ISOLATION AND CHARACTERIZATION OF TYPE III POLYKETIDE SYNTHASES FROM CHETHIKODUVELI (*Plumbago rosea* L.)

By

DHANYA RADHAKRISHNAN

(2009 - 09 - 101)

THESIS

Submitted in partial fulfilment of the

requirements for the degree of

MASTER OF SCIENCE (INTEGRATED) IN BIOTECHNOLOGY

Faculty of Agriculture

Kerala Agricultural University

M.Sc. (Integrated) BIOTECHNOLOGY COURSE

DEPARTMENT OF PLANT BIOTECHNOLOGY

COLLEGE OF AGRICULTURE

VELLAYANI, THIRUVANANTHAPURAM – 695 522

KERALA, INDIA

2014

DECLARATION

I hereby declare that the thesis entitled **"Isolation and characterization of type III polyketide synthases from chethikoduveli (*Plumbago rosea* Linn.)"** is a bonafide record of research done by me during the course of research and that the thesis has not previously formed the basis for the award of any degree, diploma, fellowship or other similar title, of any other University or Society.

Vellayani
Date: 17.12.2014

Dhanya Radhakrishnan
(2009-09-101)

JNTBGRI
Jawaharlal Nehru Tropical Botanic Garden and Research Institute
(An Institute of Kerala State Council for Science, Technology and Environment, Govt. of Kerala, National Centre of Excellence, Govt. of India)
Karimancode P. O., Palode, Thiruvananthapuram - 695 562, Kerala, India.

CERTIFICATE

Certified that this thesis, entitled **"Isolation and characterization of type III polyketide synthases from chethikoduveli (*Plumbago rosea* Linn.)"** is a record of research work done independently by **Ms. Dhanya Radhakrishnan (2009-09-101)** under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, fellowship or associateship to her.

Place: Thiruvananthapuram
Date: 18-12-2014

Dr. P. Padmesh Pillai
(Major Advisor, Advisory Committee)
Scientist-E1
Biotechnology and Bioinformatics Division
JNTBGRI, Palode
Thiruvananthapuram

Ph : Off. +91(0)472 2869226; Fax : +91(0)472 2869646; e-mail : director_tbgri@rediffmail.com; web site : www.tbgri.in

CERTIFICATE

We, the undersigned members of the advisory committee of Ms. Dhanya Radhakrishnan, a candidate for the degree of **Master of Science in (Integrated) Biotechnology** with major in Biotechnology, agree that the thesis entitled **ISOLATION AND CHARACTERIZATION OF TYPE III POLYKETIDE SYNTHASES FROM CHETHIKODUVELI (*Plumbago rosea* L.)** may be submitted by Ms. Dhanya Radhakrishnan, in partial fulfilment of the requirement for the degree.

Dr. P. Padmesh Pillai
(Chairman, Advisory Committee)
Scientist-E1
Biotechnology and Bioinformatics Division
JNTBGRI, Palode
Thiruvananthapuram

Dr. B. R. Reghunath
(Member, Advisory Committee)
Professor and Head
Department of Plant Biotechnology
College of Agriculture, Vellayani

Dr. K. Satheeshkumar
(Member, Advisory Committee)
Scientist - EII
Biotechnology and Bioinformatics Division
JNTBGRI, Palode

Dr. B. Sabulal
(Member, Advisory Committee)
Scientist - E1 & Head
Phytochemistry &
Phytopharmacology Divisions
JNTBGRI, Palode

Dr. Mathew Dan
(Member, Advisory Committee)
Scientist E1
Plant Genetic Resource Division
JNTBGRI, Palode

ACKNOWLEDGEMENT

If life is a journey and life events are milestones, I have come to an important milestone in my life – the completion of post graduation. As this five year long journey comes to a close, I would like to express my heartfelt gratitude to everyone who made this journey eventful and memorable.

First and foremost I would like to thank God for gifting me with the opportunity to meet all the wonderful people who guided me along and for providing me the chance to associate with some of the reputed institutions of Kerala like Kerala Agricultural University and Jawaharlal Nehru Tropical Botanic Garden and Research Institute.

I would like to express my boundless gratitude to my major advisor, Dr. P. Padmesh Pillai, Scientist E1, JNTBGRI, who provided me with a challenging M. Sc. project and guided me in my first research endeavor. I am grateful to him for the generosity with which he permitted me to use the facilities of Plant Molecular Biology Lab of Biotechnology and Bioinformatics Division, JNTBGRI, Palode. The dissertation provided me wide exposure to molecular biology and valuable hands on experience in all techniques performed in the lab. He was very understanding and ever ready to answer my queries.

This research experience was further enhanced by the guidance of Mr. Jayakumar, the research scholar who supervised me throughout the project. I thank him for his immense support and sincerity. He is one among the best teachers I have had. He was always welcome to my persistent questions and made it a point to clear all my doubts and apprehensions. The other research scholars were also a constant source of support. I would like to specially thank Mr. Manikandan, Mr. Binoy, Ms. Silja, Ms. Akshaya and Ms. Gayathri for their considerable help.

I gratefully acknowledge my heartfelt thanks to my beloved teachers who left no stones unturned to provide us with the best facilities and exposures during the course. Among them, first I would like to thank Dr. K. Rajmohan, Founder Director of M.Sc. (Integrated) Biotechnology Course, who, painstakingly made the multi-institutional mode a reality and for opening new avenues for us. Dr. B. R. Reghunath was very cooperative and I am thankful to him for his critical review that has immensely helped to improve this thesis. Our teachers were a huge source of inspiration and support. While Dr. K. B. Soni always challenged us with tougher tasks and encouraged us to think out of the box, Dr. Swapna Alex was always there to provide timely suggestions and valuable advice. Dr. Lekha Sreekantan was there for us to fall back on in times of crisis, Mrs. Geetha Ramesh for efficiently coordinating matters and for giving us that essential kick start whenever required and Dr. Deepa S. Nair for encouraging and cheering us on. All the other teachers of College of Agriculture and other teaching staff also played a major role in grooming us for the world out there. My heartfelt thanks to one and all. I also wish to place my sincere thanks to the non-teaching staffs, Mr. Saaji and Mr. Saju who helped us in so many different ways during the course of this programme..

I am deeply indebted to my friends, my fellow travelers for making each memory worth living over and over again. They have been a constant source of support, instant motivation and rigorous competition. I consider it my blessing to have been a part of such a well acclaimed batch.

This project would not have been possible without the whole hearted support from all the people who provided me with Plumbago rosea leaf and root samples. I gratefully acknowledge, Dr. K. Satheeshkumar, Dr. Mathew Dan, Mr. Sibi C. Varghese, Dr. Sajitha Rani, Mr. P. Dhanesan, Ms. Dhanya

Thankachan, Mr. Anand Vishnu Prakash and Mrs. Cisily Thankachan for being kind enough to provide me with samples. I am particularly indebted to Mrs. Cisily Thankachan for, she not only provided me with the sample but also took the trouble to courier the sample all the way from Waynad. Similarly, Mr. Sibi also helped greatly by providing the sample, maintaining the same in the institute for subsequent collections during the project and for providing valuable insights into various aspects related to the plant.

I am also thankful to my seniors, Mr. Ashish Subba and Ms. Rehna Augustine for their valuable guidance.

Last but not least, I would like to thank my family. Words fail to express my gratitude to my parents and siblings for always being my strength and for lifting me up during rough times and for making me what I am. I am indebted to my grandmother, aunties and uncles who lovingly cared for me during the five years of this course in the absence of my parents and never made me feel homesick.

If I have missed out anyone, I take this opportunity to gratefully acknowledge the help that had been provided, however small it may have been.

This journey has been fun and the contributions of all the people who have touched my life will always be remembered and acknowledged in the journey ahead.

Dhanya Radhakrishnan

Dedicated to My Parents

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

LIST OF PLATES CONTINUED

LIST OF APPENDICES

LIST OF ABBREVIATIONS

A_{260}	Absorbance at 260 nm wavelength
A_{280}	Absorbance at 280 nm wavelength
$A_{260/280}$	Ratio of absorbances at 260 nm to 280 nm
Asn	Asparagine
β-actin	Beta actin
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
cDNA	Complementary DNA
CHS	Chalcone Synthase
cm	Centimetre
CTAB	Cetyl trimethyl ammonium bromide
Cys	Cysteine
° C	Degree Celsius
DEPC	Diethyl pyrocarbonate
DNA	Deoxyribo nucleic acid
EDTA	Ethylene diamine tetra acetic acid
EtBr	Ethidium bromide
ExPASy	Expert Protein Analysis System
g	Gram
h	Hour
His	Histidine
HPLC	High Pressure Liquid Chromatography
ISSR	Inter-Simple Sequence Repeat
JNTBGRI	Jawaharlal Nehru Tropical Botanic Garden and Research Institute
kbp	Kilo basepair

kDa	Kilo Dalton
m	Metre
M	Molar
mg	Milli gram
min	Minute
ml	Millilitre
m*M*	Millimolar
mm	Millimetre
MSA	Multiple Sequence Alignment
NCBI	National Center for Biotechnology Information
PCR	Polymerase Chain Reaction
%	Percentage
Phe	Phenylalanine
PKS	Polyketide Synthase
p*M*	Picomolar
PVP	Polyvinyl pyrrolidone
RACE	Rapid Amplification of cDNA Ends
RLM-RACE	RNA Ligase Mediated Rapid Amplification of cDNA Ends
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolutions per minute
rRNA	Ribosomal RNA
RT-PCR	Reverse Transcription Polymerase Chain Reaction
s	Second
TBE	Tris Borate EDTA
TE	Tris EDTA

TLC	Thin Layer Chromatography
T_m	Melting temperature
µl	Microlitre
µ*M*	Micromolar

INTRODUCTION

1. INTRODUCTION

Since time immemorial, medicinal plants have been used in various health care systems like Ayurveda. *Plumbago rosea* Linn., belonging to the family Plumbaginacea, is known as chethikoduveli in Malayalam and is well known for its ethnomedicinal values. It is an esteemed remedy for leucoderma, dyspepsia, diarrhoea and leprosy (Kirthikar and Basu, 1975). The root bark of *P. rosea* accumulates a therapeutically important phytometabolite, plumbagin. Plumbagin is an orange yellow crystalline hydroxyl naphthoquinone which has been reported to possess various pharmacological properties such as anticancer (Parimala and Sachdanandam, 1993), antimalarial (Likhitwitayawuid *et al.*, 1998), antimicrobial (Didry *et al.*, 1994), cardiotonic (Itoigawa *et al.*, 1991) and antifertility (Bhargava, 1984) properties.

Due to the large number of uses, the annual requirement for plumbagin in Indian subcontinent is reported to be about 7 tonnes (Bhattacharya *et al.*, 2008). The demand for plumbagin is constantly increasing in both domestic and international markets. This has led to the overexploitation of *P. rosea* from natural habitat. Presently, this plant is categorized to be rare (Chetia and Handique, 2000). Commercial cultivation may be promoted for large scale production of the officinal part but *P. rosea* grows slowly and it may take 2-6 years to achieve the desired level of plumbagin in the mature roots of intact plants (Komariah *et al.*, 2002). On the other hand, synthetic approach of plumbagin production is not commercially promising (Ichihara *et al.*, 1980; Wurm *et al.*, 1981).

In this situation, intervention of modern biotechnological approaches such as micropropagation (Satheesh Kumar and Bhavanandan, 1988), callus culture (Heble *et al.*, 1974), cell culture (Komariah *et al.*, 2001), adventitious root cultures (Panichayupakaranant and Tewtrakul, 2002) and hairy root cultures (Gangopadhyay *et al.*, 2008) have been studied to enhance plumbagin production. However, a feasible production system that ensures uniformity and quality in *in*

vitro plumbagin content is yet to be developed. Although callus induction and proliferation systems are known to be useful for the production of natural products, the concentration of active principle is often found to be very low compared to field grown plants. This could possibly be due to inherent biochemical or genetic instability of the culture systems acquired over several passages.

This problem can be transcended by engineering metabolic pathways to increase the content of the desired secondary compound or to lower the concentration of undesired compounds or to introduce novel genes in specific steps in the pathway for biotransformation of compounds. However, it entails a thorough understanding of sequential steps involved in the synthesis of the compound of interest at the molecular level. It is presumed that plumbagin synthesis in *Plumbago* spp. is via polyketide pathway mediated by type III polyketide synthase.

Therefore, the project targets to isolate and sequence characterize type III polyketide synthase gene(s) from *Plumbago rosea* which would eventually lead to augmenting the production of plumbagin through the development of high yielding cell lines thereby helping to ensure uniformity and quality in the active principle content, in future. Furthermore, chemical variation study with respect to plumbagin content and genetic diversity using ISSR markers was also carried out in additional accessions of *Plumbago rosea.*

REVIEW of LITERATURE

2. REVIEW OF LITERATURE

In this chapter, literature concerning *Plumbago rosea*, plumbagin, biosynthetic pathway involved in plumbagin production in *P. rosea*, polyketide synthase gene and important techniques associated with the study undertaken have been reviewed.

2.1 ROSE COLOURED LEADWORT

Plumbago rosea Linn. (Plate1) (synonym *Plumbago indica* Linn.), commonly known as Chitrak or rose coloured Leadwort is a plant of immense ethno medicinal value and has been used since ancient times as a part of folk medicine. This dicotyledonous, perennial shrub, characterised by red flowers, is found growing throughout tropical India, often as a cultivated plant or as a garden escape. The root of this plant is acrid, vesicant, abortifacient and stimulant. The extract is mixed in bland oil and used externally or internally in rheumatism and paralytic afflictions. The milky juice of the plant is also used against ophthalmia and scabies (Watson and Dallwitz, 1992). It is also considered as an esteemed remedy for leucoderma, dyspepsia, diarrhoea and leprosy (Kirthikar and Basu, 1975).

P. rosea belongs to the family Plumbaginaceae which also comprises other locally recorded species such as *Plumbago capensis* Thunb. and *Plumbago zeylanica* Linn. While *P. capensis* is grown for its ornamental value, the roots of *Plumbago rosea* and *Plumbago zeylanica* are generally prescribed for medicinal purposes in Indian medicine (Warrier *et al.*, 1997). The roots of these species are the principle source of the medicinally important napthoquinone, plumbagin. Among six different species of Plumbago, *P. indica* is the richest source of plumbagin (Mallavadhani *et al.*, 2002). Hence, these are commercially used as raw material for extracting plumbagin in China and other countries (Grieve, 1995). In addition to plumbagin several other groups of phytochemicals which

include, isoshinanolone, 6-hydroxyl plumbagin, plumbaginol, leucodelphinidin and steroids like campesterol, sitosterol and stigmasterol have also been isolated from the aerial parts of this plant (Schmelzer and Gurib-Fakim, 2008)

Due to the increased market demand in both domestic and international level, over-exploitation of wild grown plants has led to the rapid decline of this plant in its natural habitat. The slow growth rate, slow attainment of desired level of plumbagin, absence of seeds and lack of fruiting stage necessitate the search for an alternative and effective source to meet with the commercial demand for plumbagin (Chetia and Handique, 2000). *P. indica* L. can be propagated both by sexual and asexual methods. However, as it produces very small amount of seed and since its seed exhibit poor germination, asexual propagation using both root and stem cutting are most commonly used (Krishnan, 2003).

In vitro culture methods are increasingly being depended upon for the sustained conservation and propagation of medicinal plants (Canter *et al.*, 2005; Sarasan *et al.*, 2006). Various *in vitro* protocols have been attempted to augment plumbagin production in *P. rosea* also. These include micropropagation (Satheesh Kumar and Bhavanandan, 1988), callus culture (Heble *et al.*, 1974), cell culture (Komariah *et al.*, 2001), adventitious root cultures (Panichayupakaranant and Tewtrakul, 2002) and hairy root cultures (Gangopadhyay *et al.*, 2008). Some of these attempts were successful.

2.1.1 General Morphology of *Plumbago rosea*

Plumbago rosea L. is a perennial evergreen shrub with long succulent roots. This half woody plant which prefers shade, frequently grows upto a height of 1.5 m. The leaves of *P. rosea* are alternate and simple, hairless, mid to deep green, and papery. It is 15 x 7 cm, oblong- elliptic, broadest towards the middle, acute at apex and attenuate at base. The petiole is short and auricles are absent at the base. Stem is erect, trailing or climbing, simple or branched from the base and tubular

with red nodes. The plant possesses an elongated spike and scarlet raceme inflorescence that is many flowered and measuring 10-30 cm long. The bracts are ovate measuring 2–3 × 1.5–2 mm with acuminate apex. The peduncle is 2–10 cm long and not glandular. The plant bears regular, bisexual flowers, exhibiting pentamery and with red corolla. The pedicel may be 0-1 mm long. The glandular calyx is tubular shaped and measures 8-9 mm long. The corolla tube measures 2.5-4.5 cm. The lobes are obovate, distinctly mucronate, purple to red in colour with rounded apex and measures 1.5-3 cm in diameter. Filaments of stamen are as long as the corolla tube, anthers are exserted just beyond the throat. Ovary is superior, ellipsoid-ovoid and indistinctly angular. Style is basally pilose and filiform. Stigma possesses 5 lobes and stigmatic glands without enlarged apex are present. Seeds and fruits are absent (Van Steenis, 1949; Saralamp *et al.*, 1996; Kurian and Sankar, 2007).

2.2 PLUMBAGIN

Plumbagin (5-hydroxyl-2-methyl-1,4-naphthoquinone) (Plate 2) is an orange yellow crystalline pigment with a molecular weight of 188.18. It has a melting point of 78-79° C and is sparingly soluble in hot water while it is completely soluble in alcohol, acetone, chloroform, benzene and acetic acid (Sigma-Aldrich, 2014).

In nature, it is predominantly produced by *Plumbaginaceae* members. *Plumbago rosea* is its richest known source with an accumulation of upto 0.9-1.0 per cent plumbagin in the tuberous roots (Dinda *et al*, 1997; Mallavadhani *et al.* 2002). Plumbagin has also been reported in other families like Droseraceae, Ebenaceae, Nepenthaceae, Polygonaceae and Euphorbiaceae (Thomson, 1971; Heubl *et al.*, 2006). Tokunaga *et al.* (2004) suggested that since many plants accumulate this naphthoquinone in significant amounts it might be responsible for important role in the plants' interaction with its environment. Studies by Villavicencio and Perez-Escandon (1992) proved that plumbagin acts as an

antifeedant agent on herbivorous insects such as *Lepidoptera.* Investigators observed that insects feeding on plants containing a critical dose of plumbagin die either immediately (Gonçalves *et al.*, 2009) or during the next ecdysis due to the inhibiton of ecdysteroid (Joshi and Sehnal, 1989) and chitin synthetase production (Kubo *et al.*, 1983). Studies have established the fungicidal activity of plumbagin against plant pathogenic species. It acts as a potent phytoalexin against parasitic plants (Bringmann *et al.*, 1999) and the inhibition of germination of seeds from other species (Spencer *et al.*, 1986; Gonçalves *et al.* 2009). From these findings Rischer *et al.* (2002) concluded the importance of the production of plumbagin and related naphthoquinones in improving the plants' fitness and conferring an adaptive advantage over other plants.

During experiments with hairy root cultures, Gangopadhyay *et al.*, (2011) observed that plumbagin is produced intracellularly and its synthesis is linked to root differentiation, therefore it may take years to achieve optimum level of plumbagin. Nonetheless, the principle source of plumbagin is still the roots of *Plumbago indica* which is collected from natural population (Mallavadani *et al.*, 2002). Due to the unsustainable nature of exploitation of natural sources, synthetic approach had been attempted as an alternative. However, this did not prove to be commercially promising (Ichihara *et al.*, 1980; Wurm *et al.*, 1981). In this situation, modern biotechnological approaches have been tried.

The annual requirement for plumbagin in Indian subcontinent is reported to be about seven tonnes (Bhattacharya *et al.*, 2008) and the demand is constantly rising due to its broad range of pharmacological activities. Detailed studies have been conducted by Nguyen *et al.* (2004) and Kuo *et al.* (2006) on the anticancer properties of plumbagin which includes anticancer activity against melanoma cell lines, breast cancer cell lines and lung cancer cells. Checker *et al.*, (2009) has shown the ability of plumbagin to inhibit T cell proliferation by blocking cell cycle progression, thereby, producing anti-inflammatory effect. Paiva *et al.* (2003) has reported its ability to produce antimalarial effect by hindering with the action

of *Plasmodium falciparum* enzyme succinate dehydrogenase. Plumbagin is established to have antimicrobial activity by various workers. Due to the action of plumbagin, the growth of *Staphyloccocus aureus* and *Candida albicans* was completely inhibited and it was also observed to be effective against *Salmonella gallinarum*, *S. typhimurium*, *Escherichia coli*, *Proteus vulgaris* and *Bacillus subtilis* (Desta, 1993; Didry *et al.*, 1994; Paiva *et al.*, 2003). Antifertility action and abortifacient activity have also been reported (Kini *et al.*, 1997 and Premkumari *et al.*, 1977).

The extraction and determination of plumbagin from *Plumbago spp.* have been reported by many researchers. Kitanov and Pashakov (1994) isolated plumbagin using HPLC technique from petroleum ether extract of *P. europaea*, while Gupta *et al.* (1993) purified the compound from the root of *P. zeylanica* L. using preparative silica gel column chromatography. Plumbagin from *Plumbago zeylanica* L. and *Plumbago indica* L. roots has also been extracted with methanol and was determined by HPLC or TLC densitometric method (Komaraiah *et al.*, 2001; Panichayupakaranant and Tewtrakul, 2002). Plumbagin extraction has been performed previously using chloroform by Soxhlet extraction by Zhong *et al.* (1984) and Pakulski and Budzianowski (1996). The plumbagin content was then estimated by TLC. Plumbagin extraction employing cold extraction with chloroform is the easiest of all the methods and comparatively less time consuming. This has been carried out by Heble *et al.* (1974). This was followed by spectrophtometric estimation of plumbagin at 415 nm. The plumbagin estimation using HPLC could then be performed using the method developed by Crouch *et al.* (1990) for the isolation of plumbagin from *in vitro* and *in vivo* grown *Drosera* species.

2.3 BIOSYNTHETIC PATHWAY OF PLUMBAGIN IN *Plumbago rosea*

Plumbagin is a hydroxyl naphthoquinone. Naphthoquinones are known to be synthesised by four different pathways in higher plants namely, Shikimate, homogentisate, the acetate-malonate and the mevalonate pathways (Jaya, 1999). Most naphthoquinones are derived from Shikimate pathway by the direct incorporation of Shikimate into the benzene ring of the naphthoquinones with the retention of the carboxy group to form various isoprenoid naphthoquinones (Dey and Harbone, 1989). However, plumbagin from *Plumbago* spp. is derived from six two carbon units via acetate-malonate pathway which is also known as the polyketide pathway. The acetate-malonate pathway which is well established in microorganisms uses L-alanine as precursor. Durand and Zenk (1974) are credited with the discovery of this route of synthesis from tracer experiments.

According to Teuscher (1970) plumbagin was also expected to be formed from Shikimate pathway followed by C-methylation in the 2-position, due to its structural similarity to juglone and menadion. However, feeding labelled Shikimate, methionine, tyrosine, phenyalanine and mevalonate to young shoots of *P. europeae* led to insignificant incorporation (0.02-0.08%) into plumbagin. On the other hand, labelled acetate and malonate gave good incorporation (0.65-3.49%), showing that the expected labelling pattern was present and that the acetate pathway was operating. From these studies it had been postulated by Durand and Zenk (1971) that the basic skeleton of the naphthoquinones is produced by a polyketide synthase involved in the polyketide pathway.

An overview of plumbagin biosynthetic pathway proposed by Rischer *et al.* (2002) is provided in Plate 3.

Plate 1. Inflorescence of *Plumbago rosea* Linn.
Photo Courtesy: Rathnayake, K. (2013)

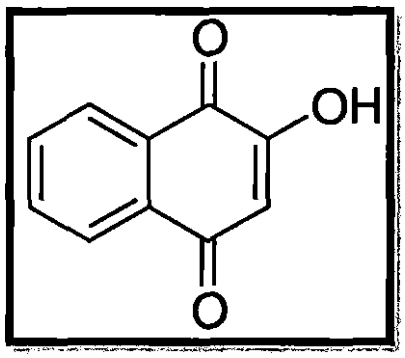

Plate 2. Chemical structure of plumbagin
Source: Sigma-Aldrich (2014)

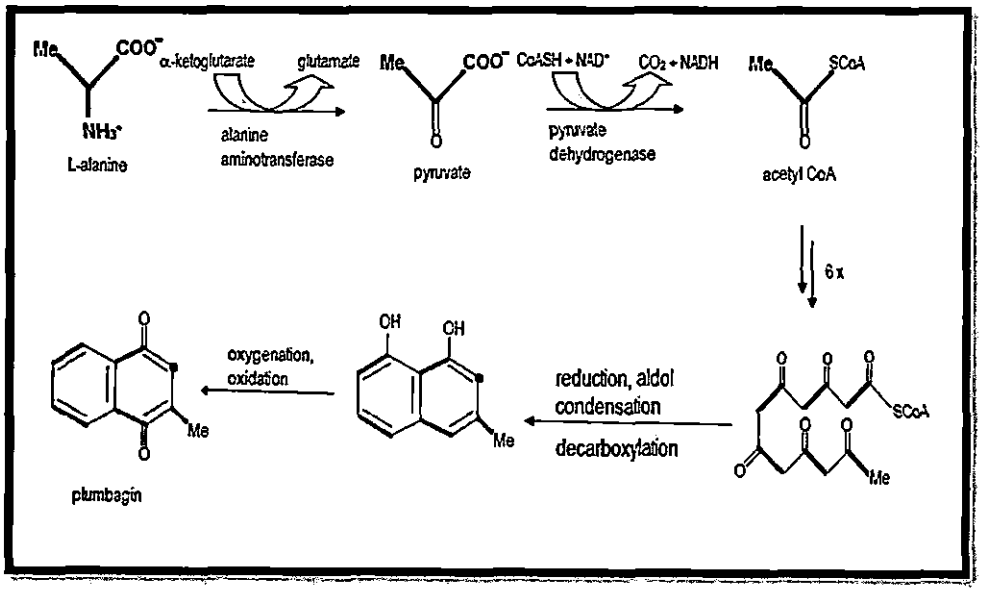

Plate 3. Proposed plumbagin biosynthetic pathway (Rischer *et al.*, 2002)

2.4 TYPE III POLYKETIDE SYNTHASES

Polyketide Synthases (PKSs) enzymes are structurally and functionally related to FAS (Fatty acid synthases). Based on their protein architecture, these enzymes can be classified into type I, type II and type III PKSs (Hopwood and Sherman, 1990; Staunton and Weissman, 2001; Fischbach and Walsh, 2006).

In plants, type III polyketide synthases (PKSs) are involved in the biosynthesis of polyketides which include a variety of secondary metabolites such as flavonoids, stilbenes, benzophenones and benzalacetone derivatives (Schröder, 1997; Flores-Sanchez and Verpoorte, 2009). Chalcone synthase (CHS, EC 2.3.1.74) which synthesizes naringenin chalcone, the precursor of flavonoids is considered as the prototype of type III PKSs. Together with stilbene synthase (STS, EC 2.3.1.95) these two enzymes are the most extensively studied at biochemical and the molecular levels from the group of type III PKSs (Schröder, 2000; Austin and Noel, 2003).

2.4.1. Type III PKSs from Different Plants

Several type III PKSs have been discovered in plants (Table 1). More than 859 nucleotide sequences of plant PKSs have been reported and documented in NCBI database and several plant PKS crystalline structures have also been characterized by investigators (Austin *et al.*, 2004; Morita *et al.*, 2008; Taguchi *et al.*, 2008).

Table 1 represents some of the Type III PKSs that have been previously reported in plants.

