Journal of Experimental Botany, Vol. 57, No. 14, pp. 3921–3932, 2006
doi:10.1093/jxb/erl146 Advance Access publication 18 October, 2006
This paper is available online free of all access charges (see http://jxb.oxfordjournals.org/open_access.html for further details)

Journal of Experimental Botany

RESEARCH PAPER

Transcriptome analysis in *Catharanthus roseus* leaves and roots for comparative terpenoid indole alkaloid profiles*

Ashutosh K. Shukla[1], Ajit K. Shasany[1], Madan M. Gupta[2] and Suman P. S. Khanuja[1,†]

[1] *Genetic Resources and Biotechnology Division, Central Institute of Medicinal and Aromatic Plants, Lucknow 226015, India*

[2] *Analytical Chemistry Division, Central Institute of Medicinal and Aromatic Plants, Lucknow 226015, India*

Received 25 May 2006; Accepted 16 August 2006

Abstract

In *Catharanthus roseus* (L.) G. Don each tissue is known to produce a distinct spectrum of terpenoid indole alkaloids. Since the invaluable antineoplastic bisindole alkaloids are restricted to the aerial parts of the plant and do not occur in its underground tissues, identification of the structural and regulatory factors operating distinctly in the shoot/leaf of the plant will be a necessity for modulation of bisindole alkaloid biosynthesis. This study aimed at elucidating the differential gene expression in the two main tissues (leaf and root) of the plant, well known for their distinct terpenoid indole alkaloid profiles. The leaf and root transcriptomes of *C. roseus* were comparatively analysed using two different approaches: (i) indirectly through construction and characterization of separate cDNA libraries; and (ii) directly through a strategically designed suppression subtractive hybridization, using the leaf and root cDNA populations as tester and driver, respectively. A total of 155 ESTs (55 and 45 from the separate leaf and root cDNA libraries, respectively, and 55 from the subtracted leaf-specific cDNA library) were subjected to homology-based classification and submitted to dbEST. The direct approach yielded an EST for *sgd* (strictosidine β-D-glucosidase) and 16 novel ESTs. *Dat* (acetyl-CoA: 4-*O*-deacetylvindoline 4-*O*-acetyl-transferase) and *sgd* transcripts could not be detected in the root system of the plant (cv. 'Dhawal') at any developmental stage (6 d, 6 weeks, or 6 months). The growth-related decrease in shoot/leaf *dat* and *sgd* transcript levels was paralleled by a concomitant decrease in shoot/leaf vindoline content.

Key words: Catharanthine, *dat*, ESTs, semi-quantitative RT-PCR, *sgd*, suppression subtractive hybridization, TIAs, vinblastine, vincristine, vindoline.

Introduction

Catharanthus roseus (L.) G. Don is one of the most extensively investigated medicinal plants and is known mainly for its pharmacologically important alkaloids (Verpoorte *et al.*, 1997). At present, the *Catharanthus* alkaloids comprise a group of about 130 terpenoid indole alkaloids (TIAs) (van der Heijden *et al.*, 2004). No other single plant species is reported to produce such a wide array of complex alkaloids (Blasko and Cordell, 1990). Wide differences have been noted in the compositions of the alkaloids isolated from the underground and aerial tissues of the plant (Mishra and Kumar, 2000; Shukla, 2005). The roots of the plant accumulate ajmalicine and serpentine, which are important components of medicines for controlling high blood pressure and other types of cardio-vascular maladies. However, the plant is particularly known for its economically important leaf-specific bisindole alkaloids, vinblastine and vincristine, which are potent antineoplastic agents (Svoboda and Blake, 1975; Neuss, 1980) and indispensable constituents of most cancer chemotherapies. They are produced *in vivo* by the condensation of vindoline and catharanthine, both of which originate from the TIA biosynthetic

* The sequence data from this study have been submitted to the dbEST database of GenBank under accession numbers CO995037-CO995086, DT378401-DT378405, DT386291-DT386318, and DT527667-DT527738.

† To whom correspondence should be addressed. E-mail: khanujazy@yahoo.com

Abbreviations: DAT, acetyl-CoA: 4-*O*-deacetylvindoline 4-*O*-acetyl-transferase; ESTs, expressed sequence tags; LTRD: leaf-tester root-driver; OMT, 16-hydroxytabersonine 16-*O*-methyltransferase; SGD, strictosidine β-D-glucosidase; TIAs, terpenoid indole alkaloids.

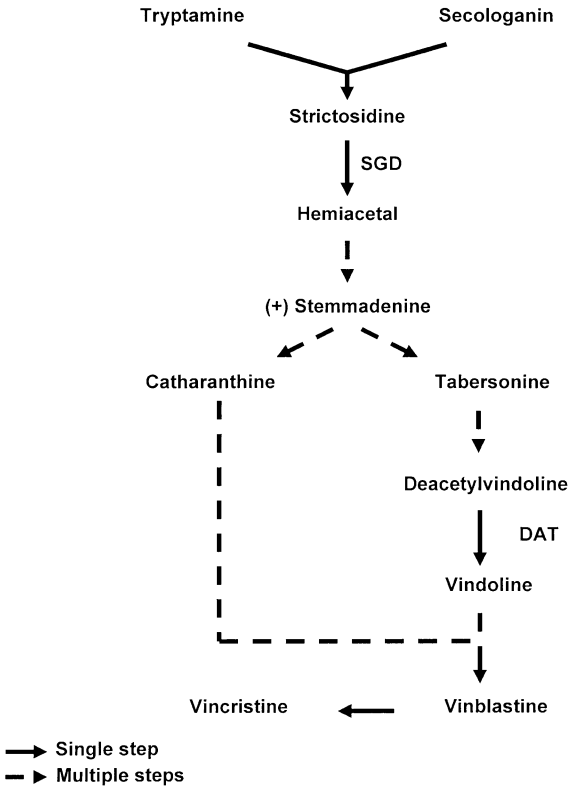

Fig. 1. The TIA biosynthetic pathway in *C. roseus*.

intermediate (+)-stemmadenine (Fig. 1). Vindoline, which is found only in the green parts of the plant and not in the roots or cell suspension cultures, is biosynthesized from the branch-point intermediate tabersonine [which is formed from (+)-stemmadenine through multiple steps] through the action of six enzymatic steps, first proposed by De Luca *et al.* (1986) and now almost completely elucidated using a combination of *C. roseus* plants and cell suspension cultures. These steps are sequentially catalysed by enzymes—T16H (tabersonine 16-hydroxylase), 16-hydroxytabersonine 16-*O*-methyltransferase (OMT), a hydroxylase (hydration), NMT (*S*-adenosyl-L-methionine: 16-methoxy-2,3-dihydro-3-hydroxy-tabersonine-*N*-methyltransferase), D4H (desacetoxyvindoline 4-hydroxylase), and acetyl-CoA: 4-*O*-deacetylvindoline 4-*O*-acetyl-transferase (DAT), of which the last three are leaf-specific. In marked contrast, catharanthine has been reported to be present in the root tissue as well as the aerial part of the plant (Deus-Neumann *et al.*, 1987).

Regulation of alkaloid biosynthesis is accomplished spatially and temporally, by restricting product biosynthesis and accumulation to particular cells and to defined stages of plant development. The presence of different alkaloids and expressed genes within the above-ground and underground tissues of the plant confirms the expression of distinct tissue-specific pathways responsible for the precise tissue- and organ-specific compartmentation of TIA biosynthesis (Laflamme *et al.*, 2001). Compartmentation of the TIA biosynthetic pathway in *C. roseus* also occurs at cellular as well as subcellular levels. Previously, in one of the most elaborate studies on multicellular compartmentation of alkaloid biosynthesis in *C. roseus*, St-Pierre *et al.* (1999) used *in situ* RNA hybridization and immunocytochemistry to establish the cellular distribution of TIA biosynthesis in the plant. *tdc* (tryptophan decarboxylase) and *ss* (strictosidine synthase) mRNAs were found to be present in the epidermis of stems, leaves, and flower buds, whereas they appeared in most protoderm and cortical cells around the apical meristem of root tips. In marked contrast, *d4h* (desacetoxyvindoline 4-hydroxylase) and *dat* mRNAs, which code for enzymes catalysing the late steps of vindoline biosynthesis, were associated with the laticifer and idioblast cells of the shoot (leaves, stems, and flower buds) but were not found in the root tissue. The significant difference in the spectrum of alkaloids in the aerial and underground tissues of the plant was also indicative of the fact that there were differences in the expression pattern of structural as well as regulatory factors in the two cases. These differences could be characterized for gene prospection through comparative transcriptome analysis of the aerial and underground tissues of the plant. Recently, some efforts have been made in this direction by using cDNA-AFLP to differentiate the root and leaf transcriptomes in *C. roseus* statistically (Shukla, 2005).

