ISOLATION AND CHARACTERIZATION OF LIGNIN BIOSYNTHETIC PATHWAY GENE(S) IN *LEUCAENA LEUCOCEPHALA*

BY
MANISH ARHA

OCTOBER 2007

ISOLATION AND CHARACTERIZATION OF LIGNIN BIOSYNTHETIC PATHWAY GENE(S) IN LEUCAENA LEUCOCEPHALA

A THESIS

SUBMITTED TO THE

UNIVERSITY OF PUNE

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

BIOTECHNOLOGY

BY

MANISH ARHA

PLANT TISSUE CULTURE DIVISION

NATIONAL CHEMICAL LABORATORY

PUNE – 411008

INDIA

OCTOBER 2007

DEDICATED TO MY GRANDPARENTS

CERTIFICATE

This is to certify that the work incorporated in the thesis entitled "**Isolation and Characterization of Lignin Biosynthetic Pathway Gene(s) in *Leucaena leucocephala***" submitted by ***Manish Arha*** was carried out under my supervision at the Plant Tissue Culture Division, National Chemical Laboratory, Pune. Materials obtained from other sources have been duly acknowledged in the thesis.

Dr. S. K. Rawal

(Research Guide)

DECLARATION BY RESEARCH SCHOLAR

I hereby declare that the thesis entitled "**Isolation and Characterization of Lignin Biosynthetic Pathway Gene(s) in *Leucaena leucocephala***", submitted for the Degree of ***Doctor of Philosophy*** to the University of Pune, has been carried out by me at Plant Tissue Culture Division, National Chemical Laboratory, Pune - 411 008, Maharashtra, India, under the supervision of Dr. S. K. Rawal (research supervisor) and Dr. B. M. Khan (co-guide). The work is original and has not been submitted in part or full by me for any other degree or diploma to any other University.

Manish Arha
(Research Scholar)

ACKNOWLEDGEMENTS

I express my sincere gratitude to my research supervisor and mentor Dr. S. K. Rawal for his guidance and invaluable help during my stay in NCL.

I thank my co-guide Dr. B. M. Khan for his guidance and encouragement throughout my work period.

Thanks are due to all members of the scientific and supporting staff of the Plant Tissue Culture Division, Biochemical Sciences Division and NCL Library.

I take this opportunity to thank Dr. R. A. Dixon and Dr. Anjan Banerjee for providing gene probe and vectors.

Many thanks are due to my labmates Sucheta, Neelima, Sushim, Noor, Sameer, Arun, Abhilash, Pallavi, Ruby, Santosh, Sumita, Kannan, Rishi, Somesh, Rajeshri, Neha, Priya, Jai, Azfar, Malini and Ulfat for the cordial atmosphere in the lab. I thank my seniors RC and Rohini for their encouragement and help during the course of this work.

I am indebted to my trainee Jay Kumar for his invaluable help during the crucial period of my work.

I take this opportunity to thank my friends Sreekanth, Nooka Raju, Suresh, Atul Thakur, Rohtas, Viku, Akku, Pankaj, Rodu, Prashant, Nishant, Abhishek, Nagraj, Shasidhar, JP, Sagar, Manje, Elan, Santosh, Chetan, Bhushan, Atul, Satya, Ambrish, Sharad, Anish, Aarohi, Nitin, SuryaPrasad, Sridevi, Sachin, Shabab, Asad, Asutosh, Sunil, Bhuban, Kiran, Shweta, Swapna, Meena, Varsha, Shyla, Anuj, Balaji, Thiru, Sony, Ashish, Shankar, Sujata, Satyendra, Arshad, Mukesh, Dileep, Venkatesh, Rajendra, Sridhar, Jitu Pandey, Sarvesh, Nilesh for their help, support, best wishes and lighter moments shared together through out my work period.

I owe my deepest gratitude to my grandparents, parents and my brother who have constantly encouraged me to excel.

I am grateful to Dr. S. Sivaram, Director, NCL for providing necessary facilities and permitting me to submit my findings in the form a thesis.

The fellowship awarded by the CSIR, India is duly acknowledged.

Manish Arha

CONTENTS

ABSTRACT		i - iii
Chapter 1	**INTRODUCTION**	1
1.1	Pulp and Paper Industry	1
1.2	Wood	1
1.3	Lignin	4
1.4	Paper making process	7
1.4.1	Mechanical Pulps	8
1.4.2	Semi-Chemical Pulps	8
1.4.3	Chemical Pulps	8
1.4.4	Biological Pulps	9
1.5	Lignin Biosynthesis and its Regulation	9
1.5.1	Phenylalanine ammonia lyase (PAL)	11
1.5.2	Cinnamate 4-hydroxylase (C4H)	13
1.5.3	Coumarate 3-hydroxylase (C3H)	14
1.5.4	4-Coumarate coenzyme A ligase (4CL)	15
1.5.5	Caffeoyl coenzyme A 3 O- methyltransferase (CCoAOMT) and Caffeate O-methyltransferase (COMT) or 5-hydroxyconiferaldehyde O-methyltransferase (AldOMT)	17
1.5.6	Cinnamoyl coenzyme A reductase (CCR)	19
1.5.7	Ferulate 5-hydroxylase (F5H) or Coniferaldehyde 5-hydroxylase (CAld5H)	20
1.5.8	Cinnamyl alcohol dehydrogenase (CAD) and Sinapyl alcohol dehydrogenase (SAD)	21
1.6	Transport of monolignols	24
1.7	Dehydrogenation	25
1.8	Polymerization	27
1.8.1	Radical generation and radical coupling	27
1.8.2	Polymerization process	27
1.8.3	Nucleation sites	28
1.9	Why *Leucaena*?	29
1.9.1	*Leucaena*	29
1.9.2	*Leucaena leucocephala* species growth forms	30

1.9.3	Fodder production and feed value of *Leucaena leucocephala*	30
1.9.4	Wood	30
1.10	Caffeoyl coenzyme A 3 O-methyltransferase (CCoAOMT; EC 2.1.1.104)	31
1.11	Why CCoAOMT from *Leucaena*?	34
1.12	References	35
Chapter 2	**MATERIALS AND METHODS**	**60**
2.1	Materials	60
2.2	Bacterial culture conditions	61
2.3	Bacterial cells transformation	62
2.3.1	*E. coli* transformation and selection	62
2.3.2	*A. tumefaciens* transformation and selection	62
2.4	The Nucleic Acids	63
2.4.1	Isolation of plasmid DNA from *E. coli* cells	63
2.4.2	Isolation of plasmid DNA from *Agrobacterium tumefaciens*	64
2.4.3	Genomic DNA Isolation	64
2.4.4	Restriction digestion of DNA	65
2.4.5	Extraction and purification of DNA from agarose gels	65
2.4.6	Total RNA Isolation	66
2.4.7	cDNA first strand synthesis by Reverse Transcription	67
2.4.8	Polymerase Chain Reaction (PCR)	68
2.4.9	Rapid Amplification of cDNA ends (RACE)	69
2.4.10	Genome Walking	73
2.4.10.1	Digestion of Genomic DNA	74
2.4.10.2	Purification of DNA	74
2.4.10.3	Ligation of Genomic DNA to GenomeWalker™ Adaptors	74
2.4.10.4	Primary PCR	75
2.4.10.5	Secondary PCR	75
2.4.11	Quantitative Real Time PCR (QPCR)	77
2.4.11.1	Pre-protocol considerations	78
2.4.11.2	Preparing the Reactions	79
2.4.11.3	PCR Cycling Programs	80

2.5		Nucleic acids blotting/hybridization	80
	2.5.1	Southern Blotting	80
	2.5.2	Slot Blot Hybridization	81
2.6		Random primer labeling	82
2.7		Hybridization	83
2.8		Protein purification	83
	2.8.1	Inclusion Bodies	83
	2.8.2	Ni-chelated affinity column	84
2.9		Sodium dodecylsulfate – Polyacrylamide gel electrophoresis (SDS -PAGE)	85
	2.9.1	Preparation of the separating gel	86
	2.9.2	Preparation of the stacking gel	87
	2.9.3	Preparation of the sample	87
	2.9.4	Loading and running the gel	87
	2.9.5	Silver staining of the gel	87
2.10		Raising Polyclonal Antibody against CCoAOMT in Rabbit and antigen affinity purification	88
	2.10.1	Affinity Column preparation for polyclonal antibody purification	88
	2.10.2	Antigen affinity purification of antibodies	89
2.11		Histology and Immunocytolocalization	90
2.12		Green Fluorescence Protein (GFP) visualization	91
2.13		Histochemical staining	91
2.14		*Agrobacterium* mediated Tobacco transformation	91
2.15		References	96
Chapter 3		**CLONING AND CHARACTERIZATION OF CAFFEOYL CoA 3- O METHYLTRANSFERASE GENE**	**100**
3.1		Introduction	100
3.2		Materials and Methods	100
3.2.1		Genomic DNA extraction and PCR	100
3.2.2		Slot Blot, Southern hybridization and random primer labeling	101
3.2.3		Bioinformatic analysis	101
3.2.4		Genome walking, cDNA synthesis, 5' and 3' RACE	101

3.3	Results and Discussion	101
3.3.1	Establishing the presence of the CCoAOMT gene in *L. leucocephala*	101
3.3.2	Slot Blot Hybridization	102
3.3.3	Southern Hybridization	102
3.3.4	Isolation of the genomic Caffeoyl CoA 3-O methyltransferase (CCoAOMT) gene	104
3.3.5	Genome Walking	111
3.3.5.1	5’ upstream walking	111
3.3.5.2	3’ downstream walking	113
3.3.6	Full length CCoAOMT genomic clone	115
3.3.7	Which 5’ UTR belongs to which CCoAOMT gene?	122
3.3.8	CCoAOMT1 and CCoAOMT2 cDNA clones	131
3.3.9	5’ and 3’ RACE	140
3.3.10	Distance Tree results	142
3.4	Conclusions	147
3.5	References	148
Chapter 4	**SPATIAL AND TEMPORAL EXPRESSION OF CCoAOMT IN *LEUCAENA LEUCOCEPHALA***	**150**
4.1	Introduction	150
4.2	Material and Methods	150
4.2.1	Cloning CCoAOMT cDNA in Expression vector	150
4.2.2	Protein expression and purification	152
4.2.3	Primary and Secondary Antibodies	152
4.2.4	Affinity purification of CCoAOMT IgG from rabbit immune serum	152
4.2.5	Histology, histochemical staining and immunocytolocalization	152
4.2.6	Semi- and absolute quantification of CCoAOMT1 and CCoAOMT2 gene expression	152
4.3	Results and Discussion	153
4.3.1	Expression of CCoAOMT in *E. coli*	153
4.3.2	Purification of polyclonal antibodies	158

4.3.3	Histology	159
4.3.4	Immunocytolocalisation	159
4.3.5	Semiquantitaive and Quantitaive Real Time PCR	160
4.4	Conclusions	166
4.5	References	167
Chapter 5	**ANALYSIS OF CCoAOMT1 AND CCoAOMT2 GENE PROMOTERS**	**169**
5.1	Introduction	169
5.2	Materials and Methods	170
5.2.1	Promoter analysis	170
5.2.2	Vector construction	170
5.2.3	*Agrobacterium tumefaciens* transformation	171
5.2.4	*Agrobacterium* mediated transformation of tobacco	171
5.2.5	Genomic DNA extraction and Polymerase Chain Reaction	171
5.2.6	Histology and Fluorescent Microscopy	171
5.3	Results and Discussion	174
5.3.1	Promoter analysis using MatInspector version 2.2	174
5.3.1.1	Common Regulatory Motifs	175
5.3.1.2	Different Regulatory Motifs	177
5.3.2	Promoter ProC1 and ProC2 directed GFP gene expression in tobacco	179
5.3.2.1	Tobacco transformation	180
5.3.2.2	GFP visualization in transformed tobacco plants	181
5.3.2.3	Integration of promoters ProC1, ProC2 and GFP in tobacco genome	181
5.3.2.4	Visualization of GFP in transverse sections of leaf midrib, stem and root of CProG1 and CProG2 transgenic tobacco plants	182
5.4	Conclusions	183
5.5	References	184
SUMMARY		**189**
List of Publications		**191**