Table1. PKS genes previously reported* in plants

Enzyme	Plant Species	References
Benzalacetone synthase (BAS), EC 2.3.1.-	*Rubus idaeus,*	Zheng and Hrazdina, 2008
	Rheum palmatum	Abe *et al.*,2001
C-methylchalcone synthase (PstrCHS2)	*Pinus strobus*	Schröder *et al.*,1998
2-pyrone synthase (2-PS)	*Gerbera hybrid*	Eckermann *et al.*, 1998
p-Coumaroyltriacetic acid synthase (CTAS)	*Hydrangea macrophylla* var. *thunbergii*	Akiyama *et al.*, 1999
Chalcone synthase (CHS), EC 2.3.1.74	*Medicago sativa*	Austin and Noel, 2003
Phlorisovalerophenone synthase (VPS), EC 2.3.1.156	*Humulus lupulus*	Okada and Ito, 2001
Isobutyrophenone synthase(BUS)	*Hypericum calycinum*	Klingauf *et al.*, 2005
Benzophenone synthase (BPS), EC 2.3.1.151	*Centaurium erythraea*	Beerhues, 1996
	Hypericum androsaemum	Liu *et al.*,2003
Pentaketide chromone synthase (PCS)	*Aloe arborescens*	Abe *et al.*, 2005
Hexaketide synthase (HKS)	*Drosophyllum lusitanicum*	Jindaprasert *et al.*, 2008
	Plumbago indica	Springob *et al.*, 2007
Aloesone synthase (ALS)	*R. palmatum*	Abe *et al.*, 2004
Octaketide synthase (OKS)	*A. arborescens*	Abe *et al.*, 2005
Stilbene synthase (STS), EC 2.3.1.95	*Arachis hypogaea*	Schöppner and Kindl, 1984
Pinosylvin synthase, EC 2.3.1.146	*Pinus strobus*	Raiber *et al.*, 1995

*Adopted from (Isvett and Verpoorte, 2009)

2.4.2. Molecular Structure of Type III PKSs

Extensive X-ray diffraction and subsequent mutational studies have been undertaken to study the structural and functional details of these enzymes (Ferrer *et al.*, 1999; Jez *et al.*, 2000; Austin *et al.*, 2004; Sankaranarayanan *et al.*, 2004). Schröder (2000) and Austin and Noel (2003) concluded that there are no significant differences in the conformation of these crystalline structures. PKSs form a symmetric homodimer with a subunit molecular mass of 40-45 kDa. Tropf *et al.* (1995) suggested that the enzyme displays a αβαβα five layered core and an independent active site is present in each monomer. The dimerization is required for activity and an allosteric cooperation between the two active sites. Ferrer *et al.* (1999) proposed that the Met137 (numbering as in *Medicago sativa* CHS) in each monomer helps to shape the active site cavity of the adjoining subunit.

Despite the functional diversity, the amino acid sequences of the plant-specific PKS proteins share common features. Comparative analyses have revealed that these proteins contain approximately 400 amino acids with highly conserved amino acid residues in the active centre. Jez *et al.* (2002) remarked that the catalytic triad, consisting of Cys164, His303 and Asn336 (numbered according to *Medicago sativa* CHS2) residues, is highly conserved in all plant-specific PKSs. The two 'gatekeeper' phenylalanines are preserved in almost all of the plant specific PKSs with a few exceptions (Abe *et al.*, 2001; Jez *et al.*, 2002). The PKS family proteins also exhibit some additional conserved amino acid residues that participate in shaping the geometry of the active site (Ferrer *et al.*, 1999).

2.4.3 Mechanism of Type III PKSs

The architecture of the active site cavity acts as a size-based filter and determines starter molecule specificity, number of elongation cycles and the stereochemistry of the cyclization reaction (Jez *et al.*, 2000; Austin *et al.*, 2004).

The basic principle of the reaction mechanism consists of the use of a starter CoA ester to perform sequential condensation reactions with two carbon units from a decarboxylated extender, usually malonyl-CoA. A linear polyketide intermediate is formed which is folded to form an aromatic ring system (Schröder, 1999).

In particular, the active site is composed of a CoA-binding tunnel, a starter substrate-binding pocket and a cyclization pocket, and the three residues: Cys164, His303 and Asn336 define this active site. Each active site is buried within the monomer and the substrates enter via a long CoA-binding tunnel. The Cys164 is the nucleophile that initiates the reaction and attacks the thioester carbonyl of the starter resulting in transfer of the starter moiety to the cysteine side chain. Asn336 orients the thioester carbonyl of malonyl-CoA near His303 with Phe215, providing a nonpolar environment for the terminal carboxylate that facilitates decarboxylation. A resonance of the enolate ion to the keto form allows for condensation of the acetyl carbanion with the enzyme-bound polyketide intermediate. Phe215 and Phe265 perform as gatekeepers (Austin and Noel, 2003). The recapture of the elongated starter-acetyl-diketide-CoA by Cys164 and the release of CoA set the stage for additional rounds of elongation, resulting in the formation of a final polyketide reaction intermediate. Later an intramolecular cyclization of the polyketide intermediate takes place (Abe *et al.*, 2003; Jez *et al.*, 2000; Jez *et al.*, 2001; Lanz *et al.*, 1991; Suh *et al.*, 2000). The GFGPG loop is a conserved region on plant PKSs that provides a scaffold for cyclization reactions (Suh *et al.*, 2000; Austin and Noel, 2003). The remarkable functional diversity of the plant PKSs derives from small modifications in the active site which greatly

influence the selection of the substrate, number of polyketide chain extensions and the mechanism of cyclization reactions. The volume of the active site cavity influences the starter molecule selectivity and limits polyketide length.

The enzyme involved in naphthoquinone biosynthesis presumably catalyzes five extensions of an acetyl-CoA starter with malonyl-CoA extender yielding a hexaketide, which subsequently undergoes two cyclizations to a naphthalene ring (Springob *et al.*, 2007).

2.4.4 Isolation and Characterisation of Hexaketide Synthases

Plant PKSs have 44-95 per cent amino acid identity and are encoded by similarly structured genes. For example, CHSs from *Petunia hybrida*, *Petroselinum hortense*, *Zea mays*, *Antirrhinum majus*, and *Hodeum vulgare*, and STS from *Arachis hypogaea* have 70-75 per cent identity on the protein level and the *CHS* and *STS* genes contain an intron at the same conserved position (Schröder *et al.*, 1988; Schröder and Schröder, 1990). The common features in amino acid sequences makes it possible to isolate new genes coding for PKS family proteins by homology-based techniques, such as using degenerate oligonucleotide primers as have been performed by Abe *et al.* (2005) and Radhakrishnan *et al.* (2009) and many other investigators.

Cloning of PKS enzyme genes that are involved in the biosynthesis of plumbagin have been previously undertaken by workers. Jindaprasert *et al.* (2008) and Springob *et al.* (2007) have cloned and expressed hexaketide synthases from *D. lusitanicum* and *P. indica* respectively in *Escherichia coli*. During *in vitro* assays using the purified recombinant PKS from *P. indica,* synthesis of several pyrones were observed. However, napthoquinones, which are natural products of PKS in the respective plants, were not observed in the *in vitro* synthesis. Therefore, it was concluded that novel PKS is involved in the biosynthesis of

naphthoquinones, and additional cofactors are probably required for the biosynthesis of these secondary metabolites *in vivo*.

Cloning of a hexaketide synthase from *P. zeylanica* (PzPKS) and its expression in tobacco plants (*Nicotiana tabacum* L.) was studied by Jadhav *et al.* (2013). However, transgenic tobacco plants were unable to synthesize plumbagin. These studies indicate that these plants lack the other components of the pathway that may be required in plumbagin synthesis (Springob *et al.*, 2007).

2.5 RNA ISOLATION

Isolation of high-quality RNA (Ribo Nucleic Acid) is a basic requirement for many molecular biology experiments. It is an essential prerequisite for gene expression studies using reverse transcription polymerase chain reaction (RT-PCR), RACE (Rapid Amplification of cDNA Ends), northern hybridization and microarray analysis (Noor Adila *et al.*, 2007).

However, most plant materials contain high levels of RNase activity, proteins, polysaccharides and polyphenols, which reduce the RNA yield and quality by co-precipitating with RNA (Tai *et al.*, 2004; Wang *et al.*, 2008).

RNA extraction protocols involve completely breaking the cells, minimising the activity of RNAses released during cell lysis and avoiding the accidental introduction of RNAses from any other source in the laboratory, because RNA is much more susceptible to degradation than DNA (Sambrook *et al.*, 1989).

A major cause of RNA degradation is ribonuclease contamination which may be endogenous or exogenous in origin. Cells and tissues, finger grease and bacteria and/or fungi in airborne dust particles are common sources of exogenous ribonuclease. To minimize exogenous ribonuclease contamination, appropriate

precautions must be followed when handling RNA (Blumberg, 1987; Welsh *et al*, 1997), and the caution must be followed in virtually all RNA-related procedures.

Wolf *et al.* (1970) reported that exogenous RNase can be destroyed by incubating labwares and electrophoresis apparatus in 0.1% DEPC (Diethyl Pyrocarbonate) at 37°C for several hours followed by autoclaving for 15-20 minutes to destroy DEPC. DEPC reacts with the ε-amino groups of lysine and the carboxylic groups of aspartate and glutamate both intra- and intermolecularly which leads to the polymerisation of ribonuclease. However, DEPC treatment is time consuming and it reacts with the adenine residues of RNA, rendering it inactive in *in vitro* translation reactions (Blumberg, 1987).

Claros and Cánovas (1999) suggested that endogenous RNase could be further avoided by preparing all the required solutions with Milli-Q autoclaved water, double autoclaving all glassware and tubes, and rinsing other plasticware with chloroform or H_2O_2, or baking for 4 h at 160°C. Since hands are usually a source of contaminating RNase, gloves must be worn. Guanidinium thiocyanate is most widely used to inhibit RNases during RNA isolation. Guanidinium is an effective inhibitor of most enzymes due to its chaotropic nature (Chomczynski and Sacchi, 1987; Sambrook *et al.*, 1989).

Numerous standard protocols have been developed for the effective isolation of high quality RNA suitable for functional genomics experiments (Chomezynski and Sacchi, 1987; Logemann *et al.*, 1987; Ainsworth, 1994). Trizol reagent was an improvement devised by Chomczynski (1993) by mixing guanidinium isothiocyanate, phenol, sodium acetate, LSS (N-lauroyl-sarcosine sodium), and other strong denaturants together. However these are often not applicable for a wide range of plant species owing to their varying content of interfering secondary metabolites and enzymes. This has resulted in the publication of a range of other RNA isolation protocols. Salzman *et al.* (1999) have developed a RNA isolation method for plant tissues containing high levels of

phenolic compounds or carbohydrates. Springob *et al.* (2007) have demonstrated the use of this method for RNA isolation from the root of *Plumbago rosea* for expression studies.

Due to the ability of RNA to bind selectively on solidified membranes some RNA isolation kits are commercially available for routine RNA isolation from plant tissues (Yamashita *et al.*, 2005). Kits are time saving as these bypass reagent preparation steps and minimise or eliminate the use of organic solvents. Many of these kits exploit silica columns which can recover even RNA molecules as small as 200 bases such as miRNA (micro RNA) and can be used for clean up purposes. These work by the binding of RNA to silica in a high salt chaotropic environment produced by diluting nucleic acid sample in guanidine reagent. Following number of washes, the purified material is eluted from the matrix under a low salt condition (Farrell, 2009).

2.6 RT-PCR

RNA is an appropriate template for gene expression studies, however it cannot serve as a PCR template. Therefore, to isolate and study a particular eukaryotic gene, transcribed at a certain time or under certain conditions, Complementary DNA (cDNA) synthesis is inevitable. cDNA is the reverse transcriptase product of mRNA and represents the coding sequence of all transcribed genes at the time of RNA isolation (Kimmel & Berger, 1987). The desired cDNA is synthesised by RT-PCR which is an *in vitro* method for enzymatically amplifying defined sequences of RNA (Rappolee *et al.* 1988). This technique combines cDNA synthesis from RNA templates with PCR to provide a rapid, sensitive method for analyzing gene expression often from a small concentration of target RNA (Linderman, 2001).

RT-PCR can also circumvent time-consuming and technically demanding cloning steps and generates full-length cDNA inserts for cloning (Borson *et al.*

1992), or arbitrarily primed enhanced sequence tag cDNA libraries (Neto *et al.* 1997).

The template for RT-PCR can be total RNA or poly $(A)^+$ selected RNA. RT reactions can be primed with random primers, oligo(dT), or a gene-specific primer (GSP). The use of mRNA-specific primers decreases background priming, whereas the use of random and oligo-dT primers maximises the number of mRNA molecules that can be analysed from a small sample of RNA. RT-PCR can be carried out either in two-step or one-step formats. In two-step RT-PCR, each step is performed under optimal conditions. cDNA synthesis is performed first in RT buffer and one tenth of the reaction is removed for PCR. In one-step RT-PCR, reverse transcription and PCR take place sequentially in a single tube under conditions optimized for both RT and PCR. The reactions are catalysed by dedicated RNA- and DNA-dependent DNA polymerases (Santos *et al*, 2004).

The two commonly used RTs are Avian Myeloblastosis Virus Reverse Transcriptase (AMV-RT) and Moloney murine leukaemia virus reverse transcriptase (MMLV-RT). Brooks *et al.* (1995) and Freeman *et al.* (1999) reported that AMV-RT is more robust than MMLV-RT, retains significant polymerisation activity up to 55°C and can help to eliminate problems associated with RNA secondary structure. In contrast, MMLV-RT and engineered derivatives (Kotewicz *et al.*,1988) have significantly less RNAse H activity than AMV-RT (Gerard *et al.*, 1998) combined with a proofreading 3'-5' exonuclease activity to reduce error rate (Arezi and Hogrefe, 2007) as a result, MMLV-RT is a better choice for the amplification of full-length high fidelity cDNA molecules (DeStefano *et al.* 1991).

The first-strand synthesis protocol generates a heterogeneous population of cDNA molecules from all available poly$(A)^+$ mRNA or total RNA, while subsequent amplification with sequence-specific primers yields a homogeneous population of the specific cDNA of interest, eliminating the need for amplification

and screening of a cDNA library. Resulting populations of cDNA can be used for microarray, conventional and real time PCR amplification.

β-Actin mRNA is used as a reference for RT-PCR assays. It was one of the first RNAs to be used as an internal standard, and it is still advocated as it is expressed at moderately abundant levels in most cell types and encodes a ubiquitous cytoskeleton protein (Kreuzer *et al.* 1999).

2.7 PCR

A basic requirement of gene expression studies is to have sufficient copies of the desired DNA fragments to work with. However, every genome usually contains only two copies of a gene which will not suffice for the study requirement. This shortcoming is overcome by PCR which is an enzymatic method of selectively replicating a specific nucleic acid sequence upto several hundred nucleotides in length (Saiki *et al.*,1988). Traditionally, PCR reactions are automatedly performed in a tube using a thermal cycler. The tube contains the reaction constituents, namely, amplification buffer with magnesium chloride, the forward and reverse primers, deoxynucleoside-5'-triphosphate (dNTP) molecules and DNA polymerase enzyme (Etebu, 2013).

The section and length of the region on the template that is to be amplified is determined by the forward and reverse primers. The oligonucleotide primers achieve this by annealing complementarily to the flanking opposite ends of the region of interest of the template DNA. The target DNA sequence is then amplified by thermostable DNA polymerase enzyme by extending the primers via a process in which the enzyme uses the deoxynucleoside-5'-triphosphate (dNTP) molecules available in the reaction matrix to build an oligonucleotide chain complementary to the template DNA. All the required reaction components are present in an amplification buffer which is most often Tris-based so as to regulate the pH of the reaction, which in turn affects the DNA polymerase activity and

fidelity. Magnesium chloride present in the buffer functions as a cofactor for the DNA polymerase enzyme. (Mullis and Faloona, 1987; Saiki *et al.*, 1988; Hill and Stewart, 1992).

Thermostable Taq DNA Polymerase extracted from thermophilic bacteria, *Thermus aquaticus* have contributed greatly to the yield, specificity, automation, and utility of PCR. This was the choice of DNA polymerase at the onset of PCR (Erlich *et al*,1991). Over the years, several other sources of polymerase enzymes have been identified. The usefulness of these polymerases differs depending on their efficiency, fidelity and cost (Bergseid, *et al.,* 1991). DNA polymerases capable of proofreading the extended sequence and rectifying errors of mismatch between the parent DNA strand and the newly formed strand are currently available (Sutton and Walker 2001).

Polymerase chain reaction entails three basic steps repeated over a number of cycles. The three fundamental steps are denaturation, annealing and extension. At first, the double stranded DNA is denatured at high temperatures (90-97°C) to produce two single-stranded DNA templates. During annealing, the primers complementarily hybridise to the DNA template strands at 50-60°C. During extension, the heat stable Taq DNA polymerase complementarily adds the available dNTPs to the end of the annealed primers at 72°C, thus, synthesising identical copies of the parent DNA double strands. This effectively doubles the DNA quantity through the third steps in the PCR cycle (Joshi and Deshpande, 2010).

As the process of denaturation, annealing, and polymerase extension is continued, the primers repeatedly bind to both the original DNA template and complementary sites in the newly synthesized strands and are extended to produce new copies of DNA. Although, the number of cycles to set for a PCR would depend on the purpose of the investigation, a maximum number of 30 cycles has been said to be advisable (Hill and Stewart, 1992). After the last cycle, samples

are usually incubated at 72°C for 5 minutes to fill in the protruding ends of newly synthesized PCR products. The end result is an exponential increase in the total number of DNA fragments which are finally represented at a theoretical abundance of 2n, where n, is the number of cycles (Gibbs, 1990 and Arnheim and Erlich,1992). Depending on the time specified for each of the three stages, PCR reactions of 30 cycles would be completed in approximately 2 hours (Etebu and Osborn, 2009).

2.8 PRIMER DESIGNING

The success and failure of a PCR reaction is determined to a large extent by the primers used for the reaction (Mullis and Faloona, 1987; Saiki *et al.*, 1988; Wilfred *et al.*, 2005; Etebu, 2008). Several computational methods have been developed for designing PCR primers for a variety of applications. However, the complexity of designing an appropriate primer varies across applications.

In many applications, the DNA sequence is known, and the design of primers is simply the identification of an appropriate segment of the known sequence. The process of designing primer for unknown members of a gene family is much more complex. These primers generally operate on conserved regions. The strategy for designing such primers is to look for conserved regions within known sequences of the gene of interest and then design the primers within those regions (Sibhatu, 2003).

Often amino acid sequences are studied for determining the conserved sequences. When amino acid sequences are used, the biological significance of sequence conservation is more apparent. The conserved sequences are identified using multiple sequence alignment programme (Higgins *et. al.*, 1988). Clustal series of programmes are widely used in molecular biology for the multiple alignments of both nucleic acid and protein sequences (Chenna *et al.*, 2003). Candidate conserved regions can be further evaluated based upon the role they

play in the function of the enzyme. Specific residues may be more or less likely to be conserved based upon whether they are stabilizing a secondary structure, binding a cofactor, or participating in a reaction at the active site. Depending upon the application, the choice of which conserved region to use can be aided with this additional information. Also, the possible codons can be predicted from the amino acid sequence allowing for the design of degenerate primers. If the correct reading frame is known for a DNA sequence, the same analyses can be made after translation of a nucleotide sequence as for an amino acid sequence. However, it is less useful to base the design on solely the nucleotide sequences, because much of the possible diversity will be missed.

Once a conserved amino acid sequence has been established, the issue of degeneracy in the genetic code is still a problem in the design of flexible primers. One strategy to get around the degeneracy issue is to synthesize many different primers for all of the possible amino acid sequences and codon usages. Using too many primers within one PCR reaction creates problems, because the concentration of each individual primer and efficiency decreases. The opposite strategy is to design one primer that has the most common amino acids or nucleotides at each position. This strategy is not appropriate for amplifying unknown sequences, because distantly related sequences will not be amplified. If the gene of interest belongs to a family of genes which share a binding site, or some such feature it is possible that the primer designed will be too general and could amplify sequences of other members of the family besides the gene of interest (Rose *et al.*, 1998).

Primers work with varying efficiencies based upon how similar they are to the target sequence. If the primer matches the target sequence perfectly, it will anneal more strongly and amplify more efficiently than if there are base pair mismatches.

Some measures have been preferred as conditions that should be adhered to in the design and selection of primers. These include avoiding sequences with 3' ends that can self-hybridize or hybridize to the 3' ends of the other primers in the PCR (forming primer dimers) (Abd-Elsalam, 2003). Suggs *et al.* (1981) suggested that the melting temperature (T_m) of both the primers should be about the same. Melting temperature given by $(T_m) = 2\times(A + T) + 4\times(G + C)$; where A, T, G and C represents the four nucleotides with adenine, thymine, guanine and cytosine respectively. The use of BLAST (Basic Local Alignment Search Tool) Analysis is recommended to ascertain potential specificity. Sheffield *et al.* (1989) and Abd-Elsalam, (2003) remarked that preferably the size of the primer should be between 20–24 bases with 50% G:C content, but, more than 3 G or C residues in the 3'-most 5 bases should be avoided. Also, primers with a G as the 3'-terminal base is avoided (Wilfred *et al.*, 2005; Etebu, 2008).

2.9 NESTED PCR

Although PCR reaction is a highly sensitive technique, the specificity, sensitivity and yield of a target DNA sequence can be further improved by nested PCR (Newton and Graham, 1994). In nested PCR an intial set of primers spaced relatively far apart is used for 15 to 30 cycles of amplification. Then an additional primer or pair of primers, located between the outer primers and specifying an internal region of the first amplified DNA product, is added, followed by 15 to 30 cycles (Linssen *et al.*, 2000). Thus, the larger fragment produced by the first round of PCR is used as the template for the second PCR. The enhanced specificity of this technique almost always eliminates any spurious non-specific amplification products. This is because, after the first round of PCR any non-specific products are unlikely to be sufficiently complementary to the nested primers to be able to serve as a template for further amplification, thus the desired target sequence is preferentially amplified.

Nested PCR protocols are useful when specimen contains only very low copy numbers of target molecules or when the PCR equilibrium favours the generation of non-specific amplification products (Sellon, 2003). However, the increased risk of contamination is a drawback of this extreme sensitivity, and great care must be taken when performing such PCRs, particularly in a diagnostic laboratory.

2.10 GENE CLONING

Cloning of PCR sequence is done within a complex DNA population. Unlike PCR, which can only amplify genes with predetermined sequences, gene cloning is not sequence specific. However, it is an inevitable step for performing further analysis on many unstudied genes.

Cloning can be classified into ligase independent and ligase dependent cloning. Ligation-independent cloning (LIC) does not involve *in vitro* ligation. It involves the generation of long (10-12 bases) protruding 3'-ends on PCR amplified DNA fragments which are annealed specifically to complementary DNA sequences of the plasmid vector, and the annealed products are subsequently transformed into competent cells. It may involve various strategies such as *in-vivo* cloning (IVC) (Oliner *et al.*, 1993), Enzymatic modification mediated LIC (Aslanidis and de Jong, 1990) and PCR-induced LIC (Shuldiner *et al.*,1990). The application of the ligase-free procedure is limited by its requirement for extra-long primer synthesis and enzymatic modification steps (Costa *et al.*,1994).

Ligase dependent method is used to covalently link the compatible ends of the DNA fragment and the linearized plasmid using ligase enzyme, thereby forming a single cyclic molecule that is capable of autonomous replication in host cells (Sambrook *et al.*, 1989). The ligase-dependent method may involve various approaches. One approach is cohesive-end ligation, in which cohesive ends on the insert and plasmid are generated by digestion with appropriate restriction

enzymes, and the complementary ends of the plasmid and insert are then joined by DNA ligase (Jung *et al.*,1993). However, for every restriction fragment having distinct cohesive ends, one linearized plasmid vector compatible with that (those) particular restriction site(s) has to be prepared. The other, less efficient approach is blunt-end ligation in which blunt-ended DNA fragments are ligated to a linearized blunt-ended plasmid (Costa and Weiner, 1994). Both of the ligase-dependent methods require multiple purification and/or enzymatic modification steps.

TA cloning is a ligase dependent method which is simple and much more efficient than blunt-ended ligation for the cloning of PCR products. Originally designed to facilitate the cloning of PCR products, it can be easily converted to a universal cloning method (Zhou *et al.*, 1995). All DNA fragments, with the exception of those with protruding 3'-ends, are directly converted to double stranded DNA molecules having 3'-A overhangs by incubation with Taq polymerase in the presence of the four dNTPs (Clark, 1988). These modified inserts are then ligated to T-vectors having a 3'-T overhang at each end (Holton and Graham, 1991). In the case of DNA fragments having protruding 3'-ends, the DNA fragments are first made blunt by incubation with T4 DNA polymerase, and are then treated with Taq polymerase. A cloning efficiency of up to 90% is routinely achieved with this method. Directional cloning is made possible by the appropriate hemi-phosphorylation of both the T-vectors and the inserts.

A number of cloning kits have been developed. StrataClone PCR cloning kit exploits the combined activities of topoisomerase I from *Vaccinia* virus and Cre recombinase from bacteriophage P1. The StrataClone PCR cloning vector mix contains two DNA arms, each charged with topoisomerase I on one end and containing a *loxP* recognition sequence on the other end. The topoisomerase-charged ends have a modified uridine (U*) overhang. *Taq*-amplified PCR products, which contain 3′-adenosine overhangs, are efficiently ligated to these vector arms through A-U* base-pairing followed by topoisomerase I-mediated

strand ligation (Shuman,1994).The resulting linear molecule (vector arm ori–PCR product–vector arm amp/kan) is then transformed, into a competent cell line engineered to transiently express Cre recombinase. Cre-mediated recombination between the vector *loxP* sites creates a circular DNA molecule (pSC-A-amp/kan) that is proficient for replication in cells growing on media containing ampicillin or kanamycin (Abremski *et al.*, 1991). The resulting pSC-A-amp/kan vector product includes a *lacZ'* α-complementation cassette for blue-white screening.

Blue white screening is important because plasmid recircularization can occur with or without integration of the insert DNA. Only a fraction of subsequently transformed bacterial cells acquires a plasmid recombined with the insert DNA. The blue/white screening allows a color discrimination of recombinants from nonrecombinants. (Miller,1992) in the presence of a histochemical dye, 5-bromo-4-chloro-3-indolyl-β-d-galactoside (X-gal) (Sambrook *et al.*,1989). These vectors used for insertion expresses the amino-terminal fragment of the *lacZ* gene product (β-galactosidase) that is capable of intra-allelic complementation (α-complementation) with the carboxy- terminal fragment encoded by appropriate host strains (Ullmann *et al*, 1967). Insertion of foreign DNA into the polycloning site of the plasmid inactivates the amino-terminal fragment of the β-galactosidase and abolishes α-complementation. Thus, cells bearing the recombinant plasmids ($LacZ^-$) would be white in presence of X-gal and can be distinguished from the blue parental vector carrying cells ($LacZ^+$).

2.11 RLM-RACE

Isolation of full-length gene transcripts is important to determine the protein coding region and study gene structure. However, isolation of novel gene sequences is often limited to expressed sequence tags (ESTs) that are short cDNA fragments that predominantly represent the 3′ end of the transcript. Incomplete cDNAs are of limited value, since the genetic information required to make a functional protein is often not present. This problem arises due to the reduced

efficiency with which reverse transcriptase reaches the cap site, especially for mRNAs that present a stable secondary structure, and to the lack of method to select the full-length cDNAs.

Frohman *et al.* (1988) remarked that currently, RACE (rapid amplification of cDNA ends) is the most effective method of cloning cDNAs ends when only partial sequences are available. RACE employs an adaptor with a defined sequence which is attached to one end of the cDNA and the region between the adaptor and the known sequences is amplified by polymerase chain reaction (PCR).

Many labs have developed significant improvements on the basic approach of RACE. Maruyama and Sugano (1994) and Shaefer (1995) reported RNA Ligase Mediated Rapid Amplification of cDNA Ends (RLM-RACE) which represents a major improvement to the classic RACE technique. RLM-RACE is designed to amplify cDNA *only* from full-length, capped mRNA, usually producing a single band after PCR.

To perform 5' RACE, total or poly(A) selected RNA is treated with Calf Intestine Alkaline Phosphatase (CIP) to remove free 5'-phosphates from molecules such as ribosomal RNA, fragmented mRNA, tRNA, and contaminating genomic DNA. The cap structure found on intact 5' ends of mRNA is not affected by CIP. The RNA is then treated with Tobacco Acid Pyrophosphatase (TAP) to remove the cap structure from full-length mRNA, leaving a 5'-monophosphate. A 45 base RNA Adapter oligonucleotide is ligated to the RNA population using T4 RNA ligase. The adapter cannot ligate to dephosphorylated RNA because these molecules lack the 5'-phosphate necessary for ligation. During the ligation reaction, the majority of the full length, decapped mRNA acquires the adapter sequence as its 5' end. A random-primed reverse transcription reaction and nested PCR then amplifies the 5' end of a specific transcript. (Suzuki, 1997; CLONTECHniques, 1996).

During 3' RACE first strand cDNA is synthesized from either total RNA or poly (A)-selected RNA, using a 3' RACE Adapter primer. The cDNA is then subjected to PCR using a 3' RACE Primer which is complimentary to the anchored adapter, and a gene specific primer. The 3' RACE primer utilizes the poly (A) tail region as an initial priming site, therefore, multiple amplification products may be synthesized, depending on the degree of specificity conferred by the gene specific primer. To generate a specific amplification product a second nested PCR may be performed as recommended by Frohman and Martin (1989) to reamplify the RACE products.

2.12 DNA ISOLATION

DNA isolation is an unavoidable procedure for molecular marker based analysis. However, various problems are encountered in the isolation and purification of high molecular weight DNA from certain medicinal and aromatic plant species. These may include the degradation of DNA due to endonucleases, isolation of highly viscous polysaccharides, inhibitor compounds like polyphenols and other secondary metabolites directly or indirectly interfering with enzymatic reactions (Weishing *et al.*, 1995). The presence of polyphenols can also reduce the yield and purity of extracted DNA (Loomis, 1974; Porebski *et al.*, 1997). Due to the inherent differences in primary and secondary metabolite composition different extraction methods are necessary since different plant groups contain diverse secondary compounds which may interfere with the isolation.