Considering the aforementioned facts, suppression subtractive hybridization was carried out in the present study to isolate leaf-specific expressed sequence tags (ESTs) from a leaf alkaloid-rich genotype of *C. roseus* (cv. 'Dhawal', US Patent No. 6,548,746; Kulkarni *et al.*, 2003) with the aim of gaining better insight into the bisindole alkaloid biosynthesis in the plant. In a strategically designed experiment, the leaf and root tissues of the plant were used as tester and driver, respectively, to generate ESTs at the plant-physiological stage responsible for high bisindole alkaloid biosynthesis. This 'direct' approach of the leaf-specific transcriptome analysis was also compared with the 'indirect' approach, where separate leaf and root cDNA libraries of *C. roseus* were created for analysis of ESTs from the two tissues.

Materials and methods

RNA isolation

For cDNA synthesis for library construction and suppression subtractive hybridization, poly $(A)^+$ mRNA was isolated from the leaf and root tissues of 6-month-old mature field-grown *C. roseus* (cv. 'Dhawal') plants (vegetatively maintained and genotypically similar) in the month of June (average day temperature invariably being $\geqslant$42 °C) using a previously described protocol (Shukla *et al.*, 2005). For semi-quantitative RT-PCR analyses, total RNA was isolated from the aerial and underground parts of the plant at the three developmental stages (6 d, 6 weeks, and 6 months) using the RNeasy Plant Mini Kit (Qiagen). The three plant developmental stages were obtained as described earlier (Shukla, 2005).

Construction of separate leaf and root cDNA libraries of C. roseus

To improve the representation of the separate leaf and root cDNA libraries and remove any bias, cloning was carried out on both plasmid as well as bacteriophage vectors. For the plasmid cDNA libraries, first-strand cDNA synthesis was primed using a *Not*I-oligo(dT)$_{15}$ primer (5′-GTGCGGCCGCTTTTTTTTTTTTTTT-3′) and the Universal Riboclone® cDNA synthesis system (Promega) was used for the library construction. Size-fractionated cDNA fragments (>400 bp) were directionally cloned into the *Eco*RI and *Not*I cloning sites of pBluescript II SK$^+$ and the ligation products were transformed into *Escherichia coli* DH5α competent cells.

For the bacteriophage cDNA libraries, first-strand cDNA synthesis was primed using a *Xho*I linker-primer (5′-GAGAGAGAGAGA-GAGAGAGAACTAGTCTCGAGTT$_{16}$T-3′) and the ZAP Express® cDNA Synthesis Kit (Stratagene) was used for the library construction. Size-fractionated cDNA fragments (>400 bp) were directionally cloned into the *Eco*RI and *Xho*I sites on the ZAP Express® vector. The ligation products were packaged *in vitro* into phage particles using the ZAP Express® cDNA Gigapack® III gold cloning kit (Stratagene). This primary library was used to infect *E. coli* strain XL1-Blue MRF′. The size of the primary leaf and root cDNA libraries was 2×10^5 pfu and 4.5×10^5 pfu, respectively. The primary cDNA libraries in ZAP Express® vector were amplified and the titre of the amplified leaf and root cDNA libraries was found to be 2.65×10^9 pfu ml^{-1} and 4.30×10^9 pfu ml^{-1}, respectively. The ExAssist™ interference-resistant helper phage was used with *E. coli* strain XLOLR for mass *in vivo* excision of the pBK-CMV phagemids from the lambda phages in both (leaf and root) the amplified ZAP Express cDNA libraries.

The selection markers for pBluescript II SK$^+$ and pBK-CMV were ampicillin and kanamycin, respectively, and the recombinants in both cases were selected by blue/white screening on X-gal/IPTG-coated plates.

Suppression subtractive hybridization and construction of leaf-specific cDNA library of C. roseus

The Clontech PCR-Select™ cDNA subtraction kit (Diatchenko *et al.*, 1996) was used for suppression subtractive hybridization to identify leaf-specific transcripts in *C. roseus* in a strategically designed experiment where the leaf and root cDNA populations served as tester and driver, respectively. All the steps were performed as per the manufacturer's guidelines. The primary and secondary PCRs to amplify the subtracted leaf-specific messages (LTRD=leaf-tester root-driver) were carried out for 27 and 15 cycles, respectively. The secondary PCR products of the subtracted leaf cDNA sample (LTRD) were cloned on pBluescript II SK$^+$ to construct a subtracted leaf-specific cDNA library of *C. roseus* using two different cloning strategies (to enhance the representation of the library). The *Rsa*I (the restriction site present at the adaptor/cDNA junction)-digested secondary PCR product (LTRD) was ligated with *Sma*I-digested pBluescript II SK$^+$ DNA for blunt-end cloning of the subtracted cDNAs (LTRDB ligation products). In another cloning approach, the *Sma*I (the restriction site present on Adaptor 1)- and *Eag*I (the restriction site present on Adaptor 2R)-restricted secondary PCR product (LTRD) was ligated with *Sma*I- and *Eag*I-restricted pBluescript II SK$^+$ DNA for blunt-cohesive site-specific cloning of the subtracted cDNAs (LTRDBC ligation products). The ligation products were used to transform *E. coli* strain DH5α-competent cells and the recombinants were selected by blue/white screening.

EST generation and analysis

Plasmid minipreps from the recombinant clones were used for sequencing the inserts on an ABI PRISM® 377 DNA sequencer (Applied Biosystems) using the ABI PRISM® Big Dye™ terminator cycle sequencing ready reaction kit (Perkin Elmer).

The inserts from the separate leaf and root cDNA libraries were sequenced from their 3′ ends using M13 forward/T7 (for clones in λZAP/pBK-CMV) and M13 reverse (for clones in pBluescript II SK$^+$) primers. Eleven (seven of leaf and four of root) cDNA clones (in pBluescript II SK$^+$), which were problematic for 3′-end sequencing (possibly due to secondary structure formation in the template) were sequenced from their 5′ ends using M13 forward primer. The inserts from the subtracted leaf-specific cDNA library were sequenced using M13 forward/T7/KS primers. The sequences obtained were screened with VecScreen (NCBI) and edited with Sequence Navigator™ v 1.0.1 software (Applied Biosystems) to remove vector, adaptor, and ambiguous sequences prior to BLAST analyses.

The edited sequences were subjected to a BLASTX analysis against the non-redundant protein database (all GenBank CDS translations+PDB+SwissProt+PIR+PRF excluding environmental samples), a BLASTN analysis against the non-redundant nucleotide database (all GenBank+EMBL+DDBJ+PDB sequences except EST, STS, GSS, environmental samples or phase 0, 1, or 2 HTGS sequences) and a BLASTN analysis against the database of GenBank+EMBL+DDBJ sequences from EST divisions (Altschul *et al.*, 1997). Both the analyses (BLASTX as well as BLASTN against the non-redundant databases) were used in conjunction to assign a putative function to an EST. Generally, there was consensus between the two types of analyses and it was possible to assign putative functions to many of the ESTs (although the true function can only be ascertained through biochemical and genetic approaches). BLASTN analysis against the EST database was only used as corroborative evidence for verification of plant origin and putative identification of ESTs. In cases where consensus could not be established between the BLASTX and BLASTN results, the BLASTX result was given weightage for classification purposes. Wherever the BLAST homologies of individual ESTs were found to be somewhat similar, the 'bl2seq' tool (NCBI) was used to align the EST sequences to check for redundancy. For functional classification of the ESTs, all ESTs were included, regardless of *E* value and redundancy. In the case of ESTs from the separate leaf and root cDNA libraries (having directionally cloned inserts), for the purpose of functional classification in all the categories (except hypothetical/unknown and miscellaneous/others categories), only BLASTX homologies to negative reading frames for ESTs sequenced on the 3′ end and to positive reading frames for ESTs sequenced on the 5′ end were considered [except for clones having inserts in reverse orientation or lacking poly (A) tails]. However, in the case of ESTs from the subtracted leaf-specific cDNA library, BLASTX homologies to both positive and negative reading frames were considered for the purpose of functional classification in all the categories since the cloning orientation of the inserts was not defined. Many clones had the potential to be classified into more than one category due to overlapping functions. However, for simplicity, those ESTs were categorized according to their most universal function.