ABBREVIATIONS

4CL	4-Coumarate coenzyme A ligase
AldOMT	5-Hydroxyconiferaldehyde O- methyltransferase
BAP	6- Benzylaminopurine
bp	Base pairs
BSA	Bovine serum albumin
C3H	Coumarate 3- hydroxylase
C4H	Cinnamate 4-hydroxylase
CAD	Cinnamyl alcohol dehydrogenase
CAld5H/ F5H	Coniferaldehyde 5- hydroxylase / Ferulate 5- hydroxylase
CCR	Cinnamoyl coenzyme A reductase
CCoAOMT	Caffeoyl coenzyme A 3-O- methyltransferase
Ci $mmol^{-1}$	Curie per milli mole
COMT	Caffeate O-methyltransferase
EDTA	Ethylene diamine tetra acetic acid disodium salt
g	grams
x g	Relative centrifugal force
g L^{-1}	grams per litre
G	Guaiacyl
GCMS	Gas Chromatograph Mass Spectroscopy
GFP	Green Fluorescence Protein
h	hour(s)
IPTG	Isopropyl β-D-thiogalactoside
Kb/Kbp	Kilobase pairs
KDa	Kilo Daltons
μL^{-1}	Micrograms per liter
μ Ci	Microcurie(s)
mg	milligrams
mM	millimolar
MCS	Multiple cloning sites
min	Minute(s)

MOPS	3-(N-Morpholino)propanesulfonic Acid
NAA	1-Napthyl aceticacid
O/N	Overnight
PAL	Phenylalanine ammonia lyase
PEG	Polyethylene glycol
PMSF	Phenyl methyl sulphonyl fluoride
rpm	Rotations per minute
S	Syringyl
s	Seconds
SAD	Sinapyl alcohol dehydrogenase
SDS	Sodium dodecyl sulphate (sodium lauryl sulphate)
U μL^{-1}	Units per microlitre
v/v	volume / volume
w/v	weight / volume
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside

ABSTRACT

World wide annual production of paper has increased more than three fold in the past forty years, amounting to a total production of 120 x 10^6 tons. In India the annual production of paper is close to 5 million tons. India's annual soft woods and pulp imports are worth rupees ten thousand crores and the projected annual short fall of paper and paper products is expected to grow to approximately 4 million MT by the year 2010. In addition, the growth of paper and paper products consumption in India is expected to be the highest in Asia during the current decade. Thus, it has become imperative to achieve self sufficiency in paper and pulp production, which may be achieved with the development of fast growing trees providing higher biomass with low lignin content per unit of land.

Wood, agricultural residues and many other plant materials which can be used for pulp and paper production consist largely of lignocellulose (a composite of mainly cellulose, hemicellulose and lignin). In the production of paper pulp (which is mainly cellulose), the hemicellulose and much of the lignin is removed using alkali. To produce bright paper, pulp may require a further bleaching treatment to remove lignin residues. This delignification process consumes large quantities of energy and hazardous chemicals. Reducing the content or changing the quality of lignin in pulp wood species without compromising the mechanical strength of the plant is desirable for paper industry. This would be beneficial both from the economical as well as environmental point of view. The results could be extrapolated to the forage crops to improve digestibility since the presence of lignin limits the ability of microorganisms to break down the cellulose and hemicellulose.

Lignin forms an integral cell wall component of all vascular plants and constitutes the second most abundant organic constituent on earth after cellulose, representing on an average of 25% of the terrestrial plant biomass (Sarkanen *et al*. 1967). Chemically, it is a complex phenolic heteropolymer of monolignols, namely coumaryl, coniferyl and sinapyl alcohols. It plays a fundamental role in conferring rigidity, strength, resistance to pathogen attack and water impermeability to the polysaccharide-protein matrix of the cell wall allowing solute conductance (Brown 1985). Despite its biological importance in plants, lignin composition, quantity and distribution is known to affect the agro-industrial utilization of plant biomass (Higuchi, 1985; Odendahl, 1994; Biermann, 1996; Baucher *et al.,* 1998).

Paper industry in India mainly uses bamboos, *Eucalyptus* sp., *Casuarina* sp. and *Leucaena* sp. as a source for paper pulp. Although all these plant species are of importance to the paper industry, *Leucaena sp.* is exclusively used in India and about 25% raw material for paper and pulp industry comes from this plant. *Leucaena* sp is a fast growing multipurpose tree adapted to a variety of soils and climatic conditions. No study has been done on Lignin Biosynthesis gene(s) so far in *Leucaena* sp. and study of these gene(s) will help in understanding the Lignin Biosynthetic Pathway in *Leucaena* sp. and its manipulation so as to meet the needs of pulp and paper industry. To meet the increasing demand of high quality wood for paper industry it is essential to provide designer plant species. It will thus be crucial to raise plantations of the *Leucaena* sp with elite materials and or genetically engineered plants that meet the demands of the pulp and the paper industry in economical and sustainable manner.
Pulp and paper industry will benefit by either reduced lignin synthesis or by alteration of the lignin components in a manner where in the sinapyl alcohol to guaiacyl alcohol ratios are in favour of sinapyl alcohol. The later of the two may be achieved by down regulating the Caffeoyl CoA 3-O-Methyl Transferase (CCoAOMT; EC 2.1.1.104) which would diversion of the flux towards the formation of sinapyl alcohol. Hence, the present study was aimed at the isolation of CCoAOMT gene(s) from *Leucaena leucocephala*, and its molecular characterization studies.

The main features of the present thesis are:

- Cloning and Characterization of Caffeoyl CoA 3 O Methyl Transferase gene
- Spatial and Temporal expression of CCoAOMT in *Leucaena leucocephala*.
- Analysis of CCoAOMT1 and CCoAOMT2 gene promoters

- **Cloning and characterization of Caffeoyl CoA 3 O- methyltransferase gene**

A PCR based approach to fish out the genomic and c-DNA clones of CCoAOMT genes from *Leucaena leucocephala* was followed. Primers were designed on the basis of consensus regions of various reported CCoAOMTs from the NCBI GenBank database. A partial CCoAOMT gene sequence was amplified and its sequence utilized to design gene specific primers. Genome walking and Rapid Amplification of cDNA ends (RACE) was performed to

fish out the full length cDNA and genomic clones of CCoAOMT. Characterization of the two genes encoding CCoAOMT 1 and CCoAOMT 2 was done.

- **Spatial and temporal expression of CCoAOMT in *Leucaena leucocephala***

CCoAOMT was immunolocalized in xylem and fibers proving its presence at the sites of active lignification. The spatial and temporal expression of the two CCoAOMT genes was studied using semi quantitative and absolute quantification *i.e.* Real Time PCR. Primers were designed so as to distinguish between two genes. The 5.8S RNA was used as the internal standard. For absolute quantification Taqman probes were used. Plants of different age *i.e.* seedling of age 1 day, 5 day, 10 day and 15 day; one season and two seasons old plants were used for temporal expression studies. Different plant part *i.e.* root, shoot, leaf and inflorescence were used for spatial expression studies.

- **Analysis of CCoAOMT1 and CCOAOMT2 gene promoters**

Bioinformatic analysis of the two partial promoter sequences was done using MatInspector 2.2. Analysis of the two promoter nucleotide sequences revealed presence of different cis-regulatory elements involved in CCoAOMT gene regulation.

Two partial promoters of CCoAOMT1 and CCoAOMT2 were used to study their role in gene expression. For this, the two promoters were cloned individually upstream of the Green Fluorescent Protein (GFP) gene with NOS 3' as the terminator. These cassettes were introduced into tobacco *via Agrobacterium tumefaciens* mediated transformation. The pattern of GFP expression under the control of these promoters was studied.

1.1 Pulp and Paper Industry

World wide annual production of paper has increased more than three fold in the past forty years, amounting to a total production of 120 x 10^6 tons. In India the annual production of paper is close to 5 million MT. India's annual soft woods and pulp imports are worth US $ 300 million and the projected annual short fall of paper and paper products is expected to grow to approximately 4 million MT by the year 2010. In addition, the growth of paper and paper products consumption in India is expected to be the highest in Asia during the current decade. Thus, it has become imperative to achieve self sufficiency in paper and pulp production, which may be achieved with the development of fast growing trees providing higher biomass with low lignin content per unit of land.

Wood, agricultural residues and many other plant materials, which can be used for pulp and paper production, consist largely of lignocellulose (a composite of mainly cellulose, hemicellulose and lignin). In the production of paper pulp (which is mainly cellulose), the hemicellulose and much of the lignin is removed using mechanical or chemical processes or a combination of both. To produce bright paper, pulp may require a further bleaching treatment to remove lignin residues. The delignification process consumes large quantities of energy and hazardous chemicals. Reducing the content or changing the quality of lignin in pulp wood species without compromising the mechanical strength of the plant is desirable for paper industry. This would be beneficial, both from the economical as well as environmental point of view. The results could be extrapolated to the forage crops to improve digestibility since the presence of lignin limits the ability of microorganisms to break down the cellulose and hemicellulose in the animal alimentary canal.

1.2 Wood

Wood characteristics vary in different types of plant. For instance, conifers (gymnosperms) produce softwood whereas angiosperms produce hardwoods. Softwoods are mainly composed of three cell types, tracheids (which play a role both in rigidity and conduction), and axial and ray parenchyma cells. Hardwoods are mainly made of fibers, vessels, and axial and ray parenchyma cells. Vessels transport water and solutes through the vascular system while fibers provide rigidity, and ray cells facilitate centripetal nutrition (Higuchi, 1997). Tracheids, vessels and fibers vary in shape and size (Table 1). The dimensions and chemical

composition of the different cell types of wood depend on genetic, developmental and environmental factors (Vallette and de Choudens, 1992). Softwoods and hardwoods differ in their pulping characteristics. Indeed, the individual cell types within wood differ in their chemical characteristics, reflecting the underlying differences in biochemistry and molecular biology that are only beginning to be appreciated.

The three major components of wood cell walls are cellulose, hemicellulose and lignin. Long molecules of cellulose provide the skeleton of the walls. Linear cellulose chains are aligned together in structures known as 'elementary fibrils' or 'protofibrils' that, in turn, associate into more complex structures called microfibrils (Fig.1 a). Microfibrils are highly organized and form distinct fibrillar cell wall layers (Delmer and Amor, 1995). Hemicelluloses and other carbohydrates provide the matrix of the cell wall. Lignin, a heterogenous hydrophobic

Table 1: Dimensions of the different cell types of Softwoods and Hardwoods (Fengel and Wegener, 1984).

Cell Type	Character	Cell Dimensions
Soft wood (***Picea abies***)		
Tracheids	Diameter	20 – 40µm
	Wall thickness	2.1 – 4.3µm
	Length	1.7 – 3.7mm
Hard wood (***Fagus sp.***)		
Fibers	Diameter	15 – 20µm
	Wall thickness	5µm
	Length	0.6 – 1.3mm
Vessels	Diameter	5 – 100µm
	Wall thickness	1µm
	Length	0.3 – 0.7mm

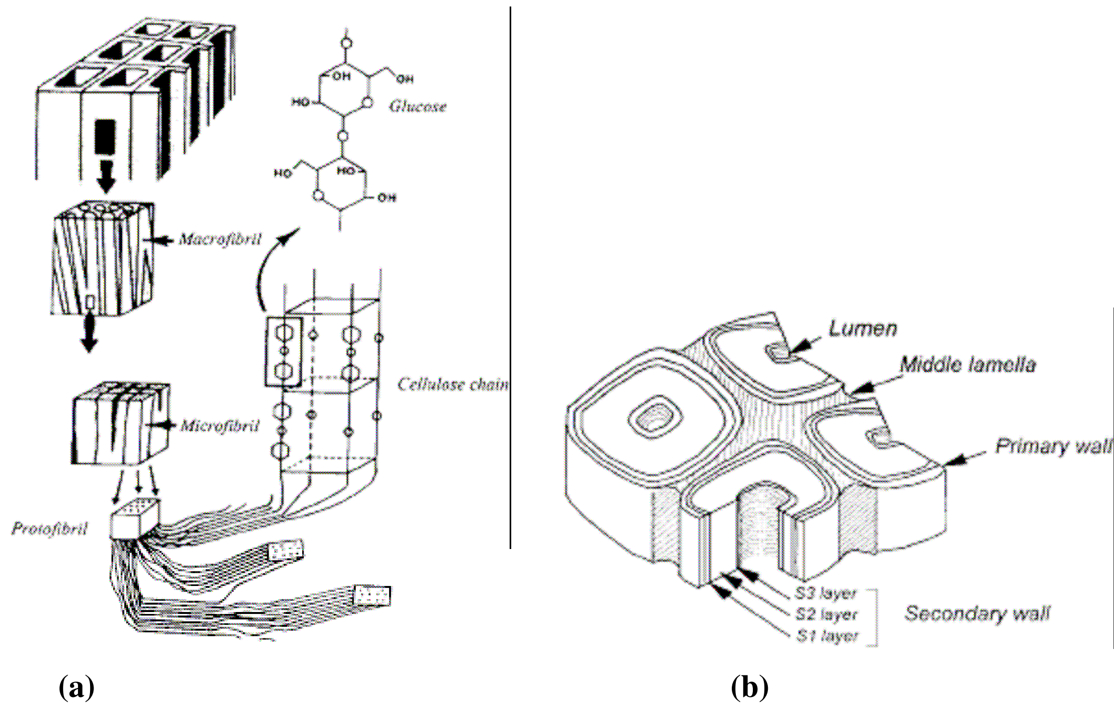


Fig. 1: (a) Organization of the cellulose skeletons in the fiber wall (Parham, 1987), (b) Structural organization of the cell walls of fibers (Petit-Conil, 1995).

phenolic polymer, encrusts the other wall components to waterproof and strengthen the wall. In a transverse plane, the parietal structure of wood cells is made of a primary and a secondary wall, the latter consisting of two or three layers, designated S1, S2, and S3 (Fig. 1 b). The various cell wall layers differ in chemical composition (Mellerowicz et al., 2001).