Murray and Thomson (1980) developed the CTAB (Cetyl Trimethyl Ammonium Bromide) method for the rapid isolation of high molecular weight plant DNA (50,000 base pairs or more in length) which is free of contaminants. The protocol uses CTAB which is an ionic detergent that lyses the plant cells and forms an insoluble complex with nucleic acids in low salt environments. The polysaccharides, phenols and other contaminants are washed away during further

steps. Raising the salt concentration solubilises the DNA and it can then be precipitated using ethanol or isopropanol. The procedure yields total cellular DNA and is ideal for the rapid isolation of small amounts of DNA from many different species and is also useful for large scale isolations (Somma, 2004).

2.13 ISSR

A DNA marker is a sequence of DNA or a gene, which is situated on a chromosome. Markers help to detect differences between individuals by displaying polymorphism, thereby helping to detect genetic differences between individual organisms or species (Collard *et al.,* 2005). DNA based molecular markers have been successful in overcoming the limitations associated with conventional markers (Weising *et al.*, 1995).

A wide range of molecular markers have been used to assess genetic diversity. These include Random Amplified Polymorphic DNA (RAPD), DNA amplification fingerprinting (DAF), Inter-simple sequence repeat (ISSR), Amplified Fragment Length Polymorphisms (AFLP), Restriction Fragment Length Polymorphisms (RFLP), Microsatellites (SSR), Sequence-characterised amplified region (SCAR), Single-Nucleotide Polymorphism (SNP) etc. (de Vicente and Fulton, 2003).

Among the different molecular markers available, Inter-Simple Sequence Repeat (ISSR) markers are dominant in nature. These make use of specific microsatellite sequence anchored either at the 3' or 5' end to amplify scorable DNA bands (Gupta *et al.*, 1994; Zeitkiewicz *et al.*, 1994). The ISSR based PCR employs longer primers which are characterised by higher annealing temperatures than Random Amplified Polymorphic DNA (RAPD) markers (Pomper *et al.* 2003), are characterised by more stringency. The ISSR markers produce high levels of polymorphism and are reproducible. These markers do not require prior target sequence knowledge and can amplify even crude DNA preparations

available in nanogram amounts. These advantages have made ISSR the alternate tool for diversity studies in many plant systems (Tsumura *et al.*, 1996 and Martin and Sanchez, 2000).

MATERIALS & METHODS

3. MATERIALS AND METHODS

The study entitled "Isolation and characterization of type III polyketide synthases from chethikoduveli (*Plumbago rosea* Linn.)" was conducted at the Biotechnology and Bioinformatics Division, Jawaharlal Nehru Tropical Botanic Garden and Research Institute (JNTBGRI), Palode, Thiruvananthapuram, during 2013-2014. Details regarding the experimental materials used and methodology adopted for the study are presented in this chapter.

3.1 ISOLATION AND CHARACTERIZATION OF TYPE III POLYKETIDE SYNTHASE GENE(S)

3.1.1 Plant Material

Total RNA was isolated from the *Plumbago rosea* accession TBG-102 that was procured from JNTBGRI, Palode, Thiruvananthapuram. Unexpanded leaves of field grown *P. rosea* were freshly harvested in liquid nitrogen immediately prior to RNA isolation. Flash freezing the sample helps to prevent RNA degradation by endogenous RNases.

3.1.2 RNA Isolation

Ribonucleases are very active, widespread and stable enzymes that are capable of causing damage and loss of RNA. Though RNase can be temporarily denatured under extreme conditions, it readily renatures. Nevertheless, precautions were taken to perform RNA isolation under RNase free condition. All the non-disposable labware including, micropipette tips, mortar, pestle and spatula were made RNase free by treating with 0.1 % diethyl pyrocarbonate (DEPC), a strong inhibitor of RNase. This was performed by soaking the labware in sufficient quantity of 0.1 % DEPC prepared in distilled water. The labware were then incubated overnight at dark in DEPC water in an isolated area. After

incubation, the DEPC water was carefully poured off. The labware were separately wrapped in aluminium foil and baked for 5 h at 180 °C to inactivate DEPC. RNase-free disposable plastic wares such as microfuge tubes that do not require any pre-treatments were also used.

Water used for RNA isolation and further analysis was also DEPC treated and autoclaved prior to use. Tris Borate EDTA (TBE buffer) (Appendix I) used for the preparation of agarose gel and tank buffer was prepared in DEPC treated water. All the chemicals used in RNA isolation were reserved for RNA application and kept separated from routinely used chemicals. A separate work area was set apart solely for RNA isolation and the work area was wiped with ethanol prior to the commencement of the isolation. Also, gloves were worn throughout the course of work in order to minimise exogenous RNase contamination and these were intermittently wiped with alcohol.

Two protocols were tried for the isolation of total RNA from *P. rosea*. TRIzol® reagent (Ambion, Life Technologies, USA) and Spectrum Plant Total RNA Kit (Sigma Aldrich, USA) were checked. Finally, based on the quality and quantity of RNA obtained, subsequent RNA isolations were performed using Spectrum Plant Total RNA Kit. RNA was isolated from 100 mg flash frozen leaf sample as per the directions in the manual. The final product was stored at -20 °C.

3.1.3 Agarose Gel Electrophoresis

The integrity and quality of the total RNA was determined by running 3 µl aliquot of RNA on 2 % agarose gel. Electrophoresis was carried out at 110 V for one hour in 1X TBE (Tris Borate EDTA) buffer. The gel was stained using ethidium bromide and analysed using gel documentation system (UVP EC3 Chemi HR 410 Imaging System, USA) using VisionWorksLS Software.

3.1.4 Spectrophotometric Analysis

The concentration of RNA was determined spectrophotometrically using Biophotometer (Eppendorf, Germany).

3.1.5 RT-PCR

3.1.5.1 First Strand cDNA Synthesis

Since RNA is unstable to work with, it was converted to cDNA by using AccuScript High Fidelity 1st Strand cDNA Synthesis Kit (Agilent, USA). High-quality intact RNA is essential for successful synthesis of full-length cDNA and yield of long RT-PCR products. Therefore, precautions were taken to minimize the potential for contamination by ribonucleases (RNases). RT-PCR was performed under RNase-free conditions wearing gloves and using sterile microcentrifuge tubes, pipette tips and RNase-free water.

First strand cDNA synthesis was performed as per the directions in the kit manual. 6µl of the isolated RNA was used for the reaction.

3.1.5.2 β-Actin Confirmation Step

The fidelity of cDNA synthesised was confirmed by amplifying the first strand cDNA synthesis product with β-actin primer to detect the presence of cDNA corresponding to β-actin gene which is constitutively expressed in all eukaryotes.

PCR reaction components and constitution

Components	Concentration	Required concentration	Volume taken (µl)
Buffer	10 X	1 X	2.5
dNTP	25 m*M*	200 µ*M*	0.2
Forward Primer	33 p*M*/µl	15 p*M*	0.5
Reverse Primer	30 p*M*/µl	15 p*M*	0.5
DNA Polymerase	2.0 Unit	1.0 Unit	0.5
cDNA			1.0
Water			1.0

The 25 µl reaction mixture contained the above mentioned constituents. The PCR (Eppendorf, Germany) programme was set with an initial denaturation at 94 °C for 2 min followed by 35 subsequent cycles with denaturation at 94 °C for 30 s, annealing at 55 °C for 30 s and extension at 72 °C for one min. This was followed by final extension at 72 °C for 7 min and the reaction was held at 4 °C for infinity.

The amplified products were resolved in 1.4 % agarose gel alongside 100bp ladder for identifying the size of the amplified product. The gel was viewed using gel documentation system.

3.1.6 Nested PCR

The cDNA product contains a heterogenous population of cDNA produced from total RNA. To synthesise cDNA that is specific for PKS gene, the cDNA product of first strand synthesis was subjected to nested PCR with gene specific

primers for PKS (Polyketide Synthase) (Abe *et al.*, 2005) namely, IKRU-1 and IKRU-2. Gradient PCR was set up to determine the annealing temperature of the primers.

3.1.6.1 Initial PCR

Initial PCR was carried out with IKRU1 primer. The 25µl reaction mixture contained the following constituents:

PCR reaction components and constitution

Components	Concentration	Required concentration	Volume taken (µl)
Buffer	10 X	1 X	2.5
dNTP	25 m*M*	200 µ*M*	0.2
IKRU-1Forward Primer	9.354 p*M*/µl	15 p*M*	1.6
IKRU-1Reverse Primer	8.127 p*M*/µl	15 p*M*	1.8
DNA Polymerase	2.0 Unit	1.0 Unit	0.5
cDNA			3.0
Water			15.4

The PCR programme was set with an initial denaturation at 94 °C for 2 min followed by 35 subsequent cycles with denaturation at 94 °C for 30 s, annealing at 42 °C or 58 °C for 30 s and extension at 72 °C for 1min. This was followed by final extension at 72 °C for 7 min and the reaction was held at 4 °C for infinity.

3.1.6.2 Nested PCR Reaction

Product of the PCR was subjected to nested PCR using primer IKRU-2.The PCR reaction mix contained the following constituents:

PCR reaction components and constitution

Components	Concentration	Required concentration	Volume taken (µl)
Buffer	10 X	1 X	2.5
dNTP	25 m*M*	200 µ*M*	0.2
IKRU-2 Forward Primer	8.04 p*M*/µl	15 p*M*	1.9
IKRU-2 Reverse Primer	6.69 p*M*/µl	15 p*M*	2.2
DNA Polymerase	2.0 Unit	1.0 Unit	0.5
cDNA			2.0
Water			15.7

The PCR programme was set with an initial denaturation at 94 °C for 2 min followed by 35 subsequent cycles with denaturation at 94 °C for 30 s, annealing at 42 °C for 30 s and extension at 72 °C for one min. This was followed by final extension at 72 °C for 7 min and the reaction was held at 4 °C for infinity.

The amplified products were resolved in 1.4 % agarose gel alongside a 100bp ladder to assess the amplification of appropriate size product.

3.1.7 Cloning

The nested PCR product was subjected to clean up with Gen Elute PCR Clean up kit (Sigma-Aldrich, USA). This was followed by the cloning of the clean up product using Strata Clone PCR Cloning kit (Agilent Technologies, USA). For

cloning, the PCR product was first ligated to the vector and the competent cells were transformed with this vector. Finally, the transformants were selected based on blue white screening. Plasmid isolation was performed using the desired recombinants and the PCR amplified insert was then sequenced.

3.1.7.1 Clean Up

The clean up was performed to purify the PCR product from other components in the reaction such as primers, nucleotides, DNA polymerase and salts. The DNA was finally eluted in 25 µl of elution solution. The eluate was analysed by performing agarose gel electrophoresis of 5 µl of sample in 1.4 % gel and by visualising in gel documentation system.

3.1.7.2 Ligation into Vector, pSC-A-amp/kan

The eluted PCR product contains 3' overhang which helps to ligate the single-stranded A-overhangs with the 5'-modified U (U*) overhangs of the StrataClone vector arms. The volume of PCR product needed for optimum ligation to the vector was determined using the following formula:

$$\text{ng of insert} = \frac{\text{ng of vector} \times \text{Kb size of insert}}{\text{Kb size of vector}} \times \frac{\text{Insert ratio}}{\text{Vector ratio}}$$

Different insert to vector ratios (1:1, 3:1, 5:1) were tried to determine the cloning efficiency.

Ligation reaction constitutents

Item	Volume (µl)
Fresh PCR product	1.35
Strata Clone Cloning Buffer	3.00
Strata Clone Vector Mix	1.00
Total volume	5.35

The contents were mixed by repeated pipetting and incubating the reaction mix at room temperature for 7 min. When incubation was completed, the tube was kept on ice.

3.1.7.3 Transformation of Competent Cells

Prior to transformation, L.B. agar plates were prepared by adding agar in the concentration 15 g/l into Luria Bertani (L.B.) broth (Appendix II). After autoclaving, the medium was brought to 50 °C and 100 µl of 50 mg/ml filter sterilized ampicillin was added to 100 ml medium (50 µg/ml of ampicillin). The contents were then dispensed in to sterile Petri plates.

For each transformation one 50 µl vial of Strata Clone Solo Pack containing competent cell was thawed on ice. The cloning reaction mixture (1 µl) was pipetted directly into the thawed Strata clone Solo Pack vial. The mixture was gently mixed by tapping and was incubated on ice for 20 min. Thereafter, the mixture was subjected to heat shock transformation at 42 °C for 90 s. The transformation mixture was again incubated on ice for 2 min. 250 µl of SOC (Super Optimal Broth with Catabolic Repressor) (Appendix III), pre warmed to 42 °C, was added to the vial. The competent cells were allowed to recover at 37 °C for 1 h with agitation at 200 rpm in a shaking incubator. 50 µl, 100 µl and 150 µl of the transformation mixture was plated on separate LB-ampicillin plates each of which had been spread with 40µl of 2 % X-gal (dissolved in dimethylformamide) and containing 50 µg/ml of ampicillin. The plates were incubated overnight at 37 °C. Next day, the plates were shifted to +4°C to prevent overgrowth of the colonies.

3.1.7.4 Selection of Transformants

The transformants were selected by blue white screening. 15 white colonies were selected from 100 µl plate and 15 colonies from 150 µl plates. The selected recombinant white colonies were cultured overnight in 5 ml LB broth containing 50 µg/ml ampicillin for plasmid isolation. Plasmids were isolated using alkali lysis method modified by Birnboim and Doly (1979). The isolated plasmid was dissolved in 20 µl Tris EDTA by gentle tapping.

The isolated plasmid DNA (1µl) was loaded on 0.8 % gel for analysing the quality and the quantity of the same. After staining with EtBr the gel was visualised using UVP gel documentation system.

3.1.7.5 Long Term Storage

A replica of the colonies used for plasmid isolation was maintained for future use by spreading 1 µl of the broth in which the colonies were cultured, on a LB plate containing 50 µg/ml ampicillin. The plate was incubated at 37 °C overnight to facilitate the growth of the colonies. On sufficient growth, the plate was stored at 4 °C for future use.

3.1.7.6 Confirmation of insertion

The presence of insert was confirmed by amplifying the isolated plasmid with T7 and T3 primers specific for the flanking sequences of the vector on either sides of the insert.

Plasmid samples were used for PCR in the following quantities as per the intensity observed in the plasmid gel picture.

Plasmid quantity used for cloning

Plasmid Quantity Used	Intensity Observed
0.5µl	good intensity bands
1.0µl	low intensity bands
1.5µl	very low intensity bands

PCR reaction components and constitution

Components	Concentration	Required concentration	Volume taken (µl)
Buffer	10 X	1 X	2.5
dNTP	25 m*M*	200 µ*M*	0.2
T3 Forward Primer	10 p*M*/ul	15 p*M*	1.5
T7 Reverse Primer	10 p*M*/ul	15 p*M*	1.5
DNA Polymerase	2.0 Unit	1.0 Unit	0.5
Plasmid DNA			0.5-1.5
Water			Make volume upto 25µl

The PCR products were loaded on 1.4 % gel along with 100bp DNA ladder and visualised after staining with ethidium bromide.

3.1.7.7 Restriction digestion

The recombinant plasmids showing variation in insert size were selected by restriction digesting the amplified plasmid samples separately with Eco RI and NotI. First the plasmid DNA was 10 times diluted and amplified by PCR as in 3.1.7.6. The plasmids showing amplification were chosen for restriction digestion. The reaction mixture for restriction digestion contained the following constituents.

Restriction digestion components and constitution

Constituents	Quantity (µl)
Sterile deionised water	7.3
10 X Restriction Enzyme Buffer	2.0
10 µg/µl Acetylated BSA	0.2
Amplified PCR product	10.0

The constituents were mixed by pipetting and 0.5 µl of 10 µg/µl of respective restriction enzyme was added and the reaction mixture was mixed again by gentle pipetting. The tubes were centrifuged for a few seconds to settle the contents to the bottom of the tubes and the tubes were incubated at 37 °C for 4 h. The restriction digestion products were loaded onto 1.4 % agarose gel and visualised after staining with ethidium bromide.

3.1.8. Sequencing Clones

Six plasmids were selected for sequencing based on variation in restriction digestion products. Sequencing was performed using Sanger's method at SciGenom, Ernakulam. Sequencing was performed from 3' and 5' end of the inserts using the T3 and T7 primers respectively.

3.1.9. Data Analysis

The sequence data was analysed using ABI sequence viewer to view the quality value (QV) of the obtained sequence. The obtained sequences were used to generate consensus sequence using BioEdit software. BLASTX (Altschul *et al.*, 1990) was performed using the consensus sequence to study similarity with already reported PKS sequences. ExPASy (Expert Protein Analysis System) (Gasteiger *et al.*, 2003) was used to generate the amino acid sequences corresponding to the consensus sequences obtained. Four of the amino acid sequences thus generated were subjected to Multiple Sequence Alignment (MSA) with amino acid sequences of CHS (Chalcone Synthase) reported from other plants. Based on the MSA data the extent of the PKS gene to be sequenced in the 3' and 5' direction was determined. Also primer designing was performed based on the conserved regions observed in the MSA carried out with amino acid sequences. The nucleotide sequences corresponding to the conserved amino acid regions were used to design RLM-RACE (RNA Ligase Mediated Rapid Amplification of cDNA Ends) primers.

3.1.10 Primer designing

Primer designing was performed manually by using the nucleotide sequences that corresponded to the amino acid regions showing similarity in MSA. Two sets of gene specific primers, each consisting of an outer primer and an inner primer were designed for 3' RLM-RACE and 5' RLM-RACE.

Primer designing was performed based on the primer designing suggestions mentioned in First Choice RLM-RACE Kit (Ambion, USA). The recommendations that were followed for primer designing include:

- Primer size: 20–27 bases
- G:C content: 50 per cent, with no secondary structure
- Placing not more than 3 G or C residues in the 3'-most 5 bases

• Avoid sequences with 3' ends that can self-hybridize or hybridize to the 3' ends of the other primer in the PCR (forming primer dimers)

• Finally, using primer design software, evaluate your gene-specific primers in combination with the corresponding RACE Primer.

These primers were synthesized and supplied by Sigma Aldrich.

3.1.11 RLM-RACE

To obtain the full length sequence of the cDNA, 5' and 3' RLM-RACE was performed using First Choice RLM-RACE Kit (Ambion, USA).

3.1.11.1 5' RLM-RACE

3.1.11.1.1 RNA Processing

RNA processing is performed to ligate an adapter at the 5' end of full length decapped mRNA. RNA processing was performed using 14.5 µl (10 µg/µl) of RNA as per the directions in the manual. The processed RNA was resuspended in 10 µl nuclease free water.

3.1.11.1.2 Reverse Transcription

The mRNA was reverse transcribed according to the kit manual. 2 µl ligated RNA was used for the reaction.

3.1.11.1.3 5' Outer RLM-RACE

RT reaction product (1 µl) and 2 µl PKS 2N primer was used for performing 5' outer RLM-RACE as per kit instructions. The following PCR programme was used for the reaction.

PCR programme for outer RLM-RACE

	Stage	Repeats	Temperature	Time
Initial denaturation	1	1	94 °C	3 min
Amplification	2	35	94 °C	30 s
			55-65 °C	30 s
			72 °C	30 s
Final extension	3	1	72 °C	7 min

15 µl of the amplified product was loaded along with a 100 bp ladder in 1.4 % agarose gel and the product with maximum amplification at suitable annealing temperature was used for the inner 5' RLM-RACE.

3.1.11.1.4 Inner 5' RLM-RACE PCR

1µl outer PCR product amplified at annealing temperature 55 °C was used for performing the inner 5' RLM-RACE as per the kit. This reaction used 2 µl of the previously designed 5' RACE gene specific inner primer (PKS 2). A gradient PCR was performed using the same programme used for outer 5' RACE with annealing temperatures ranging from 55- 65°C.

15µl amplified products were resolved in 1.4 % agarose gel to ensure the quality of the amplified product at optimum annealing temperature.

3.1.12 Gel Elution

For eluting the desired band for further downstream reactions, the inner 5' RLM-RACE with PKS 2 inner gene specific primer using the product of

3.1.11.1.3 was repeated with duplicate with all conditions as per section 3.1.11.1.4 and annealing temperature set at 60 °C.

The product of both the duplicates were pooled and PCR clean up was performed with Gen Elute Clean up kit (Sigma, USA) as in section 3.1.7.1 and 5 µl of the clean up product was loaded on 1.4 % agarose gel.

The 400 bp PCR product was eluted from the gel using Sigma gel elution kit. The final product (25 µl) was recovered and stored at -20 °C for further reactions.

3.1.13 Cloning of 5' RLM-RACE Product and Analysing Transformants

The quantity and quality of the recovered gel elution product was analysed by loading 5 µl of the same on 1.4 % agarose gel. Based on the intensity of the product the quantity of the product was estimated to be around 30 ng/µl. Therefore, 1.2 µl of the reaction product was used to clone the specific PCR product using Strata Clone PCR Cloning kit (Agilent Technologies, USA) as per section 3.1.7.

Twenty two transformed colonies were subcultured overnight as in section 3.1.7.3 in 5 ml of LB broth containing 50 µg/ml ampicillin for plasmid isolation. The plasmid was isolated as per 3.1.7.4 section and visualised in 1 % agarose gel. Based on the presence of plasmid, the required quantity was diluted and PCR amplification was performed for 19 plasmids with T7 and T3 primers as in section 3.1.7.6. The amplification products were subjected to PCR clean up as in section 3.1.7.1 and 12 samples which showed single band with sufficient intensity were selected for sequencing at SciGenom, Ernakulam.

3.1.14 Data Analysis

The sequences were analysed using ABI software. BLASTX and MSA was performed as in section 3.1.9.

3.1.15 3' RLM-RACE

Following first strand cDNA synthesis, 3' cDNA end was isolated by using 3' RLM-RACE PCR using the primer PKS 2 RACE-n1that was previously designed.

3.1.15.1 RT PCR

Reverse transcription was performed as per the instructions in First Choice RLM-RACE Kit from freshly isolated total RNA. A quantity of 2.5 µl RNA was used for the reaction. The RT-PCR product was then used for outer 3' RLM-RACE PCR.

3.1.15.2 Outer 3' RLM-RACE PCR

The reaction was performed according to the instructions in the manual. RT reaction product (1 µl) and 2 µl of previously designed 3' RACE gene-specific outer primer (PKS 2 RACE-n1) was used in the outer reaction.

A gradient PCR programme was set with an initial denaturation at 94 °C for 3 min followed by 35 subsequent cycles with denaturation at 94 °C for 30 s, annealing for 30 s at a range of temperatures from 55 °C to 57 °C and extension at 72 °C for 1min. This was followed by final extension at 72 °C for 7 min and the reaction was held at 4 °C for infinity.

The amplified product (15 µl) was loaded along with a 100 bp ladder in 1.4 % agarose gel and based on the product size the requirement to perform inner 3' RLM-RACE was determined.

3.1.16 Gel Elution

For eluting the desired band for further downstream reactions, the outer 3' RLM-RACE with PKS 2 RACE-n1 gene specific primer was repeated with duplicate with all conditions as per section 3.1.15.2 and annealing temperature set at 57 °C.

The product of both the duplicates were pooled and PCR clean up was performed with Gen Elute Clean up kit (Sigma, USA) as in section 3.1.7.1 and the clean up product was loaded on 1.4 % gel and run for 90 min .

The 700 bp PCR product was eluted from the gel using Sigma gel elution kit. A volume of 30 µl of the final product was recovered and stored at -20 °C for further reactions.

3.1.17 Cloning of 3' RLM-RACE Product and Analysing Transformants

The quantity and quality of the recovered gel elution product was analysed by loading 5 µl of the same on 1.4 % agarose gel. Based on the intensity of the product the quantity of the product was estimated to be around 6 ng/µl. Therefore, 2 µl of the reaction product was used to clone the specific PCR product using Strata Clone PCR Cloning kit (Agilent Technologies, USA) as per section 3.1.7.

20 transformed colonies were subcultured overnight as in section 3.1.7.3 in 5 ml of LB broth containing 50 µg/ml ampicillin for plasmid isolation. The plasmid was isolated as per 3.1.7.3 section and visualised in 1 % agarose gel. Based on the observed intensity of the isolated plasmid, required quantity was

diluted and PCR amplification was performed with T7 and T3 primers as in section 3.1.7.6. The amplification products were subjected to PCR clean up as in section 3.1.7.1 and 7 samples which showed single band with sufficient intensity were selected for sequencing at SciGenom, Ernakulam.

3.1.18 Data Analysis

The sequences were analysed using ABI software. BLASTX and MSA were performed as in section 3.1.9. The corresponding amino acid sequence was deduced using ExPASy and the amino acid sequence was merged with the amino acid sequence previously obtained by merging the core fragment and 5' end fragment. ProtParam (Gasteiger *et al.*, 2005) was used to analyse the molecular weight (MW) and pI (isoelectric point) of the protein, PyMOL (The PyMOL Molecular Graphics System) was used for visualising the 3D model of the protein and a phylogenetic tree was constructed using MEGA software (Tamura *et al.*, 2007). CD-Search (Marchler-Bauer, 2011) in NCBI was used for mapping conserved domains.

3.2 MOLECULAR MARKER ANALYSIS

3.2.1 Plant Material

Eight accessions of *Plumbago rosea* were assessed for genetic diversity using ISSR marker. Approximately 10 g of leaf samples were collected from each field grown *P. rosea* plants. The leaf samples were preserved in -80 °C for long term storage prior to DNA isolation.

3.2.2 Genomic DNA Isolation and ISSR

Three different protocols were tried for the isolation of DNA from *P. rosea*. First, one day (Suman *et al.*, 1999) protocol using cetyltrimethylammonium

bromide (CTAB) extraction buffer was employed. Isolation of DNA with Gen Elute Plant Genomic DNA Miniprep kit was also checked. Finally, based on the quality of the DNA obtained, Murray and Thompson (1980) method was followed for the isolation of total genomic DNA. The isolated DNA was resuspended in 100µl 1X TE and stored at -20 °C for further use.

3.2.3 Quantification of DNA

The quantity and quality of the isolated DNA was assessed using biophotometer (Eppendorf, Germany). The DNA sample was 50 times diluted by adding 1 µl of isolated DNA sample to 49 µl of 1X TE. The absorbance of the sample was read at both 260 nm and 280 nm and the ratio was calculated. An absorbance (A_{260}) of 1.0 corresponds to 50 µg DNA/ml of the solution.

$$\text{Concentration } (\mu g/\mu l) = \frac{A_{260} \times \text{Dilution Factor} \times 50}{1000}$$

3.2.4 Agarose Gel Electrophoresis

The integrity and purity of the isolated DNA was analysed by subjecting it to agarose gel electrophoresis. 1 % agarose gel containing 0.5 µg/ml ethidium bromide was prepared in 1X TBE. The wells were loaded with 3 µl of the samples with 3 µl of bromophenol blue. The gel was run at 110 V for 1 h in 1 X TBE tank buffer. The gel was visualised using gel documentation system (UVP EC3 Chemi HR 410 Imaging System, USA) using VisionWorksLS Software.

3.2.5 ISSR Analysis

The isolated genomic DNA samples were amplified using 12 ISSR primers supplied by IDT (Integrated DNA Technology, USA). The 25 µl reaction mixture contained the following constituents.

PCR reaction components and constitution

Components	Concentration	Required concentration	Volume taken (µl)
Buffer	10 X	1 X	2.5
dNTP	25 m*M*	200 µ*M*	0.2
Primer		15 p*M*	1.0
DNA Polymerase	2.0 Unit	1.0 Unit	0.5
Genomic DNA			1.0-3.5
Water			Make volume upto 25.0µl

PCR programme (Applied Biosysytems) was set with an initial cycle of denaturation at 94 °C for 2 min, annealing at 50-53 °C for 2 min and extension at 72 °C for 2 min. This was followed by 34 cycles of denaturation at 94 °C for 30 s, annealing at 50-53 °C for 1 min and extension at 72 °C for 2 min. Final extension was performed at 72 °C for 7 min.

The amplified products were resolved in 1.4 per cent agarose gel (1X TBE) and visualised using gel documentation system.

3.2.6 Genetic Data Analysis

The PCR products were scored for the presence (+) or absence (-) of bands. The binary matrix was analysed using WINBOOT software to generate a dendrogram (Yao and Nelson, 1996). Similarity matrix was generated with the same input file using WINDIST software. The genotype and allelic frequency data were used to compute the genetic diversity indices such as, Percentage of

polymorphic loci (P), observed number of alleles (na), expected number of alleles (ne), Shannon index of genetic diversity (I) and Nei's gene diversity (h) at the population level using the statistical package POPGENE 1.3 (Yeh *et al* .1997).

3.3 PLUMBAGIN EXTRACTION AND ESTIMATION

Plumbagin extraction and estimation were performed using root samples of nine *P. rosea* accessions used for the studies in section 3.1 and 3.2.

3.3.1 Plant Material Extraction

All extraction procedures were performed with powdered dried root sample. Prior to extraction, the root samples were shade dried for five days to remove moisture and the dried roots were powdered in a grinder to reduce the particle size.

