IOSR Journal of Environmental Science, Toxicology and Food Technology (IOSR-JESTFT)
e-ISSN: 2319-2402,p- ISSN: 2319-2399.Volume 11, Issue 8 Ver. II (August. 2017), PP 10-16
www.iosrjournals.org

Leaf specific Terpenoid Indole Alkaloid biosynthesis in *Catharanthus roseus* towards identification of novel structural and regulatory factors.

Farhina Pasha[1], Manal Amir
University of Tabuk, Faculty of science.Tabuk. Saudi Arabia
Corresponding author: Farhina Pasha

Abstract: *Since gene expression is tissue-specific, cDNA libraries of each tissue of an organism provides ESTs that are expressing specifically in that tissue at a particular temporal and developmental stage. Expressed Sequence tags (ESTs) can be mapped to specific chromosome locations using techniques such as fiber-FISH . Such tools would provide the genetically genomics approach of investigating plant genomes. A total of 26 white colonies were picked up from the mass-excised leaf cDNA library of C. roseus and checked for insert frequency . Three-fourth of the whites were recombinant and these were subjected to sequencing using T7 sequencing primer. Usable sequence was obtained for 12 clones. 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. Both the ESTs LZ365 and LZ418 showed similarity to protein kinase sequence in the database. However, they were not similar to each other when checked with the 'bl2seq' and CLUSTALW tools, implying that they matched with different domains of protein kinase. An EST-based approach seems to be the most attractive tool available for investigating leaf-specific Terpenoid Indole Alkaloid (TIA) biosynthesis in Catharanthus roseus towards identification of novel structural and regulatory factors.*

Key words: *C. roseus, TIA, BLAST, cDNA microarrays and ESTs*

Date of Submission: 08-08-2017 Date of acceptance: 28-08-2017

I. Introduction

Many novel chemicals such as taxanes, vinca alkaloids, podophyllotoxins and camptothecins show a range of cytotoxic activities and act as anti-tumor agents. Of particular interest are the pharmaceutically valuable bisindole alkaloids from the *Catharanthus roseus* (Madagascar periwinkle).*C. roseus* (L.) G. Don of the Apocynaceae family originally described from Madagascar has now become naturalized in several countries and has a pan-tropical distribution (*Maria Manuela R. Costa et. al. 2008*). It is widely cultivated for its ornamental features and many improved cultivars have been developed, for deriving the life saving bisindole alkaloids. It is an erect, bushy perennial herb having a height of about 75cm with a sub-woody base and profusely branched morphology. It is known to accumulate more than 130 terpenoid indole alkaloids in its different organs *(van Der Heijden et. al., 2004).* The plant was known for its medicinal values even in 50 B.C. (*Virmani et al., 1978).* However, it has shot into prominence due to the established anticancer activities of its leaf-specific bisindole alkaloids vincristine and vinblastine, which have high market value *(Shukla et.al. 2005).* Both these alkaloids are biosynthesized by the coupling of monomeric alkaloids vindoline and catharanthine. The leaf-specificity of bisindole alkaloids is due to the localization of certain enzymatic steps of the vindoline biosynthetic pathway in the thylakoid membrane of the chloroplast. NMT, which catalyses the fourth step from the pathway intermediate tabersonine to vindoline is localized in the thylakoids of the chloroplasts (*De Luca and Cutler, et.al. 1987*). Although many of the structural and regulatory factors involved in vindoline and bisindole alkaloid biosynthesis have been reported a majority of them are yet to be elucidated.