Lignin deposition is one of the final stages of xylem cell differentiation and mainly takes place during secondary thickening of the cell wall (Donaldson, 2001). Lignin deposition proceeds in different phases, each preceded by the deposition of carbohydrates, and starts at the cell corners in the region of the middle lamella and the primary wall when S1 formation has initiated. When the formation of the polysaccharide matrix of the S2 layer is completed, lignification proceeds through the secondary wall. The bulk of lignin is deposited after cellulose and hemicellulose have been deposited in the S3 layer. Generally, lignin

concentration is higher in the middle lamella and cell corners than in the S2 secondary wall (Baucher et al., 1998; Donaldson, 2001; Saka and Goring, 1985).

The three monolignols (S, G and H) are incorporated at different stages of cell wall formation. Typically, H units are deposited first, followed by G units, and S units still later in angiosperms (Donaldson, 2001; Terashima et al., 1995). Lignin in vessels is generally enriched in G units, whereas lignin in fibers is typically enriched in S units (Saka and Goring, 1985). A large proportion of S units is also found in secondary walls of ray parenchyma (Fergus and Goring, 1970). In gymnosperms, the lignin deposited in compression wood is enriched in H units (Timell, 1986). The difference in timing of monolignol deposition is associated with variations in lignin condensation in the individual cell wall layers, as shown by immunocytochemistry with antibodies raised against pure H, pure G, or mixed GS synthetic lignins (Chabannes et al., 2001; Joseleau and Ruel, 1997). Lignin deposition and the relative incorporation of the different monolignols into the polymer, are spatially and temporally regulated. The mechanisms controlling this process are not yet fully resolved but are likely governed by the interplay between the spatio-temporal expression of monolignol biosynthetic genes, the kinetics of monolignol delivery to the cell wall and the chemistry of monolignol coupling to the growing polymer in the complex macromolecular environment of the cell wall.

1.3 Lignin

The name lignin is derived from Latin *lignum* meaning wood. After cellulose, lignin is most abundant natural (terrestrial) organic polymer and a major constituent of wood. It forms an integral cell wall component of all vascular plants, representing on an average of 25% of the terrestrial plant biomass. Lignin content is higher in softwoods (27-33%) than in hardwood (18-25%) and grasses (17-24%). The highest amounts of lignin (35–40 %) occur in compression wood on the lower part of branches and leaning stems (Fengel and Wegner, 1984; Sarkanen and Ludwig, 1971). Lignin does not occur in algae, lichens or mosses (Nimz and Tutschek, 1977), whereas the “lignins” of bark differ in their structure from typical wood lignins (Zimmermann et al., 1985). The ability to synthesize lignin has been essential in the evolutionary adaptation of plants from an aquatic environment to land and provides crucial structural integrity to the cell wall and stiffness and strength of the stem (Chabannes et al.,

2001; Jones et al., 2001). Lignin is predominantly synthesized and deposited in the secondary cell wall of specialized cells such as xylem vessels, tracheids and fibers. It is also deposited in minor amounts in the periderm where in association with suberin it provides a protective role against pathogens (Sarkanen and Ludwig 1971). In addition, lignin waterproofs the cell wall, enabling transport of water and solutes through the vascular system.

Though lignin has been studied for more than a century, many aspects of its biosynthesis remain unresolved. The monolignol biosynthetic pathway has been redrawn many times and remains a matter of debate (Dixon et al., 2001; Humphreys and Chapple, 2002). Likewise, the biochemical processes leading to dehydrogenation of the monolignols in the cell wall and their polymerization and deposition are fields of active discussion (Davin and Lewis, 2000; Hatfield and Vermerris, 2001; Lee et al., 1997; Rouhi, 2001; Sederoff et al., 1999).

Lignins are complex racemic aromatic heteropolymers synthesized from the dehydrogenative polymerization of monolignols, namely coumaryl, coniferyl and sinapyl alcohol monomers differing in their degree of methoxylation (Freudenberg and Neish, 1968) (Fig. 2 a-c). These monolignols produce, respectively, *p*-hydroxyphenyl (**H**), guaiacyl (**G**), and syringyl (**S**) phenylpropanoid units when incorporated into the lignin polymer. The amount and composition of lignins vary among taxa, cell types and individual cell wall layers, and are influenced by developmental and environmental cues (Campbell and Sederoff, 1996). Dicotyledonous angiosperm (hardwood) lignins consist principally of **G** and **S** units and traces of **H** units, whereas gymnosperm (softwood) lignins are composed mostly of **G** units with low levels of **H** units. Lignins from grasses (monocots) incorporate **G** and **S** units at omparable levels, and more **H** units than dicots (Baucher et al., 1998).

Lignification is the process by which H, G and S units are linked together via radical coupling reactions (Sarkanen and Ludwig, 1971; Freudenberg and Neish, 1968). The main "end-wise" reaction couples a new monomer (usually a monolignol and usually at its β position) to the growing polymer, giving rise to different structures (Fig. 2 d-f) all of which are β-linked. Coupling between preformed lignin oligomers results in units linked 5–5 and 5–O–4 (Fig. 2 g, h). The coupling of two monolignols is a minor event, with resinol (β–β) units (Fig. 2 i) or cinnamyl alcohol end groups (Fig. 2 j) as the outcome.

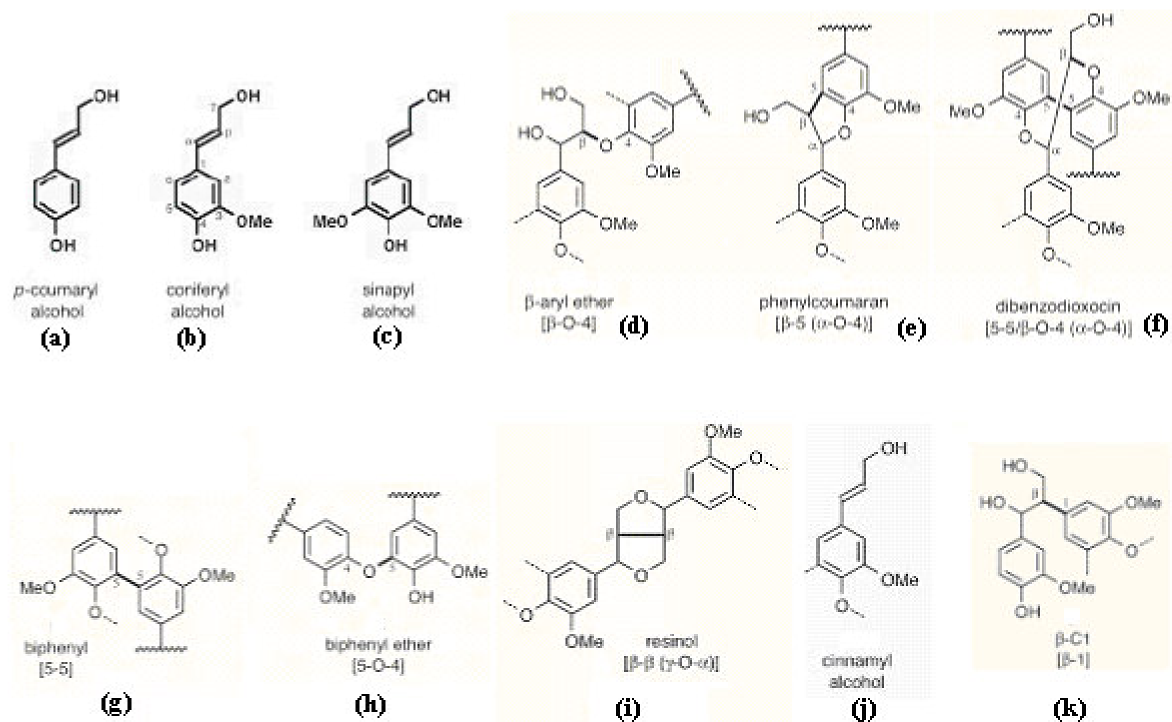


Fig. 2: Monolignols *p-* coumaryl (a), coniferyl (b) and sinapyl alcohol (c). Different structures formed due to "end wise" coupling (d-f). Coupling between preformed lignin oligomers results in units linked 5–5 and 5–O–4 (g, h). Coupling of two monolignols form resinol β–β units (i) or cinnamyl alcohol (j) end groups (Boerjan et al., 2003).

Monolignol dimerization and lignin are substantially different processes (Adler, 1997), explaining why lignification produces frequencies of the various units that are different from those produced by dimerization or bulk polymerization *in vitro* (Fig. 3).

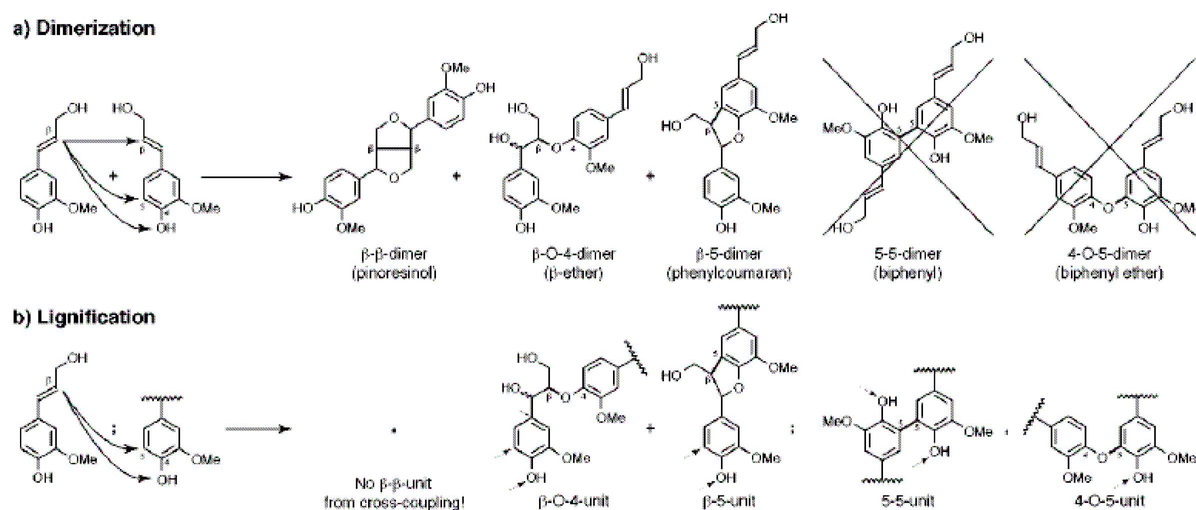


Fig. 3: Lignification differs substantially from simple dimerization of monolignols. (*a*) Dimerization of coniferyl alcohol produces only three dimers, in each of which at least one of the coniferyl alcohols is coupled at its β position. The 5–5 and 5–O–4 dimers (crossed out) do not arise in any significant way from monomer dimerization reactions. The new bond formed by the radical coupling reaction is noted in bold. (*b*) Cross-coupling of coniferyl alcohol with a G unit gives only two main products, explaining why there are more β-ethers formed during lignification than in monolignol dimerization. Coupling of preformed oligomers is the source of most of the 5–5- and 5–O–4 units. Sites of further coupling reactions during lignification are indicated by arrows (Boerjan et al., 2003).