One gram of powdered plant sample was extracted with 20 ml chloroform with intermittent shaking at every 15 min. The solvent was changed every 30 min. This cycle was repeated six times until the loss of yellow coloration of the solvent indicating dissolved plumbagin. The used solvent was pooled to obtain 120ml of crude extract. The extract was concentrated to 10 ml using rotavapor (Heidolf, Germany).

3.3.2 TLC

The extracts of plants that had been used for the study undertaken in section 3.2. was used to perform TLC to determine the R_f value of plumbagin and to study purity. TLC was performed on a 20cm x 20cm piece of aluminium backed TLC Silica gel 60-F_{254} plate (MERCK, Germany). The TLC plate was spotted with 10 µl of plumbagin standard (prepared by dissolving 1mg plumbagin powder (Sigma Aldrich, USA) in 20µl of chloroform) and 100µl of each

plumbagin extract. Thereafter, the plate was left undisturbed to dry. The plate was then developed using benzene.

The R_f (retardation factor) value was calculated as per the formula:

$$R_f \text{ value} = \frac{\text{distance travelled by substance}}{\text{distance travelled by solvent front}}$$

The identity of the compound was confirmed by the R_f value obtained by co-chromatography with the authentic sample.

3.3.3 UV-VIS Spectrophotometry

The plumbagin content in the crude extracts used for studies 3.1 and 3.2 were estimated by spectrophotometry at 415 nm using XP 3001 Xplorer, UK. A caliberation curve was established from an authentic sample at concentration range of 10-35 μg/ml. The absorbance was plotted against the concentration of plumbagin and the concentration of unknown solution was computed from the calibration graph using the software UV Win 5 software v 5.0.5.

3.3.4 Preparative TLC

The plumbagin content in the sample used for the study in section 3.1 was confirmed using HPLC as the method is more accurate. To purify the crude sample for HPLC, preparative TLC was performed on glass backed silica plate. The silica plate was prepared by coating clean glass slides with silica gel-G-60 slurry (5 g silica gel-G-60 (Hi Media, India) powder for TLC was mixed with 10 ml distilled water and was used for coating 3 glass slides with an approximate thickness of 1.0 mm). The plates were allowed to dry and were activated for 1 h in hot air oven at 60 °C. The crude extract (200 μl) to be purified was spotted in a line across the plate. The plate was then developed in benzene. The band

containing plumbagin was scraped out from the TLC plate and eluted into 2ml chloroform.

3.3.5 HPLC

Plumbagin content in the sample TBG-102 was determined by HPLC using semi-preparative Gilson 321 series (Gilson, France) following the protocol developed by (Crouch *et al.*, 1990). Plumbagin was separated on a reversed phase using Kromacil C-18 column [250 x 4.6 mm, 5 µm particle size, (VDS Optilab Chromatographietechnik GmbH, Wiesenweg, Berlin)]. Column temperature was maintained at 25 °C. The mobile phase was run at 1.0 ml/min flow rate using methanol: acetic acid (80:20 v/v) adjusted to pH 3.5 with triethylamine. Plumbagin was detected at 270 nm with UV-VIS 156 detector. The injection volume for the sample was 10 µl.

Quantitative analysis was performed by comparing the peak area with the authentic sample of plumbagin from Sigma Aldrich, USA using Unipoint TM LC system software version 5.1 (Gilson Inc. France). The retention time of plumbagin was also noted.

RESULTS

4. RESULTS

The results of the study entitled "Isolation and characterization of type III polyketide synthases from chethikoduveli (*Plumbago rosea* Linn.)" carried out at the Biotechnology and Bioinformatics Division of JNTBGRI, Palode during 2013-2014 are presented in this chapter.

4.1 ISOLATION AND CHARACTERIZATION OF TYPE III POLYKETIDE SYNTHASE GENE(S)

4.1.1 RNA Isolation

Total RNA was isolated from flash frozen unexpanded leaves of *P. rosea* accession TBG-102. Spectrum Plant Total RNA Kit yielded the best result as compared to TRIzol® reagent method. On 2 % agarose gel electrophoresis, the isolated RNA was visualized as two intact bands, corresponding to 28S rRNA (ribosomal RNA), 18S rRNA. No apparent RNA degradation or genomic DNA contamination were observed, showing good quality RNA. The intensity of 28S rRNA band was approximately double the intensity of 18S rRNA band (Plate4.A). This protocol yielded total RNA with a concentration of 1194.3 µg/ml and an absorbance of 1.55 at $A_{260/280}$.

4.1.2 RT-PCR

The reverse transcription of the extracted RNA was carried out for synthesizing the first strand cDNA using oligo dT primer (Table 2). This was followed by nested PCR with degenerate primers IKRU-1 and IKRU-2 (Table 2).

First strand cDNA synthesis yielded cDNA with a concentration of 856.9 µg/ml and $A_{260/280}$ of 1.72. The fidelity of the cDNA was ensured using PCR amplification with β-actin primers (Table 2) which yielded an amplicon of expected size, approximately 200bp (Plate 4.B). This confirmed that the cDNA which forms the template for the amplification is intact.

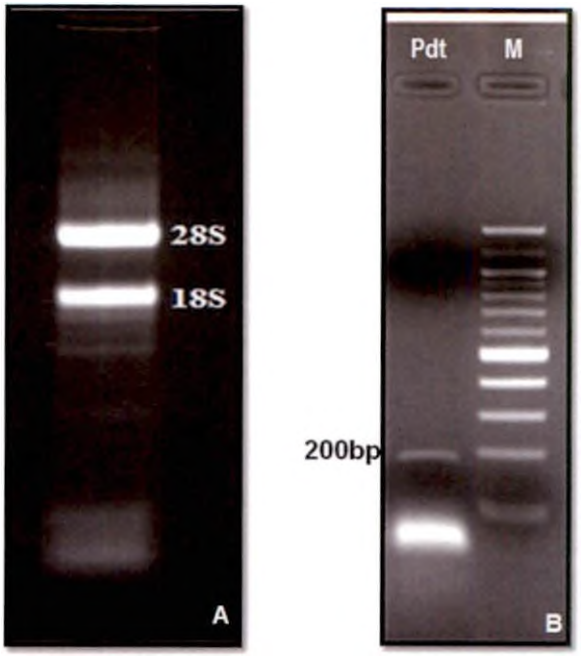

Plate 4. ***P.rosea*** **RNA isolation: (A) RNA (B) cDNA synthesised from RNA using β actin primer**

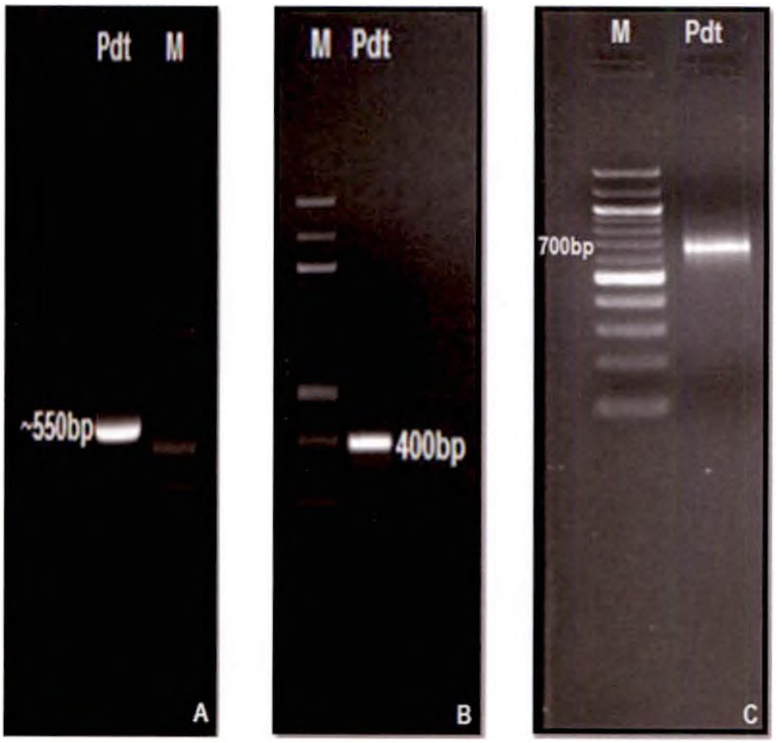

Plate 5. PCR amplicons obtained using: (A) IKRU primers (B) 5' RLM-RACE primers (C) 3' RLM-RACE primers

Plate Description: Pdt- Product **M- Marker**

Table 2: Details of primers used for isolation and characterization of *Plumbago rosea* type III polyketide synthase gene(s)

Sr No	PCR Product	Primer Code	Primer Description	Sequence (5'→3')	Template	~ Amplicon Size (kb)	Annealing Temperature (°C)
1.	β-actin product	βA-F	β actin forward primer	CCA GGC TGT TCA GTC TCT GTA	First strand cDNA	0.20	55
2.		βA-R	β actin reverse primer	CGC TCG GTA AGG ATC TTC ATC			
3.	Insert confirmation products	$T3\text{-}P_F$	T3 forward primer	GCA ATT AAC CCT CAC TAA AGG	Plasmids with insert	Core: 0.75	60
						5': 0.60	
4.		$T7\text{-}P_R$	T7 reverse primer	TAA TAC GAC TCA CTA TAG GG		3': 0.90	

Table 2: continued

Sr No:	PCR Product	Primer Code	Primer Description	Sequence (5'→3')	Template	~ Amplicon Size (kb)	Annealing Temperature (°C)
5.	PKS Partial cDNA	Oligo dT	Oligo dT	Oligo dT_{18}	RNA	0.55	42
6.		IKRU-F1	IKRU1 forward primer	RAR GCI ITI MAR GAR TGG GGI CA	First strand cDNA		
7.		IKRU-R1	IKRU1 reverse primer	TCI AYI GTI ARI CCI GGC CRA A			
8.		IKRU-F2	IKRU2 forward primer	GCI AAR GAY ITI GCI GAR AAY AA	Product of IKRU1		
9.		IKRU-R2	IKRU2 reverse primer	CCC MWI TCI ARI CCI TCI CCI GTI			

Table 2: continued

Sr No:	PCR Product	Primer Code	Primer Description	Sequence (5'→3')	Template	~ Amplicon Size (kb)	Annealing Temperature (°C)
10.	5' RLM-RACE product	5' Ad outer	5'adapter specific outer primer	GCTGATGGCGATGAATGAACACTG	Processed RNA	0.40	55
11.		PKS 2N*	5'PKS gene specific outer primer	GGAACTCACTCTTCTGGATAGCGC ACC			
12.		5' Ad inner	5' adapter specific inner primer	CGCGGATCCGAACACTGCGTTTGC TGGCTTTGATG	Product of Outer 5' RLM-RACE		60
13.		PKS 2*	5'PKS gene specific inner primer	AGAGAAGATGAAGGCTACCAGGC AGG			
14.	3' RLM-RACE product	3'Ad outer	3'adapter specific outer primer	GCGAGCACAGAATTAATACGACT	RT PCR Product	0.70	57
15.		PKS 2 RACE-n1*	3'PKS gene specific outer primer	GGAGATGGTGCTGCCGCGATCATC			
16.		3'Ad inner	3' adapter specific inner primer	CGCGGATCCGAATTAATACGACTC ACTATAGG	Product of Outer 3' RLM-RACE	Not performed	Not performed
17.		PKS 2 RACE-n2*	3'PKS gene specific inner primer	ATCCCCGAAGTCGAGAAGCCCTTG			

*Manually designed primers for RLM-RACE studies

Nested PCR of the first strand cDNA with IKRU1 primers followed by IKRU2 primers were performed after optimizing the annealing temperature of the primers to 42 °C. Nested PCR yielded an amplicon with good intensity and of approximately 550 bp (Plate 5.A).

4.1.3 Cloning

The cleaned up amplified product (5 µl) was visualized by agarose gel electrophoresis using 1.4 % gel and the amplicon was concluded to have an approximate concentration of 30 ng/µl DNA. 1.35 µl of this product was then cloned to pSC-A-amp/kan vector. Transformed colonies were observed as small white colonies while non transformed colonies were blue after overnight incubation of the LB ampicillin plates that were inoculated with the transformation mixture (Plate 6.A).

29 colonies were selected and cultured overnight in LB ampicillin broths. Plasmids with varying intensities were isolated from these colonies. Plasmids were observed as three distinct bands representing relaxed circular form, linear form and superhelical forms of plasmid (Plate 7.A). In subsequent PCR confirmation of the insert, P21 was omitted due to lack of intensity of plasmid. 25 plasmids produced expected 750 bp amplicon on PCR amplification with T7 and T3 primers (Plate 8. A).This confirmed the presence of insert in these plasmids.

The samples to be sequenced were selected based on a restriction digestion treatment of the 25 PCR amplification product in two separate reactions using Eco RI and NotI. Based on the size difference of digested products, absence of multiple bands and intensity of the amplicon, 6 clones were selected for sequencing (Plate 9. A).

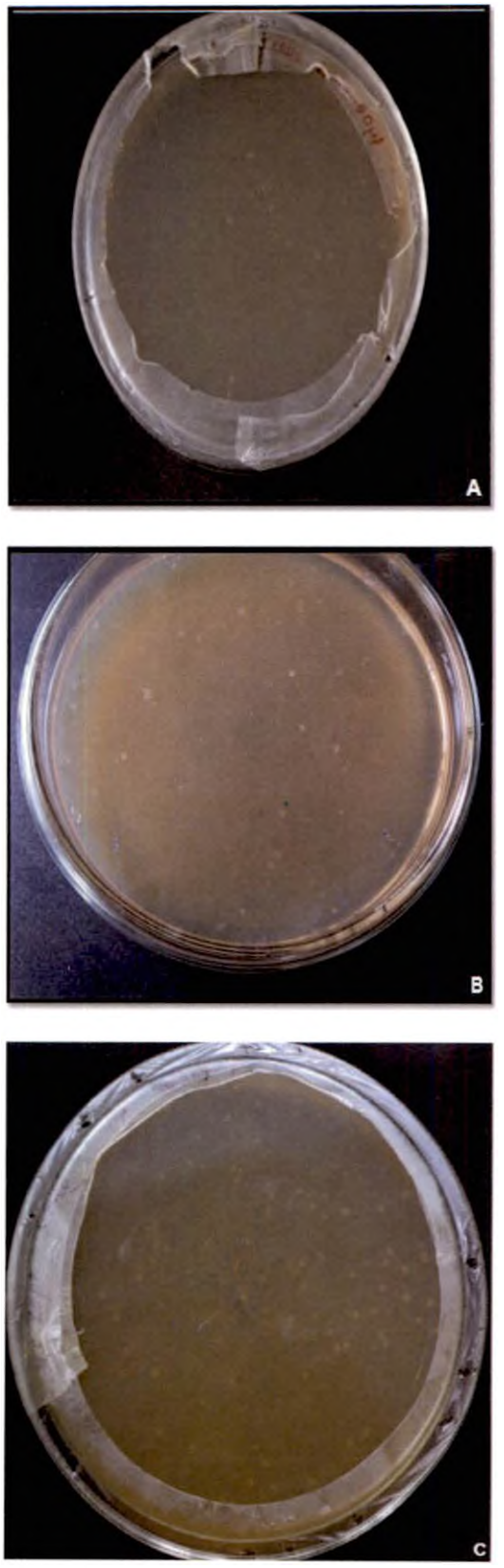

Plate 6. LB ampicillin plates with clones carrying inserts of (A) Core fragment (B) 5' RLM-RACE product (C) 3' RLM-RACE product

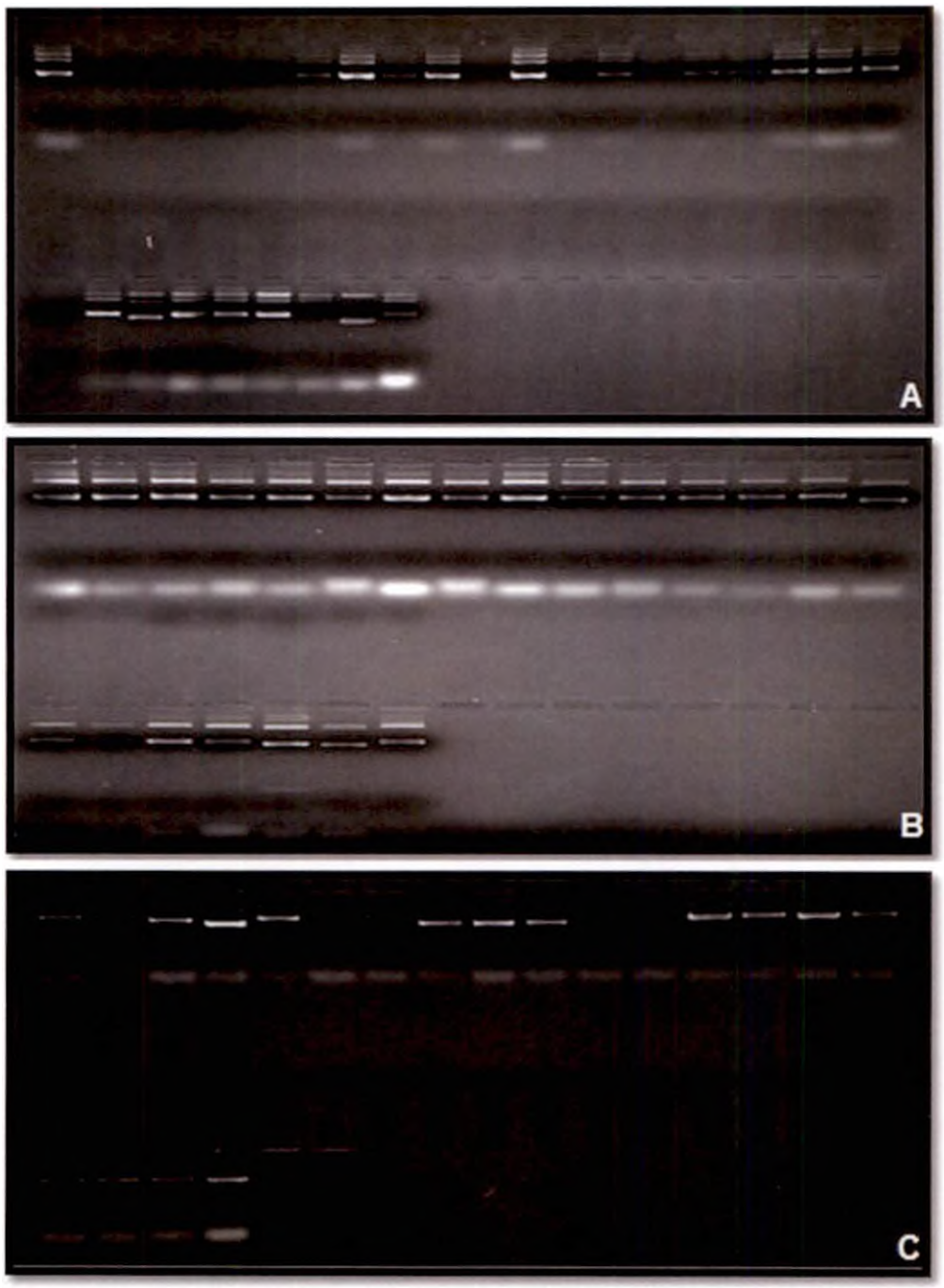

Plate 7. Plasmids containing inserts of (A) core fragment (B) 5' RLM-RACE product (C) 3' RLM-RACE product

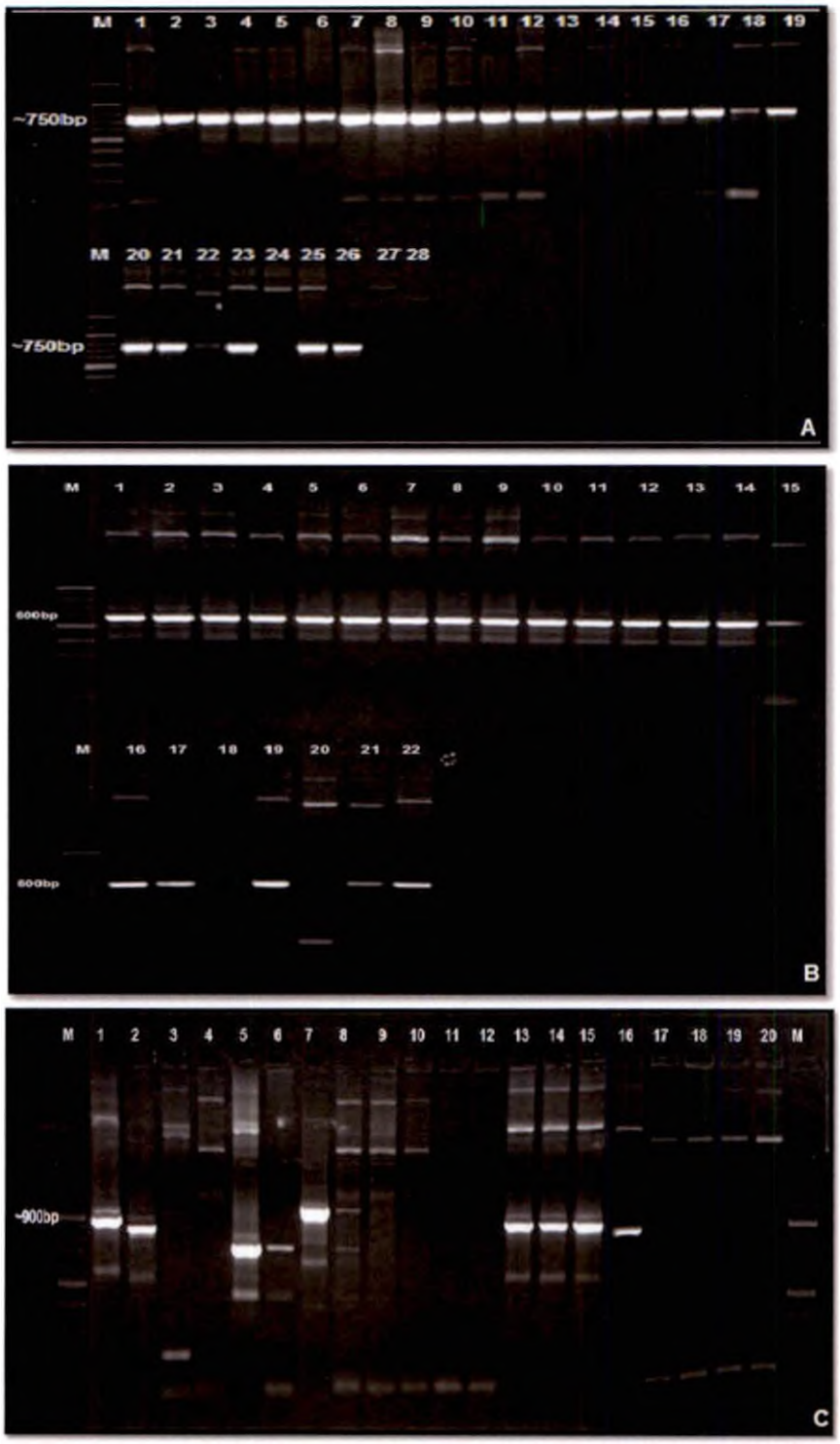

Plate 8. T7-T3 amplified plasmids containing: (A) core fragment (B) 5' RLM-RACE product (C) 3' RLM-RACE product

Plate Description
M: Marker
1-28: Plasmids amplified with T7-T3 primers

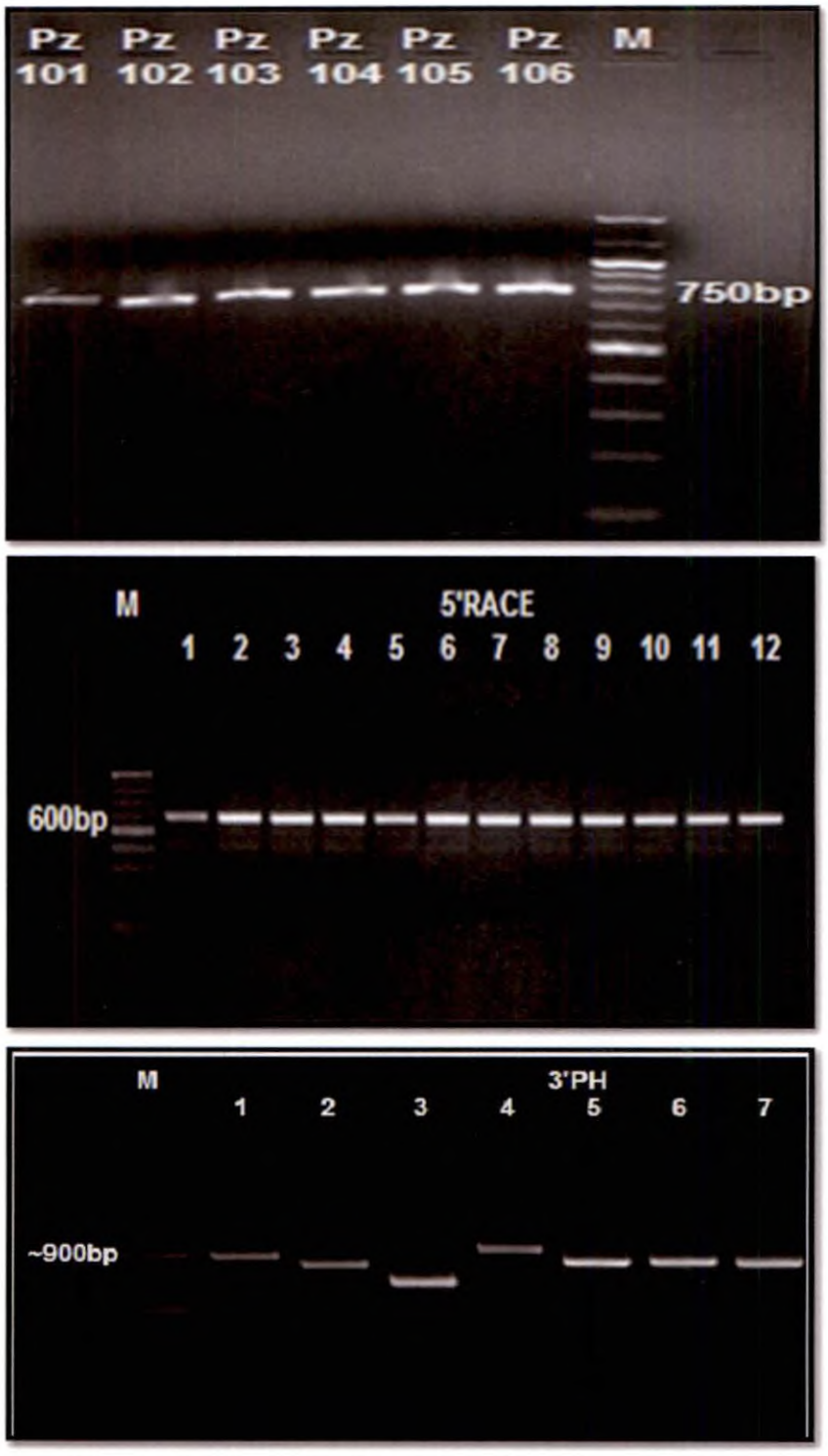

Plate 9. Amplicons selected for sequencing representing: (A) core fragment (B) 5' end (C) 3'end

Plate Description
M: Marker
Pz 101- 106: Core fragment sequences
5'RACE 1-12 : 5' end sequences
3' PH 1-7: 3' end sequences

4.1.4 Data Analysis

On sequencing, a 588bp cDNA representing the core sequence was obtained (Table3). BioEdit software was used to generate consensus sequence for each of the sequences and to perform Clustal W based multiple sequence alignment. One of the sequences showed considerable variation from the other five partial cDNA nucleotide sequences (Figure 1)

Characterization of the cDNA clones was done through homology search with BLASTX (nucleotide query to protein database search) program of NCBI (National Centre for Biotechnology Information). Since all the sequences shared a high degree of similarity, randomly one of the consensus sequences was used for performing BLASTX analysis. The deduced amino acid sequence showed 92 % identity with CHS sequence gene reported from *Polygonum cuspidatum*, 91 % identity with *Fallopia multiflora* and 90 % identity with *Fagopyrum tataricum*.

Four of the sequences were translated to corresponding amino acid sequence using ExPASy (Expert Protein Analysis System). Multiple Sequence Alignment (MSA) of the translated sequences with previously reported CHS sequences retrieved from GenBank database showed considerable similarity from 181st to 376th amino acid (Figure 2). From the multiple sequence alignment it was inferred that approximately 166 amino acids in the 3' direction and approximately 22 amino acids in the 5' direction were to be deduced from further RLM-RACE studies.