Semi-quantitative RT-PCR analysis

Semi-quantitative RT-PCR (Skrypina *et al.*, 2003) was used to measure the transcript abundance of *sgd* (strictosidine β-D-glucosidase) and *dat* in the aerial and underground parts of the plant at three different plant developmental stages (6 d, 6 weeks, and 6 months). Exactly equal amounts (250 ng) of total RNA (as determined spectrophotometrically as well as by ethidium bromide staining, where rRNA was used as the equal loading control), isolated from the different samples to be compared, were taken for first strand cDNA synthesis using the THERMOSCRIPT™ RT-PCR System (GibcoBRL). The reverse transcription reaction was primed with oligo (dT)$_{20}$ primer using 15 units of THERMOSCRIPT reverse

transcriptase as per the manufacturer's guidelines. One-tenth of the reaction products were used as the template in the PCR reactions. Gene-specific sense and antisense primers (with introduced restriction enzyme sites) were designed for *sgd* (sense 5′-TCATAAGCTTAAATAAGAAGAGAAATA-3′ and antisense 5′-AAAAGTTAAAAGAAAAGGTCGACAG-3′) and *dat* (sense 5′-GTCTAGAGATGGAGTCAGGAAAAATAT-3′ and antisense 5′-GGTCGACGTTAATTAGAAACAAATTGA-3′) employing the Primer Premier version 4.04 software using the gene sequences available in the GenBank database (AF112888 and AF053307, respectively) and synthesized on an ABI 392 DNA/RNA synthesizer (PE Applied Biosystems). In the PCR step, after the initial standardization with various numbers of cycles (25, 28, and 35) it was determined that 28 cycles provided the optimal result for product quantification for both the genes in the exponential range of PCR, and this condition was used for all the quantification assays. Each PCR reaction was performed with sense (25 ng) and antisense (25 ng) gene-specific primers along with 0.6 units of *Taq* DNA polymerase in a reaction volume of 25 µl. For both the genes *sgd* and *dat*, the thermal cycler was programmed as 94 °C for 5 min; 28 cycles of 94 °C for 1 min, 49 °C for 1 min, 72 °C for 2 min; 72 °C for 5 min; 4 °C for an infinite period. The amplicons were analysed on a 1.2% agarose gel. The fluorescent ethidium bromide-stained bands corresponding to the PCR-products were quantified (by comparing with known amounts of standard DNA) using a Gel documentation system (Image Master® VDS) and the Image Master 1D Elite software version 3.00 (Pharmacia Biotech). The full-length *sgd* gene obtained through RT-PCR (from the mature leaf sample of a 6-month-old plant) was cloned on pUC19 and end-sequenced to validate its identity.

Determination of alkaloid content

The alkaloid content in the plant tissues was determined as described earlier (Shukla, 2005). Alkaloid extraction was carried out as per the validated protocol of Shukla *et al*. (1997) and HPLC analyses of individual alkaloids (vincristine, vinblastine, catharanthine, and vindoline) were carried out as described earlier (Gupta *et al.*, 2005).

Results

It has been shown earlier that bisindole alkaloid content in the *C. roseus* leaves is dependent on tissue maturity (Naaranlahti *et al.*, 1991). Under northern Indian conditions the crop of *C. roseus* reaches maturity in the summer months beginning in June. In the present study, conditions (covering plant developmental stage, time of harvest, and type of leaf) favouring high bisindole alkaloid biosynthesis by the plant were defined on this basis and were chosen for poly $(A)^+$ mRNA isolation. The roots of genotype 'Dhawal' have been found to contain undetectable to trace amounts of catharanthine, while the economically important bisindoles (vincristine and vinblastine), along with their precursors (vindoline and catharanthine), are primarily located in the leaves of the plant (Shukla, 2005). Therefore this genotype provides an excellent plant-system for gene prospection in relation to leaf-specific TIA biosynthesis. Based on this study covering TIA profiling in *C. roseus* (Shukla, 2005), middle level leaf pairs of 6-month-old mature field-grown plants of cv. 'Dhawal' were harvested in the month of June (with average day temperature invariably being ⩾42 °C) for poly $(A)^+$ mRNA isolation. For differential expression studies, root tissue was also harvested from the same plants under the same conditions for poly $(A)^+$ mRNA isolation. The leaf and root tissue transcriptomes of *C. roseus* were comparatively analysed using ESTs generated by two different approaches: (i) indirectly through separate cDNA libraries; and (ii) directly through a strategically designed suppression subtractive hybridization, which used the leaf and root cDNA populations as tester and driver, respectively. The results obtained are described below.

ESTs from independent leaf and root cDNA libraries at the same stage

Separately prepared leaf and root cDNA libraries from the metabolically defined stage (for high bisindole biosynthesis) of the plant were analysed and found to be highly representative, as was evident by EST sequencing. A total of 120 randomly selected insert-containing clones (60 each from the leaf and root cDNA libraries) were sequenced. Among these, definable sequences were obtained for 100 clones (which provided 55 leaf and 45 root ESTs). Out of these 100 sequence-characterized clones, only two (LP15F and LP22R) were found to have identical inserts and all the remaining were distinct. The 13 functional categories, which could be defined to classify the putatively identified ESTs, encompassed the whole range of cellular functions and were: primary metabolism, cell wall/membrane-related, gene expression/RNA metabolism related, stress/pathogenesis induced, signal transduction related, defence/senescence related, cell division cycle/DNA metabolism related, protein synthesis/processing related, photosynthesis, secondary metabolism, miscellaneous/others, hypothetical/unknown proteins, and those showing no significant similarity in the databases (putative new genes).

Functional classification of leaf ESTs

The 55 putatively identified leaf ESTs could be classified into 12 out of the 13 functional categories defined above (Fig. 2). The complete description of the annotated leaf ESTs clustering into the various functional categories is summarized in Table 1. Interestingly, 22% of the ESTs were found to be potentially new genes having no similarity in the public databases, while hypothetical proteins having no defined function constituted 18% of the leaf EST set. Only one leaf EST (LZ159) among them showed homology to a *C. roseus* gene sequence from the nucleotide database, which was for cyclophilin, a protein probably involved in protein processing. The 12 ESTs showing no significant similarity in the public databases were potentially novel genes and further characterization is required to define them functionally. The absence of any EST related to secondary metabolism indicated the high complexity of the leaf tissue transcriptome and the low abundance of such genes in a randomly drawn mRNA population. This fact suggested the need for a subtractive hybridization approach to identify

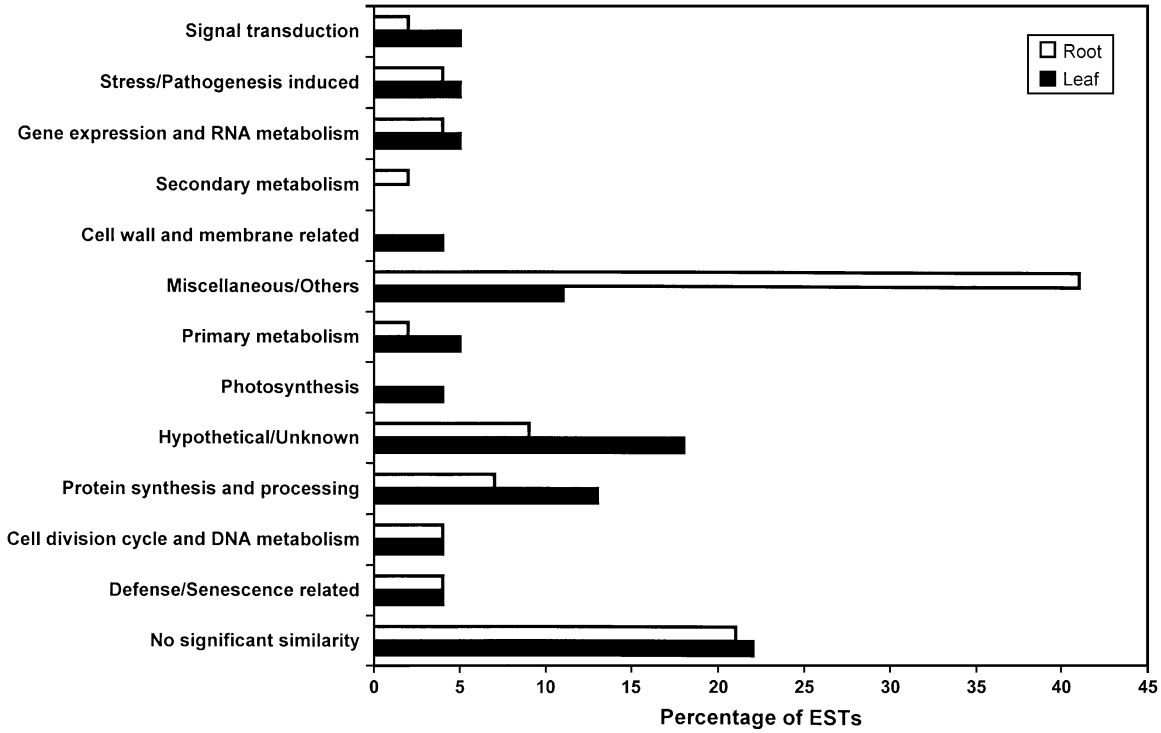

Fig. 2. Comparative distribution of the functionally classified ESTs from the *C. roseus* leaf and root cDNA libraries.

leaf-specific ESTs involved in secondary metabolism after removal of the highly abundant transcripts by using an appropriate driver (such as the root tissue). Of course, the lower number of ESTs than that which may be required to represent the leaf tissue transcriptome absolutely is not ruled out.

Functional classification of root ESTs

The 45 putatively identified root ESTs could be classified into 11 out of the 13 defined functional categories (Fig. 2). The complete description of the annotated root ESTs clustering into the various functional categories is summarized in Table 2. As expected, none of the root ESTs was found to be related to photosynthesis. The miscellaneous category in the case of root ESTs contained a greater variety of functions than in the case of leaf ESTs. Similar to the case of leaf ESTs (where 22% ESTs had no similarity in public databases), 21% of root ESTs showed no similarity in the public databases and probably represented new genes. Only 9% of root ESTs showed homology to hypothetical proteins as opposed to 18% in the case of leaf ESTs.