1.4 Paper making process

Two major categories of processes exist for the production of paper pulp: chemical and mechanical. The chemical process uses chemicals to remove lignin from fiber cell walls to obtain long and flexible fibers that consist of polysaccharides only, whereas mechanical

pulping process focuses on the mechanical separation of fibers without the removal of lignin. Though the pulp yield is highest in mechanical pulping, the pulp has limited bleachability and the paper becomes yellow as it ages. In contrast, chemical pulping yields individual intact fibers free from lignin that can interact with other fibers via hydrogen bonds, making very strong paper.

Nowadays, the Kraft process is the most widely used chemical procedure for the production of paper. However, this process is gradually being replaced by thermomechanical (TMP) and chemithermomechanical (CTMP) pulping methods that give higher pulp yields and consume less water.

In 2000, the world wood pulp production has been estimated at 135,852 thousands of metric tons, from which 6.3% are of mechanical pulp, 15.6% TMP pulp, 3.6% semimechanical pulp, and 74.5% chemical pulp (Food and Agriculture Organization, 2001).

1.4.1 Mechanical Pulps

Mechanical pulping processes may be based on any one of the following processes: stone ground wood (SGW) process, the pressure ground wood (PGW) process, the refiner mechanical pulp (RMP) procedure, the thermomechanical process (TMP) process, the chemimechanical (CMP) and the chemithermomechanical (CTMP)processes, and bleaching of pulp

1.4.2 Semi-Chemical Pulps

Semi-chemical pulps are essentially mechanical pulps that have been mildly pretreated with sodium sulfite (Na_2SO_3) and sodium carbonate (Na_2CO_3) to partially remove lignin and hemicelluloses. This is followed by bleaching of the pulp.

1.4.3 Chemical Pulps

Chemical pulps represent the largest part of the world pulp market and are used to produce nearly all paper and board grades (Food and Agriculture Organization, 2001). This may involve alkaline or acidic pulping processes and a final bleaching step. The aim of chemical pulping is to dissolve and remove the lignin from the fiber wall and to separate the fibers at the middle lamella without mechanical damage Gierer (1985).

1.4.4 Biological Pulps

Biopulping is achieved by treatment of wood chips with lignin-degrading fungi prior to pulping. Such a pretreatment reduces energy requirements and improves the paper strength. In addition to the economical benefits, the biopulping process is not harmful to the environment, because only benign materials are used and no additional waste streams are generated (Reid, 1991; Akhtar et al., 1998a; Highley and Dashek, 1998). Biomechanical pulping is on the way toward being commercialized (Akhtar et al., 1998b; Breen and Singleton, 1999). A fungal pretreatment in chemical pulping has also been shown to be beneficial for subsequent pulping processes because part of the lignin is removed or modified (Akhtar et al., 1998a).

1.5 Lignin Biosynthesis and its Regulation

For the last two decades, there has been a great deal of interest in cloning and characterization of the genes controlling monolignol biosynthesis in order to understand monolignol biosynthetic pathways in trees and other plants. A number of reviews have been done about the advancements of monolignol biosynthesis pathways (Whetten and Sederoff, 1995; Whetten et al., 1998; Humphreys and Chapple, 2002; Boerjan et al., 2003). There is enormous variation in lignin content and composition among plant species, tissues, cell types, and even developmental stages and environmental conditions play a role. Data from the studies using different plant materials display many agreements as well as certain disagreements. Thus, it is debatable whether lignin biosynthesis in all plants follows the exact same pathway or not. To date, most of the genes for monolignol biosynthesis have been identified and characterized in various plant species. A summarized picture of the main and possible monolignol biosynthesis pathways of wood formation in trees is shown in Fig.5. The genes involved in the pathway, the reaction catalyzed by them and their regulation will be discussed.

Natural variations in lignin content and composition observed between different plants, tissues and cell types (Grand *et al*., 1985; 1983; Monties, 1998; Wu *et al*., 1992; Campbell and Sederoff, 1996; Buxton and Redfearn, 1997; Sederoff *et al*., 1999; Donaldson, 2001) as well as occurrence of natural mutants (Kuc and Nelson, 1964; Ralph *et al*., 1997; Halpin *et*

al., 1998) suggest that human induced changes in lignin content and composition through genetic engineering approaches is a realistic possibility.

Attempts have been made to genetically modify plant species with the aim of lignin reduction. Conceptually, genetic engineering of lignin can be accomplished at three levels of control, namely synthesis of monolignols, its transport and polymerization at the site of deposition (Fig.4) as the overall rate of lignification is regulated not only by the monolignol biosynthesis, but also by the coordinated transport, storage, mobilization and polymerization of monolignol precursors to the cell wall.

During the past decade, significant headway has been made in the cloning and subsequent expression of sense and antisense constructs of a number of genes pertaining to monolignol biosynthesis, transport and polymerisation (Whetten *et al*. 1998; Whetten and Sederoff 1991; Boudet *et al*. 1995; Boerjan *et al*. 2003; Boudet *et al*. 1997; Baucher *et al*. 2003; Boudet 2000, Sederoff 1999;; Grima-Pettenati and Goffner 1999; 2000).

LEVELS OF CONTROL FOR
GENETIC MANIPULATION OF LIGNIN

1
Synthesis of monolignols

2
Transport of monolignols
from the site of synthesis
to the site of polymerization

3
Polymerization of monolignols
to give final product

Fig. 4: Levels of control for manipulation of lignin biosynthesis. The control level 1 involves manipulation of genes for PAL, C4H, COMT or AldOMT, CCoAOMT, F5H, 4CL, CCR and CAD, level 2 involves the genes for UDPG-GT and CBG, while level 3 includes POD, LAC and PO

The transgenics have produced unexpected findings leading to a profound reappraisal of our understanding of the phenylpropanoid "metabolic grid". The enzymes and their genes involved in lignin biosynthesis and related findings of gene regulation of lignin biosynthesis pathway genes is discussed in the following.

1.5.1 Phenylalanine ammonia-lyase (PAL)

Monolignol biosynthesis is considered to start from phenylalanine. The enzyme Phenylalanine ammonia-lyase, PAL, that catalyzes the conversion of phenylalanine to trans-cinnamic acid, is the initial step towards monolignol biosynthesis and other phenolic secondary plant metabolites. Genes encoding PAL have been studied in *Populus* species (Osakabe et al., 1995; Kao et al.; 2002), loblolly pine (Whetten and Sederoff, 1991) and other plant species (Jones, 1984; Ohl et al., 1990; Leyva et al., 1992; Bate et al., 1994; Hatton et al., 1995; Kumar and Ellis, 2001). *PAL* exists as a multiple member gene family and the individual members can be involved in different metabolic pathways as suggested by their expression patterns in association with certain secondary compounds accumulated in specific tissue or developmental stage. The expression studies of two *PAL* genes suggested that one is associated with condensed tannin metabolism and the other with monlignol biosynthesis (Kao et al., 2002). In the Arabidopsis genome, four *PAL* genes were identified and were phylogenetically classified into two groups based on sequence similarity (Raes et al., 2003). The biochemical activity of all known *PAL*s is verified to specifically catalyze deamination of phenylalanine, but genetic and physiological function may vary among different *PAL* members.

The expression of *PAL* genetic function is controlled by various genetic circuits and signaling pathways. The *cis*-element structures in *PAL* gene promoters can be part of the molecular circuit that directs a variety of the *PAL* genetic and physiological functions. In some *PAL* promoters, conserved AC *cis*-elements of box P, box A, and box L are identified for regulating the phenylpropanoid genes expression related to monolignol biosynthesis (Cramer et al., 1989; Lois et al., 1989; Logemann et al., 1995). In other *PAL* member promoters, the *cis*-elements of H box and G box are found (Cramer et al., 1989; Lois et al., 1989; Osakabe et al., 1995; Leyva et al., 1992; Raes et al., 2003). Many other *cis*-elements

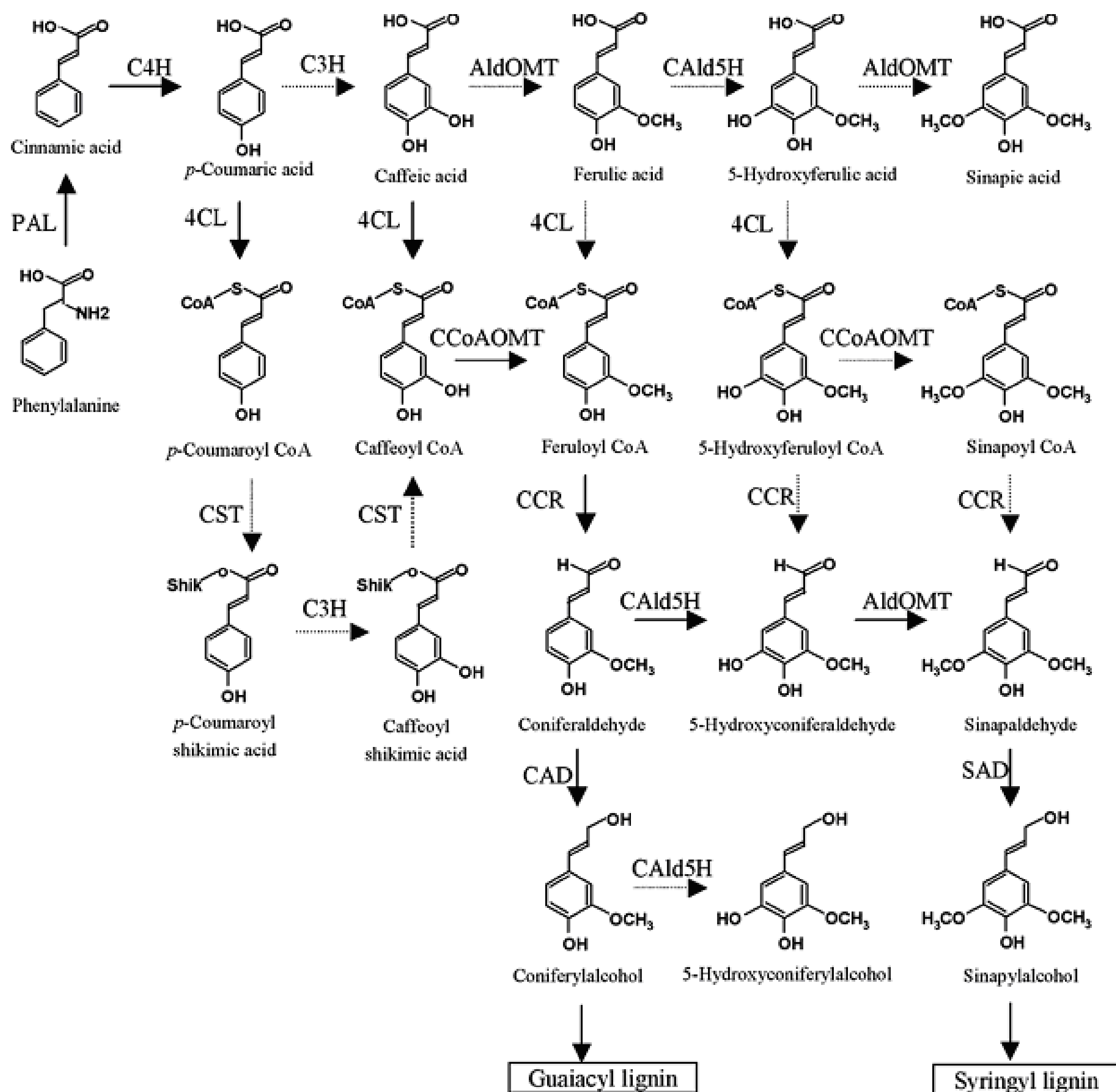


Fig. 5: An overview of the monolignol biosynthesis pathways. The main pathways are indicated in solid line arrows and possible pathways in dotted line arrows. *PAL*, phenylalanine ammonia-lyase; *C4H*, cinnamic acid 4-hydroxylase; *C3H*, *p*-coumarate 3-hydroxylase; *CST*, hydroxycinnamoyl- CoA:shikimate hydroxycinnamoyltransferase; *4CL*, 4-coumaroyl-CoA 3-hydroxylase; *CCoAOMT*, caffeoyl-CoA *O*-methyltransferase; *CCR*, cinnamoyl coenzyme A reductase; *CAld5H*, coniferyl aldehyde 5-hydroxylase; *AldOMT*, 5-hydroxyconiferyl aldehyde *O*-methyltransferase; *CAD*, cinnamyl alcohol dehydrogenase; *SAD*, sinapyl alcohol dehydrogenase (Li et al., 2006).