Expressed sequence tags (ESTs) are typically single-pass and partial sequences from cDNA clones (*Adams et al., 1991),* representing the transcriptome. They have provided high impetus for bioprospection of secondary metabolism-related genes that are involved in the biosynthesis of secondary metabolites of commercial importance. This had been traditionally difficult due to lack of complete knowledge of many of the biosynthetic pathways (*Shelton et al., 2002*). Since gene expression is tissue-specific, cDNA libraries of each tissue of an organism provides ESTs that are expressing specifically in that tissue at a particular temporal and developmental stage. ESTs can be mapped to specific chromosome locations using techniques such as fiber-FISH (*Horelli-Kuitunen et al., 1999*). Such tools would provide the genetical-genomics approach of investigating plant genomes. ESTs are also a useful resource for designing probes for cDNA microarrays used to study up- and down-regulation of gene expression. In view of the above mentioned facts and developments, an

EST-based approach seems to be the most attractive tool available for investigating leaf-specific TIA biosynthesis in *C. roseus* towards identification of novel structural and regulatory factors. In this direction the present study was carried out with the following objectives:

1. Screening of a leaf cDNA library of *C. roseus* in λ ZAP.
2. DNA sequencing of inserts from the recombinant clones to establish ESTs.
3. Homology-based analysis of the obtained ESTs to identify their putative functions.

II. Materials and Methods

1. **Bacterial strains used:** XLOLR Strain (Stratagene) **(GENOTYPE:** Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 thi-1 recA1 gyrA96 relA1 lac [F′ proAB lacIq ZΔM15 Tn10 (Tetr)] Su– (nonsuppressing) λ R (lambda resistant)
2. **Vectors used:** pBK-CMV (Stratagene)
3. **Composition of the buffers and solutions** used are given in **Table 1**. All the solutions were prepared in deionized MilliQ water (Millipore). Solutions were sterilized by autoclaving according to requirement. Solutions that were not recommended for autoclaving were prepared in autoclaved deionized MilliQ water.
4. **Culture media-**The composition of the various culture media is given in the **Table 2**
5. **Enzymes, kits, antibiotics and chemicals-** The enzymes, kits, antibiotics and other chemicals used have been described in **Table 3.**
6. **Software's used-**Various software's and databases used during the course of experimentation are Blast X, Blast N, VecScreen **(**NCBI **) and** Sequence analysis software V 5.2(Applied Biosystem)
7. **Screening of mass-excised leaf cDNA library of C. roseus:** The mass-excised leaf cDNA library of C. roseus was prepared. To titer the excised phagemids, 1ul of this mass-excised library was combined with 200ul of XLOLR cells (OD600 of 1) in a microcentrifuge tube and incubated for 15 minutes at 37° C. Forty microliter of 5X NZY broth (Table 2)was added to the microcentrifuge and incubated at 37° C for 45 minutes for allowing sufficient expression of the kanamycin-resistance gene product prior to plating of selective medium. *(Shukla et.al 2005)*

7.Blue white selection for recombinants- The excised cell mixture (100ul) was plated on LA-kanamycin agar plates containing 50ug/ml kanamycin and coated with 50 ul of X-gal (20mg/ml) and 12 ul of IPTG (24mg/ml) and the plates were incubated overnight at 37° C. (Table 1.)White colonies were picked up for plasmid preparations representing XLOLR Strain (Stratagene).

8. Plasmid isolation by modified alkaline lysis method- White colonies were inoculated in 5ml of LB media in sterile screw caps and incubated overnight at 37° C in a shaker at 200 rpm. Small aliquots (1.5 ml) of the culture were centrifuged at 6000 rpm for 5 minutes at 4° C to pellet the cells. The supernatant was removed by aspiration. The bacterial cells were again suspended in 200ul of solution A [containing 2 ul of RNaseA (10mg/ml)]. Then 300ul of freshly prepared solution B was added to it and the contents were mixed by inversion and incubated on ice for 5 minutes. It was followed by addition of 300ul solution C was and incubation on ice for 5 minutes. The microcentrifuge tube was spinned at 12000 rpm for 10 minutes at room temperature to pellet down the cell debris. The supernatant was transferred into a clear microcentrifuge tube. Chloroform (400ul) was added and mixed by inversion for 30 seconds and centrifuged at 10,000 rpm for 1 minute. The upper aqueous layer was transferred to another microcentrifuge tube. The chloroform step was repeated once again. Equal volume of isopropanol was added to the upper aqueous layer and mixed by inversion and spinned at 12,000 rpm for 10 minutes at room temperature. The isopropanol was removed by aspiration and the pellet was washed with 500ul of 70% ethanol. The pellet was vacuum dried for 10 minutes and the dried pellets were dissolved in 32ul of deionised water.(*Maria et.al 2008)*