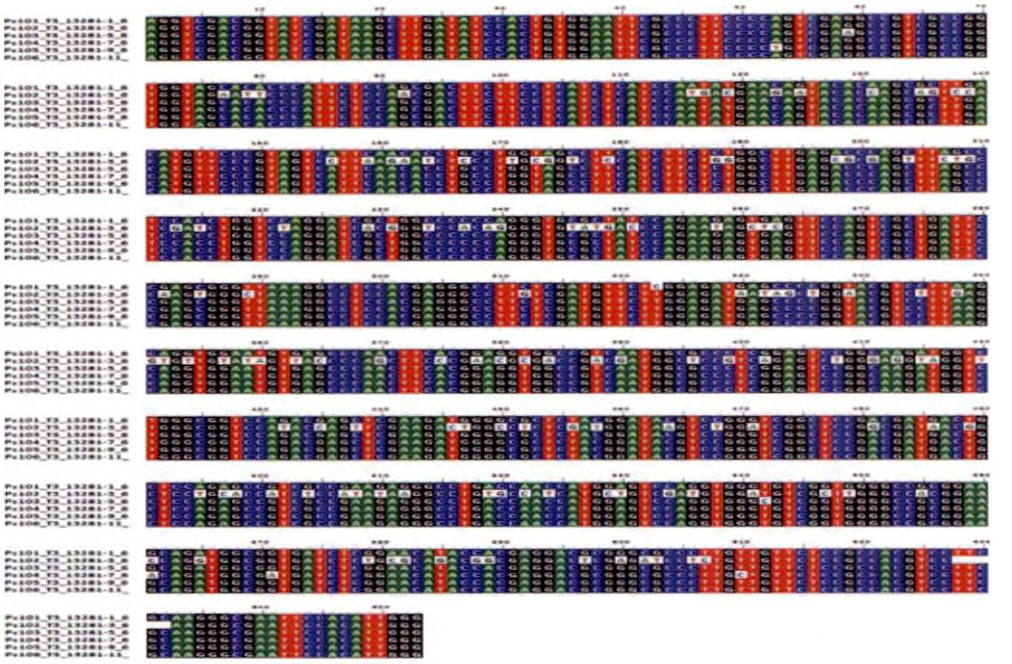

Figure1. Nucleotide sequence alignment of sequences representing core fragment

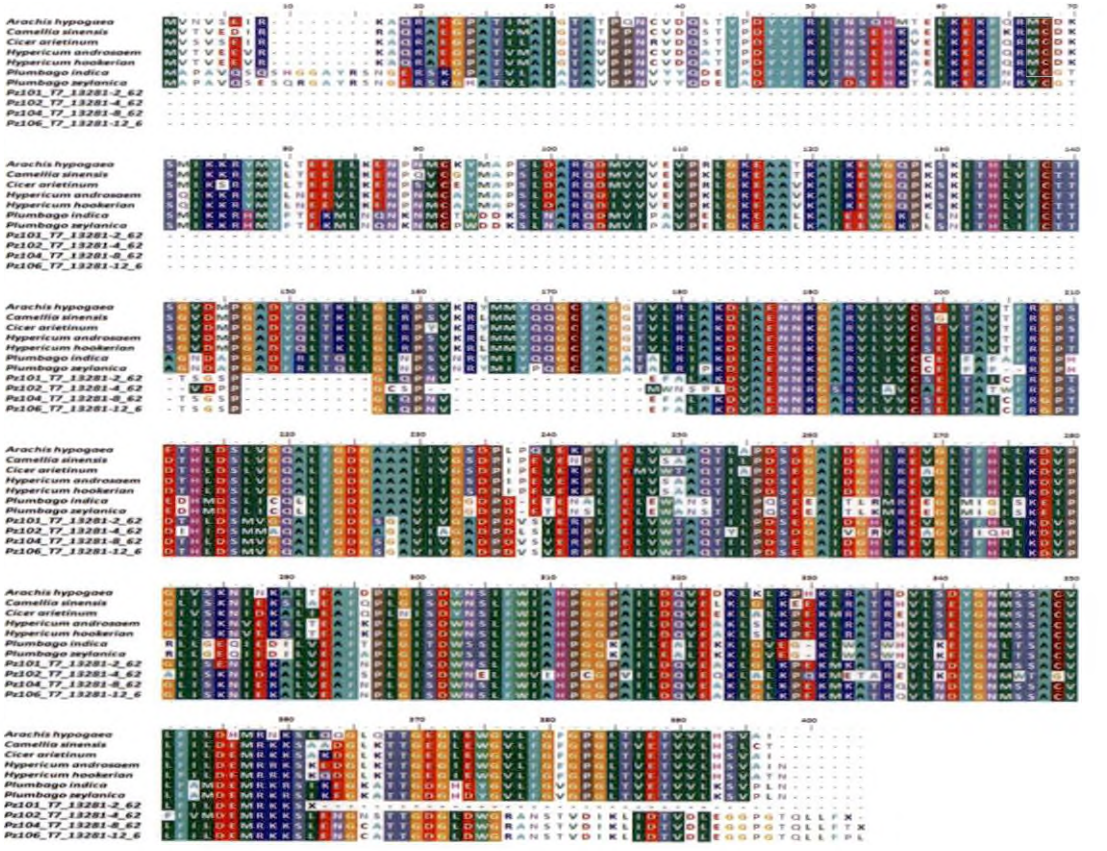

Figure 2. Amino acid sequence of core fragment aligned with reported PKS sequences

Table 3: Sequences of clones and full length *P. rosea* PKS sequence

Sr No:	Name of clone	Sequence length (bp)	Sequence (5'→3')
1.	PKS core partial cDNA (Pz101)	588	GAACTAGTGGATCCCCCGGGCTGCAGCCCAATGTGGAATTCGCCCTTGCGAAG GACGTGGCGGAGAACAACAAGGGCGCCCGCGTCCTCGTGGTATGTTCCGAGAT CACCGCCATCTGCTTCCGCGGCCCCACCGACACCCACCTGGACTCCATGGTTGG TCAGGCCCTCTTCGGCGACGGCTCTGGAGCAGTCATCGTCGGAGCCGACCCGG ACGTATCCGTAGAGCGCCCCATCTTCGAGCTCGTCTGGACCGCCCAGACCATCT TGCCGGACTCCGAGGGCGCCATTGACGGCCACCTCCGTGAAGTTGGGCTCACAT TCCACCTGCTTAAGGACGTACCGGGGCTCATCTCCGAGAACATCGAGAAGGCC CTCGTGGAGGCCTTTAACCCGCTCGGAATCAGCGACTGGAACTCACTCTTCTGG ATAGCGCACCCTGGGGGGCCAGCGATCCTGGACCAGGTGGAGGCTAAGCTGGG TCTCAAGCCAGAGAAGATGAAGGCTACCAGGCAGGTGTTGAATGACTACGGGA ACATGTCAAGCGCTTGTGTACTGTTCATATTGGATGAGATGAGGAAGAAGTCGC

Table 3 continued

Sr No:	Name of clone	Sequence length (bp)	Sequence (5'→3')
2.	5' RLM-RACE product (5'RACE1)	549	CCTCGAGGTCGACGGTATCGATAAGCTTGATATCCACTGTGGAATTCGCCCTTC GCGGATCCGAACACTGCGTTTGCTGGCTTTGATGAAAGGCGTATATGTATTAAT AGCTAGGCAAAGCAAGCCCCACCAAGCCTGCCTGCCAAACAACCCCCACCCCC CGCCCCGCCTAATTAAGATGGATGATTATTGATCGTAATTTAAGAAAAAAAAAT GGATAAGGTACGGAATCCAATTAATTATTAATTAATGGCCGGAACGCTGTGCA GCACGACAGTCTCAACGGTAAGCCCAGGCCCGAACCCAAACAGCACACCCCAG TCCAATCCTTCACCCGTAGTAGCACACCCACTCTCCAGCGACTTCTTCCTCATCT CATCCAATATGAACAGTACACAAGCGCTTGACATGTTCCCGTAGTCATTCAACA CCTGCCTGGTAGCCTTCATCTTCTCTAAGGGCGAATTCCACATTGGGCTGCAGC CCGGGGGATCCACTAGTTCTAGAGCGGCCGCACCGCGGGAGCTCCAATTCGCC CTATAGTGAGTC
3.	3' RLM-RACE product (3'PH5)	603	ATGGTGACCGTGGAAGAAGTCAGGGGAGCATACCGCAGCAAGGCGCAGAGGG CCGAGGGTCCGGCCACCGTGATGGCCATCGGAACCGCGGTCCCGCCCAACTGC GTTGACCAAGCGACGTACCCCGACTACTATTTCCGTATCACCAACAGCGAGCAC AAGGCTGAGCTCAAGGAGAAGTTCCAACGCATGTGCGACAAGTCCCAAATCAA GAAGCGTTACATGTACTTGAACGAGGAGGTCTTGAAGGAGAATCCCAATATGT GTGCTTACATGGCACCTTCTCTGGATGCTAGGCAAGACATTGTAGTGGTTGAAG TGCCTAAATTGGGTAAAGAGGCCGCAGTTAAGGCCATCAAGGAATGGGGCCAG CCCAAGTCCAAGATCACCCACTTGGTCTTTTGCACCACTAGTGGAGTGGACATG CCCGGGGCCGACTACCAGTTGACCAAGCTATTGGGCCTCCGCCCATCGGTCAAG CGTCTCATGATGTACCAGCAGGGCTGCTTCGCCGGGGGCACGGTCCTCCGTCTC GCCAAGGACCTCGCCGAGAACAACAAGGGTGCACGAGTGCTCGTCGTTTGCTC GGAGATCACGGCCGTA

Table 3 continued

Sr No:	Name of clone	Sequence length (bp)	Sequence (5'→3')
4.	Full length cDNA sequence of PKS from *Plumbago rosea*	1197	ATGGTGACCGTGGAAGAAGTCAGGGGAGCATACCGCAGCAAGGCGCAGAGGG CCGAGGGTCCGGCCACCGTGATGGCCATCGGAACCGCGGTCCCGCCCAACTGC GTTGACCAAGCGACGTACCCCGACTACTATTTCCGTATCACCAACAGCGAGCAC AAGGCTGAGCTCAAGGAGAAGTTCCAACGCATGTGCGACAAGTCCCAAATCAA GAAGCGTTACATGTACTTGAACGAGGAGGTCTTGAAGGAGAATCCCAATATGT GTGCTTACATGGCACCTTCTCTGGATGCTAGGCAAGACATTGTAGTGGTTGAAG TGCCTAAATTGGGTAAAGAGGCCGCAGTTAAGGCCATCAAGGAATGGGGCCAG CCCAAGTCCAAGATCACCCACTTGGTCTTTTGCACCACTAGTGGAGTGGACATG CCCGGGGCCGACTACCAGTTGACCAAGCTATTGGGCCTCCGCCCATCGGTCAAG CGTCTCATGATGTACCAGCAGGGCTGCTTCGCCGGGGGCACGGTCCTCCGTCTC GCCAAGCTTGCGAAGGACGTGGCGGAGAACAACAAGGGCGCCCGCGTCCTCGT GGTATGTTCCGAGATCACCGCCATCTGCTTCCGCGGCCCCACCGACACCCACCT GGACTCCATGGTTGGTCAGGCCCTCTTCGGCGACGGCTCTGGAGCAGTCATCGT CGGAGCCGACCCGGACGTATCCGTAGAGCGCCCCATCTTCGAGCTCGTCTGGAC CGCCCAGACCATCTTGCCGGACTCCGAGGGCGCCATTGACGGCCACCTCCGTGA AGTTGGGCTCACATTCCACCTGCTTAAGGACGTACCGGGGCTCATCTCCGAGAA CATCGAGAAGGCCCTCGTGGAGGCCTTTAACCCGCTCGGAATCAGCGACTGGA ACTCACTCTTCTGGATAGCGCACCCTGGGGGGCCAGCGATCCTGGACCAGGTG GAGGCTAAGCTGGGTCTCAAGGAGAAGATGAAGGCTACCAGGCAGGTGTTGAA TGACTACGGGAACATGTCAAGCGCTTGTGTACTGTTCATATTGGATGAGATGAG GAAGAAGTCGCTGGAGAGTGGGTGTGCTACTACGGGTGAAGGATTGGACTGGG GTGTGCTGTTTGGGTTCGGGCCTGGGCTTACCGTTGAGACTGTCGTGCTGCACA GCGTTCCGGCCATTAATTAA

4.1.5 Primer Designing

Primer designing was performed manually by using the nucleotide sequences that corresponded to the amino acid regions showing similarity in MSA. Two sets of nested gene specific primers, each consisting of an outer primer and an inner primer were designed for 3' RLM-RACE and 5' RLM-RACE (Table 2). PKS 2N and PKS 2 were designed for the outer and inner 5' RLM-RACE respectively. PKS 2 RACE-n1 and PKS 2 RACE-n2 were designed for the outer and inner 3' RLM-RACE respectively. However, only PKS 2 RACE-n1 primer was used for the 3' RLM-RACE as outer 3' RLM-RACE was not performed. The details of the primers designed are given in Table 2.The annealing temperatures of the primers were standardized during the respective RLM-RACE reactions by setting gradient conditions.

4.1.6 5'RLM-RACE

The processed RNA was used for performing 5' RLM-RACE. Outer gene specific primer, PKS 2N primer was used for the outer 5' RLM-RACE. The primer annealing temperature was standardized to be 55°C by providing a gradient condition during the annealing step.

The outer 5' RLM-RACE product was subjected to inner 5' RLM-RACE using the inner gene specific primer, PKS 2. The annealing temperature of this nested primer was standardized to be 60°C by providing a gradient condition during the annealing step. 5' RLM-RACE yielded a good intensity amplicon of approximately 400bp.

4.1.7 Cloning of 5' RLM-RACE Products

The 400 bp amplicon was estimated to have an approximate concentration of 30 ng/µl DNA by gel electrophoresis following gel elution (Plate 5.B). 1.2 µl of this product was then cloned to pSC-A-amp/kan vector. 22 transformed colonies that appeared as small white colonies in LB Ampicillin plates (Plate 6.B) were subjected to plasmid isolation following overnight culturing in LB ampicillin

broth. 21 plasmids were successfully isolated (Plate 7.B). On amplification with T7 and T3 primers 19 plasmids were confirmed to contain the insert as these yielded 600bp amplicons (Plate 8.B).12 samples which showed good intensity 600bp single band were selected for sequencing (Plate 9.B).

4.1.8 Data Analysis

The approximately 549 bp long obtained sequences were subjected to multiple sequence alignment. The sequences showed high degree of sequence similarity. Since all the sequences were highly similar, one of the sequences (Table 3) was randomly chosen for BLASTX analysis which revealed significant homology with already reported CHS genes such as Chalcone synthase from *Fagopyrum esculentum* and *Fagopurum tataricum*. The aminoacid sequence corresponding to 5' end was obtained using ExPASy. The longest ORF was merged with the already obtained aminoacid sequence of the core cDNA. This yielded a 232 amino acid long *Plumbago rosea* PKS sequence. On performing MSA of this sequence with already reported CHS sequences it was inferred that 166 amino acids were to be sequenced in the 3' end (Figure 3).

4.1.9 3' RLM-RACE

Following first strand cDNA synthesis, 3' cDNA end was isolated by using 3' RLM-RACE. Only the outer 3' RLM RACE was performed as this yielded the product of desired size. Outer 3' RLM RACE was performed using previously designed gene specific primer, PKS 2 RACE-n1. Out of the three temperatures, 55 °C, 56 °C and 57 °C tested, the optimum annealing temperature of the primer was concluded to be 57 °C based on the intensity of the products obtained. Outer 3' RLM RACE yielded a number of low intensity bands of different sizes, but the 700 bp band was chosen as it is closest to the expected cDNA size that is to be sequenced (Plate 5.C). The amplicon was eluted from the gel.

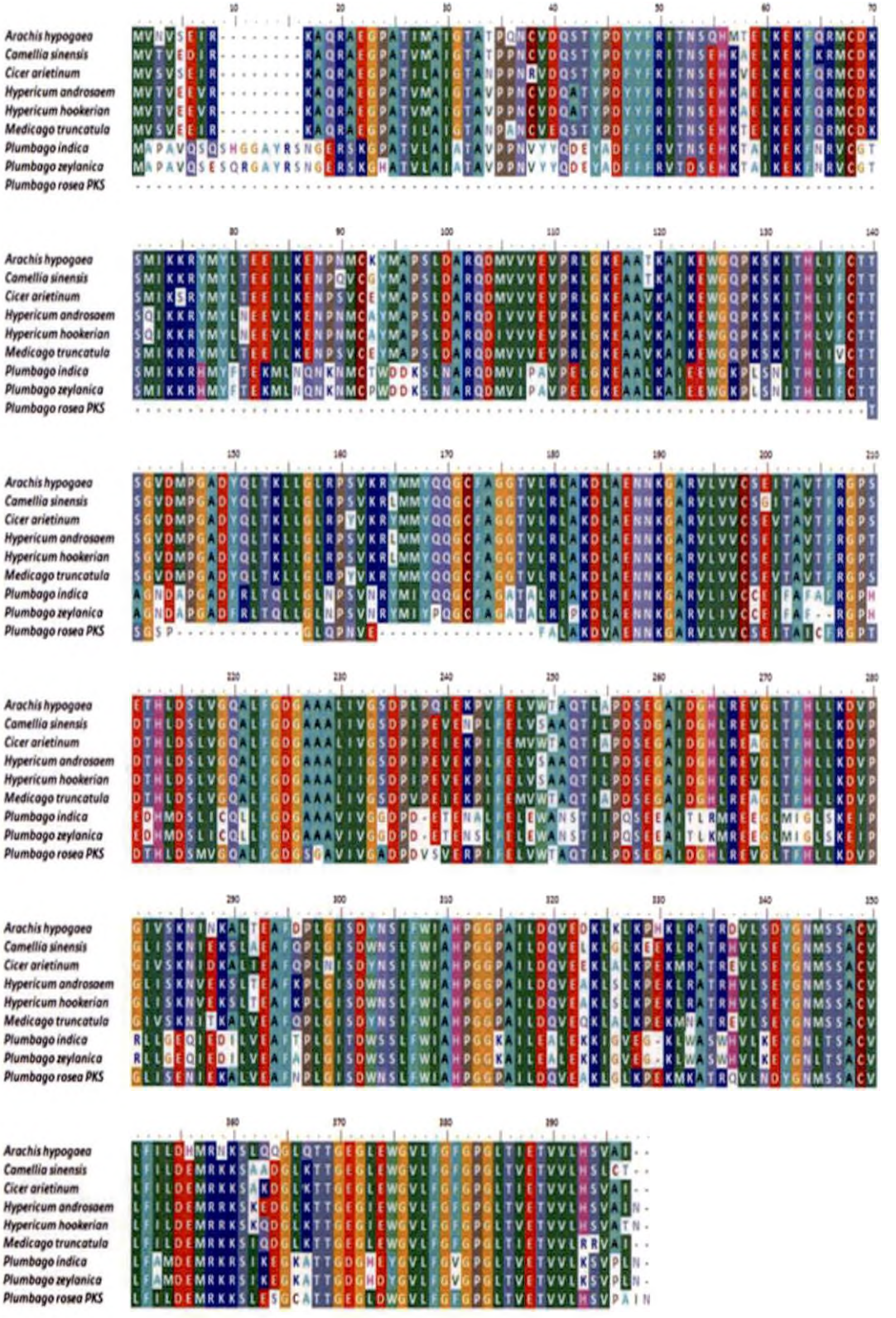

Figure 3. Multiple sequence alignment of *Plumbago rosea* PKS (core and 5' end) with other CHS superfamily sequences

4.1.10 Cloning of 3' RLM-RACE products

The 700 bp band was estimated to have an approximate concentration of 20 ng/µl . This product (2 µl) was then cloned to pSC-A-amp/kan vector (Plate 6.C). Plasmids were isolated from 20 transformed colonies that appeared as small white colonies (Plate 7.C). 7 plasmids yielded amplicons of different sizes when amplified with T7 and T3 primers (Plate 8.C). These 7 samples which showed good intensity amplicons were selected for sequencing (Plate 9.C).

4.1.11 Data Analysis

On multiple sequence alignment, the sequences were found to be highly dissimilar with only sequences 3'PH5, 3'PH6 and 3'PH7 sharing similarity. BLASTX analysis, revealed that only the 3'PH5 (Table 3) sequence showed high similarity i.e 98 % identity with aromatic polyketide synthase from *Hypericum hookerianum*, naringenin chalcone synthase from *Hypericum androsaemum* and chalcone synthase from *Hypericum sampsonii* (Figure 4). The corresponding amino acid sequence was deduced from the nucleotide sequence followed by aligning and merging with previously deduced core fragment and 5' end amino acid sequences to obtain full length amino acid sequence (Figure 5).

The ORF was seen to encode protein of 398 amino acid with a calculated molecular weight of 43.5 kDa and a theoretical pI (isoelectric point) of 5.89.The BLASTX analysis of the full length cDNA sequence showed 89 % identity with naringenein chalcone synthase from *Hypericum androsaemum*, chalcone synthase from *Hypericum sampsonii* and aromatic polyketide synthase from *Hypericum hookerianum*. The CD-Search in NCBI was used to map conserved domains such as active site, product binding domain, malonyl CoA binding site and dimer interface. The PKS modeling was done using Swiss-Model server with *Medicago sativa* CHS (pdb accession no.: 1d6f) as template and was visualized with PyMOL (Plate 10). A phylogenetic tree showing the relationship of the deduced sequence with previously reported CHS sequences was generated using MEGA software (plate11). The sequence obtained from *P. rosea* was found to be grouped

with PKS sequence previously reported from *Hypericum androsaemum* CHS *and* *Hypericum hookerianum* PKS.

The obtained full length cDNA sequence is represented in Table 3

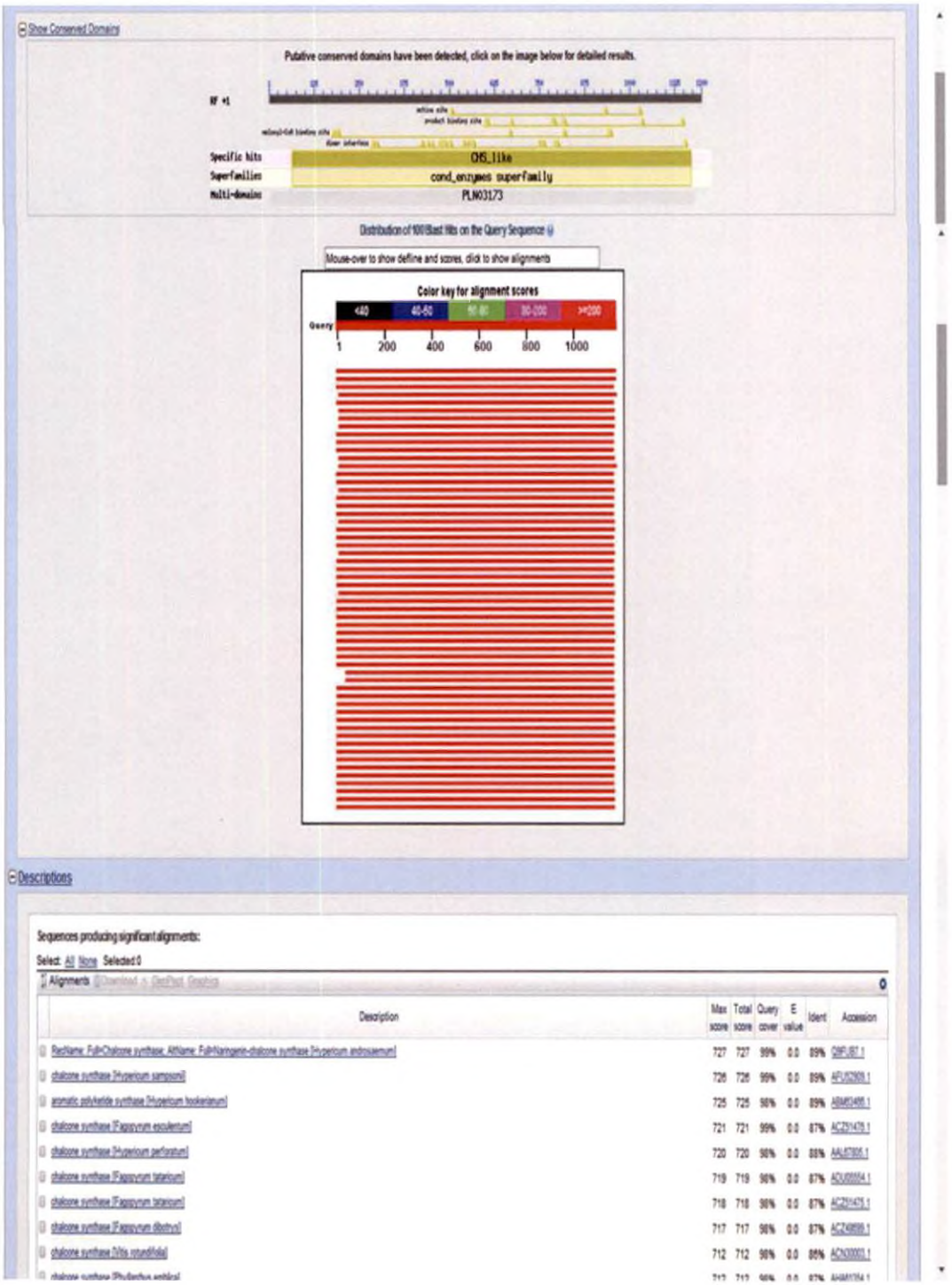

Figure 4. BLASTX result and conserved domains of full length *Plumbago rosea.*

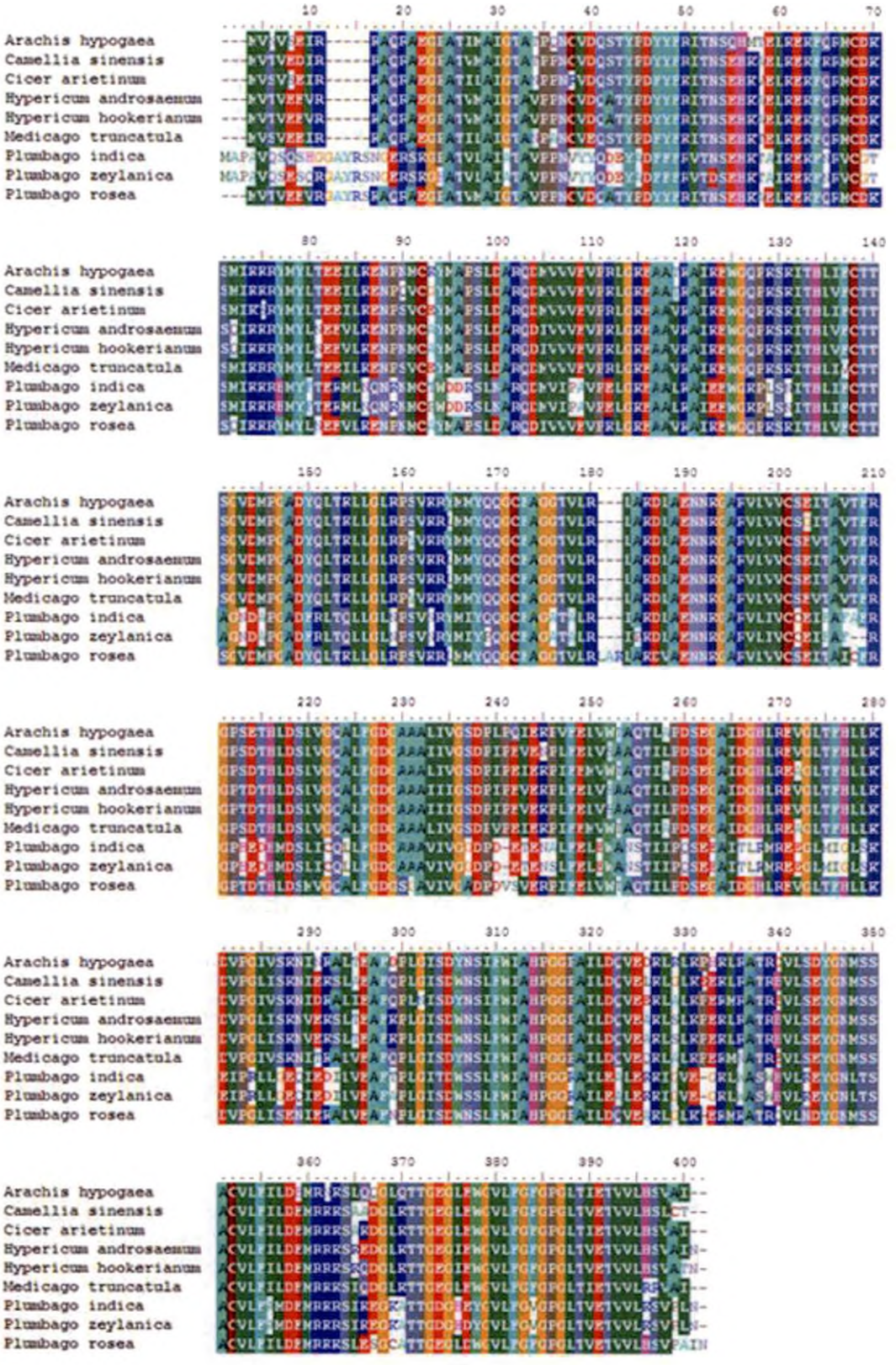

Figure 5. Multiple sequence alignment of full length amino acid sequence of *Plumbago rosea* PKS with other CHS superfamily sequences

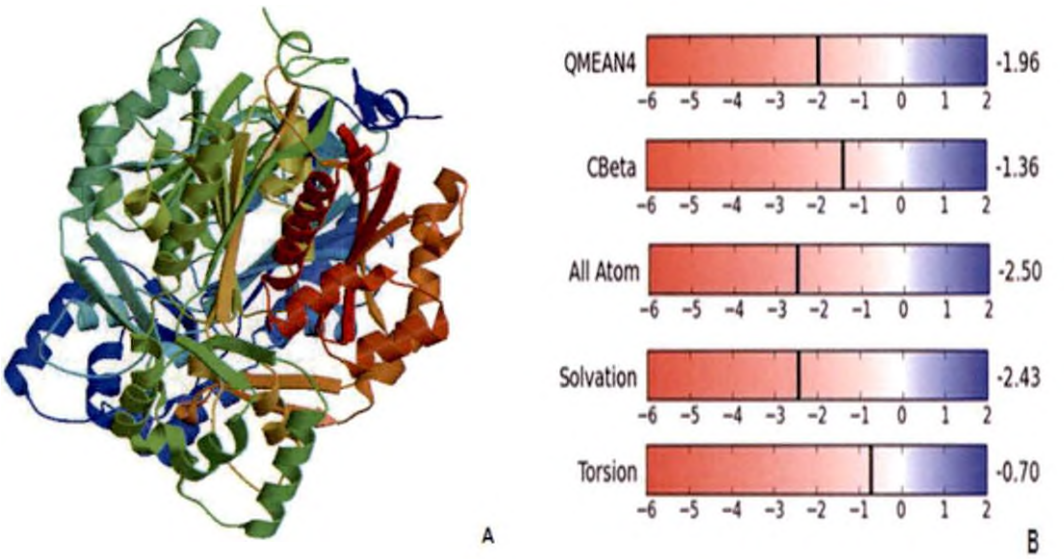

Plate 10. Tertiary structure of *Plumbago rosea* PKS (A) 3D structure (B) Z scale of protein parameters

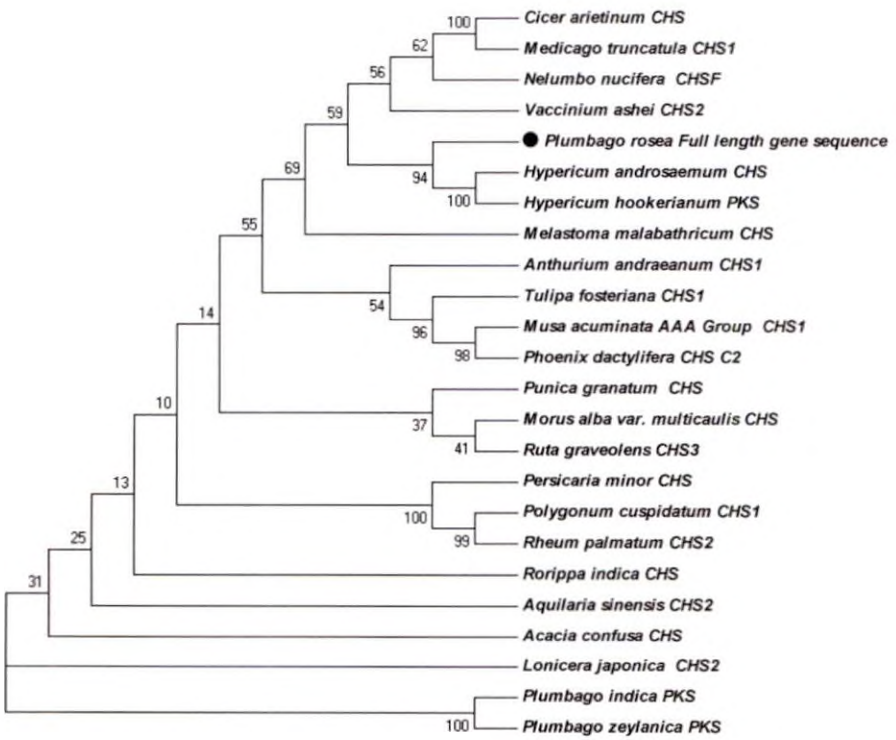

Plate 11. Phylogenetic tree showing relationship of *Plumbago rosea* PKS with previously reported CHS superfamily genes.