have been identified in members of the *PAL* gene family, however, their function and regulation remains to be studied. *PAL* gene expression has been suppressed by 85% and >98% in transgenic plants with resultant 52% (Sewalt et al., 1997) and 70% (Korth et al., 2001) reduction in Klason lignin content, respectively. Lignin monomeric composition, determined by pyrolysis GC-MS, was characterized by a lower proportion of G units and a 1.7-fold increase in S/G ratio (Sewalt et al., 1997). Because PAL catalyzes the first step of the phenylpropanoid pathway, reduction of its activity results in a wide range of abnormal phenotypes. The transgenic plants were stunted, had curled leaves, and had thinner cell walls in the secondary xylem with less lignin than those of the control (Elkind et al., 1990; Bate et al., 1994). These plants were also more susceptible to the fungal pathogen *Cercospora* nicotianae (Maher et al., 1994). A slight increase in Klasone lignin and dry matter content was observed in the stem of *PAL*-over expressing plants (Howles et al., 1996; Korth et al., 2001). Over expression of *PAL* did not lead to changes in lignin composition as determined by pyrolysis GC-MS (Sewalt et al., 1997), but to a decrease in the amount of S units, yielding a reduction in the S/G ratio when lignin was analyzed by thioacidolysis (Korth *et al*., 2001).

1.5.2 Cinnamate 4-hydroxylase (C4H)

The conversion of cinnamate to *p*-coumarate is catalyzed by C4H. C4H is a cytochrome P450- dependent monooxygenase, belonging to the *CYP73* family. Similar to PAL, C4H is thought to be involved in a number of secondary metabolism pathways in addition to monolignol biosynthesis as *p*-coumarate is an intermediate for biosynthesis of many secondary compounds (Croteau et al., 2000). Multiple *C4H* gene members are identified in many plant species, however, only one *C4H* is known in the Arabidopsis genome (Raes et al., 2003). The expression study of two *C4H* members in quaking aspen indicated that one is strongly expressed in developing xylem tissues and the other is more active in leaf and young shoot tissues. In other species, *C4H* gene is expressed in a variety of tissues and the expression is induced by wounding, light, pathogen attacks and other biotic and abiotic stimuli (Bell-Lelong et al., 1997; Raes et al., 2003). The mechanisms that regulate the genetic function of *C4H* gene and its family members are yet unknown.

In transgenic tobacco plants, C4H activity was altered by expressing the alfalfa class I *C4H* (CYP73A3) (Sewalt et al., 1997; Blount et al., 2000) or the French bean class II *C4H* (CYP73A15) (Blee et al., 2001) genes in sense or antisense orientation (class I and class II *C4H* share approximately 60% similarity). Overexpression of class I *C4H* had no effect on Klasone lignin, nor on the S/G ratio. In contrast, a 76% reduction in total *C4H* activity led to a 63% decrease in Klason lignin and a modification of the lignin monomeric composition. The amount of S units was strongly reduced and S/G decreased by over 90% (Sewalt et al., 1997). Similarly, a reduction by 90% of C4H activity by down-regulation of the class II *C4H* resulted in a 27% decreased lignin content and one tobacco transgenic line had a decreased S/G ratio (Blee et al., 2001). In transgenic tobacco modified to reduce *C4H* expression, decreased *PAL* activity has also been shown (Blount et al., 2000). Regulation may be mediated by pathway intermediates, for example, cinnamic acid may act as a feedback regulator of the phenylpropanoid pathway.

1.5.3 Coumarate 3-hydroxylase (C3H)

Early biochemical evidence suggested that the reaction coumarate to caffeate is catalyzed by a nonspecific phenolase, but that suggestion did not receive much support in other studies (Stafford and Dresler, 1972; Boniwell and Butt, 1986; Kojima and Takeuchi, 1989; Petersen et al. 1999). Recently, the gene encoding *p*-coumarate 3-hydroxylase (C3H) was cloned and an alternative pathway proposed based on the enzyme activity of *CYP98A3* gene from Arabidopsis (Schoch et al., 2001; Franke et al., 2002 a; Nair et al., 2002). The proposed alternative suggested that the hydroxylation at the 3- position of the aromatic ring of cinnamic acid does not directly occur on *p*-coumarate, instead, *p*-coumarate is first converted to *p*-coumaroyl CoA ester by 4-cinnamoyl-CoA ligase (4CL), then the CoA ester group of *p*-coumaroyl CoA is exchanged by hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase (CST) to form *p*-coumaroyl shikimic acid which serves as a substrate of *C3H* to produce caffeoyl shikimic acid. Subsequently caffeoyl shikimic acid reverts back to caffeoyl CoA to push metabolism towards the biosynthesis of monolignols. Among tree species, a *CYP98* cDNA was cloned from sweetgum and aspen (Osakabe et al., 1999), but the postulated genetic and biochemical functions in monolignol biosynthesis have not been demonstrated for its role in wood formation.

The gene encoding *p*-coumarate 3-hydroxylase (*C3H*) has only recently been cloned by two independent research groups. Using a functional genomics approach, Schoch et al. (2001) identified CYP98A3 as a possible candidate for *C3H*. In parallel, by screening *Arabidopsis* mutants under UV light, Franke *et al*. (2002a) isolated the reduced epidermal fluorescence 8 (*ref8*) mutant. By positional cloning, the *REF8* gene was identified as the cytochrome P450-dependent monooxygenase *CYP98A3*. The *ref8* mutant had collapsed xylem vessels, a higher cell wall degradability and a higher susceptibility to fungal colonization (Franke et al., 2002b), associated with the accumulation of *p*-coumarate esters instead of sinapoylmalate and with a reduction in lignin content of 60-80%. A range of analyses showed that lignin composition was dramatically altered, being almost entirely madeup of *p*-coumaryl alcohol units (Franke et al., 2002b).

1.5.4 4-Coumarate oenzyme A ligase (4CL)

Genetic and biochemical functions of 4-Coumarate Coenzyme A ligase (*4CL*) genes have been clearly demonstrated in association with monolignol biosynthesis (Lewis and Yamamoto, 1990; Lee et al., 1997; Hu et al., 1998, 1999; Harding et al., 2002). *4CL* genes usually exist as a multi-gene family. Different expression patterns of *4CL* members are found in herbaceous and tree species. Four *4CL* genes were detected in the Arabidopsis genome and the expression of each member was regulated differentially in tissues and development stages (Raes et al., 2003). In aspen trees, two *4CL* genes were cloned and their expression clearly distinct, with one in epidermal and leaf tissue and the other specifically in developing xylem tissue (Hu et al., 1998; Harding et al., 2002). Furthermore, the enzymatic activities of *4CL* members from aspen, loblolly pine, tobacco, soybean, Arabidopsis, and many other species were found to have distinct substrate specificities (Voo et al., 1995; Zhang and Chiang, 1997; Hu et al., 1998; Lindermayr et al., 2003; Schneider et al., 2003; Hamberger and Hahlbrock, 2004). Whether the substrate specificity of the *4CL* members relates to different metabolic pathways is unknown. As the *4CL* catalytic kinetics vary among species, it is also likely that the mainstream pathway mediated by *4CL* may not be exactly the same in all plant species or tissues. Nevertheless, monolignol biosynthesis is tightly controlled by *4CL*. Suppression of *4CL* expression through antisense technology has demonstrated the effectiveness of reducing total lignin content (Lee et al., 1997; Hu et al., 1999; Li et al., 2003). In aspen, suppression of

4CL expression led to more than 55% lignin reduction in wood. Thus, technology aimed at *4CL* suppression could be applied to plant genetic modification for better fiber production and other utilizations.

Transgenic plants with reduced *4CL* activity have been produced in tobacco (Kajita et al., 1996, 1997), *Arabidopsis* (Lee et al., 1997), and aspen (Hu et al., 1999; Li et al., 2003). In tobacco, reduction of *4CL* by over 90% resulted in 25% less lignin. In poplar and *Arabidopsis* with a >90% reduced *4CL* activity, lignin content was reduced by 45–50%. In tobacco, the low *4CL* activity was associated with browning of the xylem tissue (Kajita et al., 1996). In transgenic aspen down-regulated for *4CL*, Hu et al. (1999) also detected an increase in non lignin alkali-extractable wall-bound phenolics (p-coumaric acid, caffeic acid, and sinapic acid), which were not incorporated into the lignin polymer. However, no difference in lignin S/G composition for *Arabidopsis* and tobacco was observed. Discrepancy between the results published by Kajita et al. (1997) and Hu et al. (1999) is that the transgenic tobacco lines with the most severe reduction in lignin content (25%) were characterized by a collapse of vessel cell walls and reduced growth (Kajita et al., 1997), whereas the transgenic poplars with a 45% reduction in lignin content had a normal cell morphology and a higher growth rate than the control (Hu et al., 1999). The increased level of hydroxycinnamic acids as non-lignin cell wall constituents has been suggested to contribute to the cell wall strength in transgenic poplar (Hu et al., 1999). Because several 4CL isozymes exist with different cell-specific expression, down-regulation of several or all isozymes simultaneously may perturb metabolite levels other than those involved in lignin, with a secondary effect on growth as a consequence. Antisense inhibition of *4CL* in aspen trees led to a 15% increase in cellulose content. These results suggest that lignin and cellulose deposition are regulated in a compensatory fashion and that a reduced carbon flow toward phenylpropanoid biosynthesis increases the availability of carbon for cellulose biosynthesis (Hu et al., 1999; Li et al., 2003).

A combinatorial down-regulation of *4CL* along with an over expression of *F5H* in xylem has been achieved by co-transformation of two *Agrobacterium* strains in aspen (Li et al., 2003). Additive effects of independent transformation were observed, in particular a 52% reduction in lignin content associated with a proportional increase in cellulose and a higher S/G ratio.

1.5.5 Caffeoyl coenzyme A 3-O- methyltransferase (CCoAOMT) and Caffeate O-methyltransferase (COMT) or 5-hydroxyconiferaldehyde O-methyltransferse (AldOMT)

In monolignol biosynthesis, methylation is required at the 3-carbon and at the 5-carbon positions on the aromatic ring of the monolignol unit. The 3-carbon methylation leads to guaicyl unit formation and methylations on the 3- and 5-positions results in a syringyl unit. In early studies, it was postulated that two types of methyltransferases were necessary for methylations (Higuchi, 1997). Mono-functional *O*-methyltransferase to methylate the 3-position and therefore controlled the G monolignol unit biosynthesis, and the bi-functional *O*-methyltransferase to catalyze both 3- and 5-methylations leading to S monolignol unit biosynthesis. However, molecular studies have revealed that there are two kinds of genes encoding for the enzymes that specifically catalyze the 3- and 5-methylation, respectively (Ye et al., 1994; Li et al., 1999, 2000; Chen et al., 2001). Biochemical evidence from the studies using tree material suggests that the 3- and 5-methylations occur at different biosynthesis stages. The 3-methylation occurs on the CoA ester intermediate while the 5-position is methylated at the aldehyde intermediate (Osakabe et al., 1995; Li et al., 2000). The two genes are: one encoding a *CCoAOMT* and the other encoding a 5-*O*-methyltransferase that preferably methylates 5-hydroxyconiferaldehyde. The 5- *O*-methyltransferase was thought to methylate caffeic acid and was named *COMT* accordingly. However, more recent evidence indicates that the methylation catalyzed by COMT enzyme basically occurs at 5-hydroxyconiferaldehyde, therefore the enzyme was renamed *AldOMT* (Li et al., 2000). In addition to these two types of OMTs involved in the monolignol biosynthesis in angiosperms, another OMT (named AEOMT) that can methylate both hydroxycinnamic acids and hydroxycinnamoyl CoA esters was found in the gymnosperm loblolly pine (Li et al., 1997). However, identification of *AldOMT* in gymnosperms has not been reported. It appears that *CCoAOMT* gene plays a predominant role in the gymnosperm lignin biosynthesis. In loblolly pine *CCoAOMT* was detected as a single copy and specifically expressed in developing xylem (Li et al., 1999).