Purification of plasmid by PEG method for automated DNA sequencing- Eight microliter of 4M NaCl and 40ul of sterile 13% PEG 8000 were added to the plasmid solution and incubated for 20 minutes on ice. The plasmid was than pelleted at 12,000 rpm for 15 minutes at 4°C. The plasmid DNA pellet obtained was washed with 500ul of 70% ethanol by centrifuging at 12,000 rpm for 5 minutes at 4° C, vacuum dried and resuspended in 20ul of sterile deionised water. The yield was checked on 0.8% agarose gel.(*Feliciello I et.al* 1993)

Insert frequency analysis in the selected whites through restriction digestion of plasmid DNA- Around 300-500 ng of plasmid DNA of the whites was taken in a microcentrifuge tube. Two microliter of 10X restriction buffer was added to the tube. This was followed by addition of 0.5ul of Eco RI (10U/ul) and Xho I (10U/ul) each to the tube (Table 3). The volume was finally made up to 20ul using sterile deionised water. The microcentrifuge tube was incubated for 2 hours at 37° C in a waterbath. The result was analyzed on 0.8% agarose gel. **Primer synthesis-**T7 sequencing primer was synthesized using 392 DNA/RNA Synthesizer

(Applied Biosystem) using phosphoramidite chemistry for oligonuleotide synthesis. The primer synthesized was purified using oligonuleotide purification cartridge following the manufacturer's guidelines.

Automated DNA sequencing of insert-containing recombinants-The plasmid DNA of the selected leaf cDNA clones was subjected to automated DNA sequencing. The dideoxy method of DNA sequencing was used in which 2' 3'-dideoxynucleotide analogue of a nucleotide base when incorporated at the 3' end of a growing chain, caused termination of chain elongation selectively at A, T, G, or C because the chain lacked a 3' hydroxy group. The ABI Prism Big Terminator Cycle Sequencing Ready Reaction Kit was used for setting up the cycle sequencing PCR reaction using the manufacturer's guidelines.To set up the cycle sequencing reaction, 150ng plasmid DNA template, 1.6 pmoles of T7 sequencing primer, 4ul of Terminator Ready mix and sterile deionised water (to make up to 10ul) were added to the PCR tube. The components of the tube were mixed and transferred to a thermal cycler (*Gene Amp PCR system 9600, Perkin Elmer*) programmed for 25 cycles of 96° C for 10seconds, 50° C for 5 seconds, 60° C for 4 minutes, and finally 4° C for infinity.The reaction product was purified to remove the unincorporated dye terminators, which if present, obscure the data in the early part of the sequence. The purification was started by addition of master mix A (10ul of sterile deionised water and 2ul of 0.5M EDTA). This was followed by the addition of master mix B (50ul of absolute ethanol and 2ul of 3M Sodium Acetate pH 4.6) to the cycle sequencing reaction and mixing the components by vortexing. The tubes were placed on ice of 15 minutes and then centrifuged at 3000X g for 30 minutes at room temperature. The supernatant was aspirated and 500ul of 70% ethanol was added to the pellet. The pellet was washed at 3000xg for 5 minutes. The ethanol was removed and the pellet was dried for 20 minutes in a lyophilizer at room temperature. The dried pellet was dissolved in 12ul of HI-DI formamide (highly deionized formamide) by vortexing and sample was transferred to 96 well sequencing plate. The plate was denatured at 96° C for two minutes using Perkin Elmer Gene Amp 9600 System and immediately snap-chilled. The sample tray was placed on an autosampler. The results were analyzed using Applied Biosystems Sequence analysis software v 5.2.