ESTs from the subtracted leaf-specific cDNA library

Since the target of the study was to identify TIA biosynthesis-related genes being expressed specifically/abundantly in the leaf tissue as compared with the root tissue, the highly specific suppression subtractive hybridization approach utilizing the leaf and root cDNA populations as tester and driver, respectively, proved to be the method of choice. A total of 60 randomly selected insert-containing clones from the subtracted leaf-specific cDNA library were sequenced and 55 definable sequences were obtained (which provided 55 leaf-specific ESTs).

Redundancy in the subtracted leaf-specific cDNA library

Clones in the pairs LTRDBC21/LTRDBC37, LTRDBC80a/LTRDBC85a, LTRDBC80b/LTRDBC85b, and LTRDBC88/LTRDBC91 had similar inserts. Clones LTRDBC11 and LTRDBC26 had similar insert sequences but the latter clone was longer. The insert of clone LTRDBC14 had sequence similarity with that of clone LTRDBC12b on the same strand but on the opposite strand with those of clones LTRDBC11 and LTRDBC26. Inserts of clones LTRDBC19 and LTRDB8a had sequence similarity on opposite strands. Six clones had two inserts cloned in the same vector molecule (as a cloning artefact) as determined by restriction enzyme site analysis and BLAST results of the sequences. In such clones, each individual insert was considered a separate EST (LTRDBC4a and LTRDBC4b; LTRDBC12a and LTRDBC12b; LTRDBC18a and LTRDBC18b; LTRDBC80a and LTRDBC80b; LTRDBC85a and LTRDBC85b; and LTRDB8a and LTRDB8b) due to different BLAST results obtained for each cloned insert. As an exception, although clone LTRDBC34 had an internal *Eag*I site, BLAST results indicated that it was genuinely present in the insert and was not a cloning artefact.

Table 1. *Putative functions associated with leaf ESTs in different functional categories*

Primary metabolism
Aminotransferase (LZ87)
Enolase (LZ143)
Aspartyl aminopeptidase (LP24R)
Cell wall and membrane related
Alpha-1,4-glucan-protein synthase [UDP-forming] (UDP-glucose:protein transglucosylase) (UPTG) (amylogenin); membrane protein (LP3R)
Putative membrane protein [*Corynebacterium glutamicum* ATCC 13032] (LP37R)
Gene expression and RNA metabolism related
RSZp22 splicing factor (LZ51)
RNA polymerase sigma factor (LP6F)
Zinc finger protein (LP13R)
Stress/pathogenesis induced
Galactinol synthase, isoform GolS-1 [*Ajuga reptans*] (LZ127)
ADP-ribosylation factor (LP7R)
Pathogenesis-related protein PR10A [*Datisca glomerata*] (LP12F)
Signal transduction related
Receptor-like protein kinase 4 RLK4; serine/threonine kinase-like protein [*Arabidopsis thaliana*] (LZ133)
Systemin receptor SR160 [*Oryza sativa* (japonica cultivar-group)]; receptor protein kinase [*Arabidopsis thaliana*] (LP18R)
Protein kinase Xa21 [*Oryza sativa* (japonica cultivar-group)] (LP19R)
Defence/senescence related
Cysteine proteinase inhibitor; multicystatin (LZ4)
Cystatin; cysteine proteinase inhibitor (LZ210)
Cell division cycle and DNA metabolism related
Replication protein E1 [human papillomavirus type 23] (LZ117)
DNA polymerase A family protein (LP17R)
Protein synthesis and processing related
Ubiquitin-conjugating enzyme (LZ31)
Chloroplast translation elongation factor EF-Tu (LZ58)
Cyclophilin (LZ159)
Phenylalanyl-tRNA synthetase class IIc family protein (LP9R)
60S ribosomal protein L15 (LP15F)
60S ribosomal protein L15 (LP22R)
60S ribosomal protein L29 [*Panax ginseng*]; ripening regulated protein DDTFR19 [*Lycopersicon esculentum*=*Solanum lycopersicum*] (LP32R)
Photosynthesis
Chlorophyll *a*/*b*-binding protein of LHCII type I (LZ73)
Photosystem I (PS I) reaction centre protein subunit II (psaD) (LP2R)
Miscellaneous/others[a]
Asparagine-rich protein [*Plasmodium falciparum* 3D7] (LZ132; +3)
gp9 [enterobacterium phage Mu] (LZ155)
Phosphatidylinositolglycan related [*Arabidopsis thaliana*] (LP5R)
Putative glycine-*N*-acyltransferase [*Bos taurus*] and arylacetyltransferase [*Bos taurus*] (LP21R; +2)
ATP-dependent RNA helicase, putative; protein id: At1g35530.1 [*Arabidopsis thaliana*] [*Dictyostelium discoideum*] (LP29R; +2)
ATP-dependent helicase [*Synechococcus* sp. WH 8102] (LP43R); ribulose-bisphosphate carboxylase (EC 4.1.1.39) large chain precursor—rice chloroplast (LP43R; +1)
11. Hypothetical/unknown proteins[a]
LZ14, LZ48, LZ53, LZ55, LZ72, LZ104, LZ109, LZ204, LP8F, LP16F
12. No significant similarity (probably new genes)
LZ16, LZ21, LZ40, LZ46, LZ65, LZ131, LP10F, LP11F, LP25R, LP26R, LP38R, LP45R

[a] This category contains positive homologies to 3′-end-sequenced ESTs and negative homologies to 5′ end-sequenced ESTs also.

Table 2. *Putative functions associated with root ESTs in different functional categories*

Primary metabolism
NADH dehydrogenase subunit 1 (RP9R)
Gene expression and RNA metabolism related
RING finger protein 15 (zinc finger protein RoRet) (tripartite motif-containing protein 38) (RZ31)
RNA polymerase beta II subunit (RZ151)
Stress/pathogenesis induced
Phi-1-like phosphate-induced protein (RZ160)
Phi-1-like phosphate-induced protein (RZ169)
Signal transduction related
Receptor activity modifying protein 1 [*Mus musculus*] (RZ110)
Defence/senescence related
Thioredoxin peroxidase (RZ152)
Ethylene-forming-enzyme-like dioxygenase; leucoanthocyanidin dioxygenase-like protein (RP8F)
Cell division cycle and DNA metabolism related
Nucleoid DNA-binding-like protein (RZ86)
Actin-depolymerizing factor (RP3R)
Protein synthesis and processing related
Aspartic proteinase (RZ40)
60S ribosomal protein L3 (RZ67; +3)[b]
Cyclophilin or peptidylprolyl isomerase (EC 5.2.1.8) (RZ140)
Secondary metabolism
O-methyltransferase; flavonoid *O*-methyltransferase (RZ41)
Miscellaneous/others[a]
Non-LTR retroelement reverse transcriptase [*Oryza sativa* (japonica cultivar-group)] (RZ35)
CLCA4 [*Pan troglodytes*] (RZ39)
Cell wall-associated hydrolase [*Vibrio vulnificus* CMCP6] (RZ43; +1)
COG1197: transcription-repair coupling factor (superfamily II helicase) [*Streptococcus suis* 89/1591] (RZ43)
Haemagglutinin [influenza A virus (A/duck/NY/191255-79/02(H5N2))] (RZ47; +3)
FabG protein [*Mannheimia succiniciproducens* MBEL55E] (RZ54; +2)
Similar to abnormal cell LINeage LIN-41, heterochronic gene; *Drosophila* dappled/vertebrate TRipartite Motif protein related; B-box zinc finger, filamin and NHL repeat-containing protein (123.8 kD) (lin-41) [*Apis mellifera*] (RZ56; +1); zinc finger protein [*Xenopus laevis*] (RZ56; +1)
Nuclear protein (4B256) [*Caenorhabditis elegans*] (RZ58)
UTP–glucose-1-phosphate uridylyltransferase [*Schizosaccharomyces pombe*] (RZ64; +1)
Zinc-binding dehydrogenase [*Bacillus megaterium*] (RZ69)
Phytosulfokine peptide precursor (RZ78)
Pc-fam-1 protein, putative [*Plasmodium chabaudi*] (RZ87)
COG1197: transcription-repair coupling factor (superfamily II helicase) [*Streptococcus suis* 89/1591] (RZ88); fructosyltransferase [*Paenibacillus polymyxa*] (RZ88); sucrose/fructan hydrolase (EC 3.2.1.–)–[*Bacillus polymyxa*] (RZ88)
Spore germination protein [*Oceanobacillus iheyensis* HTE831] (RZ96)
COG1396: predicted transcriptional regulators [*Burkholderia cepacia* R1808] (RZ97; +3); NADH-ubiquinone oxidoreductase subunit [*Sinorhizobium meliloti*] (RZ97; +1)
GCN5-related *N*-acetyltransferase (GNAT) family protein [*Arabidopsis thaliana*] (RZ104)
Dem protein [*Lycopersicon esculentum*=*Solanum lycopersicum*] (RZ139)
NAC2 protein [*Oryza sativa* (japonica cultivar-group)] (RZ155)
2,3-Biphosphoglycerate-independent phosphoglycerate mutase-related/phosphoglyceromutase-related [*Arabidopsis thaliana*] (CR11F; −3); gp25 [bacteriophage phi-BT1] (RP11F)
10. Hypothetical/unknown proteins[a]
RZ57, RZ85, RZ109, RZ130
11. No significant similarity (probably new genes)
RZ37, RZ73, RZ84, RZ111, RZ114, RZ148, RZ183, RP1F, RP17F

[a] This category contains positive homologies to 3′ end-sequenced ESTs and negative homologies to 5′ end-sequenced ESTs also.
[b] Typical EST [without poly (A) tail] showing positive reading frame homology when sequenced from the 3′ end.