Down-regulation of COMT activity has been achieved using either antisense or sense transgenes in tobacco (Dwivedi et al., 1994; Ni et al., 1994; Atanassova et al., 1995), poplar (Van Doorsselaere et al., 1995; Tsai et al., 1998; Jouanin et al., 2000) and alfalfa (Guo et al.,

2001). In all three species, drastic reductions in the lignin S/G ratio were apparent and an unusual phenolic compound 5-hydroxyconiferyl alcohol (5OHG) was present in the polymer (Atanassova et al., 1995; Van Doorsselaere et al., 1995; Tsai et al., 1998; Lapierre et al., 1999; Jouanin et al., 2000; Guo et al., 2001; Marita et al., 2003). In the lignin of the transgenic poplars described by Jouanin et al. (2000), the level of 5OHG units even exceeded that of S units. In COMT suppressed alfalfa, β-β, β-1, and β-5 linkages involving S units, were absent (Guo et al., 2001), whereas in *COMT*-down-regulated poplar, free phenolic groups in β-O-4-linked G units were less abundant (Lapierre et al., 1999). In *COMT*-suppressed poplar, reduced lignin content has been reported (Jouanin et al., 2000) as well as no change in lignin amount in poplar (Van Doorsselaere et al., 1995) and aspen (Tsai et al., 1998). Reports describing *COMT*-suppression in tobacco also differ on whether lignin content is (Ni et al., 1994) or is not (Dwivedi et al., 1994; Atanassova et al., 1995) reduced. Despite these discrepancies, the data from all of the *COMT*-suppressed tobacco and poplar plants indicate that COMT plays a predominant role in determining the incorporation of S units into the lignin polymer. In alfalfa, a reduction in *COMT* activity affected both the content of G and S units (Guo et al., 2001; Marita et al., 2003). In accordance with these results, Parvathi et al. (2001) found that in alfalfa COMT is also involved in the methylation of caffeoyl aldehyde.

Down regulation of *CCoAOMT* affected the Klason lignin content by 12–50% in transgenic tobacco (Zhong et al., 1998; Pincon et al., 2001a), alfalfa (Guo et al., 2001; Marita et al., 2003) and poplar (Meyermans et al., 2000; Zhong et al., 2000). In tobacco and poplar, the decreased lignin content was due to reduction of both G and S units as determined by pyrolysis GC-MS (Zhong et al., 1998, 2000) or thioacidolysis (Meyermans et al., 2000). Because, the decrease in G units was more pronounced, the S/G ratio increased (Zhong et al., 1998; Meyermans et al., 2000). In contrast, the S unit amount was not reduced in transgenic alfalfa (Guo et al., 2001) nor transgenic tobacco (Pincon et al., 2001a). In contrast to the transgenic poplars, which were not affected in growth or morphology, the transgenic tobacco plants down-regulated for *CCoAOMT* had collapsed vessel walls (probably because of the reduced lignin content) and altered growth and flower development (Pincon et al., 2001a).

Simultaneous down-regulation of both COMT and CCoAOMT in tobacco (Zhong et al., 1998; Pincon et al., 2001a) and alfalfa (Guo et al., 2001) resulted in combinatorial and/or

additive effects. There was a greater reduction in Klasone lignin content in tobacco (Pincon et al., 2001a) but not in alfalfa (Guo et al., 2001). In both species, the lignin S/G ratio was reduced although in tobacco this was due to decreases in both G and S units (Zhong et al., 1998), whereas only S units decreased in alfalfa (Guo et al., 2001).

1.5.6 Cinnamoyl coenzyme A reductase (CCR)

The reduction of cinnamoyl CoA esters to cinnamaldehydes is the first metabolic step committed to monolignol formation. This step is catalyzed by CCR. Many studies of CCR activity indicated that five cinnamoyl-CoA esters *viz.* *p*-coumaroyl- CoA, caffeoyl-CoA, feruloyl-CoA, 5-hydroxyferuloyl-CoA and sinapoyl-CoA, could be used as substrate (Wengenmayer et al., 1976; Luderitz and Grisebach, 1981; Sarni et al., 1984; Goffner et al., 1994). The CCR enzyme purified from *Eucalyptus* xylem tissue was active toward *p*-coumaroyl-CoA, feruloyl-CoA, caffeoyl-CoA and sinapoyl-CoA with approximately equal affinity (Goffner et al., 1994). Similar to the native protein, the recombinant *Eucalyptus* CCR protein was also demonstrated to be active with the substrates *p*-coumaroyl-CoA, feruloyl-CoA and sinapoyl-CoA (Lacombe et al., 1997). Recently, the characterization of aspen CCR recombinant protein indicated that CCR selectively catalyzed the reduction of feruloyl-CoA from the five cinnamoyl CoA esters (Li et al., 2005). When CCR and CCoAOMT were coupled together, the linked reactions constitute the pathways from caffeoyl-CoA ester to coniferaldehyde (Fig. 5). In addition, the results also suggested that the neighboring CCoAOMT and CCR enzymes require different pH environments and compartmentalization *in vivo*. The *CCR* genes in various species appear as a multiple members family. In the *Populus* genome, there exist 8 CCR-homolog or *CCR*-like gene sequences.

Transgenic tobacco (Piquemal et al., 1998; Ralph et al., 1998; O'Connell et al., 2002) and *Arabidopsis* (Goujon et al., 2003b) down-regulated for *CCR*, are characterized by an approximate 50% decrease in Klason lignin. The lignin S/G ratio was increased (mainly because of a decrease in the G unit amount) in transgenic tobacco and variable, depending on the growth conditions, in transgenic *Arabidopsis*. A change in the lignin structure was also indicated by the higher amount of alkali-labile material that could be released from the extractive-free lignin polymer of the transgenic lines (O'Connell et al., 2002). The transgenic plants with the lowest *CCR* activity and 50% reduced lignin had abnormal phenotypes, such

as collapsed vessels, stunted growth, and abnormal leaf development. Important alterations in the fiber cell walls were observed, such as a loosening in the arrangement of the cellulose microfibrils, that resulted in reduced cell wall cohesion (Pincon et al., 2001b; Goujon et al., 2003b). Also an increased amount of tyramine ferulate (Figure 5), an unsual component of tobacco cell walls that is probably a sink for feruloyl-CoA, was incorporated into the lignin of the CCR-down-regulated tobacco plants (Ralph et al., 1998). A *CCR* mutant, designated irregular xylem (*irx4*), has been identified in *Arabidopsis* (Jones et al., 2001*).* Like the *CCR*-down-regulated tobacco and *Arabidopsis* described above, this mutant is characterized by a 50% reduced lignin content collapse of the vessels, and an altered growth and morphology.
By crossing transgenic tobacco down-regulated for *COMT* (Atanassova et al., 1995) with tobacco down-regulated for *CCR* (Piquemal et al., 1998), a simultaneous reduction in *COMT* and *CCR* expression was achieved (Pincon et al., 2001b).

1.5.7 Ferulate 5-hydroxylase/ Coniferaldehyde 5-hydroxylase (F5H/CAld5H)

A necessary step to biosynthesize S monolignols is hydroxylation at the 5-position on the aromatic ring of cinnamic intermediates. This reaction was thought to occur using ferulic acid as the substrate and catalyzed by F5H, which is encoded by a P450 protein gene belonging to *CYP84* family. Although forward genetics evidence demonstrated that *F5H* gene is essential for S-lignin formation in Arabidopsis (Meyer et al., 1996), it was unable to identify the intermediate on which the 5-hydroxylation biochemically occurs. Homologous genes have been cloned from a number of tree species. The biochemical function of this P450 gene was first demonstrated by expressing a sweetgum *CYP84* gene in yeast (Osakabe et al., 1999). The biochemical data suggest that the *CYP84* protein catalyzes 5-hydroxylation using coniferaldehyde, instead of the postulated ferulic acid, as a substrate to produce 5-hydroxyconiferaldehyde. Thus, *F5H* is actually a *CAld5H*. The 5-hydroxylation of coniferaldehyde was further confirmed with an Arabidopsis *CYP84* recombinant protein (Humphreys et al., 1999). According to the biochemical function of this *CYP84* gene, it was suggested that the S-monolignol biosynthesis pathway is branched out from a guaicyl intermediate at coniferaldehyde. Consistent with this view, 5-hydroxyconiferaldehyde is then methylated by *COMT* or *AldOMT* as described above. The genetic function of *CYP84* is also demonstrated through a reverse genetics approach by over expression of the gene, which

leads to the intensified S units in lignin (Franke et al., 2000; Li et al., 2003). Because the lignin with higher percentages of S-unit has a potentially significant value in the pulping economy (Chang and Sarkanen, 1973), over expression of *CAld5H* gene in trees has great potential to produce desirable wood material for fiber production.

An *Arabidopsis* mutant deficient in *F5H* (*fah1*) has been described and it produced lignin deficient in S units (Chapple et al., 1992) with a consequently increased frequency of phenylcoumaran (β-5) and biphenyl (5-5) linkages (Marita et al., 1999). When *Arabidopsis F5H* was overexpressed from the *C4H* promoter in the mutant, a lignin almost entirely composed of S units linked by β-O-4 linkages was produced (Meyer et al., 1998; Marita et al., 1999). The proportion of S units in the lignin of these plants was the highest ever reported for any plant (Ralph, 1996). Similarly, lignin of tobacco and poplar transformed with the same chimeric gene was enriched in S units (Franke et al., 2000). Li *et al*. (2003) overexpressed a sweetgum *F5H* (*Cald5H*) under the control of a xylem-specific promoter (*Pt4CL1P*) in transgenic aspen and reported a 2.5-fold increase in the S/G ratio and no changes in lignin content. An accelerated maturation/lignification of stem secondary xylem cells was noted in these *F5H* over expressing plants (Li et al., 2003). A 25–35% reduction in Klasone lignin content was observed in *F5H*-overexpressing *Arabidopsis* (Marita et al., 1999) and tobacco (Franke et al., 2000).

1.5.8 Cinnamyl alcohol dehydrogenase (CAD) and Sinapyl alcohol dehydrogenase (SAD)

In gymnosperm wood, coniferyl alcohol is the major monolignol unit while both coniferyl alcohol and sinapyl alcohol are monolignols in angiosperm wood. The last metabolic step forming these monolignols is reduction of coniferaldehyde and sinapaldehyde. CAD is suggested to catalyze multiple cinnamyl alcohol formations from their corresponding cinnamaldehydes (Lewis and Yamamoto, 1990; Whetten and Sederoff, 1995; Whetten et al., 1998). In loblolly pine, *CAD* is a single copy gene and its mutation leads to abnormal lignin formation in wood (MacKay et al., 1997; Lapierre et al., 2000). When the *Populus* tree was studied for monolignol biosynthesis in wood-forming tissue, in addition to *CAD*, it was found in aspen that another gene, its sequence similar to but distinct from *CAD*, is also associated with lignin biosynthesis (Li et al., 2001). The biochemical characterization of the

recombinant protein encoded by this gene indicated that the enzymatic activity has specific affinity toward sinapaldehyde, therefore it was named SAD. Compared with SAD enzyme kinetics, CAD showed a catalytic specificity towards coniferaldehyde instead. The catalytic specificities of the two enzymes have recently been further verified in protein structure analysis (Bomati and Noel, 2005). Furthermore, it was demonstrated that the expression of *CAD* is associated with G-lignin accumulation while *SAD* was associated with S-lignin formation during xylem differentiation (Li et al., 2001). The evidence from molecular, biochemical and cellular characterizations strongly suggest that *CAD* is involved in G-monolignol biosynthesis and *SAD* in S-monolignol biosynthesis in aspen wood formation. However, a recent genetic study using an Arabidopsis model system suggests a broad CAD function for both G- and S-lignin biosynthesis in the herbaceous species (Sibout et al., 2005). Nevertheless, more evidence connecting the biochemical function to its genetic role may be required in order to completely understand how *CAD* and *SAD* genes play a role in monolignol biosynthesis during wood formation.