III. Results and Discussion

A total of 26 white colonies were picked up from the mass-excised leaf cDNA library of *C. roseus* and checked for insert frequency (Figure1). Three-fourth of the whites were recombinant and these were subjected to sequencing using T7 sequencing primer. Usable sequence was obtained for 12 clones. 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 (*Altschul et al., 1997*). The basic local alignment search tool (BLAST) available at http://www.ncbi.nlm.nih.gov/BLAST/ was used to analyse the EST sequences. 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 investigation done counter to non-dispensable nucleotide database (all GenBank + EMBL + DDBJ + PDB sequences except EST, STS, GSS, environmental samples or phase 0, 1 or 2 HTGS sequences) as well as a BLASTN done in contrast to GenBank + EMBL + DDBJ database sequences starting with EST . Both the studies were used in combination to assign a fictional task to an EST. The comparison of BLASTN to EST was used only as supportive evidence for plant origin and tentative 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 purpose. If there was any similarities in BLAST homologies with EST, the 'bl2seq' tool (NCBI) was done. The 'Score' values (S) revealed the "significance" for the match amongst the dual sequences. The initial raw score was standardized for "bit score" S', (set of units). Additional parameter used is 'Expect' value (E), indicating the number of hits. This can be "projected" while searching any database. It showed a decreasing tend with 'Score' (S) assigned to a match between two sequences. Higher significance between the two matched was evident by the lower E –value(near to zero).Details of the result of blast analyses are given in Table 4. Both the ESTs LZ365 and LZ418 showed similarity to protein kinase sequence in the database. However, they were not similar to each other when checked with the 'bl2seq' and CLUSTALW tools, implying that they matched with different domains of protein kinase. None of the ESTs showed "no significant similarity" in the BLASTN analysis. Leaf EST LZ362 showed similarity to chloroplast ribosomal protein (CL28), which was as per expectation. EST LZ374 was similar to ribosomal protein S26 from pea. EST LZ445 showed similarity to cytochrome P450 mono-oxygenase from maize. These enzymes play key roles in plant secondary metabolism and hence the function of this EST needs further elucidation. Interestingly, much information could not be derived from the BLAST analysis of other EST sequences. Even for those ESTs where some hints about the putative functions of the genes could be obtained, the true functions could be ascertained only through molecular genetics approaches.

Acknowledgment:
The Authors are thankful to Dr. Zahid Khorshid Head of the Biology Department and Supervisor Dr Aisha Alatawi , University of Tabuk Tabuk Saudi Arabia for their encouragement and support.

References

[1]. Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde B, Moreno RF, Kerlavage AR, McCombie WR and Venter JC: 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**, 1651-1656.

[2]. Altschul SF1, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ. :1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997 Sep 1;25(17):3389-402.

[3]. De Luca V, Balsevich J, Tyler RT, Eilert U, Panchuk BD and Kurz WGW: 1986. Biosynthesis of indole alkaloids: Developmental regulation of the biosynthetic pathway from tabersonine to vindoline in *Catharanthus roseus*. *J Plant Physiol* **125**, 147-156.

[4]. De Luca V and Cutler AJ :1987. Subcellular localization of enzymes involved in indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiol* **85**, 1099-1102.

[5]. Feliciello I and Chinali G. :1993 A modified alkaline lysis method for the preparation of highly purified plasmid DNA from Escherichia coli. Anal Biochem. ;212(2):394-401.

[6]. Horelli-Kuitunen N, Aaltonen J, Yaspo ML, Eeva M, Wessman M, Peltonen L and Palotie A: 1999. Mapping ESTs by fiber-FISH. *Genome Res* 9: 62-71.