Functional classification of the leaf-specific ESTs

The subtracted leaf-specific ESTs could be classified only into 10 functional categories (Fig. 3; Table 3). As expected, a large proportion (22%) of the ESTs sampled grouped in the chloroplast/photosynthesis-related category. Presence of a high number of ESTs related to stress/pathogenesis (15%) and defence/senescence (11%) indicated that the leaf tissue, rather than the root tissue, is primarily responsible for the overall plant defence and stress responses. This is in accordance with the rationale that the leaf tissue being at a greater risk from herbivory and pests, is primarily responsible for defensive response via production of antifeedent secondary metabolites (including certain TIAs).

The most significant finding of the direct approach in the current study was the isolation of an EST for strictosidine β-D-glucosidase (*sgd*) (LTRDB35), a key branch-point enzyme in the TIA biosynthetic pathway, and 16 ESTs with no significant homology in public databases, which could provide a useful starting point for isolation and identification of novel leaf-specific genes. The key to the differential gene expression in the leaf as compared with the root lies in the function of these genes and it may also account for the variance in the alkaloid profiles of the two tissues.

Direct approach versus indirect approach

There was good agreement between the results obtained using both approaches. None of the functions associated with the root ESTs was obtained in the subtracted leaf-specific cDNA library. Besides, genes like cyclophilin, which were present in both of the tissues, root as well as leaf, were not detected in the leaf-specific cDNA library. The functions that were found in the unsubtracted as well as the subtracted leaf cDNA library were pathogenesis-related proteins, galactinol synthase, and chlorophyll *a*/*b*-binding protein. The crucial advantage gained by the direct approach over the indirect approach was the identification of the EST for the secondary metabolism-related gene *sgd*. However, further work is required to establish any resemblance between the 16 ESTs obtained by the direct approach (subtracted leaf-specific) and the 12 ESTs obtained by the indirect approach (unsubtracted) from the leaf tissue, which probably denote novel gene functions.

*Leaf specific/prominent transcript (*dat *and* sgd*) abundance in relation to the* in planta *TIA biosynthetic potential*

Targeted metabolite analysis was combined with semi-quantitative RT-PCR-based transcript profiling for leaf specific/prominent genes, *dat* (reported in literature) and *sgd* (from the direct approach in the present study). Transcript profiling of the two genes suggested an extensive effect of plant developmental stage on their transcription (Figs 4, 5), which correlated well with the observed shifts in alkaloid (including specific TIAs) biosynthesis in the plant. The growth-related decrease in shoot/leaf *dat* and *sgd* transcript levels was paralleled by a concomitant decrease in vindoline content of *C. roseus* cv. 'Dhawal' (Table 4). Moreover, *dat* and *sgd* transcripts could not be detected in the root system of the plant at any of the developmental stages (6 d old, 6 months old, or 6 weeks old), which correlated well with the absence of vindoline in the roots in all plant developmental stages. This result was as per

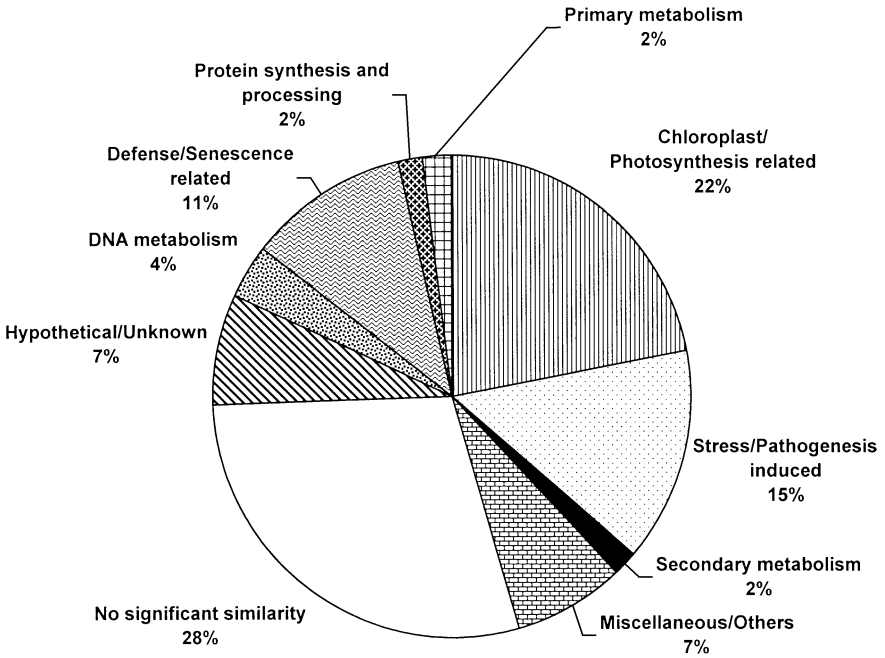

Fig. 3. Functional classification of the subtracted leaf-specific ESTs of *C. roseus*.

Table 3. *Putative functions associated with the subtracted leaf-specific ESTs in different functional categories*

Primary metabolism
Serine hydroxymethyltransferase/Glycine hydroxymethyltransferase (LTRDB68)
Stress/pathogenesis induced
Pathogenesis-related protein (LTRDBC11; LTRDBC14; LTRDBC26; LTRDBC12b; LTRDB46)
Metallothionein-like protein [*Citrus unshiu*] (LTRDB56)
Galactinol synthase GolS-1 [*Ajuga reptans*] (LTRDB66)
Ozone-responsive stress-related protein [*Arabidopsis thaliana*] (LTRDB69)
Defence/senescence related
Senescence-associated protein (*Pisum sativum*) (LTRDBC22; LTRDBC88; LTRDBC91; LTRDB17)
Cysteine protease (LTRDB26)
Senescence-associated protein-like [*Oryza sativa*] (LTRDB42)
DNA metabolism related
DNA-directed DNA polymerase (EC2.7.7.7) I [*Thermus aquaticus*] (LTRDBC34; LTRDBC39)
Protein synthesis and processing related
Elongation factor 2 [*Oryza sativa*] (LTRDB24)
Chloroplast/photosynthesis related
Ycf2 protein (LTRDBC1; LTRDB55)
Photosystem I subunit VIII [*Zea mays*] (LTRDBC19; LTRDB8a)
Rubisco small subunit [*Coffea arabica*] (LTRDBC21; LTRDBC37)
Ycf15 protein [*Panax ginseng*] (LTRDBC72)
ATP synthase beta subunit [*Rheum rhaponticum*] (LTRDBC87)
PsbC (LTRDB4)
Chlorophyll *a*/*b*-binding protein type I (LTRDB6)
Rubisco activase [*Malus*×*domestica*] (LTRDB29)
Photosystem I iron-sulphur centre subunit VII (LTRDB49)
Secondary metabolism
Strictosidine beta-glucosidase [*Catharanthus roseus*] (LTRDB35)
Miscellaneous/others
Cysteine desulfurase [*Plasmodium falciparum*] (LTRDBC28)
Lipid transfer protein 2 [*Lycopersicon pennellii*] (LTRDBC38)
Protein-tyrosine phosphatase [*Cytophaga hutchinsonii*] (LTRDBC80b; LTRDBC85b)
Hypothetical/unknown proteins
LTRDBC4a; LTRDBC7; LTRDBC12a; LTRDBC89
No significant similarity (probably new genes)
LTRDBC13; LTRDBC18a; LTRDBC27; LTRDBC63; LTRDBC80a; LTRDBC85a; LTRDB14; LTRDB8b; LTRDB22; LTRDB37; LTRDB40; LTRDB61; LTRDB67; LTRDB70; LTRDBC4b; LTRDBC18b

logical expectation for DAT, the last enzyme in the vindoline biosynthetic pathway, but it was interesting to note that SGD, acting much earlier in the pathway at the crucial post-strictosidine branch-point, also followed a similar trend. This reinforced the belief that vindoline biosynthesis is restricted to the aerial/green tissue of the plant. Lack of or a low level of catharanthine accumulation in the roots of *C. roseus* as observed for cv. 'Dhawal' was not unusual, as previously workers have also come across such cultivars (Toivonen *et al.*, 1989). As a negative control for vindoline and consequently bisindole alkaloid biosynthesis, catharanthine-producing cell suspension cultures of *C. roseus* (cv. 'Dhawal') were found to express *sgd* but not *dat* (data not shown). Considering this in conjunction with the case of the root tissue transcriptome, this clearly demonstrated that cellular differentiation is absolutely required for *dat* expression, whereas it is not stringently mandatory for *sgd* expression.

