucleic Acids Res 25:3389–3402

Anisimova M, Gascuel O (2006) Approximate likelihood ratio test for branches: a fast, accurate and powerful alternative. Syst Biol 55:539–552

Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783–795

Broomhead AJ, Rahman MA, Dewick PM, Jackson DE, Lucas JA (1991) Matairesinol as precursor of *Podophyllum* lignans. Phytochem 30:1489–1492

Canel C, Moraesb RM, Dayana FE, Ferreirab D (2000) Podophyllotoxin. Phytochem 54:115–120

Choudhary DK, Kaul BL, Khan S (1998) Cultivation and conservation of *Podophyllum hexandrum*—an overview. J Med Aromat Plant Sci 20:1071–1073

Conesa A, Götz S (2008) Blast2GO: a comprehensive suite for functional analysis in plant genomics. Int J Plant Genom 2008:1–13

Conesa A, Götz S, Garcia-Gomez JM, Terol T, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21:3674–3676

Davin LB, Wang HB, Crowell AL, Bedgar BL, Martin DM, Sarkanen S, Lewis NG (1997) Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. Science 275:362–366

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36(Web Server issue):465–469

Dereeper A, Audic S, Claverie JM, Blanc G (2010) BLAST-EXPLORER helps you building datasets for phylogenetic analysis. BMC Evol Biol 10:8

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. J Biol Chem 271:29473–29482

Farkya S, Bisaria VS, Srivastava AK (2004) Biotechnological aspects of the production of the anticancer drug podophyllotoxin. Appl Microbiol Biotechnol 65:504–519

Figgitt DP, Denever SP, Dewick PM, Jackson DE, Willians P (1989) Topoisomerase II: a potential target for novel antifungal agents. Biochem Biophys Res Commun 160:257–262

Fukuda Y, Osawa T, Namiki M, Ozaki T (1985) Studies on antioxidative substances in sesame seed. Agric Biol Chem 49:301–306

Gang DR, Kasahara H, Xia ZQ, Vander Mijnsbrugge K, Bauw G, Boerjan W, Van Montagu M, Davin LB, Lewis NG (1999) Evolution of plant defense mechanisms. Relationships of phenylcoumaran benzylic ether reductases to pinoresinol-lariciresinol and isoflavone reductases. J Biol Chem 274:7516–7527

Goswami R, Punja ZK (2008) Molecular and biochemical characterization of defense responses in ginseng (*Panax quinquefolius*) roots challenged with *Fusarium equiseti*. Physiol Mol Plant Pathol 72:10–20

Götz S et al (2008) High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res 36:3420–3435

Götz S et al (2011) B2G-FAR, a species centered GO annotation repository. Bioinformatics 27:919–924

Huang X, Madan A (1999) CAP3: a DNA sequence assembly program. Genome Res 9:868–877

Iqbal MJ, Yaegashi S, Ahsan R, Shopinski KL, Lightfoot DA (2005) Root response to *Fusarium solani* f. sp. glycines: temporal accumulation of transcripts in partially resistant and susceptible soybean. Theor Appl Genet 110:1429–1438

Jackson DE, Dewick PM (1984) Aryltetralin lignans from *Podophyllum hexandrum* and *Podophyllum peltatum*. Phytochem 23:1147–1152

Jaggi M, Kumar S, Sinha AK (2011) Overexpression of an apoplastic peroxidase gene *CrPrx* in transgenic hairy root lines of *Catharanthus roseus*. Appl Microbiol Biotechnol 90:1005–1016

Jantasuriyarat C, Gowda M, Haller K, Hatfield J, Lu G, Stahlberg E, Zhou B, Li H, Kim H, Yu Y, Dean RA, Wing RA, Soderlund C, Wang GL (2005) Large-scale identification of expressed sequence tags involved in rice and rice blast fungus interaction. Plant Physiol 138:105–115

Katayama T, Davin LB, Chu A, Lewis NG (1993) Novel benzylic ether reductions in lignan biogenesis in *Forsythia intermedia*. Phytochem 33:581–591

Kumar S, Dutta A, Sinha AK, Sen J (2007) Cloning, characterization and localization of a novel basic peroxidase gene from *Catharanthus roseus*. FEBS J 274:1290–1303

Lata C, Sahu PP, Prasad M (2010) Comparative transcriptome analysis of differentially expressed genes in foxtail millet (*Setaria italica* L.) during dehydration stress. Biochem Biophys Res Commun 393:720–727

Molog GA, Empt U, Petersen M, Van Uden W, Pras N, Alfermann AW (2001) Deoxypodophyllotoxin 6-hydroxylase, a cytochrome P450 monooxygenase from cell cultures of *Linum flavum* involved in the biosynthesis of cytotoxic lignans. Planta 214:288–294

Moraes RM, Bedir E, Barrett H, Burandt C Jr, Canel C, Khan IA (2002) Evaluation of *Podophyllum peltatum* accessions for podophyllotoxin production. Planta Med 68:341–344

Oliva A, Moraes RM, Watson SB, Duke SO, Dayan FE (2002) Aryltetralin lignans inhibit plant growth by affecting the formation of mitotic microtubular organizing centers. Pest Biochem Physiol 72:45–54

Peebles CA, Hughes EH, Shanks JV, San KY (2009) Transcriptional response of the terpenoid indole alkaloid pathway to the overexpression of ORCA3 along with jasmonic acid elicitation of *Catharanthus roseus* hairy roots over time. Metab Eng 11:76–86

Puranik S, Jha S, Srivastava PS, Sreenivasulu N, Prasad M (2011) Comparative transcriptome analysis of contrasting foxtail millet cultivars in response to short-term salinity stress. J Plant Physiol 168:280–287

Rahman MA, Dewick PM, Jackson DE, Lucas JA (1990) Biosynthesis of lignans in *Forsythia intermedia*. Phytochem 29:1841–1846

Raina SK, Wankhede DP, Jaggi M, Singh P, Jalmi SK, Raghiram B, Sheikh AH, Sinha AK (2012) CrMPK3, a mitogen activated protein kinase from *Catharanthus roseus* and its possible role in stress induced biosynthesis of monoterpenoid indole alkaloids. BMC Plant Biol 12:134

Schmitt J, Petersen M (2002) Influence of methyl jasmonate and coniferyl alcohol on pinoresinol and matairesinol accumulation in a *Forsythia intermedia* suspension culture. Plant Cell Rep 20:885–889

Schroder HC, Merz H, Steffen R, Muller WEG, Sarin PS, Trumm S, Schulz J, Eich E (1990) Differential in vitro anti-HIV activity of natural lignans. Z Naturforsch C 45:1215–1221

Sinha AK, Jaggi M, Raghuram B, Tuteja N (2011) Mitogen-activated protein kinase signaling in plants under abiotic stress. Plant Signal Behav 6:196–203

van der Fits L, Memelink J (2000) ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. Science 289:295–297

Van Uden W, Pras N, Visser JF, Malingré TM (1989) Detection and identification of podophyllotoxin produced by cell cultures derived from *Podophyllum hexandrum* Royle. Plant Cell Rep 8:165–168

Xia ZQ, Costa MA, Proctor J, Davin LB, Lewis NG (2000) Dirigent mediated podophyllotoxin biosynthesis in *Linum flavum* and *Podophyllum peltatum*. Phytochem 55:537–549

Xia ZQ, Costa MA, Pelissier HC, Davin LB, Lewis NG (2001) Secoisolariciresinol dehydrogenase purification, cloning, and functional expression. Implications for human health protection. J Biol Chem 276:12614–12623