[7]. Maria Manuela R. Costa2, Frederique Hilliou2, Patrícia Duarte2, Luís Gustavo Pereira, Iolanda Almeida, Mark Leech, Johan Memelink, Alfonso Ros Barceló and Mariana Sottomayor. :2008 Molecular Cloning and Characterization of a Vacuolar Class III Peroxidase Involved in the Metabolism of Anticancer Alkaloids in *Catharanthus roseus* Plant Physiology vol. 146 no. 2 403-417

[8]. Shelton D, Leach D, Baverstock P and Henry R: 2002. Isolation of genes involved in secondary metabolism from *Melaleuca alternifolia* (Cheel) using expressed sequence tags (ESTs). *Plant Sci.* **162**, 9-15.

[9]. 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.

[10]. Van der Heijden R, Jacobs DI, Snoeijer W, Hallard D, and Verpoorte R: 2004. The *Catharanthus* alkaloids: pharmacognosy and biotechnology. *Curr Med Chem* **11**, 607-628.

[11]. Virmani OP, Srivastava GN and Singh P: 1978. *Catharanthus roseus* – the tropical periwinkle. *Indian Drugs* (Sep), 231-252.

Table1: Buffers and solutions used

Buffers/Solutions	**Composition/Preparation**
(A) Stock solutions	
1M Tris-Cl (pH 8.0)	Dissolved 121.1 g Tris base in 800 ml deionized water, adjusted pH to 8.0 with concentrated HCl, finally adjusted the volume to 1 litre with deionized water and autoclaved.
0.5M EDTA (pH 8.0)	Dissolved 186.1 g disodium ethylenediaminetetraacetate.$2H_2O$ in 800 ml deionized water, stirred vigorously using a magnetic stirrer, adjusted pH to 8.0 with NaOH, finally adjusted the volume to 1litre with deionized water and autoclaved.
10N NaOH	Dissolved 40 g molecular biology grade sodium hydroxide pellets in 80 ml sterile deionized water in a sterile bottle and finally adjusted the volume to 100 ml with sterile deionized water.
10% SDS	Dissolved 10 g molecular biology grade sodium lauryl sulphate in 80ml sterile deionized water in a sterile bottle and finally adjusted the volume to 100 ml with sterile deionized water.
Ethidium bromide (10 mg/ml)	Dissolved 100 mg ethidium bromide powder in 10 ml sterile deionized water in a sterile dark coloured bottle and stored at room temperature.
X-gal (20 mg/ml)	Dissolved 100 mg X-gal powder in 5 ml N, N'-dimethyl formamide in a dark coloured bottle and stored at -20^0C protected from light.
IPTG (24 mg/ml)	Dissolved 120 mg IPTG powder in 5 ml of deionized water, filter sterilized using a 0.22 µm filter unit and stored in aliquots at -20^0C.
50X TAE	Dissolved 242 g Tris base in 500 ml deionized water, added 57.1 ml of glacial acetic acid and 100 ml of 0.5M EDTA (pH 8.0) to it. Adjusted the volume to 1 litre with deionized water and autoclaved.
6X Loading buffer	Dissolved 50 mg bromophenol blue and 50 mg xylene cyanol in 14 ml sterile deionized water, added 6 ml of sterile glycerol, mixed thoroughly and stored in aliquots at 4ºC.
Kanamycin (50 mg/ml)	Dissolved 250 mg kanamycin monosulfate salt in 5 ml deionized water, filter sterilized using a 0.22 µm filter unit and stored in 500 µl aliquots at -20^0C.
Tetracycline (10 mg/ml)	Dissolved 50 mg tetracycline hydrochloride in 5 ml deionized water, filter sterilized using a 0.22 µm filter unit and stored in 500 µl aliquots at -20^0C protected from light.
RNaseA (10 mg/ml)	Dissolved 5 mg RNaseA powder in 0.5 ml sterile water in a microcentrifuge tube having a pin-hole in its cap and kept on a boiling water bath for 20 minutes to destroy any contaminating DNase. Stored at -20ºC.
(B) Plasmid isolation by modified alkaline lysis/PEG precipitation method	
Solution A	25mM Tris (pH 8.0), 10mM EDTA (pH 8.0)
Solution B	0.2N NaOH, 1% SDS. Always prepared a fresh solution.
Solution C	3M potassium acetate (pH4.8). Dissolved 29.442 g potassium acetate in 60 ml deionized water, adjusted pH to 4.8 with glacial acetic acid, finally made up the volume to 100 ml with deionized water and autoclaved.
13% PEG 8000	Dissolved 13 g PEG in 80 ml deionized water, adjusted volume to 100 ml with deionized water and autoclaved.
4M NaCl	11.7 g of NaCl was dissolved in 50 ml of deionised water and autoclaved.
(C) cDNA library	
10mM $MgSO_4$	Dissolved 2.465 g $MgSO_4.7H_2O$ in 1 litre deionized water and autoclaved.
10% (w/v) Maltose	Dissolved 1 g maltose in 10 ml deionized water. Filter sterilized using a 0.22 µm filter unit and stored in aliquots at 4^0C.