Discussion

In the case of the separate leaf cDNA library the presence of an EST for a ubiquitin-conjugating enzyme was interesting since these enzymes have roles in post-translational modifications by selective protein breakdown via the ubiquitin-dependent proteolytic pathway. A ubiquitin-mediated degradation of tryptophan decarboxylase during plant development has also been reported (Fernandez and De Luca, 1994), suggesting that TIA biosynthesis regulation involves transcriptional mechanisms and post-translational ubiquitin-dependent proteolysis. This has been further substantiated by the cloning of a novel cDNA (*Crubie2*) encoding a ubiquitin-conjugating enzyme E2 from *C. roseus* (Sibéril *et al.*, 2002). Expression of *Crubie2* is reportedly repressed in developing organs and down-regulated by cytokinin, suggesting that a decrease in the ubiquitin-dependent proteolytic pathway may take part in the regulation of alkaloid biosynthesis in *C. roseus* suspension cultures. In a related development, RNA interference (RNAi) has been used to turn off a cytokinin signalling circuitry in *C. roseus* by Papon *et al.* (2004). Gene expression-related factors obtained among the leaf ESTs carry much significance in the light of recent research findings. For example, expression of two RNA polymerase sigma factors in maize has been found to be light-inducible and tissue-specific, and it is believed that they may play a role in differential gene expression during chloroplast biogenesis (Tan and Troxler, 1999). Besides, zinc finger proteins have been implicated in activation of the pathogen-defence response in plants (Oh *et al.*, 2005), realization of jasmonic acid hormonal responses (van der Krol *et al.*, 1999), and in transcriptional activation or repression of tissue-specific and light-regulated gene expression in plants (Yanagisawa and Sheen, 1998). Recently, Pauw *et al.* (2004*a*) have demonstrated that zinc finger proteins even act as transcriptional repressors of alkaloid biosynthesis genes in *C. roseus*. They performed a yeast one-hybrid screening with the elicitor-responsive part of the *tdc* promoter to identify three members of the Cys_2/His_2-type (transcription factor IIIA-type) zinc finger protein family from *C. roseus* (ZCT1, ZCT2, and ZCT3), which bind in a sequence-specific manner to the *tdc* and *ss* promoters *in vitro* and repress their activity in *trans*-activation assays. In addition, the ZCT proteins can repress the activating activity of APETALA2/ethylene response-factor domain transcription factors, the ORCAs, on the *ss* promoter. In view of these facts, the molecular function of the transcription, splicing, and sigma factors obtained in the leaf cDNA library of *C. roseus* needs to be explored further in relation to TIA biosynthesis (specially the late steps in vindoline biosynthesis, which require differentiated green tissue as well as light). Apart from regulation through transcription factors, signal transduction by protein kinases is an important component of elicitor-induced alkaloid biosynthesis

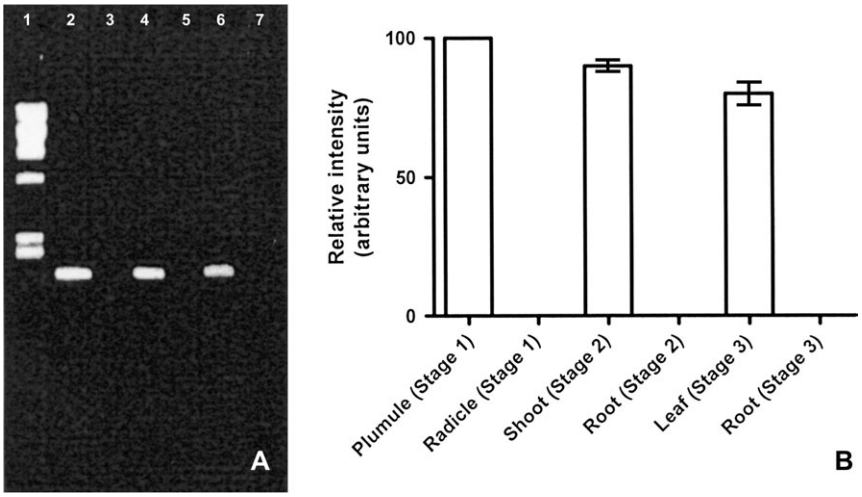

Fig. 4. Variation in *sgd* transcript abundance in *C. roseus* at different developmental stages. (A) Lane 1, DNA molecular weight marker (λDNA/*Hin*dIII fragments); lane 2, plumule (stage 1; 6-d-old germinating seedling); lane 3, radicle (stage 1; 6-d-old germinating seedling); lane 4, shoot (stage 2; 6-week-old seedling); lane 5, root (stage 2; 6-week-old seedling); lane 6, leaf (stage 3; 6-month-old mature plant); lane 7, root (stage 3; 6-month-old mature plant). (B) The data presented for relative transcript abundance is the mean ±standard deviation of three independent experiments. The relative intensities of the transcripts are expressed as a percentage of the most intense sample (which was taken as 100).

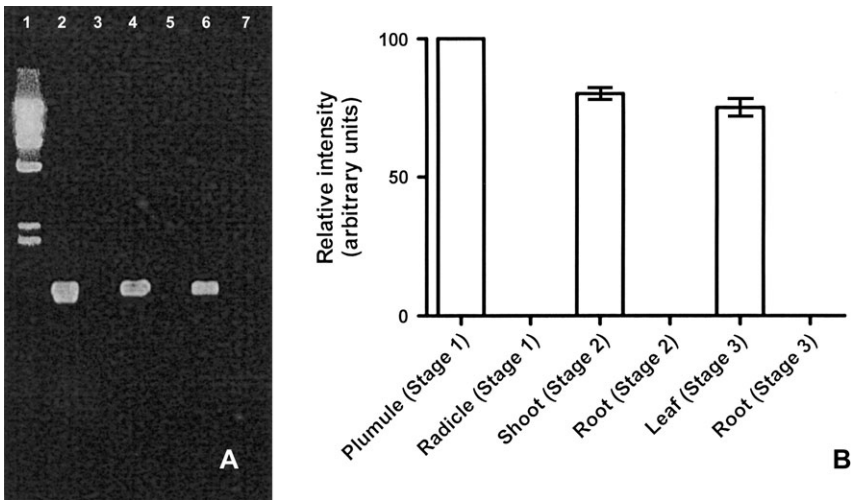

Fig. 5. Variation in *dat* transcript abundance in *C. roseus* at different developmental stages. (A) Lane 1, DNA molecular weight marker (λDNA/*Hin*dIII fragments); lane 2, plumule (stage 1; 6-d-old germinating seedling); lane 3, radicle (stage 1; 6-d-old germinating seedling); lane 4, shoot (stage 2; 6-week-old seedling); lane 5, root (stage 2; 6-week-old seedling); lane 6, leaf (stage 3; 6-month-old mature plant); lane 7, root (stage 3; 6-month-old mature plant). (B) The data presented for relative transcript abundance is the mean ±standard deviation of three independent experiments. The relative intensities of the transcripts are expressed as a percentage of the most intense sample (which was taken as 100).

in plants, as demonstrated for TIA biosynthesis in *C. roseus* (Menke *et al.*, 1999) and benzophenanthridine alkaloid biosynthesis in *Sanguinaria canadensis* (Mahady *et al.*, 1998). The RLK (receptor-like protein kinases) group of plant kinases has attracted much attention recently because of their likely roles in the perception and transduction of external signals. Schulze-Muth *et al.* (1996) characterized a novel RLK from *C. roseus* and identified a threonine residue, which was required for both autophosphorylation and the activity of the kinase to phosphorylate other proteins.

Interestingly, the root cDNA library contained an EST for *O*-methyltransferase/flavonoid *O*-methyltransferase (RZ41), which is a secondary metabolism-related enzyme. This was surprising because the root is an underground plant organ and it is also known that the expression of the key enzyme in the precursor formation, chalcone synthase, is strictly light-dependent (Kaltenbach *et al.*, 1999). However, recently, Cacace *et al.* (2003) identified CrOMT2, a flavonoid OMT that was expressed in dark-grown cell cultures of *C. roseus* (which do not synthesize flavonoids) and co-purified with 16-hydroxytabersonine

Table 4. *Chronomics of TIA biosynthesis in* C. roseus *(cv. 'Dhawal')*

The data presented for alkaloid content (% dry weight) is the mean of 10 replicates where standard deviation was between 5% and 19% of the mean values. The data presented for relative transcript abundance is the mean of three independent experiments where the standard deviation was between 2% and 5% of the mean values. The relative transcript abundance has been expressed as a percentage of the sample with maximum transcript abundance (which was taken as 100).