4.2 MOLECULAR MARKER ANALYSIS

4.2.1 DNA Isolation

Leaves of 8 accessions (Table 4) of *P. rosea* were used for genomic DNA isolation. The agarose gel electrophoresis of the isolated DNA in 1 % gel showed the presence of good quality genomic DNA bands. Further, the absorbance reading of the extracted genomic DNA by using spectrophotometric method revealed good quality and quantity of DNA (Table 4).

Table 4: Quality and quantity of isolated genomic DNA

Sr. No:	Accession No:	Location	Absorbance (A_{260} nm)	Absorbance (A_{280} nm)	A_{260} / A_{280}	DNA Yield ($\mu g \mu l^{-1}$)
1.	P1	JNTBGRI	0.202	0.108	1.92	0.52
2.	P2	JNTBGRI	0.184	0.104	1.77	0.46
3.	P3	JNTBGRI, Itti Achuthan Garden	0.113	0.065	1.75	0.28
4.	P4	College of Agriculture, Vellayani	0.333	0.175	1.90	0.83
5.	P5	Chengannur	0.560	0.282	1.99	1.40
6.	P6	Trivandrum	1.665	0.788	2.11	4.16
7.	P7	Ernakulam	0.060	0.040	1.52	0.15
8.	P8	Waynad	0.058	0.029	2.02	0.15

4.2.2 ISSR Profiling

The ISSR assays were carried out using the isolated genomic DNA and twelve primers which produced consistent amplification pattern. Details of the ISSR primers used are provided in table 5. The study used 17-18bp long primers some of which were anchored at 3' end for increased specificity. The genomic DNA isolated was used as template for ISSR assays. The primers altogether generated 71 scorable bands, out of which 70 were polymorphic thereby showing 98.6 % polymorphism. The number of products generated by these arbitrary primers were found to range from 3-14, with primer 847 producing the maximum number of 14 bands (Plate 12.B) and primers 815 and 829 producing the minimum number of 3 bands. On an average the primers generated 5 bands and the mean number of polymorphism per primer is 6. Only primer 844 revealed 100 % monomorphism across the accessions (Plate 12.A).

A dendrogram was generated using WINBOOT software. The Chengannur and Ernakulam accessions were found to be the most related while the sample from Waynad formed an outlier (Plate 13)

Similarity matrix of the accessions were analyzed based on Dice's coefficient. The similarity matrix (Table 6) developed using the WINDIST showed that similarity index ranges from 0.32 to 0.82 with mean value of 0.6 thereby suggesting high levels of genetic variability in the species.

The genotype and allelic frequency data were used to compute the genetic diversity indices such as, % of polymorphic loci (P), observed number of alleles (na), expected number of alleles (ne), Shannon index of genetic diversity (I) and Nei's gene diversity (h) at the population level using the statistical package POPGENE 1.3 (Table 7). The accessions included in this study showed relatively high level of genetic diversity h= 0.26 and I = 0.38 respectively. Accessions P3 and P7showed the highest level of genetic diversity h= 0.31, I= 0.44 and h= 0.31,

I=0.45 respectively, while accession P1 exhibited the lowest (h=0.19,I=0.28).The mean value of observed number of alleles and expected number of alleles are 1.67 and 1.47 respectively. The mean number of polymorphic loci is 8 and the mean percentage polymorphic loci is 66.67%.

Table 5: Sequences of primers used for ISSR analysis

Sl. No.	Primer Code	Primer Sequence (5' → 3')	T_m (°C)	GC %	No. of bands	No. of polymorphic bands
1	811	CAC AGA GAG AGA	52	53	4	4
2	815	CTC TCT CTC TCT	52	53	3	3
3	820	GTG TGT GTG TGT	52	53	5	5
4	823	TCT CTC TCT CTC	52	53	7	7
5	826	ACA CAC ACA CAC	52	53	7	7
6	827	ACA CAC ACA CAC	52	53	6	6
7	829	TGT GTG TGT GTG	52	53	3	3
8	835	AGA GAG AGA GAG	53	53	5	5
9	836	AGA GAG AGA GAG	50	47	5	5
10	841	GAG AGA GAG AGA	53	53	5	5
11	844	CTC TCT CTC TCT	53	53	7	6
12	847	CAC ACA CAC ACA	53	53	14	14
		Total Number of Bands			71	70

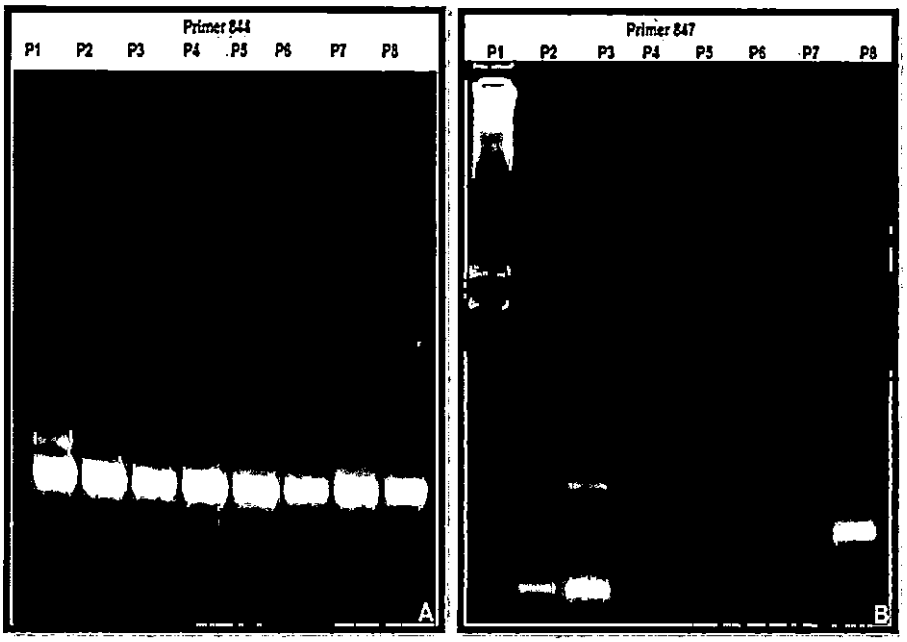

Plate 12. Accessions P1-P8 amplified using ISSR primers: (A) 844 (B) 847

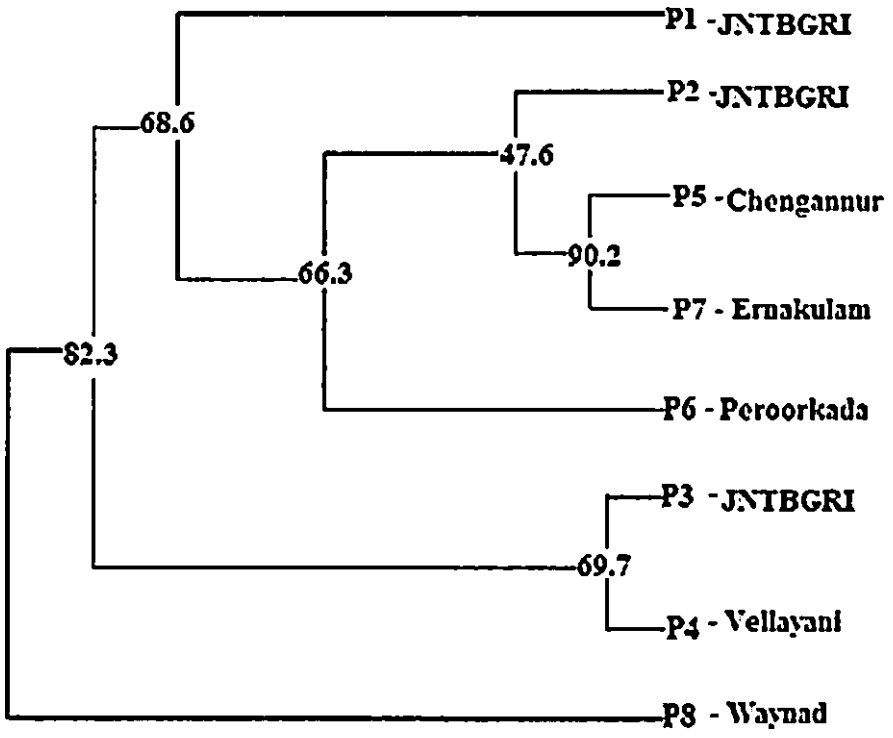

Plate 13. Dendrogram showing relationship between accessions used for ISSR study

Table 6: Similarity matrix of *Plumbago rosea* ISSR data.

P1	1.00							
P2	0.63	1.00						
P3	0.34	0.49	1.00					
P4	0.50	0.50	0.60	1.00				
P5	0.58	0.73	0.61	0.61	1.00			
P6	0.54	0.66	0.48	0.42	0.71	1.00		
P7	0.61	0.71	0.64	0.61	0.82	0.69	1.00	
P8	0.37	0.35	0.32	0.41	0.36	0.44	0.37	1.00
	P1	P2	P3	P4	P5	P6	P7	P8

Table 7: Genetic diversity indices of *Plumbago rosea* accessions

Population	na	ne	h	I	No: of polymorphic loci	% of Polymorphic loci
P1	1.5	1.33	0.19	0.28	6	50.00
P2	1.67	1.43	0.25	0.36	8	66.67
P3	1.75	1.57	0.31	0.44	9	75.00
P4	1.58	1.47	0.26	0.37	7	58.33
P5	1.58	1.46	0.25	0.36	7	58.33
P6	1.67	1.47	0.26	0.38	8	66.67
P7	1.75	1.55	0.31	0.45	9	75.00
P8	1.83	1.44	0.28	0.43	10	83.33
Mean	1.67	1.47	0.26	0.38	8	66.67

4.3 PLUMBAGIN EXTRACTION AND ESTIMATION

Plumbagin extraction and estimation were performed using root samples of nine *P. rosea* accessions used for the studies in section 3.1 and 3.2.

4.3.1 TLC

TLC was performed to determine the R_f (retardation factor) value of plumbagin and to study purity of the crude extract. R_f value was 0.64 in benzene solvent system which corresponds to the reported value.The extracts were observed to be devoid of impurities as only a single band could be observed on TLC. Also on comparitive analysis it could be inferred from the TLC plate that P3 had higher plumbagin content as compared to the other accessions based on the brightness of the plumbagin spotted (Plate14.A).

4.3.2 Plumbagin Estimation by Spectrophotometry at 415nm

The plumbagin content in the crude extracts were estimated by spectrophotometry at 415nm (Table 8). The maximum plumbagin content was 0.7 % in P3 accession collected from JNTBGRI and the minimum content was 0.2 % in P7 accession collected from Ernakulam. P1, P5, P8 and TBG-102 were found to have similar plumbagin content (0.6%) while P2 and P6 were found to contain similar content (0.4%).

4.3.3 HPLC

Plumbagin content in the accession TBG-102 was further confirmed by HPLC and was determined to be 0.505 %.The retention time of plumbagin was estimated to be 3.6 min (Plate 15)

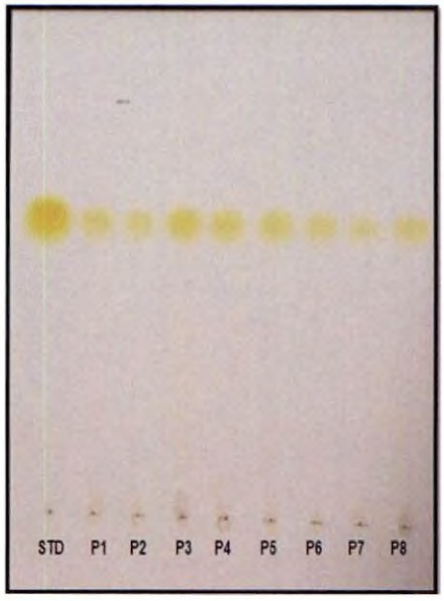

Plate 14. TLC of plumbagin standard (STD) and plumbagin extracts from accessions P1-P8

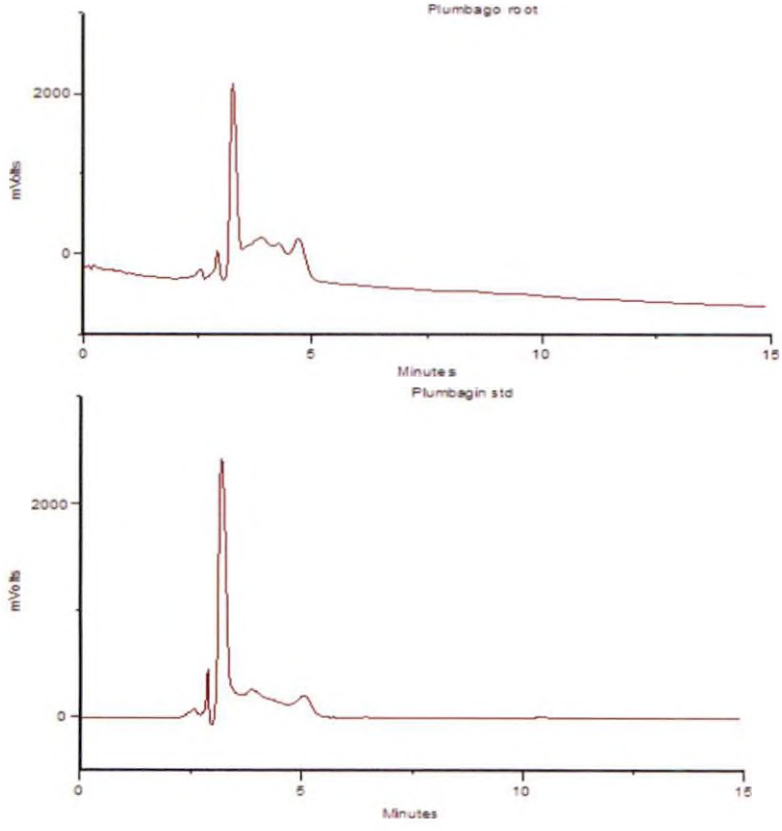

Plate 15. HPLC graph showing retention time of plumbagin extract and plumbagin standard

Table 8. Plumbagin content in *P. rosea* accession estimated by spectrophotometry

Sr. No.	Accession	Concentration [µg/ml]	Absorbance at 415.00 nm	Percentage (%)
1	P1	293.145	4.356	0.586
2	P2	191.738	2.849	0.383
3	P3	353.100	5.247	0.706
4	P4	329.347	4.894	0.658
5	P5	314.139	4.668	0.628
6	P6	210.310	3.125	0.420
7	P7	125.256	1.861	0.240
8	P8	287.358	4.270	0.574
9	TBG-102	311.380	4.627	0.575

DISCUSSION

5. DISCUSSION

Plants produce a wide variety of secondary metabolites which do not play any essential role in cellular metabolism but are involved in the adaptation of the plant to its environment. These array of metabolites have been exploited since ancient time as flavours, fragrances, colorants, antibiotics and drugs. Polyketides form a group of structurally and functionally diverse secondary metabolites that have a wide range of applications with particular importance in pharmaceutical industry. Plumbagin is one such polyketide that has been of immense pharmaceutical interest due to its potential medicinal activities such as anticancer (Parimala and Sachdanandam, 1993), antimalarial (Likhitwitayawuid *et al.*, 1998), antimicrobial (Didry *et al.*, 1994), cardiotonic (Itoigawa *et al.*, 1991) and antifertility (Bhargava, 1984) activities. The demand for plumbagin is escalating due to these medicinal values. In nature, plumbagin has been reported to be synthesised by plants belonging to many families, however, *Plumbago rosea* roots are known to be the richest natural source of this bioactive compound. Due to the increased collection of this slow growing plant from its natural habitat, the plant is at the risk of extinction and is currently categorized as rare. In this context, biotechnological interventions that could help to enhance the yield of this compound in a commercially viable manner would be of utmost importance to human welfare as well as for plant conservation. Elucidation of the biosynthetic pathway involved in the production of the compound is inevitable for comprehensive studies and for developing strategies for commercial production. In this respect, very few studies have been carried out to uncover the biosynthetic pathway of plumbagin production in *Plumbago rosea*.

Plumbagin was expected to be formed from Shikimate pathway followed by C-methylation in the 2-position, due to its structural similarity to juglone and menadion (Teuscher, 1970). But, tracer experiments conducted by Durand and Zenk (1971) with acetate and malonate in young shoots of *Plumbago europea*, revealed that acetate-malonate pathway was operating during plumbagin

biosynthesis. Therefore, it was postulated that the basic skeleton of these naphthoquinones is produced by a polyketide synthase involved in the polyketide pathway. This route of synthesis of plumbagin was further confirmed by Durand and Zenk (1974) with the help of tracer experiments carried out in *Drosophylum lusitanicum*.

Subsequently, very few investigations have been undertaken to reveal the polyketide synthase genes involved in plumbagin biosynthesis. The foremost studies in this direction were performed by cloning of PKS from *Plumbago indica* (Springob *et al.*, 2007) and *Drosophyllum lusitanicum* (Jindaprasert *et al.*, 2008). Springob *et al.* (2007) cloned a cDNA encoding type III polyketide synthase involved in naphthoquinone synthesis from the root of *P. indica*. The recombinant enzyme when expressed in *Escherichia coli*, resulted in alpha pyrone hexaketides which however, did not fold into naphthalene derivatives. Based on these studies it was suggested that the novel PKS is involved in the biosynthesis of naphthoquinones, and additional cofactors may be required for the biosynthesis of these secondary metabolites *in vivo*. Similarly, when recombinant DluHKS, isolated from *Drosophyllum lusitanicum* by Jindaprasert *et al.* (2008) was expressed in *Escherichia coli*, the expected products, naphthalenes, were not produced. From the hexaketide alpha-pyrone that was produced, it was concluded that the recombinant enzyme provides the backbone of these secondary metabolites. It was proposed that the absence of naphthalenes *in vitro* could be due to the absence of an accessory protein that is involved in the *in vivo* pathway. Cloning of a hexaketide synthase from *P. zeylanica* (PzPKS) and its expression in tobacco plants (*Nicotiana tabacum* L.) was also studied by Jadhav *et al.* (2013). However, the transgenic tobacco plants were unable to synthesize plumbagin. These studies indicate that these plants also lack the other components of the pathway that may be required in plumbagin synthesis (Springob *et al.*, 2007).

In the light of these studies, the present study was undertaken to clone type III PKS gene(s) that may be present in *Plumbago rosea* L. syn *P. indica*. This

study is different from the study conducted by Springob *et al.* (2007) in the method used for the cloning of full length cDNA. Springob *et al.* (2007) isolated the *Pin*PKS from cDNA library by screening the library with a *Pin1* obtained by first strand cDNA synthesis followed by amplification with previously reported (Samappito *et al.,* 2003; Helariutta *et al.*, 1995) degenerate primers. The full length cDNA in the current study was isolated by first strand cDNA synthesis and subsequent amplification using degenerate primers reported for CHS (Abe *et al.*, 2005) followed by 3' and 5' RLM-RACE.

PKSs have evolved by gene duplication and divergence by mutations, thereby providing plants with an adaptative differentiation (Tropf *et al.*, 1994; Durbin *et al.*, 2000; Lukacin *et al.*, 2001). Therefore, *Plumbago rosea* may also possess multiple families of PKSs and with different enzymatic activity as have been reported in other plants. (Rolfs and Kindl, 1984; Zheng *et al.*, 2001; Samappito *et al.*, 2002; Matousek *et al.*, 2006). These multiple genes may be responsible for the co-existence of different metabolites within the same plant as in the case of *Plumbago rosea* which is known to possess multiple compounds. This presence of families of PKSs in one single species emphasizes the importance of their characterization to understand their functional divergence, to completely elucidate biosynthetic pathways and to pave way for commercial productions based on metabolic engineering. Furthermore, chemical variation with respect to plumbagin content and genetic diversity using ISSR markers was also carried out in available accessions of *Plumbago rosea.*

5.1 ISOLATION AND CHARACTERIZATION OF TYPE III POLYKETIDE SYNTHASE GENE(S)

5.1.1 RNA Isolation

Both TRIzol® reagent method and Spectrum Plant Total RNA Kit were tried for RNA isolation. However, TRIzol® reagent yielded highly degraded RNA.

Normally, RNA degradation results due to too much manipulation of the sample or the thawing of the sample in the absence of TRIzol reagent or because the work environment or TRIzol reagent are RNase contaminated. The first two possibilities were ruled out because the quick grinding of the sample in sufficient quantity of liquid nitrogen and immediate transfer of the same to TRIzol reagent could not reduce RNA degradation. The sterility of the work environment was ensured by adopting stringent hygienic conditions. The TRIzol reagent was also changed and freshly opened reagent was tried, nevertheless, the degradation persisted. Intact RNA was finally isolated from snap frozen samples using Spectrum Plant Total RNA Kit. Therefore it can be concluded that the degradation might have been due to some endogenous substance which was effectively eliminated by the components of the kit. The isolated RNA was visualized as two intact bands, corresponding to 28S rRNA, 18S rRNA. No apparent RNA degradation or genomic DNA contamination were observed, showing good quality RNA.

5.1.2 Isolation and Characterisation of PKS

The RNA was used to synthesise cDNA. The fidelity of the cDNA was ensured using β-actin primers which are used as internal standard as, β-actin encodes a ubiquitous cytoskeleton protein and it is expressed at moderately abundant levels in most cell types (Kreuzer *et al.* 1999). Nested PCR was performed with degenerate primers designed by (Abe *et al.*, 2005) based on the conserved sequences of previously reported CHS genes. On performing nested PCR a 550bp amplicon was observed which was obtained by amplification of IKRU1 products with IKRU2 primers.

The amplicon was cloned into pSC-A-amp/kan vector and the presence of insert was confirmed by amplification of isolated plasmids with T7 and T3 primers which yielded a 750 bp product. The amplicon consists of the 550bp

insert in addition to the 100bp sequences complementary to T7 and T3 primers found flanking the insert. Six of these clones were sequenced.

The clone representing the core partial cDNA was 588bp in size after the trimming subsequent to sequencing. Trimming was performed to eliminate the extreme end sequences obtained after sequencing which have relatively low quality value (QV). However, inspite of trimming, some of the primer sequences would have remained resulting in increase of insert size 550bp to 588bp. On multiple sequence alignment of the 6 sequences, the sequence Pz102 showed nucleotide differences scattered along the length of the sequence. This could either be due to sequencing error or due to genuine variation shown within the clone. On analysing the QV value of the sequence the nucleotides showing variation were observed to have high QV value which helped to eliminate the possibility of sequencing error. Therefore it can be concluded that the variations are inherent in the sequence.

Since all the sequences were highly similar, one of the nucleotide sequence was randomly chosen to perform BLASTX analysis. On comparing with already reported sequences, the deduced amino acid sequence showed 92 % identity with CHS sequence gene reported from *Polygonum cuspidatum*, 91 % identity with *Fallopia multiflora* and 90% identity with *Fagopyrum tataricum*. Randomly four of the nucleotide sequences, namely, Pz101, Pz102, Pz104 and Pz106 were translated into corresponding amino acid sequences and multiple sequence alignment was performed with already reported CHS sequences. The amino acid sequences showed considerable similarity from 181th to 376th amino acid when aligned with previously reported CHS sequences. Pz 101 was observed to be shorter as compared to the other sequences. From the multiple sequence alignment it was inferred that approximately 166 amino acids in the 3' direction and approximately 22 amino acids in the 5' direction were to be deduced from further RLM-RACE studies.

To perform RLM-RACE, primer designing was performed manually by using the nucleotide sequences that corresponded to the amino acid regions showing similarity in MSA. Two sets of nested gene specific primers, each consisting of an outer primer and an inner primer were designed for 3' RLM-RACE and 5' RLM-RACE. PKS 2N and PKS 2 were designed for the outer and inner 5' RLM-RACE respectively. PKS 2 RACE-n1 and PKS 2 RACE-n2 were designed for the outer and inner 3' RLM-RACE respectively. However, only PKS 2 RACE-n1 primer was used for the 3' RLM-RACE as outer 3' RLM-RACE was not performed.

The processed RNA when subjected to 5' RLM-RACE analysis resulted in an intense and discrete amplicon of 400bp. The amplicon was cloned and PCR confirmation of the insert with T3-T7 primers, showed a 600bp product. On multiple sequence alignment the sequences showed a high degree of sequence homology. Since all the sequences were highly similar one of the sequence was randomly chosen for BLASTX analysis which revealed significant homology with already reported CHS genes such as Chalcone synthase from *Fagopyrum esculentum* and *Fagopurum tataricum*. Since all the sequences were highly similar one of the sequence was randomly chosen and translated to corresponding amino acid sequence. The longest ORF was merged with the amino acid corresponding to core amino acid sequence to obtain a 232 amino acid long sequence. On performing MSA of this sequence with already reported CHS amino acid sequences it was inferred that 166 amino acids were to be sequenced in the 3' end.