Transgenic plants with reduced *CAD* activity have been produced in tobacco (Halpin et al., 1994; Hibino et al., 1995; Stewart et al., 1997; Yahiaoui et al., 1998), poplar (Baucher et al., 1996) and alfalfa (Baucher et al., 1999), whereas *CAD* mutants exist in pine (MacKay et al., 1997), maize (Halpin et al., 1998) and *Arabidopsis* (Sibout et al., 2003). An unusual monomer, dihydroconiferyl alcohol, was shown to be incorporated into the lignin of the pine *CAD* mutant and accounted for 30% of the lignin compared to only 3% in wild-type lignin (Ralph et al., 1997). Accordingly, higher amounts of arylpropane-1,3- diol structures, arising from dihydroconiferyl alcohol, have been found in the lignin of the pine *CAD* mutant (Ralph et al., 1999b, 2001a). In contrast, no dihydroconiferyl alcohol has been found in the lignin of transgenic angiosperms down-regulated for *CAD*, such as tobacco and poplar (Ralph et al., 1998).

A higher amount of cinnamaldehydes has been detected in the lignin of *CAD*-down-regulated tobacco (Halpin et al., 1994: Ralph et al., 1999a, 2001a), poplars (Kim et al., 2002), pine *CAD* mutant (Ralph et al., 1997) and in the *Arabidopsis Atcad-D* mutant (Sibout et al., 2003). However, only a slighty lower Klasone lignin content was measured in the wood of transgenic poplar lines downregulated for *CAD* (Lapierre et al., 1999; Pilate et al., 2002), in the pine *CAD* mutant (MacKay et al., 1997) and in the *Arabidopsis Atcad-D* mutant (Sibout

et al., 2003). The lignin of plants with low *CAD* activity was more extractable in alkali (Halpin et al., 1994; Baucher et al., 1996; Bernard-Vailh´e et al., 1996; Yahiaoui et al., 1998; MacKay et al., 1999).

The S/G ratio of the lignin of transgenic tobacco (Ralph et al., 1998) and transgenic alfalfa (Baucher et al., 1999) was reduced, suggesting that in these plants the uncondensed S structures are more affected than their G analogs.

These data are in apparent conflict with the recent proposal that *SAD*, and not *CAD*, is involved in S lignin biosynthesis in angiosperms (Li et al., 2001).

A simultaneous down-regulation of *CAD* and *CCR* has been achieved by crossing homozygous transgenic lines in which either *CAD* (Halpin et al., 1994) or *CCR* (Piquemal et al., 1998) was downregulated (Chabannes et al., 2001b). The lignin content was decreased by approximately 50% in tobacco with 32% of wild-type *CCR* activity and 12% of wild-type *CAD* activity. The phenotype of the double transformants was normal with only slight alterations in the vessel shape, showing that, similarly to the results of Zhong et al. (1998) and Hu et al. (1999), plants can also tolerate important reductions in lignin content. A simultaneous suppression of *COMT* (to 24% of wild-type level), *CCR* (to18%of wild-type evel), and *CAD* (to 4% of wild-type level) was achieved in tobacco by a single chimeric construct, consisting of partial sense sequences for the three different genes. The transgenic lines were stunted and had characteristics of *COMT*, *CCR*, and *CAD* suppression in lignin; for example, the xylem was red (indicative of *CAD* suppression), contained collapsed vessels (indicative of *CCR* suppression), and had reduced staining for S lignin (indicative of *COMT* suppression) (Abbott et al., 2002).

In analysis of functional genes involved in monolignol biosynthesis of angiosperm, three genes, *CAld5H*, *AldOMT* and *SAD*, control a line of consecutive metabolic steps and constitute a pathway toward S-monolignol biosynthesis. These three genes have not been known to be present in gymnosperm species that do not synthesize S-lignin. Gymnosperm wood is primarily comprised of tracheid elements, but angiosperm wood contains two types of thickened secondary wall cells, vessel element and fiber cells. Apparently the fiber cell is evolved along with occurrence of angiosperm species. It is known that G-monolignol units are dominant in tracheids and vessels and S units predominate fiber cells. It can be postulated that monolignol biosynthesis pathway evolution may be correlated with cell type

specification in the course of plant evolution; however, this hypothesis remains to be verified. It is believed that lignin is polymerized at the outside of the plasma membrane in secondary cell walls. Thus, monolignols that are synthesized inside plasma membrane need to be transported across plasma membranes for polymerization. Based on biochemical and cellular evidence, it has been suggested that laccases and peroxidases may be two types of possible enzymes involved in lignin polymerization (Bao et al., 1993; Christensen et al., 1998; Ostergaard et al., 2000). However, convincing genetic evidence to support this suggestion is lacking. As to how lignin is polymerized from monomers, has been debated for long. Whether lignin is polymerized from monomer units randomly or in a guided way remains an enigma (Ralph et al., 2004; Davin and Lewis, 2005). A gene encoding a dirigent protein was cloned and the biochemical results suggested the dirigent protein might play a role in guiding a stereo-specific lignin polymerization (Davin et al., 1997). This hypothesis still remains to be confirmed. On the other hand, monolignol glucosides are found in developing xylem tissues in many species but currently it is unclear whether the monolignol glucosides are used as intermediates for monolignol storage or for transportation crossing the plasma membrane (Dharmawardhana et al., 1995; Meyermans et al., 2000; Steeves et al., 2001; Tsuji and Fukushima, 2004). Overall, although there are various studies on the process of monolignol cross-membrane transportation and lignin polymerization, the genetic and molecular evidence is elusive regarding what chemical format is taken for the transportation and the mechanisms of how monolignols are transported to the outside of membrane where they are polymerized into lignin.

1.6 Transport of monolignols

After their synthesis, the lignin precursors are transported to the cell wall where they are oxidized and polymerized. In gymnosperms and some angiosperms, monolignol 4–*O*–β-D-glucosides accumulate to high levels in the cambial tissues (Steeves et al., 2001). It has been hypothesized that these monolignol glucosides are storage or transport forms of the monolignols and that a uridine diphosphate glucose (UDPG) coniferyl alcohol glucosyl transferase (Steeves et al., 2001), together with coniferin-β-glucosidase (CG), may regulate storage and mobilization of monolignols for lignan and/or lignin biosynthesis (Dharmawardhana et al., 1995; Samuels et al., 2002). Whether these glucosides are

transported via Golgi-derived vesicles or through direct plasma membrane pumping by ABC transporters is still unknown (Samuels et al., 2002).

According to early studies with radiolabeled monolignol precursors, lignification of the cell wall has been hypothesized to proceed after cell death (Pickett-Heaps, 1968). These findings are now supported by experiments with the zinnia cell system showing that lignification of tracheary elements that have undergone programmed cell death still progresses by supply of monolignols from the surrounding xylem parenchyma cells (Hosokawa et al., 2001).

1.7 Dehydrogenation

After transport of the monolignols to the cell wall, lignin is formed through dehydrogenative polymerization of the monolignols (Christesen et al., 2000). The dehydrogenation to monolignol radicals has been attributed to different classes of proteins, such as peroxidases, laccases, polyphenol oxidases, and coniferyl alcohol oxidase. Which of these enzymes or a combination thereof are responsible for the dehydrogenation of the monolignols in planta and whether monolignol oxidation occurs through redox shuttle-mediated oxidation are still unclear (Onnerud et al., 2002)?

Although peroxidases are believed to catalyze the final condensation of cinnamyl alcohols in the formation of lignin, no definitive proof has been presented yet for the involvement of any specific peroxidase isozyme *in vivo*, mainly because of the high number of genes that encode peroxidases (Tognolli *et al*., 2002) and the typically low substrate specificities of these enzymes. Peroxidases use hydrogen peroxide (H_2O_2) to oxidize their substrates (Christesen et al., 2000). How H_2O_2 is generated in the cell wall is still a matter of debate. Evidence is emerging for a role for an NADPH oxidase in lignifying tissues, which would supply H_2O_2 for monolignol oxidation (Ogawa et al., 1997; Ros-Barcela et al., 2002).Peroxidase itself is able to generate H_2O_2 from a variety of reducing substrates, such as cysteine, glutathione, NADPH, ascorbate, and indole-3-acetic acid (Bolwell et al., 1995, Christesen et al., 2000; Ferrer et al., 1990).

Laccases (*p*-diphenol: O_2 oxidoreductases) are copper-containing, cell wall–localized glycoproteins that are encoded by multigene families in plants. In contrast to peroxidases, laccases consume O_2 instead of H_2O_2 to oxidize the monolignols. Laccases of a variety of species are expressed in lignifying cells (Bao et al., 1993; Driouich et al., 1992; Ranocha et

al., 1999; Sterjiades et al., 1992). The precise role played by laccases in lignification is not yet understood, but there is correlative evidence that laccase and oxygen participate in the polymerization of monolignols. For instance, when peroxidase is inhibited either in the absence of H_2O_2 or in the presence of H_2O_2 scavengers (catalase and superoxide dismutase), coniferyl alcohol is still oxidized and O_2 consumed in tobacco xylem (McDougall *et al*., 1994). Because laccases operate in the absence of toxic H_2O_2, these enzymes could be involved in the early stages of lignification (Sterjiades *et al*., 1993). Five divergent laccase genes have been cloned and characterized from poplar (Ranocha *et al*., 1999, 2000).

Both anionic and cationic peroxidases have been implicated in lignification based on their affinity for coniferyl alcohol, their location in the cell wall and their expression in lignified tissue (M"ader and F¨ussl, 1982; Lagrimini et al., 1987; El Mansouri et al., 1999). Nevertheless, no change in lignin content was obvious in transgenic tobacco plants that were deficient in the major anionic peroxidase (Lagrimini et al., 1997a). However, transgenic poplar with a 44% reduction in the activity of a stem-specific anionic peroxidase (PRXA3a) had a 21% reduced lignin content and a higher content in β-O-4 linked (uncondensed) structures in lignin (Yahong et al., 2001). The over expression of peroxidase genes in transgenic poplar (PXP 3-4; Christensen et al., 2001a, 2001b) and in tobacco (spi 2; Elfstrand et al., 2002) resulted in 800-fold and 5-fold increased total peroxidase activity, respectively. No effects of the genetic modification on the overall phenotype, the Klasone lignin content, or the recovery yield of G and S units were identified. In addition, the tobacco plants with higher peroxidase activity were more susceptible to *Phytophthora parasitica*, but allowed less growth of *Erwinia carotovora.* In contrast, independently produced transgenic tobacco lines with a 10-fold higher peroxidase activity were characterized by increased lignin content in leaves, stems, and roots (Lagrimini, 1991; Chabbert et al., 1992). No changes in lignin monomeric composition were detected, but the amount of monomers involved in β-O-4 linkages was lower in the transgenic lines, suggesting that the lignin was more condensed (Chabbert et al., 1992). Overproduction of an anionic peroxidase in tomato led to an increase in lignin content in leaves and fruit (Lagrimini et al., 1993; El Mansouri et al., 1999).

Transgenic *Liriodendron* (Dean *et al*., 1998) and poplar (Ranocha et al., 2000, 2002) down-regulated in laccase had neither altered phenotype nor any change in lignin amount or S/G composition. However, in transgenic poplar down-regulated for one of the laccase genes

(*lac3*), the walls of xylem cells were irregular in contour when compared with the control and had adhesion defects either at the primary cell wall of adjacent cells or within the secondary cell wall of a given cell (Ranocha et al., 2002).

1.8 Polymerization

1.8.1 Radical generation and radical coupling

After their dehydrogenation, the radicals, which are relatively stable owing to electron delocalization that provides single-electron density to the side-chain β position, are coupled. The most important reaction is cross-coupling to the growing polymer to extend the complex three-dimensional lignin network (Fig. 3). But, such coupling reactions are radical quenching. Each extension of the polymer requires new radicals on each of the two coupling partners. Radicals on the growing lignin polymer are thought to be generated by radical transfer from monolignols or other intermediaries. (Takahama and Oniki, 1994).