Plant developmental stage	Tissue	Relative transcript abundance		Total alkaloid (%)	Vincristine (%)	Vinblastine (%)	Catharanthine (%)	Vindoline (%)
		sgd	*dat*					
Stage 1	Plumule	100	100	8.65	nd[a]	nd	nd	0.0080
6-d-old germinating seedling	Radicle	0	0	9.34	nd	nd	nd	nd
Stage 2	Seedling shoot	90	80	3.28	nd	nd	nd	0.0065
6-week-old greenhouse seedling	Seedling root	0	0	6.42	nd	nd	nd	nd
Stage 3	Mature leaf	80	75	2.56	0.0003	0.0003	0.0036	0.0022
6-month-old mature plant growing in the field	Mature root	0	0	2.60	nd	nd	nd	nd

[a] nd, Not detected.

O-methyltransferase. The root cDNA library also included ESTs having homology to 1-aminocyclopropane-1-carboxylate oxidase, which catalyses the final step in the biosynthesis of the plant signalling molecule ethylene that is formed in response to wounding and stress, as well as during senescence (Alexander and Grierson, 2002), and thioredoxin peroxidase, which is involved in antioxidative defence and reduces H_2O_2 and alkyl hydroperoxides with electrons provided by thioredoxin (Berberich *et al.*, 1998). This gains significance as ethylene has been found to stimulate indole alkaloid accumulation in *C. roseus* suspension cultures (Yahia *et al.*, 1998; Papon *et al.*, 2005). Besides, Zhao *et al.* (2001) found that the activity of protective enzymes (superoxide dismutase, catalase, and peroxidase) involved in H_2O_2 metabolism increases to high levels in fungal elicitor-treated cell cultures of *C. roseus* due to the oxidative burst preceding indole alkaloid biosynthesis. However, Pauw *et al.* (2004*b*) have shown that the production of reactive oxygen species is not related to the activation of genes involved in TIA biosynthesis.

In the case of the subtracted leaf-specific cDNA library, the presence of ESTs for ozone-responsive proteins such as metallothionein-like protein (Etscheid *et al.*, 1999) is expected, since the leaf tissue rather than the root tissue provides the first line of defence against aerial environmental stresses. The presence of cysteine protease is also interesting as this defence/senescence-related enzyme has been demonstrated recently to possess a dual role in tomato—as an enzyme in the cytoplasm and as a transcription factor in the nucleus where it induces the expression of 1-aminocyclopropane-1-carboxylic acid synthase required for ethylene biosynthesis (Matarasso *et al.*, 2005).

Prominence of sgd *in the leaf tissue: implications in the post-strictosidine TIA biosynthetic pathway*

SGD is a key enzyme of the TIA biosynthetic pathway, functioning at the crucial branch-point, which determines the fate of the metabolic flux in the post-strictosidine pathway. According to previous studies, the highest SGD expression (transcript abundance as well as enzyme activity) in the *C. roseus* plant has been found in the leaves, although, it also occurs in the stem, root, and flower, but at a lower level (Geerlings *et al.*, 2000). By contrast, the present study demonstrated that *sgd* is a leaf-specific gene and that detectable *sgd* transcript levels in the root tissue are absent at all three plant developmental stages (6 d old, 6 months old, and 6 weeks old). This may be attributed to the expression pattern in the plant genotype used in the study (cv. 'Dhawal'), or alternatively different isoforms of the enzyme might be occurring in the leaf and root tissues wherein the leaf-specific gene form could be detected in the suppression subtractive hybridization experiment. A third, and more likely possibility, is the presence of very low, undetectable, *sgd* transcript levels in the root tissue, by contrast to the high levels in the leaf tissue. In any case, it is evident that SGD expression is prominent in the leaf tissue vis-à-vis the root tissue.

SGD catalyses the deglucosylation of strictosidine, the first step toward the various types of TIAs (Fig. 1). It is thus logical to hypothesize that the key to tissue- as well as genus-specific diversification in TIA biosynthesis must be at the glucosidase or in the steps directly after this enzyme in the biosynthetic pathway (Verpoorte *et al.*, 1997). This hypothesis gains support from the results of the studies of Stevens *et al.* (1992) on suspension cultures of three genera (*Catharanthus*, *Cinchona*, and *Tabernaemontana*), which suggested that SGD expression was regulated differentially in each genus. Possibly the enzyme induces a different type of rearrangement of the aglycone in different cases, thus playing a decisive role in steering the biosynthesis in the direction of a specific type of alkaloid. Other possibilities include the presence of different isoforms of the enzyme, the hydrolysed alkaloid moiety *in vivo* being rapidly passed on to another—as yet unidentified—enzyme for further transformation, or chemical conversion of the strictosidine

aglycone taking place under tissue/genus-specific conditions *in vivo* (Luijendijk *et al.*, 1996). Further comparative research on this particular step may provide the answer to the question—how do the different classes of TIAs arise in different plant species or different tissues of the same species?

Conclusions

The developmental clock of TIA biosynthesis in *C. roseus* integrates/unifies the myriad temporal signals received from throughout the plant, which affect distinct alkaloid profiles in the above-ground and underground organs of the plant at various developmental stages. Future research effort required on *C. roseus* demands exploitation of modern genomic and proteomic approaches to unravel the regulatory networks operating in the biosynthesis of TIAs. The present study based on transcriptomic approaches has provided a host of novel ESTs, which need to be further characterized to ascertain their prospective roles in TIA biosynthesis. There is a high probability that one or more of the leaf-specific novel genes isolated in the present study might code for a regulatory factor responsible for the leaf-specific expression of vindoline biosynthesis, and it could provide a starting point for future research directed favourably towards modulation/engineering of the TIA biosynthetic pathway. The key to the differential gene expression in the leaf as compared with the root lies in the function of these genes and it may also account for the variance in the alkaloid profiles of the two tissues. It is also evident from the present study that a larger number of ESTs should be analysed for detection of rare secondary metabolism-related genes.

Another crucial consideration for better understanding of the TIA biosynthetic pathway, both at the regulatory as well as the structural levels, in future research schemes will be the analysis of a wide range of intermediates and secondary metabolites rather than the final products alone, as it will provide information about the distribution of metabolic flux around key branch-points (Morgan and Shanks, 2002). Strictosidine, which is the common precursor of over 3000 monoterpenoid indole alkaloids, is one such example. Similarly, the alkaloids of *C. roseus* roots include oxidized 19-*O*-acetylated forms of tabersonine that are not intermediates of vindoline biosynthesis and may represent a potential competitive pathway for tabersonine in this underground organ (De Luca and Laflamme, 2001). Although most interest lies in the post-strictosidine pathway leading towards vindoline via tabersonine, the level of the post-strictosidine metabolic flux into all the branching pathways needs to be measured if TIA biosynthesis is to be studied in totality. Complete understanding of the spatial and temporal regulation of TIA biosynthesis in *C. roseus* is still elusive due to a few missing links. These missing links will define the future research effort on this plant.

Acknowledgements

We are grateful to the Council of Scientific and Industrial Research (CSIR), India, and the Department of Biotechnology, India, for the financial support extended by them. We also thank Dr RN Kulkarni, CIMAP Resource Centre, Bangalore for cv. 'Dhawal' seeds and Ms Chhavi Agnihotri for the artwork in this paper.

References

Alexander L, Grierson D. 2002. Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. *Journal of Experimental Botany* **53,** 2039–2055.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25,** 3389–3402.

Berberich T, Uebeler M, Feierabend J. 1998. Cloning of a cDNA encoding a thioredoxin peroxidase (TPx) homolog from winter rye (*Secale cereale* L.) (Accession No. AF076920). (PGR98-167). *Plant Physiology* **118,** 711–712.

Blasko G, Cordell GA. 1990. Isolation, structure elucidation, and biosynthesis of the bisindole alkaloids of *Catharanthus.* In: Brossi A, Suffness M, eds. *The alkaloids*, Vol. 37. San Diego, CA: Academic Press, 1–76.

Cacace S, Schröder G, Wehinger E, Strack D, Schmidt J, Schröder J. 2003. A flavonol *O*-methyltransferase from *Catharanthus roseus* performing two sequential methylations. *Phytochemistry* **62,** 127–137.

De Luca V, Balsevich J, Tyler RT, Eilert U, Panchuk BD, Kurz WGW. 1986. Biosynthesis of indole alkaloids: developmental regulation of the biosynthetic pathway from tabersonine to vindoline in *Catharanthus roseus. Journal of Plant Physiology* **125,** 147–156.

De Luca V, Laflamme P. 2001. The expanding universe of alkaloid biosynthesis. *Current Opinion in Plant Biology* **4,** 225–233.

Deus-Neumann B, Stöckigt J, Zenk MH. 1987. Radioimmunoassay for the quantitative determination of catharanthine. *Planta Medica* **53,** 184–188.

Diatchenko L, Lau Y-FC, Campbell AP, *et al.* 1996. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences, USA* **93,** 6025–6030.

Etscheid M, Klümper S, Riesner D. 1999. Accumulation of a metallothionein-like mRNA in Norway spruce under environmental stress. *Journal of Phytopathology* **147,** 207–213.