Following first strand cDNA synthesis, 3' cDNA end was isolated by using 3' RLM-RACE. Outer 3' RLM-RACE was performed using previously designed gene specific primer, PKS 2 RACE-n1. Outer 3' RLM RACE yielded a number of low intensity bands of different sizes, but the 700bp band was chosen as it is closest to the expected cDNA size that is to be sequenced. As suggested in the kit, inner 3' RLM RACE was not performed since the expected product was

obtained and further reactions would lead to reduction in product size. The amplicon was cloned and sequenced following the confirmation of insert.

The nucleotide sequence 3' PH5 on BLASTX analysis showed high degree of similarity with aromatic polyketide synthase from *Hypericum hookerianum*, naringenin chalcone synthase from *Hypericum androsaemum* and chalcone synthase from *Hypericum sampsonii*. The translated amino acid sequence was merged with the partial sequence obtained in the previous sections. The ORF was seen to encode protein of 398 amino acid with a calculated MW of 43.5kDa and a theoretical pI of 5.89.The BLASTX analysis of the full length cDNA sequence showed 89 % identity with naringenein chalcone synthase from *Hypericum androsaemum*, chalcone synthase from *Hypericum sampsonii* and aromatic polyketide synthase from *Hypericum hookerianum*. The CD-Search in NCBI was used to map conserved domains. It was observed that the active site consisting of the catalytic triad of CHS, Cys-His-Asn were conserved in corresponding positions. Also, eleven residues that form the product binding domain, eight residues composing malonyl CoA binding site and 31 residues forming dimer interface were observed to be conserved in the protein sequence elucidated. The PKS modeling was done using Swiss-Model server with *Medicago sativa* CHS (pdb accession no.: 1d6f) as template and visualized with PyMOL. The PyMOL data showed the presence of 40 α helices and 42 β strands with 67 turns. A phylogenetic tree showing the relationship of the deduced sequence with previously reported CHS sequences was generated using MEGA software. The sequence obtained from *P. rosea* was found to be grouped with PKS sequence previously reported from *Hypericum androsaemum* CHS *and Hypericum hookerianum* PKS.

As the sequence analysis strongly indicates that the isolated full length cDNA belongs to typical CHS class of genes, there could be additional PKSs of non-CHS class involved in the synthesis of plumbagin.

5.2 MOLECULAR MARKER ANALYSIS

5.2.1 DNA Isolation

Three different protocols were tried for the isolation of DNA from *P. rosea*. Murray and Thompson (1980) method was selected over one day (Suman *et al.*, 1999) protocol using cetyltrimethylammonium bromide (CTAB) extraction buffer and Gen Elute Plant Genomic DNA Miniprep kit, as it helped to reduce the RNA contamination and yielded higher concentration of DNA as compared to the other two methods. The $A_{260/280}$ of the isolated DNA was found to vary between 1.52 to 2.11. One of the major problems encountered in DNA isolation was high phenolics content which was eliminated by freshly adding 1.2% PVP (Polyvinyl pyrrolidone) to extraction buffer prior to DNA isolation. Also the RNA contamination was reduced by increasing the quantity of RNase enzyme from 5µl to 8µl followed by intermittent mixing during the 30 minutes incubation at 37°C.

5.2.2 ISSR Based Diversity Analysis

Plumbago rosea is over exploited for its root which is considered as its officinal part. Since the root is harvested destructively, collections from natural habitats have led to a depletion of the natural populations. This study which used 12 arbitrary primers and 8 accessions could help to understand the genetic diversity that exists among these accessions and thereby throw light on the diversity among the wild populations.

The ISSR based PCR employs longer primers which are characterised by higher annealing temperatures than Random Amplified Polymorphic DNA (RAPD) markers (Pomper *et al.* 2003). As a result, these markers are characterised by higher stringency and reproducibility. Usually, ISSR primers based on di-nucleotide repeats reveal high polymorphism (Blair et al. 1999; Joshi *et al.* 2000; Nagaoka and Ogihara 1997). The primers altogether generated 71

scorable bands, out of which 70 were polymorphic thereby showing a high degree of 98.6% polymorphism. The coefficient of genetic similarity ranged from 0.32 to 0.82 with mean value of 0.60 which suggests high variability. Other indices such as, percentage of polymorphic loci (P), observed number of alleles (na), expected number of alleles (ne), Shannon index of genetic diversity (I) and Nei's gene diversity (h) at the population level also demonstrated similar interpretations. High variability is normally associated with species which are highly outcrossing in nature. However *P. rosea* is propagated vegetatively due to lack of seeds. So the high levels of polymorphism detected in these vegetatively propagated accessions of *Plumbago rosea* may be attributed to the new gene combinations that the plant might have acquired in order to better adapt to changing environmental conditions.

A dendrogram was generated using WINBOOT software. The samples of *Plumbago rosea* clustered broadly under two major groups with an outlier. The first group consisted of two accessions of JNTBGRI clustered along with accessions from Chengannur, Ernakulam and Peroorkada with confidence limit of 47.6 to 90.2%. The second cluster consisted of one JNTBGRI accession clustered to the accession from Vellayani with a confidence limit of 69.7%. The accession from the distant Wayanad was found to form an outlier.

It can be concluded that conservation strategies may be designed to preserve the naturally present genetic diversity because loss of genetic diversity could lead to decrease in a species ability to survive (Ellstrand and Elam,1993;Milligan et al 1994,Reisch et al 2003). This study is significant as this is the first study to use ISSR molecular markers to evaluate inter population diversity of *Plumbago rosea*, which will be useful for further investigations.

5.3 PLUMBAGIN EXTRACTION AND ESTIMATION

Plumbagin extraction was performed using cold extraction with chloroform, which is the easiest of all the methods and comparatively less time consuming (Heble *et al.*, 1974).The extracts were used to perform TLC to determine the R_f (retention/ retardation factor) value of plumbagin and to study the purity of the concentrates. No impurities could be observed in the TLC samples and the R_f value was found to be 0.64which corresponded to previously reported value for benzene based solvent system. Spectrophotometric estimation of plumbagin at 415nm revealed that plumbagin content was found to vary between 0.2 % to 0.7 %. CSIR (1959) has reported the content to vary between 0.9-1%. In this study the maximum (0.7 %) was observed in the P3 accession collected from JNTBGRI and the least plumbagin content was estimated in P3 accession collected from Ernakulam. This variation could be due to the variation in locality, influence of edaphic factors like soil condition or seasonal influence that have been reported to influence plumbagin content besides genetic factors (CSIR, 1959). The relatively low content of plumbagin could be due to the collection of the roots during rainy conditions as it has been reported that plumbagin content is high in the roots of older plants grown in dry soil.

The HPLC data of TBG-102 was found to correlate with the spectrophotometric estimation. The plumbagin content in the crude extract of the sample was estimated to be 0.575 % when subjected to spectrophotometry while it was slightly less (5.05%) as expected when estimated with HPLC. This was because, HPLC is more sensitive off the two methods and the extract was quantified after preparative TLC based purification. The retention time and peaks generated by the sample as estimated from the HPLC data was similar to that obtained for the authentic sample, suggesting that the cold extraction method used for the plumbagin extraction was successful in isolating plumbagin without significant impurities.

On comparing the genetic diversity indices with plumbagin estimation data it was observed that P2 and P6 which showed identical genetic diversity indices also showed similar plumbagin content. Similar condition was reported in P4 and P5. However, this could not be considered as a correlation because, P3 and P7 which showed highly dissimilar plumbagin content also showed identical genetic diversity indices. Also, P1 and P8 which showed similar plumbagin content did not show identical genetic diversity indices. Therefore no significant correlation could be established between the genetic diversity data and plumbagin content. This suggests that the loci studied using ISSR markers did not influence plumbagin production in the plant.

SUMMARY

6. SUMMARY

The study entitled "Isolation and characterization of type III polyketide synthases from chethikoduveli (*Plumbago rosea* Linn.)" was conducted at the Biotechnology and Bioinformatics Division, JNTBGRI, during 2013-2014. The objective of the study was to isolate and sequence characterize type III polyketide synthase gene(s) involved in the biosynthesis of plumbagin in *Plumbago rosea.* Furthermore, chemical variation with respect to plumbagin content and genetic diversity using ISSR markers was also carried out in available accessions of *Plumbago rosea.*

RNA isolation from *Plumbago rosea* was carried out using Spectrum Plant Total RNA Kit which was found to produce higher yield and intact RNA as compared to TRIzol® reagent method which consistently yielded protein contaminated and degraded RNA.

The RNA was reverse transcribed and PCR amplified using already reported degenerate primers designed based on conserved regions of CHS genes. The resulted amplicon of 550bp was cloned and sequenced to yield the core sequence of PKS. This was followed by 5' and 3' RLM-RACE analysis using primers designed based on the core sequence. 5' and 3' RACE yielded 400bp and 700bp amplicons respectively. These were also cloned and sequenced. The full length cDNA sequence was obtained by comparing and aligning the sequences of 5' RACE, 3' RACE and core fragment.

The full length cDNA was found to be 1,197bp coding for 398 amino acids. The protein was calculated to have a molecular mass of 43.5kDa and a theoretical pI of 5.89. BLASTX analysis of the full length cDNA showed 89% identity with naringenin chalcone synthase from *Hypericum androsemum,* chalcone synthase from *Hypericum sampsonii* and aromatic polyketide synthase from *Hypericum hookerianum.*

DNA isolation was carried out in 8 accessions of *P. rosea* using Murray and Thompson (1980) method since it helped to reduce the RNA contamination and yielded higher concentration of DNA as compared to one day (Suman *et al.*, 1999) protocol using CTAB extraction buffer and Gen Elute Plant Genomic DNA Miniprep kit. Addition of PVP to extraction buffer prior to isolation was found to effectively counter the high phenolics content. Also the RNA contamination was significantly reduced by increasing the quantity and duration of RNase treatment.

ISSR based genetic diversity analysis revealed high degree of polymorphism (98.6 % polymorphism). The coefficient of genetic similarity ranging from 0.32 to 0.82 with mean value of 0.60 indicates high variability. Other indices such as, percentage of polymorphic loci (P), observed number of alleles (na), expected number of alleles (ne), Shannon index of genetic diversity (I) and Nei's gene diversity (h) at the population level also demonstrated similar interpretations.

Plumbagin extraction was performed using easy and comparatively less time consuming cold extraction method using chloroform. The extracts were used to perform TLC to determine the R_f value of plumbagin and to study the purity of the concentrates. No impurities could be observed during TLC and the R_f value was concluded to be 0.64. Spectrophotometric estimation of plumbagin at 415nm showed variation in plumbagin content ranging from 0.2 % to 0.7 %.

The type III PKS gene isolation and characterisation was an important step in studying the genes involved in controlling the molecular mechanism of plumbagin production. The sequence analysis strongly indicates that the isolated full length cDNA belongs to typical CHS class of genes and there could be additional PKSs of non-CHS class involved in the synthesis of plumbagin. The ISSR based genetic diversity study was a pioneering study in genetic diversity analysis in *Plumbago rosea.*

REFERENCES

7. REFERENCES

Abd-Elsalam, K.A. 2003. Bioinformatic tools and guideline for PCR primer design. *Afr. J. Biotechnol.* 2(5): 91-95.

Abe, I., Oguro, S., Utsumi, Y., Sano, Y., and Noguchi, H. 2005. Engineered biosynthesis of plant polyketides: chain length control in an octaketide-producing plant type III polyketide synthase. *J. Am. Chem. Soc.* 127: 12709-12716.

Abe, I., Sano, Y., Takahashi, Y., and Noguchi, Y. H. 2003. Site-directed mutagenesis of benzalacetone synthase: the role of Phe215 in plant type III polyketide synthases. *J. Biol. Chem.* 278: 25218-25226.

Abe, I., Takahashi, Y., Morita, H., and Noguchi, H. 2001. Benzalacetone synthase. A novel polyketide synthase that plays a crucial role in the biosynthesis of phenylbutanones in *Rheum palmatum*. *Eur. J. Biochem.* 268: 3354-3359.

Abe, I., Utsumi, Y., Oguro, S., Morita, H., Sano, Y., and Noguchi, H. 2005. A plant type III polyketide synthase that produces pentaketide chromone. *J. Am. Chem. Soc.* 127: 1362-1363.

Abe, I., Utsumi, Y., Oguro, S., and Noguchi, H. 2004. The first plant type III polyketide synthase that catalyzes formation of aromatic heptaketide. *FEBS. Lett.* 562: 171-176.

Abremski, K., Hoess, R., and Sternberg, N. 1983. Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination. *Cell* 32(4):1301-1311.

Ainsworth, C. 1994. Isolation of RNA from floral tissue of *Rumex acetosa* (Sorrel). *Pl. Mol. Biol. Rep.* 12: 198-203

Akiyama, T., Shibuya, M., Liu, H.M., and Ebizuka,Y. 1999. p-Coumaroyltriacetic acid synthase, a new homologue of chalcone synthase, from *Hydrangea macrophylla* var. *thunbergii*. *Eur. J. Biochem.* 263: 834-839.

Arezi, B. and Hogrefe, H.H. 2007. *Escherichia coli* DNA polymerase III epsilon subunit increases Moloney murine leukemia virus reverse transcriptase fidelity and accuracy of RT-PCR procedures. *Anal. Biochem.* 360(1): 84-91.

Arnheim, N. and Erlich, H. 1992. Polymerase chain reaction strategy. *Annu. Rev. Biochem.* 61(XIV): 131-156

Aslanidis, C. and de Jong, P. J. 1990. Ligation-independent cloning of PCR product (LIC-PCR). *Nucl. Acids Res.* 18: 6069-6074.

Austin, M.B., Bowman, M. E., Ferrer, J. L., Schröder, J., and Noel, J. P. 2004b. An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. *Chem. Biol.* 11: 1179-1194.

Austin, M.B., Izumikawa, M., Bowman, M.E., Udwary, D. W., Ferrer, J.L., Moore, B., and Noel, J. P. 2004a. Crystal structure of a bacterial type III polyketide synthase and enzymatic control of reactive polyketide intermediates. *J Biol Chem* 279: 45162-45174.

Austin, M.B. and Noel, J.P. 2003. The chalcone synthase superfamily of type III polyketide synthases. *Nat. Prod. Rep.* 20: 79-110.

Beerhues, L. 1996. Benzophenone synthase from cultured cells of *Centaurium erythraea*. *FEBS. Lett.* 383: 264-266.

Bergseid, M., Scott, B., Mathur, S., Nielson, K., Shoemaker, D., and Mathur, E. 1991. A high fidelity thermostable DNA polymerase isolated from *Pyrococcus furiosus*. *Strategies* 4: 34-35.

Bhargava, S.K. 1984. Effects of plumbagin on reproductive function of male dog. *Indian J. Exp. Biol.* 22: 153-156.

Bhattacharya, R., Ray, A., Gangopadhyay, M., and Bhattacharya, S. 2008. Study of *in vitro* conservation of *Plumbago indica*- a rare medicinal plant. *Plant Cell Biotechnol. Mol. Biol.* 8: 39-46.

Birnboim, H.C. and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7: 1513-1523.

Blair, M.W., Panaud, O., and McCouch, S.R. 1999. Intersimple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 98: 780-792.

Blumberg. 1987. Creating a ribonuclease-free environment. *Methods Enzymol.* 152: 20-24.

Borson, N.D., Sato, W.L., and Drewes, L.R. 1992. A lock-docking oligo (dT) primer for 5' and 3' RACE PCR. *PCR Methods Appl.* 2: 144-148.

Bringmann, G., Schlauer, J., Rückert, M., Wiesen, B., Ehrenfeld, K., Proksch, P., and Czygan, F.C. 1999. Host-derived acetogenins involved in the incompatible parasitic relationship between *Cuscuta reflexa* (Convolvulaceae) and *Ancistrocladus heyneanus* (Ancistrocladaceae). *Plant Biology* 1: 581-584.

Brooks, E.M., Sheflin, L.G., and Spaulding, S.W. 1995. Secondary structure in the 3'UTR of EGF and the choice of reverse transcriptases affect the detection of message diversity by RT-PCR. *BioTechniques* 19: 806-815.

Canter, P.H., Thomas, H., and Ernst, E. 2005. Bringing medical plants into cultivation opportunities and challenges for biotechnology. *Trends Biotechnol.* 23: 180-185.

Checker, R., Sharma, D., Sandur, S.K., Khanam, S., and Poduval, T.B. 2009. Anti-inflammatory effects of plumbagin are mediated by inhibition of NF-kappaB activation in lymphocytes. *Int. Immunopharmacol.* 9: 949-958.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., and Thompson, J.D. 2003. Multiple sequence alignment with the clustal series of programs. *Nucleic Acids Res.* 31: 3497-3500

Chetia, S. and Handique, P.J. 2000. A high frequency *in vitro* shoot multiplication of *Plumbago indica*- A rare medicinal plant *Curr. Sci.* 78: 1187-1189.

Chomczynski, P. 1993. A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *BioTechniques* 15: 532-537.

Chomczynski and Sacchi. 1987. Single step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.*162: 156-159.

Clark, J.M. 1988. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* 16: 9677-9686.

Claros, M.G. and Cánovas, F.M. 1999. RNA isolation from plant tissues: a practical experience for biological undergraduates. *Biochemical Education* 27: 110-113

CLONTECHniques. Technical Bulletin. January,1996, pp. 2-4.

Collard, B.C.Y., Jahufer, M.Z.Z., and Pang, E.C.K. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assissted selection for crop improvement: The basic concepts. *Euphytica* 142: 169-196.

Costa, G. L., Grafsky, A., and Weiner, M.P. 1994. Cloning and analysis of PCR-generated DNA fragments. *PCR Meth. Appl.* 3: 338-345.

Costa, G.L. and Weiner, M.P. 1994. Polishing with T4 or Pfu polymerase increases the efficiency of cloning of PCR fragments. *Nucleic Acids Res.* 22: 2423.

Crouch, I.J., Finnie, J.F., and Staden, J. 1990. Studies on the isolation of plumbagin from *in vitro* and *in vivo* grown *Drosera* species. *Plant Cell Tiss. Org. Cult.*21: 79-82.

CSIR. 1959. The wealth of Indian-Raw Materials Vol. VIII. Council of Scientific and Industrial Research, New Delhi, 162-163.

de Vicente, M.C., and Fulton, T. 2003. *Using molecular marker technology in studies on plant genetic diversity.* [Online] IPGRI, Rome, Italy and Institute for Genetic Diversity, Ithaca, New York, USA. Available at: http://www2.bioversityinternational.org/Publications/Molecular_Markers_Volume_1_en/index.asp [Accessed 4 May 2010].

Desta, B. 1993. Ethiopian traditional herbal drugs. Part II. Antimicrobial activity of 63 medicinal plants. *J. Ethnophramacol.* 39: 129-139.

DeStefano, J.J., Buiser, R.G., Mallaber, L.M., Myers, T.W., Bambara, R.A., and Fay, P.J. 1991. Polymerization and RNase H activities of the reverse transcriptases from avian myeloblastosis, human immunodeficiency, and Moloney murine leukemia viruses are functionally uncoupled. *J. Biol. Chem.* 266: 7423-7431.

Dey, P.M. and Harbone, J.B. 1989. Phenolic metabolism. Methods in Plant Biochemistry. In: Dey, P.M. and Harborne, J.B. (ed.), Academic Press.London. pp.413.

Didry, N., Dubrevil, L., and Pinkas, M. 1994. Activity of anthraquinonic and naphthoquinonic compounds on oral bacteria. *Pharmazie.* 49: 681-683.

Dinda, B., Hajra, A.K., and Chel, G. 1997. Naphthoquinones of *Plumbago* species a review. *J. Indian Chem. Soc.*74: 974-979.

Durand, R. and Zenk, M.H. 1971. Biosynthesis of plumbagin (5-hydroxy-2-methyl-1, 4-naphthoquinone) via the acetate pathway in higher plants. *Tetrahedron Lett.* 32: 3009-3012.

Durand, R. and Zenk, M.H. 1974. Enzymes of the homogentisate ring-cleavage pathway in cell suspension cultures of higher plants. *Phytochemistry* 13: 1483-1492.

Durbin, M.L., McCaig, B., and Clegg, M.T. 2000. Molecular evolution of the chalcone synthase multigene family in the morning glory genome. *Plant Mol. Biol.* 42: 79-92.

Eckermann, C., Schröder, G., Schmidt, J., Strack, D., Edrada, R. A., Helariutta, Y., Elomaa, P., Kotilainen, M., Kilpelainen, I., Proksch, P., Teeri, T.H., and Schröder, J. 1998. New pathway to polyketides in plants. *Nature* 396: 387-390.

Ellstrand, N.C. and Elam, D.R. 1993. Population genetic consequences of small population size: implications for plant conservation. *Ann.Rev.Ecol.Syst.*24: 217-242.

Erlich, H. A., Gelfand, D., and Sninsky, J. J. 1991. Recent Advances in the Polymerase Chain Reaction. *Science* 252: 1643-1651.

Etebu, E. 2008. Detection and quantification of the pea footrot disease pathogen (*Nectria haematococca*) in agricultural soils: A potential model for disease prediction. PhD thesis. The University of Sheffield, Sheffield, UK, 258pp.

Etebu, E. 2013. Potential panacea to the complexities of Polymerase Chain Reaction (PCR). *Ad. Sci. Tech.* 13: 53-59

Etebu, E. and Osborn, A.M. 2009. Molecular assays reveal the presence and diversity of genes encoding pea footrot pathogenicity determinants in *Nectria haematococca* and in agricultural soils. *J. Appl. Microbiol.* 106 (5): 1629-1639

Farrell, R.E. 2009. *RNA Methodologies: Laboratory Guide for Isolation and Characterization.* Academic Press, Massachusetts, pp. 71-73

Ferrer, J.L., Jez, J.M., Bowman, M.E., Dixon, R.A., and Noel, J.P. 1999. Structure of chalcone synthase and the molecular basis of plant polyketide biosynthesis. *Nat. Struct. Biol.* 6: 775-784.

Fischbach, M.A. and Walsh, C.T. 2006. Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: Logic, machinery, and mechanisms. *Chem. Rev.* 106: 3468-3496.

Flores-Sanchez, I.J. and Verpoorte, R. 2009. Plant polyketide synthases: A fascinating group of enzymes. *Plant Physiol. Biochem.* 47: 167-174.

Freeman,W.M., Walker, S.J., and Vrana, K.E. 1999. Quantitative RT-PCR: pitfalls and potential. *BioTechniques* 26:112-115.

Frohman, M.A. 1990. *PCR Protocols: A Guide to Methods and Applications.* In: Innis, M. A., Gelfand, D. H., Sninsky, J. J., and White, T. J. (eds), Academic Press, San Diego,28p.

Frohman, M.A., Dush, M.K., and Martin, G.R. 1988. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc Natl Acad Sci USA* 85: 8998-9002

Frohman, M.A. and Martin, G.R. 1989. Rapid amplification of cDNA ends using nested primers. *Techniques* 1: 165-173.

Gangopadhyay, M., Dewanjee, S., and Bhattacharya, S. 2011. Enhanced plumbagin production in elicited *Plumbago indica* hairy root cultures. *J. Biosci. Bioeng.* 111: 706-710.

Gangopadhyay, M., Sircar, D., Mitra, A., and Bhattacharya, S. 2008. Hairy root culture of *Plumbago indica* as a potent source for harvesting plumbagin. *Biol. Planta* 52: 533-537.

Gerard, C.J., Olsson, K., Ramanathan, R., Reading, C., and Hanania, E.G. 1998. Improved quantitation of minimal residual disease in multiple myeloma using real-time polymerase chain. *Cancer Res.* 58(17): 3957-3964.

Gibbs, R.A. 1990. DNA amplification by the polymerase chain reaction. *Anal. Chem.* 62: 1202-1214.

Gonçalves, S., Ferraz, M., and Romano, A. 2009. Phytotoxic properties of *Drosophyllum lusitanicum* leaf extracts and its main compound plumbagin. *Sci. Hortic.*122: 96-101.

Grieve, M. 1995. A modern herbal *Plumbago* [on-line]. Available: http://www.botanical.com/botanical/mgmh/p/plumba54.html [25 Dec. 2001].

Gupta, M., Chyi, Y.S., Romero Severson, J., and Owen, J.L. 1994. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. *Theor. Appl. Genet.* 89: 998-1006.

Gupta, M.M., Verma, R.K., Uniyal, G.C., and Jain, S.P. 1993. Determination of plumbagin by normal phase high performance liquid chromatography. *J. Chromatogr.* 637: 209-212.

Heble, M.R., Narayanaswammi, S., and Chadha, M.S. 1974. Tissue differentiation and plumbagin synthesis in variant cell strains of *Plumbago zeylanica. Plant Sci. Lett.* 2: 405-409.

Helariutta,Y., Elomaa, P., Kotilainen, M., Griesbach, R.J., Schröder, J., and Teeri, T. H. 1995. Chalcone synthase-like genes active during corolla development are differentially expressed and encode enzymes with different catalytic properties in *Gerbera hybrida* (Asteraceae). *Plant Mol. Biol.* 28: 47-60.

Heubl, G., Bringmann, G., and Meimberg, H. 2006. Molecular phylogeny and character evolution of carnivorous plant families in Caryophyllales-revisited. *Plant Biol.* 8: 821-830.

Higgins, D.G. and Sharp, P.M. 1988. CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* 73: 237-244

Hill, P.J. and Stewart, G.S.A.B. 1992. The polymerase chain reaction in molecular and micro-biology. *Biotechn. Genet. Eng. Rev.* 10: 343-378.

Holton, T.A. and Graham, M.W. 1991. A simple and efficient method for direct cloning of PCR products using ddT-tailed vectors. *Nucleic Acids Res.* 19: 1156.

Hopwood, D.A. and Sherman, D.H. 1990. Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. *Annu. Rev. Genet.* 24: 37-66.

Ichihara, A., Ubukata, M., and Chadba, M.S. 1980. Synthesis of plumbagin by the retro- Diels- Alder reaction *Agric. Biol. Chem.* 44: 211-213.

Isvett, J.F.S. and Verpoorte, R. 2009. Plant polyketide synthases: a fascinating group of enzymes. *Plant Physiol. Biochem.* 47: 167-174.

Itoigawa, M., Takeya, K., and Furukawa, H. 1991. Cardiotonic action of plumbagin on guinea-pig papillary muscle. *Planta Med.* 57: 317-319.

Jadhav, S., Phapale, P.B., Thulasiram, H.V., and Bhargava, S. 2013. Polyketide synthesis in tobacco plants transformed with a *Plumbago zeylanica* type III hexaketide synthase. *Phytochemistry* 98: 92-100.

Jaya, S. 1999. Development of plant tissue culture of *Plumbago rosea.* Linn. for enhanced production of secondary metabolites. PhD thesis, Cochin University of Science and Technology, Cochin, 18p.

Jez, J.M., Austin, M.B., Ferrer, J.L., Bowman, M.E. Schro¨der, J., and Noel, J.P. 2000. Structural control of polyketide formation in plant-specific polyketide synthases. *Chem. Biol.* 7: 919-930.

Jez, J.M., Bowman, M.E., and Noel, J.P. 2002. Expanding the biosynthetic repertoire of plant type III polyketide synthases by altering starter molecule specificity. *Proc Natl Acad Sci USA.* 99: 5319-5324.

Jez, J.M., Ferrer, J.L., Bowman, M.E., Austin, M.B., Schröder, J., Dixon, R.A., and Noel, J.P. 2001. Structure and mechanism of chalcone synthase-like polyketide synthases, *J. Ind. Microbiol. Biotechnol.* 27: 393-398.

Jez, J.M., Ferrer, J.L., Bowman, M.E., Dixon, R.A., and Noel, J.P. 2000. Dissection of malonylcoenzyme: a decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochemistry* 39: 890-902.

Jindaprasert, A., Springob, K., Schmidt, J., De-Eknamkul, W., and Kutchan, T.M. 2008. Pyrone polyketides synthesized by a type III polyketide synthase from Drosophyllum lusitanicum. *Phytochemistry.* 69: 3043-3053.

Joshi, M. and Deshpande, J. D. 2010. Polymerase chain reaction: methods, principles and application. *IJBR.* 1(5): 81-97.

Joshi, N.K. and Sehnal, F. 1989. Inhibition of ecdysteroid production by plumbagin in *Dysdercus cingulatus. J. Insect Physiol.* 35: 737-741.

Joshi, S.P., Gupta, V.S., Aggarwal, R.K., Ranjekar, P.K., and Brar, D.S. 2000. Genetic diversity and phylogenetic relationship as revealed by inter-simple

sequence repeat (ISSR) polymorphism in the genus *Oryza*. *Theor. Appl. Genet.* 100: 1311-1320.

Jung, V., Pestka, S.B., and Pestka, S. 1993. Cloning of polymerase chain reaction-generated DNA containing terminal restriction endonuclease recognition sites. *Method Enzymol.* 208: 357- 362.