(D) Automated DNA sequencing	
3M Sodium acetate (pH 4.6)	Dissolved 6.15 g molecular biology grade anhydrous sodium acetate in 15 ml deionized water, adjusted pH to 4.6 with glacial acetic acid, finally adjusted the volume to 25 ml and autoclaved.
(E) Automated DNA synthesis of oligonucleotides primer	
Dilute ammonium hydroxide	Diluted liquid ammonia solution with deionized water (1:10; v/v).
2% TFA	Mixed 0.1 ml TFA with 4.9 ml deionized water.
20% Acetonitrile	Prepared by mixing 1 ml of acetonitrile with 4 ml of deionized water.

Table 2: Bacterial culture media used in the experiments

Medium	**Composition**
Luria broth (LB)	Dissolved 10 g tryptone, 5 g yeast extract, 10 g NaCl in 800 ml deionized water. Adjusted pH to 7.0 with 5N NaOH, finally adjusted volume to 1 litre with deionized water and autoclaved.
LB–Kanamycin Agar (per Liter)	1 liter of LB agar ,Autoclave , Cool to 55°C , Add 50 mg of filter-sterilized kanamycin .Pour into petri dishes (~25 ml/100-mm plate)
Luria agar (LA)	Added 16 g agar to 1 litre LB before autoclaving and finally sterilized by autoclaving.
5X NZY Broth	Dissolved 0.25 g of NaCl, 0.1 g of MgSO4, 0.25 g of Yeast extract and 0.5 g casein hydrolysate in 8 ml deionised water , adjusted pH to 7.5 with NaOH, made up volume to 10 ml with deionised water and autoclaved.
LA-Tetracycline	Prepared 1 litre LA medium, autoclaved, cooled to 55ºC, added 1.25 ml of filter sterilized tetracycline (10 mg/ml), and poured into plates. Avoided exposure to light.
LA-Kanamycin	Prepared 1 litre LA medium, autoclaved, cooled to 55ºC, added 1 ml of filter sterilized kanamycin (50 mg/ml), and poured into plates.

Table 3: Enzymes, kits, antibiotics, HPLC columns, chemicals, nucleic acids and primers.

Item	**Description**	**Source**
Restriction enzymes	EcoRI, XhoI	Roche
Antibiotics	Tetracycline, Kanamycin	Sigma
Kit	Big dye terminator cycle sequencing ready mix v3.1	Applied Biosystems
Chemicals	acetonitrile, glacial acetic acid, EtBr, absolute ethanol, sodium acetate, agarose, agar, bromophenol blue, tryptone, glycerol, Na_2EDTA, IPTG, potassium acetate, maltose, casein hydrolysate, DMSO, DMF, sodium lauryl sulphate, PEG 8000, blue dextran, $MgSO_4.7H_2O$, NaCl, sodium hydroxide, Tris base, yeast extract, xylene cyanol, X-gal, chemicals for automated DNA synthesis	Himedia/ Merck/ SRL/ Sigma/ Perkin Elmer/ Rankem/ Qualigens * From CIMAP repository
Nucleic acids	λ DNA**,** DNA molecular weight marker (λ DNA/ HindIII + EcoRI fragments), pBK-CMV	Bangalore Genei/ Stratagene
Primer	T7 primer: 5'-GTAATACGACTCACTATAGGGC-3'	Synthesized in lab