Fernandez JA, De Luca V. 1994. Ubiquitin-mediated degradation of tryptophan decarboxylase from *Catharanthus roseus. Phytochemistry* **36,** 1123–1128.

Geerlings A, Ibanez MM-L, Memelink J, van der Heijden R, Verpoorte R. 2000. Molecular cloning and analysis of strictosidine β-D-glucosidase, an enzyme in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus. Journal of Biological Chemistry* **275,** 3051–3056.

Gupta MM, Singh DV, Tripathi AK, Pandey R, Verma RK, Singh S, Shasany AK, Khanuja SPS. 2005. Simultaneous determination of vincristine, vinblastine, catharanthine and vindoline in leaves of *Catharanthus roseus* by high performance liquid chromatography. *Journal of Chromatographic Science* **43,** 450–453.

Kaltenbach M, Schröder G, Schmelzer E, Lutz V, Schröder J. 1999. Flavonoid hydroxylase from *Catharanthus roseus*: cDNA heterologous expression, enzyme properties and cell-type specific expression in plants. *The Plant Journal* **19,** 183–193.

Kulkarni RN, Baskaran K, Chandrashekara RS, Khanuja SPS, Darokar MP, Shasany AK, Uniyal GC, Gupta MM, Kumar S. 2003. 'Dhawal', a high alkaloid producing periwinkle plant. US Patent No. 6,548,746.

Laflamme P, St-Pierre B, De Luca V. 2001. Molecular and biochemical analysis of a Madagascar periwinkle root-specific minovincinine-19-hydroxy-*O*-acetyltransferase. *Plant Physiology* **125,** 189–198.

Luijendijk TJC, Nowak A, Verpoorte R. 1996. Strictosidine glucosidase from suspension cultured cells of *Tabernaemontana divaricata. Phytochemistry* **41,** 1451–1456.

Mahady GB, Liu C, Beecher CW. 1998. Involvement of protein kinase and G proteins in the signal transduction of benzophenanthridine alkaloid biosynthesis. *Phytochemistry* **48,** 93–102.

Matarasso N, Schuster S, Avni A. 2005. A novel plant cysteine protease has a dual function as a regulator of 1-aminocyclopropane-1-carboxylic acid synthase gene expression. *The Plant Cell* **17,** 1205–1216.

Menke FLH, Parchmann S, Mueller MJ, Kijne JW, Memelink J. 1999. Involvement of the octadecanoid pathway and protein phosphorylation in fungal elicitor-induced expression of terpenoid indole alkaloid biosynthetic genes in *Catharanthus roseus. Plant Physiology* **119,** 1289–1296.

Mishra P, Kumar S. 2000. Emergence of periwinkle *Catharanthus roseus* as a model system for molecular biology of alkaloids: phytochemistry, pharmacology, plant biology and *in vivo* and *in vitro* cultivation. *Journal of Medicinal and Aromatic Plant Sciences* **22,** 306–337.

Morgan JA, Shanks JV. 2002. Quantification of metabolic flux in plant secondary metabolism by a biogenetic organizational approach. *Metabolic Engineering* **4,** 257–262.

Naaranlahti T, Auriola S, Lapinjoki SP. 1991. Growth-related dimerization of vindoline and catharanthine in *Catharanthus roseus* and effect of wounding on the process. *Phytochemistry* **30,** 1451–1453.

Neuss N. 1980. The spectrum of biological activities of indole alkaloids. In: Phillipson JD, Zenk MH, eds. *Indole and biogenetically related alkaloids.* London: Academic Press, 293–313.

Oh S-K, Park JM, Joung YH, Lee S, Chung E, Kim S-Y, Yu SH, Choi D. 2005. A plant EPF-type zinc finger protein, *CaPIF1*, involved in defence against pathogens. *Molecular Plant Pathology* **6,** 269–285.

Papon N, Bremer J, Vansiri A, Andreu F, Rideau M, Creche J. 2005. Cytokinin and ethylene control indole alkaloid production at the level of the MEP/terpenoid pathway in *Catharanthus roseus* suspension cells. *Planta Medica* **71,** 572–574.

Papon N, Vansiri A, Gantet P, Chenieux JC, Rideau M, Creche J. 2004. Histidine-containing phosphotransfer domain extinction by RNA interference turns off a cytokinin signalling circuitry in *Catharanthus roseus* suspension cells. *FEBS Letters* **558,** 85–88.

Pauw B, Hilliou FAO, Martin VS, *et al.* 2004*a.* Zinc finger proteins act as transcriptional repressors of alkaloid biosynthesis genes in *Catharanthus roseus. Journal of Biological Chemistry* **279,** 52940–52948.

Pauw B, van Duijn B, Kijne JW, Memelink J. 2004*b.* Activation of the oxidative burst by yeast elicitor in *Catharanthus roseus* cells occurs independently of the activation of genes involved in alkaloid biosynthesis. *Plant Molecular Biology* **55,** 797–805.

Schulze-Muth P, Irmler S, Schröder G, Schröder J. 1996. Novel type of receptor-like protein kinase from a higher plant (*Catharanthus roseus*). *Journal of Biological Chemistry* **271,** 26684–26689.

Shukla AK. 2005. Molecular studies on biosynthesis of shoot alkaloids in *Catharanthus roseus* (L.) G. Don. PhD thesis, Department of Biochemistry, University of Lucknow, India.

Shukla AK, Shasany AK, Khanuja SPS. 2005. Isolation of poly $(A)^+$ mRNA for downstream reactions from some medicinal and aromatic plants. *Indian Journal of Experimental Biology* **43,** 197–201.

Shukla YN, Rani A, Kumar S. 1997. Effect of temperature and pH on the extraction of total alkaloids from *Catharanthus roseus* leaves. *Journal of Medicinal and Aromatic Plant Sciences* **19,** 430–431.

Sibéril Y, Thiersault M, Nepumoceno G, Doireau P, Gantet P. 2002. Cloning of a cDNA encoding an E2 ubiquitin-conjugating enzyme from *Catharanthus roseus*: expression analysis in plant organs and in response to hormones in cell suspensions. *Journal of Experimental Botany* **53,** 149–150.

Skrypina NA, Timofeeva AV, Khaspekov GL, Savochkina LP, Beabealashvilli RS. 2003. Total RNA suitable for molecular biology analysis. *Journal of Biotechnology* **105,** 1–9.

Stevens LH, Schripsema J, Pennings EJM, Verpoorte R. 1992. Activities of enzymes involved in indole alkaloid biosynthesis in suspension cultures of *Catharanthus*, *Cinchona* and *Tabernaemontana* species. *Plant Physiology and Biochemistry* **30,** 675–681.

St-Pierre B, Vazquez-Flota FA, De Luca V. 1999. Multicellular compartmentation of *Catharanthus roseus* alkaloid biosynthesis predicts intercellular translocation of a pathway intermediate. *The Plant Cell* **11,** 887–900.

Svoboda GH, Blake DA. 1975. The phytochemistry and pharmacology of *Catharanthus roseus* (L) G. Don. In: Taylor WI, Farnsworth NR, eds. *The catharanthus alkaloids.* New York, NY: Marcel Dekker, 45–83.

Tan S, Troxler RF. 1999. Characterization of two chloroplast RNA polymerase sigma factors from *Zea mays*: photoregulation and differential expression. *Proceedings of the National Academy of Sciences, USA* **96,** 5316–5321.

Toivonen L, Balsevich J, Kurz WGW. 1989. Indole alkaloid production by hairy root cultures of *Catharanthus roseus. Plant Cell, Tissue and Organ Culture* **18,** 79–93.

van der Heijden R, Jacobs DI, Snoeijer W, Hallard D, Verpoorte R. 2004. The *Catharanthus* alkaloids: pharmacognosy and biotechnology. *Current Medicinal Chemistry* **11,** 607–628.

van der Krol AR, van Poecke RMP, Vorst OFJ, Voogt C, van Leeuwen W, Borst-Vrensen TWM, Takatsuji H, van der Plas LHW. 1999. Developmental and wound-, cold-, dessication-, ultraviolet-B-stress-induced modulations in the expression of the petunia zinc finger transcription factor gene *zpt2-2. Plant Physiology* **121,** 1153–1162.

Verpoorte R, Van der Heijden R, Moreno PRH. 1997. Biosynthesis of terpenoid indole alkaloids in *Catharanthus roseus* cells. In: Cordell GA, ed. *The alkaloids*, Vol. 49. San Diego, CA: Academic Press, 221–299.

Yahia A, Kevers C, Gaspar T, Chenieux J-C, Rideau M, Creche J. 1998. Cytokinins and ethylene stimulate indole alkaloid accumulation in cell suspension cultures of *Catharanthus roseus* by two distinct mechanisms. *Plant Science* **133,** 9–15.

Yanagisawa S, Sheen J. 1998. Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. *The Plant Cell* **10,** 75–89.

Zhao J, Hu Q, Guo Y-Q, Zhu W-H. 2001. Elicitor-induced indole alkaloid biosynthesis in *Catharanthus roseus* cell cultures is related to Ca^{2+} influx and oxidative burst. *Plant Science* **161,** 423–431.