Rathnayake, K. 2013. *Plumbago rosea* [Photograph].Retrived 23 August, 2014, from Herbal plants Sri Lanka blogspot http://herbalplantslanka. blogspot.in/2013/05/rath-nitulplumbago-indica_10.html

Kimmel, A.R. and Berger, S.L.1987. Preparation of cDNA and the generation of cDNA libraries. *Methods Enzymol.* 152: 307-316.

Kini, D.P., Pandey, S., and Shenoy, B.D. 1997. Antitumor and antifertility activities of plumbagin controlled release formulations. *Indian J. Exp. Biol.* 35: 374-379.

Kirthikar, K.R. and Basu, B.D. 1975. Plumbagineae. In: Blatter, B., Caius, J.F., and Mhaskar, K.S. (eds.) *Indian Medicinal Plants*. M/s Periodical Experts, Delhi, India, pp. 1465-1471.

Kitanov, G.M. and Pashankov, P.P. 1994. Quantitative investigation on dynamics of plumbagin in *Plumbago indica* L. root and herb by HPLC. *Pharmazie.* 49: 642.

Klingauf, P., Beuerle, T., Mellenthin, A., El-Moghazy, S.A.M., Boubakir, Z., and Beerhues, L. 2005. Biosynthesis of the hyperforin skeleton in *Hypericum calycinum* cell cultures. *Phytochemistry* 66: 139-145.

Komaraiah, P., Kavi Kishor, P.B., and Ramakrishna, S.V. 2001. Production of plumbagin from cell cultures of *Plumbago rosea* L. *Biotechnol. Lett.* 23: 1269-1272.

Komariah, P., Naga Amrutha, R., Kavi Kishor, P.B., and Ramakrishna, S.V. 2002. Elicitor enhanced production of plumbagin in suspension cultures of *Plumbago rosea* L. *Enzyme Microb. Technol.* 31: 634-639.

Kotewicz, M.L., Sampson, C.M., D'Alessio, J.M., and Gerard, G.F. 1988. Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. Nucleic Acids Res. 16: 265-277.

Kreuzer, K.A., Lass, U., Landt, O., Nitsche, A., Laser, J., Ellerbrok, H., Pauli, G., Huhn, D., and Schmidt, C.A. 1999. Highly sensitive and specific fluorescence reverse transcription-PCR assay for the pseudogene-free detection of beta-actin transcripts as quantitative reference. *Clin. Chem.* 45: 297-300.

Krishnan, S. 2003. *Plumbago indica.* Available: http://Butterfly Pavilion Horticulture.htm [10 Jan 2003].

Kubo, I., Uchida, M., and Klocke, J. 1983. An insect ecdysis inhibitor from the African medicinal plant *Plumbago capensis* (Plumbaginaceae); a naturally occurring chitin synthetase inhibitor. *Agric. Biol. Chem.* 47: 911-913.

Kuo, P.L., Hsu, Y.L., and Cho, C.Y. 2006. Plumbagin induces G_2-M arrest and autophagy by inhibiting the AKT/mammalian target of rapamycin pathway in breast cancer cells. *Mol. Cancer Ther.* 5(12): 3209-3221.

Kurian, A. and Sankar, M.A. 2007. *Medicinal Plants.* Horticulture Science Series, New India Publishing Agency, New Delhi, pp.143-123.

Lanz, T., Tropf, S., Marner, F.J., Schröder, J., and Schröder, G. 1991. The role of cysteines in polyketide synthases: site-directed mutagenesis of resveratrol and chalcone synthases, two key enzymes in different plant-specific pathways. *J. Biol. Chem.* 266: 9971-9976.

Likhitwitayawuid, K., Kaewamatawong, R., Ruangrungsi, N., and Krungkrai, J. 1998. Antimalarial naphthoquinones from *Nepenthes thorelii*. *Planta Med.* 64: 237-241.

Linderman, J.R. and Greene, A.S. 2001. Distribution of agiotensin II receptor expression in the microcirculation of striated muscle. *Microcirculation* 8(4): 275-81

Linssen, B., Kinney, R.M., Aguilar, P., Russell, K.L., Watts, D.M., Kaaden, O.R., and Pfeffer, M. 2000. Development of reverse transcription- PCR assay specific for detection of equine encephalitis viruses. *J. Clin. Microbiol.* 38: 1527-1535.

Liu, B., Falkenstein-Paul, H., Schmidt, W., and Beerhues, L. 2003. Benzophenone synthase and chalcone synthase from *Hypericum androsaemum* cell cultures: cDNAcloning, functional expression and site-directed mutagenesis of two polyketide synthases. *Plant J.* 34: 847-855.

Logemann, J., Schell, J., and Willmitzer, L. 1987. Improved method for the isolation of RNA from plant tissues. *Anal. Biochem.* 163: 16-20.

Loomis, M.D. 1974. Overcoming problems of phenolics and quinines in the isolation of plant enzymes and organelles. *Methods Enzymol.* 31: 528-544

Lukacin, R., Schreiner, S., and Matern, U. 2001. Transformation of acridone synthase to chalcone synthase. *FEBS. Lett.* 508: 413-417.

Mallavadani, U.V., Sahu, G., and Muralidhar, J. 2002. Screening *Plumbago* species for bioactive marker plumbagin. *Pharm. Biol.* 40: 508-511.

Martin, J.P. and Sanchez Yelamo, M.D. 2000. Genetic relationships among species of the genus *Diplotaxis* (Brassicaceae) using inter-simple sequence repeat markers. *Theor. Appl. Genet.* 101: 1234-1241.

Maruyama, K. and Sugano, S. 1994. Oligo-capping: a simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* 138: 171-174

Matousek, J., Vrba, L., Skopek, J., Orctova, L., Pesina, K., Heyerick, A., Baulcombe, D., and De Keukeleire, D. 2006. Sequence analysis of a "true' chalcone synthase (*chs H1*) oligofamily from hop (*Humulus lupulus* L.) and PAP1 activation of *chs* H1 in heterologous systems. *J. Agric. Food Chem.* 54:7606-7615.

Miller, J.H. 1992. A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria. CSH Laboratory Press, Cold Spring Harbor, NY.

Milligan, B.G., Leebens-Mack J., and Strand, A.E.1994. Conservation genetics: beyond the maintenance of marker diversity. *Mol.Ecol.*3: 423-435.

Morita, H., Tanio, M., Kondo, S., Kato, R., Wanibuchi, K., Noguchi, H., Sugio, S., Abe, I., and Kohno, T. 2008. Crystallization and preliminary crystallographic analyses of a plant type III polyketide synthase that produces benzalacetone. *Acta. Crystallogr. F.* 64: 304-306.

Mullis, K.B. and Faloona, F.A. 1987. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol.* 155: 335- 350

Murray, M.G. and Thomson, W.F. 1980. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8: 4321-4325.

Nagaoka, T. and Ogihara, Y. 1997. Applicability of inter-simple sequence repeat polymorphisms in wheat for use as DNA markers in comparison to RFLP and RAPD markers. *Theor. Apppl. Genet.* 94: 597-602.

Neto, E.D., Harrop, R., Correa-Oliveira, R., Wilson, R.A., Pena, S.D., and Simpson, A.J. 1997.Minilibraries constructed from cDNA generated by

arbitrarily primed RT-PCR: an alternative to normalized libraries for the generation of ESTs from nanogram quantities of mRNA. *Gene* 186: 135-142.

Newton, C.R. and Graham, A. 1994. *PCR*. BIOS Scientific Publishers Limited, Oxford, 23p.

Nguyen, A.T., Malonne, H., Duez, P., Fastre, R. V., Vanhaelen, M., and Fontaine, J. 2004. Cytotoxic constituents from *Plumbago zeylanica. Fitoterapia* 75: 500-504.

Noor Adila, A.K., Farah Diba, A.B., Zamri, Z., Wan Mohtar, W.Y., Aidil, A.H., Mahadi, N.M., Murad, A.M.A. 2007. Comparison of methods for isolating high quality DNA and RNA from an oleaginous fungus *Cunninghamella bainieri* strain 2a1. *Malays. J. Microbiol.* 3: 7-13

Okada, Y. and Ito, K. 2001. Cloning and analyses of valerophenone synthase gene expressed specifically in lupulin gland of Hop (*H. lupulus* L.). *Biosci. Biotechnol. Biochem.* 65:150-155.

Oliner, J.D., Kinzler, K.W., and Vogelstein, B. 1993. *In vivo* cloning of PCR products in *E. coli. Nucl. Acids Res.* 21: 5192-5197.

Paiva, S.R., de Figueiredo, M.R., Aragão, T.V., and Kaplan, M.A.C. 2003. Antimicrobial activity *in vitro* of plumbagin isolated from *Plumbago* species. *Mem. Inst. Oswaldo Curz.* 98: 959-961.

Pakulski, G. and Budzianowski, J. 1996. Ellagic acid derivatives and naphthoquinones of *Dionaea muscipula* from *in vitro* cultures. *Phytochemistry* 4: 775-778.

Panichayupakaranant, P. and Tewtrakul, S. 2002. Plumbagin production by root cultures of *Plumbago rosea. Electron. J. Biotechnol.* 5: 228-232.

Parimala, R. and Sachdanandam, P. 1993. Effect of plumbagin on some glucose metabolizing enzymes studied in rats in experimental hepatoma. *Mol. Cell. Biochem.* 12: 59-63.

Pomper, K. W., Crabtree, S., Brown, S.P., Jones, S.C., Bonney, T.M., Layne, D.R. 2003. Assessment of genetic diversity of pawpaw varieties with inter-simple sequence repeat markers. *J. Amer. Soc. Hort. Sci.* 128: 521-525.

Porebski, S., Bailey, L.G., and Baum, B.R. 1997. Modification of CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Mol. Biol. Reptr.* 15: 8-1.

Premkumari, P., Rathinam, K., and Santhakumari, G. 1977. Antifertility activity of plumbagin. *Ind. J. Med. Res.* 65: 829-838.

Radhakrishnan, E.K., Sivakumar, K.C., and Soniya, E.V. 2009. Molecular Characterization of Novel form of Type III Polyketide Synthase from *Zingiber Officinale* Rosc. and its Analysis using Bioinformatics Method. *J. Proteomics Bioinform.* 2: 310-315.

Raiber, S., Schröder, G., and Schröder, J. 1995. Molecular and enzymatic characterization of two stilbene synthases from Eastern white pine (*Pinus strobus*): a single Arg/His difference determines the activity and the pH dependence of the enzymes. *FEBS. Lett.* 361: 299-302.

Rappolee, D.A., Mark, D., Banda, M.J., and Werb, Z. 1988. Wound macrophages express TGF-alpha and other growth factors *in vivo*: analysis by mRNA phenotyping. *Science* 241: 708-712.

Reisch, C., Poschlod, P., Wingender, R. 2003. Genetic variation of *Saxifraga paniculata* Mill. (Saxifragaceae): molecular evidence for glacial relict endemism in central Europe. *Biol.J.Linn.Soc.* 80: 11-21.

Rischer, H., Hamm, A., and Bringmann G. 2002. *Nepenthes insignis* uses a C2-portion of the carbon skeleton of l-alanine acquired via its carnivorous organs, to build up the allelochemical plumbagin. 2002. *Phytochemistry* 59: 603-609.

Rolfs, C.H. and Kindl, H. 1984. Stilbene synthase and chalcone synthase: Two different constitutive enzymes in cultured cells of *Picea excelsa. Plant Physiol.* 75: 489-492.

Rose, T., Schultz, E., Henikoff, J., Pietrokovski, S., McCallum, C., and Henikoff, S. 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res.* 26(7):1628-1635.

Saiki, R., Gelfand, D., Stoffel, S., Scharf, S., Higuchi, R., Horn, G., Mullis, K., and Erlich, H. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239 (4839): 487-491

Salzman, R.A., Fujita, T., Salzman, K.Z., Hasegawa, P.M., and Bressan R.A. 1999. An improved RNA isolation method for plant tissues containing high levels of phenolic compounds or carbohydrates. *Plant Mol. Biol. Rep.* 17: 11-17.

Samappito, S., Page, J., Schmidt, J., De-Eknamkul, W., and Kutchan, T.M. 2002. Molecular characterization of root-specific chalcone synthases from *Cassia alata. Planta* 216: 64-71.

Samappito, S., Page, J.E., Schmidt, J., De-Eknamkul, W., and Kutchan, T.M. 2003. Aromatic and pyrone polyketides synthesized by a stilbene synthase from *Rheum tataricum. Phytochemistry* 62: 313-323.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. CSH Laboratory Press, Cold Spring Harbor, NY.

Sankaranarayanan, R., Saxena, P., Marathe, U., Gokhale, R.S., Shanmugam, V. M., and Rukmini, R. 2004. A novel tunnel in mycobacterial type III polyketide synthase reveals the structural basis for generating diverse metabolites. *Nat. Struct. Mol. Biol.* 11: 894-900.

Santos, C.F.D., Sakai, V.T., Machado, M.A.D.A.M., Schippers, D.N., and Greene, A.S. 2004. Reverse transcription and polymerase chain reaction: principles and applications in dentistry. *Journal of Applied Oral Science 12:*1-11.

Saralamp, P., Chuakul, W., Temsiririrkkul, R., and Clayton, T. 1996. *Medicinal Plants in Thailand.* Amarin Printing and Publishing Public Co., Ltd., Bangkok, 30p.

Sarasan, V., Cripps, R., Ramsay, M., Atherton, C., McMichen, M., Prendergast, G., and Rowntree, J. 2006. Conservation: *In vitro* of threatened plants-progress in the past decade. *In Vitro Cell Dev. Biol. Plant* 42: 206-214.

Satheesh Kumar, K., and Bhavanandan, K. V. 1988. Micropropagation of *Plumbago rosea* Linn. *Plant Cell Tiss. Organ Cult.* 15: 275-278.

Schmelzer, G.H. and Gurib-Fakim, A. 2008. Plant Resources of Tropical Africa 11, Medicinal plants 1, PROTA foundation, Wageningen, Netherland.

Schröder, J. 1997. A family of plant-specific polyketide synthases: Facts and predictions. *Trends Plant Sci.* 2: 373-378.

Schröder, J. 1999. The chalcone/stilbene synthase-type family of condensing enzymes. In: Sankawa, U. (ed.), *Comprehensive Natural Products Chemistry* (Vol. 1). Elsevier Science, Oxford, pp. 749-771.

Schröder, J. 2000. The family of chalcone synthase-related proteins: Functional diversity and evolution. In: Romeo, J.T., Ibrahim, R. K., Varin, L., and De Luca, V. (eds), *Evolution of Metabolic Pathways* (Vol. 34.). Pergamon Press, Amsterdam, pp. 55-89.

Schröder, G., Brown, J.W.S., and Schröder, J. 1988. Molecular analyses of resveratrol synthase: cDNA, genomic clones and relationship with chalcone synthase. *Eur. J. Biochem.* 172: 161-169.

Schröder, J., Raiber, J.S., Berger, T., Schmidt, A., Schmidt, J., Soares-Sello, A. M., Bardshiri, E., Strack, D., Simpson, T.J., Veit, M., and Schröder, G. 1998. Plant polyketide synthases: a chalcone synthase-type enzyme which performs a condensation reaction with methylmalonyl-CoA in the biosynthesis of C-methylated chalcones. *Biochemistry* 37: 8417-8425.

Schröder, J. and Schröder, G. 1990. Stilbene and chalcone synthases: Related enzymes with functions in plant-specific pathways. *Z Naturforsch C.* 45: 1-8.

Schöppner, A. and Kindl, H. 1984. Purification and properties of a stilbene synthase from induced cell suspension cultures of peanut. *J. Biol. Chem.* 259: 6806-6811.

Sellon, R.K. 2003. Update on molecular techniques for diagnostic testing of infectious disease. *Vet. Clin. N. Am-Small. 33*(4): 677-693.

Shaefer, B. 1995. Revolution in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. *Analytical Biochem.* 227: 255-273.

Sheffield, V.C., Cox, D.R., Lerman, L.S., and Myers, R.M. 1989. Attachment of a 40 base pair G+C rich sequence (GC-clamp) to genomic DNA fragments by polymerase chain reaction results in improved detection of single base changes. *Proc. Natl. Acad. Sci.* 86: 232-236.

Shuldiner, A.R., Scott, L.A., and Roth, J. 1990. PCR-induced (ligase-free) subcloning: a rapid reliable method to subclone polymerase chain reaction (PCR) products. *Nucl. Acids Res.* 18: 1920

Shuman, S. 1994. Two classes of DNA end-joining reactions catalyzed by vaccinia topoisomerase I. *J Biol Chem.* 269(51): 32678-84.

Sibhatu, M. B. 2003. Cloning and Sequencing of Glucosyltransferase (GT) Genes from Grapefruit (*Citrus paradisi*) Seedlings; Searching for 7-O-GT. M.Sc. thesis. East Tennessee State University, p.164

Sigma-Aldrich. 2014. Plumbagin Structure [on-line]. Available Source: https://www.sigma-aldrich.com [30 July 2014].

Somma, M. 2004. Extraction and purification of DNA (Session 4). – In: Querci M., Jermini, M. and Van Den Eede, G. (eds.), Training course on – the analysis of food samples for the presence of genetically modified organisms: user manual.

Spencer, G.F., Tjarks, L.W., England, R.E., and Seest, E.P. 1986. The effect of naturally occurring naphthoquinones on velvetleaf (*Abutilon theophrasti*) germination. *J. Nat. Prod.* 49: 530–533.

Springob, K., Samappito, S., Jindaprasert, A., Schmidt, J., Page, J.E., De-Eknamkul, W. and Kutchan, T.M. 2007. A polyketide synthase of *Plumbago indica* that catalyzes the formation of hexaketide pyrones. *FEBS. Journal* 274: 406-417.

Staunton, J. and Weissman, K.J. 2001. Polyketide biosynthesis: A millennium review. *Nat. Prod. Rep.* 18: 380-416.

Suggs, S.V., Wallace, R.B., Hirose, T., Kawashima, E.H., and Itakura, K. 1981. Use of synthetic oligonucleotides as hybridization probes: isolation of cloned cDNA sequences for human beta 2-microglobulin. *Proc. Natl. Acad. Sci. U.S.A.* 78: 6613-6617

Suh, D.Y., Kagami, J., Fukuma, K., and Sankawa, U. 2000. Evidence for catalytic cysteine–histidine dyad in chalcone synthase. *Biochem. Biophys. Res. Commun.* 275: 725-730.

Suman P.S., Shasany, A.K., Darokar, M.P., and Kumar, S. 1999. Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. *Plant Mol. Biol. Rep.*17: 1-7.

Sutton, M.D. and Walker, G.C. 2001. Managing DNA polymerases: coordinating DNA replication, DNA repair, and DNA recombination. *Proc. Natl. Acad. Sci.* 98 (15): 8342-8349

Suzuki, Y., Yoshitomo, K., Maruyama, K., Suyama, A., and Sugano, S. 1997. Construction and characterisation of a full length enriched amd a 5'- end enriched cDNA library. *Gene* 200: 149-156.

Taguchi, C., Taura, F., Tamada, T., Shoyama, Y., Shoyama, Y. Tanaka, H., Kuroki, R., and Morimoto, S. 2008. Crystallization and preliminary X-ray diffraction studies of polyketide synthase-1 (PKS-1) from *Cannabis sativa*. *Acta. Crystallogr. F.* 64: 217-220.

Tai, H. H., Pelletier, C., and Beardmore, T. 2004. Total RNA isolation from *Picea mariana* dry seed. *Plant Mol. Biol. Rep.* 22: 93a-93e.

Teuscher, E. 1970. *Pharmakognosie*. Part II. Akademie Verlag. Berlin. pp 245.

Thomson, R.H. *Naturally occurring napthoquinones*. 2nd ed. London: Academic Press, 1971.

Tokunaga, T., Takadab, N., and Ueda, M. 2004. Mechanism of antifeedant activity of plumbagin, a compound concerning the chemical defense in carnivorous plant. *Tetrahedron Lett.* 45: 7115-7119.

Tropf, S., Kärcher, B., Schröder, G., and Schröder, J. 1995. Reaction mechanisms of homodimeric plant polyketide synthases (stilbene and chalcone

synthase): a single active site for the condensing reaction is sufficient for synthesis of stilbenes, chalcones and 6'-deoxychalcones. *J. Biol. Chem.* 270: 7922-7928.

Tropf, S., Lanz, T., Schröder, J., and Schröder, G. 1994. Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. *J. Mol. Evol.* 38: 610-618.

Tsumura, Y., Ohba, K., and Strauss, S.H. 1996. Diversity and inheritance of inter-simple sequence repeat polymorphisms in Douglas fir (*Pseudotsuga menziesii*) and sugi (*Cryptomeria japonica*). *Theor. Appl. Genet.* 92: 40-45.

Ullmann, A., Jacob, F., and Monod, J. 1967. Characterization by *in vitro* complementation of a peptide corresponding to an operator-proximal segment of the β-galactosidase structural gene of *Escherichia coli*. *J. Mol.Biol. 24*: 339-343.

Van Steenis, C.G.G.J. 1949. Plumbaginaceae. In: C.G.G.J., Van Steenis (ed.), *Flora Malesiana.* P. Noordhoff, Groningen, pp.107-112.

Villavicencio, M.A. and Perez-Escandon, B.E. 1992. Plumbagin activity (from *Plumbago pulchella* Boiss. Plumbaginaceae) as a feeding deterrent for three species of Orthoptera. *Folia Entomol. Mex.* 86: 191-198.

Wang, X., Tian, W., and Li, Y. 2008. Development of an efficient protocol of RNA isolation from recalcitrant tree tissues. *Mol. Biotechnol.* 38: 57-64.

Warrier, P.K., Nambiar V.P.K., and Ramankutty C. 1997.*Indian Medicinal Plants, A Compendium of 500 Species.* In: Warrier, P.K., Nambiar V.P.K, and Ramankutty C. (eds.),Vol. IV. Madras, India, Orient Longman Ltd., 321p.

Watson, L. and Dallwitz, M.J. 1992. *The Families of Flowering Plants: Descriptions, Illustrations, Identification and Information Retrieval.*

Version: 19th August 1999. Available: http://biodiversity.uno.edu/delta/ [Febuary 20, 2003].

Weishing, K., Nybom, H., Wolff, K., and Meyer, W. 1995. *DNA isolation and purification. In: DNA Fingerprinting in Plants and Fungi*, pp44-59.CRC Press, Boca Raton, Florida.

Weising, K., Nybom, H., Wolff, K., and Kahl, G. 2005. *DNA fingerprinting in plants: principles, methods, and applications* (2nd Ed.). Taylor & Francis, Florida, pp. 1-20.

Welsh, M.J., Kaufman, P.B., and Zhang, H.H. 1997. *Methods in Gene Biotechnology*. CRC Press, New York, pp. 247-250.

Wilfred, F., Röling, M., and Head, I.M. 2005. Prokaryotic systematics: PCR and sequence analysis of amplified 16S rRNA genes. In: Osborn, A.M. and Smith, C.J. (eds.), *Molecular Microbial Ecology* (1st Ed.). Taylor and Francis, UK. pp. 25-65.

Wolf, B., Lesnaw, J.A., and Reichmann, M.E. 1970. A mechanism of the irrevesible inactivation of bovine pancreatic ribonuclease by diethylpyrocarbonate. *Eur. J. Biochem.* 13: 519-525.

Wurm, G., Geres, U., and Schimdt, H. 1981. 1, 4-Naphtoquinones: A new synthesis of plumbagin. *Arch. Pharm.* 314: 1055-1056.

Yamashita, Y., Sakurai, T., Kuno, N., Uchida, K., and Yokobayshi, T. 2005. RNA extraction method, RNA extraction reagent, and method for analyzing biological materials. United States Patent No.981521.Available: http://www.patents.com/us-20050123965.html [07 August 2014].

Yao, I.V. and Nelson, R.J. 1996. WinBoot: a program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-

based dendrograms. IRRI Disc. Ser. No. 14. Int. Rice Res. Inst., Manila, Philippines.

Yeh, F. C., Yang, R. C., Boyle, T. B. J., Ye, Z. H., and Mao, J. X .1997. POPGENE, the User- Friendly Shareware for Population Genetic Analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada.

Zheng, D. and Hrazdina, G. 2008. Molecular and biochemical characterization of benzalacetone synthase and chalcone synthase genes and their proteins from raspberry (*Rubus idaeus* L.). *Arch. Biochem. Biophys.* 470: 139-145.

Zheng, D., Schröder, G., Schröder, J., and Hrazdina, G. 2001. Molecular and biochemical characterization of three aromatic polyketide synthase genes from *Rubus idaeus. Plant Mol. Biol.* 46: 1-15.

Zhong, S.M., Waterman, P.G., and Jad, J.J. 1984. Naphthoquinones and triterpenes from African *Diospyrus* species. *Phytochemistry* 23: 1067-1072.

Zhou, M.Y., Clark, S.E., and Gomez-Sanchez, C.E. 1995. Universal cloning method by TA strategy. BioTechniques. 19: 34-35.

Zietkiewicz, E., Rafalski, A., and Labuda, D. 1994. Genome fingerprinting by simple sequence repeat (SSR) – anchored polymerase chain reaction amplification. *Genomics* 20: 176-183.

APPENDICES

APPENDIX I

Tris Borate EDTA (T.B.E.) Buffer (20X) for 100ml solution

Tris base	21.6 g
Boric acid	11 g
0.5 M EDTA (pH 8.0)	8 ml

Make upto 500ml with distilled water (DEPC treated water for RNA isolation)

APPENDIX II

Luria Bertani (L. B.) Broth for 100ml

Peptone/ Tryptone	1 g
Yeast extract	0.5 g
NaCl	1 g

Adjust pH to 7 with 5N NaOH

APPENDIX III

Super Optimal Broth with Catabolic Repressor (S.O.C.) for 1L

1. SOB Broth

Tryptone	20 g
Yeast extract	5 g
NaCl	0.5 g

Autoclave SOB broth

2. Add 10ml each filter sterilized 1 M $MgCl_2$ and 1 M $MgSO_4$ to autoclaved SOC

3. Per 100 ml of SOC broth add freshly prepared 2ml filter sterilized 20 % (w/v) glucose

APPENDIX IV

1 X Tris EDTA (T. E.) Buffer for 100ml

1 M Tris HCl	1 ml
0.5 M EDTA	0.2 ml

Made upto 100ml with distilled water

ABSTRACT

ISOLATION AND CHARACTERIZATION OF TYPE III POLYKETIDE SYNTHASES FROM CHETHIKODUVELI (*Plumbago rosea* L.)

By

DHANYA RADHAKRISHNAN

(2009-09-101)

ABSTRACT

Submitted in partial fulfillment of the requirement

for the degree of

MASTER OF SCIENCE (INTEGRATED) IN BIOTECHNOLOGY

Faculty of Agriculture

Kerala Agricultural University

M.Sc. (INTEGRATED) BIOTECHNOLOGY COURSE

DEPARTMENT OF PLANT BIOTECHNOLOGY

COLLEGE OF AGRICULTURE

VELLAYANI, THIRUVANANTHAPURAM – 695 522

KERALA, INDIA

2014

ABSTRACT

The study entitled "Isolation and characterization of type III polyketide synthases from chethikoduveli (*Plumbago rosea* Linn.)" was conducted at the Biotechnology and Bioinformatics Division, JNTBGRI, Palode, Thiruvananthapuram, during 2013-2014. The objective of the study was to isolate and sequence characterize type III polyketide synthase gene(s) involved in the biosynthesis of plumbagin in *Plumbago rosea*. This knowledge would eventually help in augmenting the production of plumbagin through the development of high yielding cell lines. Furthermore, chemical variation with respect to plumbagin content and genetic diversity using ISSR markers was also carried out in available accessions of *Plumbago rosea*.

Nested degenerate primers specific to conserved regions of previously reported CHSs were used to amplify the cDNA synthesised from the reverse transcribed total RNA isolated from TBG-102. The amplicon was cloned and sequenced. The full length cDNA was elucidated using 5' and 3' RLM-RACE followed by cloning and sequencing. The full length cDNA was found to be 1,197 bp coding for 398 amino acids. The protein was calculated to have a molecular mass of 43.5 kDa and a theoretical pI of 5.89. BLASTX analysis of the full length cDNA showed 89 % identity with naringenin chalcone synthase from *Hypericum androsemum*, chalcone synthase from *Hypericum sampsonii* and aromatic polyketide synthase from *Hypericum hookerianum*. Simialr interpretation was drawn from the constructed phylogentic tree.In brief, it is concluded that the isolated full length cDNA belongs to typical CHS family and therefore, additional non-CHS class of PKSs would be involved in plumbagin synthesis.

Genetic diversity analysis with ISSR markers revealed high degree of polymorphism as the primers generated 71 scorable bands, out of which 70 were polymorphic (98.6 % polymorphism). Thereby, the accessions were concluded to have high genetic diversity.

Chemical analysis in these accessions showed variation in plumbagin content from 0.2 % to 0.7 %. The observed variation in plumbagin content may be due to the prevailing high genetic diversity in these accessions and variation in edaphic factors. The plumbagin content in TBG-102 was estimated using HPLC and spectrophotometric analysis. Both methods produces comparable results. Therefore cold extraction method was found to be reproducible.