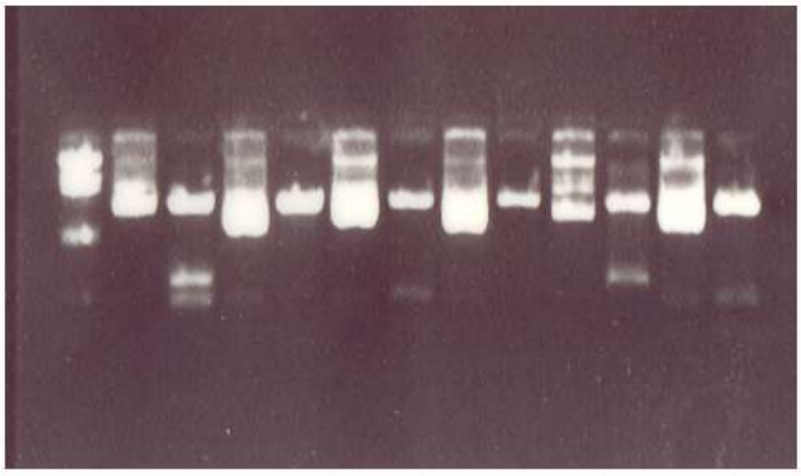

Figure.1 : Restriction analysis of whites from the leaf cDNA library of C. roseus
Lane 1: DNA marker (lambda DNA / Hind III fragments).
Lane 2-13 : pairs of samples showing the unrestricted (left) and the restricted (right) plasmid DNA.

Table 4 : Blast Analysis of EST Sequences

Sr. No	Clones	Length of Read Sequences (bp)	Blast N			Blast X			Consensus putative similarity
			Similarity	Score	E-value	Similarity	Score	E-value	
1	LZ359	520	*Taeniopygia guttata* clone 0061P0018F10 vesicle-associated membrane protein 4 variant 2-like mRNA Zebrafish DNA sequence from clone DKEY-177P2 in linkage group 20, complete sequence	46.1 bits (23) 46.1 bits (23)	0.12 0.12	unknown protein [*Arabidopsis thaliana*] phosphatidate cytidylyl transferase (cdsA) [*Treponema pallidum* subsp. *pallidum* str.*Nichols*]	82.0 bits (201) 37.4 bits (85)	1e-14 0.30	
2	LZ362	594	*N. tabacum* mRNA for chloroplast ribosomal protein CL28	186 bits (94)	6e-44	50S ribosomal protein L28, chloroplast precursor (CL28) [*Nicotiana tabacum*]	176 bits (446)	3e-46	Chloroplast ribosomal protein CL28 [*Nicotiana tabacum*]
3	LZ365	643	*Nicotiana tabacum* NtPERK1 mRNA for PERK1-like protein kinase, partial cds *Arabidopsis thaliana* ATP binding / protein kinase/ protein serine/threonine kinase/ protein-tyrosine kinase (AT3G24550) mRNA	192 bits (97) 95.6 bits (48)	1e-45 2e-16	PERK1-like protein kinase [*Nicotiana tabacum*] ATP binding / protein kinase/ protein serine/threonine kinase/ Protein - tyrosine kinase [*Arabidopsis thaliana*]	277 bits (709) 257 bits (656)	2e-73 3e-67	Protein kinase [*Nicotiana tabacum* and *Arabidopsis thaliana*]
4	LZ374	649	*Lycopersicon esculentum* cDNA, clone: FC14BF02, HTC in fruit *Pisum sativum* ribosomal protein S26 (RPS26) mRNA, complete cds	291 bits (147) 216 bits (109	2e-75 7e-53	Ribosomal protein S26 [*Pisum sativum*]	136 bits (342)	8e-31	Ribosomal protein S26 [*Pisum sativum*]
5	LZ376	571	*Oryza sativa* (japonica cultivar-group) *Arabidopsis thaliana* unknown protein (AT1G29040) mRNA, complete cds	60.0 bits (30) 44.1 bits (22)	9e-06 0.53	Os01g0613300 [*Oryza sativa* (japonica cultivar-group)] unknown protein [*Arabidopsis thaliana*]	131 bits (329) 135 bits (339)	4e-34 8e-37	

6	LZ390	653	*Medicago truncatula* clone mth2-47n14, complete sequence *Nicotiana tabacum* NtEIG-E80 mRNA, complete cds	97.6 bits (49) 95.6 bits (48)	5e-17 2e-16	NtEIG-E80 [*Nicotiana tabacum*] photoassimilate-responsive protein PAR-1b-like protein [*Arabidopsis thaliana*]	227 bits (579) 184 bits (467)	3e-58 3e-45	
7	LZ397	113	Homo sapiens BAC clone RP11-1252L15 from 7, complete sequence Human DNA sequence from clone RP4-710L4 on chromosome Xq11.2-12 Contains the 5' end of the MTMR8 gene for myotubularin related protein 8, a SHC (Src homology 2 domain containing) transforming protein 1 (SHC1) pseudogene and a CpG island	46.1 bits (23) 44.1 bits (22)	0.022 0.089	No significant similarity found			
8	LZ418	645	*Arabidopsis thaliana* ATP binding / kinase/ protein kinase/ protein serine/threonine kinase (AT2G25760) mRNA	224 bits (113)	3e-55	ATP binding / kinase/ protein kinase/ protein serine/threonine kinase [*Arabidopsis thaliana*]	342 bits (876)	1e-92	Protein kinase [*Arabidopsis thaliana*]
9	LZ421	482	Zebrafish DNA sequence from clone DKEYP-72A3 in linkage group 17, complete sequence *Oryza sativa* (japonica cultivar-group)	46.1 bits (23) 44.1 bits (22)	0.11 0.44	L-ornithine oxygenase [*Aspergillus oryzae*]	32.3 bits (72)	7.4	
10	LZ424	529	*Zea mays* clone EL01N0202C08.c mRNA sequence	56.0 bits (28)	1e-04	No significant similarity found			
11	LZ443	671	*Lotus japonicus* genomic DNA, chromosome 4, clone:LjT37A01, TM0244, complete sequence	67.9 bits (34)	4e-08	putative calcium-transporting ATPase [*Oryza sativa* (japonica cultivar-group)]	70.1 bits (170)	8e-11	
12	LZ443	671	*Lotus japonicus* genomic DNA, chromosome 4, clone:LjT37A01, TM0244, complete sequence	67.9 bits (34)	4e-08	putative calcium-transporting ATPase [*Oryza sativa* (japonica cultivar-group)]	70.1 bits (170)	8e-11	
13	LZ445	640	*Nerium oleander* 26S ribosomal RNA gene, partial sequence	1189 bits (600)	0.0	cytochrome P450 like_TBP [*Nicotiana tabacum*] probable cytochrome P450 monooxygenase - maize	210 bits (535) 197 bits (500)	3e-53 4e-49	

Farhina Pasha. "Leaf specific Terpenoid Indole Alkaloid biosynthesis in Catharanthus roseus towards identification of novel structural and regulatory factors." IOSR Journal of Environmental Science, Toxicology and Food Technology (IOSR-JESTFT) , vol. 11, no. 8, 2017, pp. 10–16.