Similar radical transfer mechanisms can be envisioned between the monolignols and the growing polymer, *i.e.* the monolignols may act as the radical shuttles. When a monolignol radical encounters a polymer radical, it may cross-couple with it, but when the polymer is not electron-deficient, radical transfer may occur and the monolignol will diffuse back to the peroxidase/laccase to be reoxidized. Alternatively, redox shuttles, such as a Mn^{2+}/Mn^{3+} system (Onnerud et al., 2002), may be involved.

1.8.2 Polymerization Process

The actual process of lignin polymer formation occurs without the rigid biochemical controls seen in the biosynthesis of the precursor monolignols, giving rise to a unique class of polymers. Lignins are racemic (Ralph et al., 1999), deriving from radical coupling reactions under chemical (but no apparent biochemical control) between phenolic radicals in an essentially combinatorial fashion.

The accepted model for lignin polymerization, based on simple chemically controlled combinatorial coupling reactions, was recently challenged by Devin and Lewis (2000). It was proposed that the macromolecular assembly of lignin is not based on “random coupling” of monolignols. The new theory arose from the discovery of a class of dirigent proteins implicated in lignan biosynthesis (Davin et al., 1997). The first such dirigent protein

discovered guided the dimerization of coniferyl alcohol radicals to produce an optically active lignan, pinoresinol. The corresponding gene was cloned and shown to encode a cell wall–localized protein. The finding was extrapolated to lignification, suggesting that such proteins would logically be responsible for specifying the exact structure of the lignin polymer, bringing lignins in line with proteins and polysaccharides that are more carefully biosynthesized (Devin and Lewis, 2000; Lewis, 1999).

1.8.3 Nucleation sites

Lignin is first deposited in the middle lamella and the cell corners of the primary wall after the formation of the secondary wall has started, nucleation sites, from which the lignin polymers can grow. The nature of these nucleation sites is unknown. Ferulates, conjugated to polysaccharides, and their dehydrodimers are well established. Evidence that ferulates and diferulates may act as attachment sites for monolignols has been shown (Ralph et al., 1995). Given that the middle lamella and the cell corners are rich in Ca^{2+} pectate (Carpita and Gibeut, 1993) and are the first sites to be lignified, Ca^{2+} pectate-bound peroxidases may conceivably play a role in the spatial control of lignin deposition, and changes in Ca^{2+} and H^+ concentrations may modulate the location of these peroxidases (Carpin et al., 2001).

The negatively charged pectins are also good binding sites for polyamines (Carpin et al., 2001) and, hence, may be suitable sites for H_2O_2 generation by polyamine oxidases (Moller and McPherson, 1998). Pectin-binding peroxidases and polyamine oxidases may act locally in the early stages of lignin deposition both for H_2O_2 generation and oxidation of monolignols, cinnamic acids bound to polysaccharides or polyamines, or aromatic residues on certain proteins, such as glycine-rich proteins (Keller et al., 1989).

As described above, many different transgenic plants and a few mutants are now available with altered lignin content, altered lignin composition/ structure, or both. Whether or not these changes in cell wall biochemistry could have advantages in industrial operations, such as pulping, can only be determined experimentally. A few such experiments have been performed. To put the results into context, we will first describe briefly the different methods and parameters important in the production of pulp and paper before reviewing the impact that specific genetic engineering can have on pulping properties.

1.9 Why *Leucaena*?

Paper industry in India mainly uses bamboos, *Eucalyptus* sp., *Casuarina* sp. and *Leucaena* sp. as a source for paper pulp. Selection of the species depends upon availability, price and acceptability by any one given industrial unit. In bamboo growing countries, like India, the proportionate use of bamboos and hardwood species is in the ratio of 15:85. Although all these plant species are of importance to the paper industry, *Leucaena sp.* is exclusively used in India and about 25% of raw material for pulp and paper industry comes from this hard wood tree. To meet the increasing demand of high quality wood for paper industry it is essential to provide designer plant species. However, as a safeguard for the future no plant should be harvested from areas that may challenge sustainability. It will thus be crucial to raise plantations of the plant species with elite materials and or genetically modified plants that meet the demands of the pulp and the paper industry in economical and sustainable manner.

1.9.1 *Leucaena*

Leucaena is a native of Central America has been naturalized pan-tropically. Members of the genera are vigorous, drought tolerant, highly palatable, high yielding, rich in protein and grow in a wide range of soils (Jones, 1979; Hughes, 1998). However, these attributes are limited by the occurrence of anti-nutritive factors in the fodder, such as tannins and mimosine (Jones, 1979; Hegarty et al., 1964b; Hammond et al., 1989 a, b).

Leucaena occupies 2 to 5 million hectares of land worldwide (Brewbaker and Sorenson, 1990). They are recognized as some of the fastest growing and most useful trees in the tropics (NRC, 1984).

Leucaena is represented by 22 species. Of these, 6 are intraspecific taxa and 2 are widespread spontaneous hybrids. Most of the species are diploid 2n=52 or 56. However, 4 species are tetraploid 2n = 4X=104 or 112) (Hughes, 1998). *L. leucocephala* is a member of the genus related to the other species within the Mimosoideae sub-family, its subspecies and other related genera.

Family: Leguminosae
Sub family: Mimosideae
Sub tribe: Mimoseae
Genus: *Leucaena*
Species: *leucocephala*
Sub species: Glabrata (Rose; S. Zarate); Ixtahuacana (Hughes) and Leucocephala (Benth) Var. Peru and Cunningham
Related genera: *Desmanthus*; *Schleinitzia*; *Calliandropsis*; *Neptunia; Alantsilodendron*; *Gagnebina*; *Dichrostachys*; and *Kanaloa*

1.9.2 *Leucaena leucocephala* species growth forms

L. leucocephala has three growth forms that are commonly found in tropics (Hughes, 1998). The first and common type is a shrubby free shading form known as the Hawaiian *Leucaena*, which is weedy and low yielding. The second growth form is giant type, which is a tall tree of about 20 meters, with large leaves, pod and seeds, and almost without branches. It is native of Central America and Mexico, and produces twice the biomass of the common type. It is suitable for timber, wood product and industrial fuel. The last is Peru type, which is a multi-branched, semi-erect, medium height (about 10 meters) and the most productive form.

1.9.3 Fodder production and feed value of *Leucaena leucocephala*.

Leucaena leucocephala produces 6 to 18 tones of forage dry matter per hectare in the wet season and 2 to 3 tones of dry matter per hectare in the dry season. All parts of *Leucaena* are edible to livestock (Shelton, 1995). *Leucaena* leaves contain about 19 to 26 percent crude protein. They are rich in source of carotene and vitamins. The provitamin-A in *Leucaena* is among the highest ever recorded in a plant specimens (Jones, 1994). *Leucaena* is rich in calcium and phosphorus, but deficient in sodium and iodine. The *in- vivo* digestibility is 50 to 70 percent (Norton and Poppi, 1995). The presence of mimosine, a tyrosine analogue, a toxic non protein amino acid limit the use of this plant as a forage (Allison et al., 1990; Gupta and Atreja, 1999; Tangendjaja and Willis, 1980)

1.9.4 Wood

Leucaena wood has a thin bark which is about 8% dry matter at the age of 5 years. The sapwood is yellow-white, while the heartwood is yellow to reddish brown. Bole wood has a

specific gravity of 0.54 at the age of 6 to 8 years. This is similar to the density, tensile, compression, bending and shear strength of oak, ash, birch and sugar maple. It is fine – textured and workable. It absorbs preservatives, and can be treated against termites (Pottinger and Huges,1995).

Leucaena wood is among the best hardwoods for the paper and rayon making. It produces pulp that is high in holocellulose, low in silica, ash, lignin, alcohol-benzene soluble and hot water soluble. Pulp yield is 50 to 52%. Its short fibre is suitable for rayan production (Pottinger and Hughes, 1995). Wood from giant *Leucaena* has a heating value of 4640 Kcal.kg^{-1} at the age of 2 to 4 years, and 7000 Kcal.kg^{-1} at the age of 8 years which is equivalent to 70% of the heating value of fossil fuel. In power generation, wood replaces fossil fuel in generating electricity and for the production of charcoal for producer gas generators that power vehicles, boats and irrigation pumps.

1.10 Caffeoyl CoA 3 O-methyltransferase (CCoAOMT; EC 2.1.1.104)

S-adenosyl-L-methionine methyltransferases (EC 2.1.1. -) are key enzymes in the phenylpropanoid, flavanoid and many other metabolic pathways in plants. The enzymes Caffeate 3-O methyltransferase (COMT; EC 2.1.1.68) and Caffeoyl CoA 3-O methyltransferase (CCoAOMT; EC 2.1.1.104) control the degree of methoxylation in lignin precursors *i.e. p*- hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units of lignin (Higuchi, 1990; Ralph et al., 1998; Boerjan et al., 2003). As discussed earlier until recently, the methylation reactions at the C3 and C5 hydroxyl functions of the lignin precursors were thought to occur mainly at the cinnamic acid level by bi-functional COMT. However, the association of CCoAOMT expression with lignification (Pakusch *et al*., 1991; Ye *et al*., 1994; Ye and Varner, 1995; Ye, 1997; Martz *et al*., 1998; Chen *et al*., 2000) and the observation that down-regulation of COMT preferentially affected the amount of S units suggested the existence of an alternative pathway for the methylation of the lignin precursors at the hydroxycinnamoyl-CoA level and specific O-methyltransferase *i.e* CCoAOMT converts caffeoyl CoA into feruloyl CoA and 5-hydroxyferuloyl CoA into sinapyl CoA (Martz et al., 1998).

The enzyme CCoAOMT supports the biosynthesis of guaiacyl lignin (reviewed in Boudet et al., 2000; Inoue et al., 2000; Guo et al., 2000). This enzyme also appears to be encoded by a

small gene family in various angiosperms for example tobacco (Martz et al., 1998), poplar (Chen et al, 2000), maize (Laura et al., 1999) etc.

Though an alternative pathway for the CCoAOMT catalyzed methylation of the lignin precursors at the hydroxycinnamoyl-CoA level was a new finding but Neish (1968) put forward a lignin biosynthetic pathway in which the carboxyl group was first activated on cinnamic acid or on *p*-coumaric acid, and the subsequent reactions including methylation occurred on the ester forms. This pathway was proposed based on the following lines of evidence. First, ester forms of p-coumaric acid, caffeic acid, and ferulic acid were detected in plants (El-Basyouni et al., 1964; El-Basyouni and Neish, 1966; Brown, 1966, 1969). Second, the free acids would tend to be insoluble in the acidic environment, so the ester forms would be preferred intermediates (Brown, 1969). However, this proposed methylation pathway using the ester forms instead of the acid forms was largely neglected since then.

CCoAOMT was first identified in parsley and carrot cultured cells challenged with fungal elicitors (Matern et al., 1988; Kuhnl et al., 1989; Pakusch et al., 1989). CCoAOMT from parsley and carrot specifically methylates caffeoyl CoA (Fig.1) but not caffeic acid, and its identification led to questions on the conventional view of methylation reactions on the acid forms (Kuhnl et al., 1989). They proposed roles for CCoAOMT in the biosynthesis of phenylpropanoid-derived compounds such as lignin. However, this proposal was not recognized because CCoAOMT in the cultured cells induced by fungal elicitors was thought to be a defense response (Pakusch et al., 1989) and there was no correlative evidence available to suggest that CCoAOMT expression was associated with lignification. The importance of CCoAOMT in lignin biosynthesis was not realized until the finding that CCoAOMT was specifically expressed in all lignified cells in Zinnia (Ye et al., 1994).

CCoAOMT is distributed ubiquitously in plants. After identification of the elicitor-inducible CCoAOMT in parsley and carrot, Schmitt et al.(1991) purified CCoAOMT, sequenced its peptides and cloned CCoAOMT cDNA from parsley using probes based on the peptide sequence. Based on their homology to the parsley CCoAOMT, CCoAOMT cDNAs and genomic DNAs have been isolated from a number of other plants such as zinnia (Ye et al., 1994), quaking aspen (Meng and Campbell, 1995), grape vine (Busam et al., 1997), tobacco (Busam et al., 1997; Martz et al., 1998), alfalfa (Inoue et al., 1998), eucalyptus (De Melis et al., 1998), pine (Li et al., 1999), poplar (Chen et al, 2000), maize (Laura et al., 1999) etc.

CCoAOMTs from different plants share high sequence identity at the amino acid level, but they show little sequence similarity to COMT except for the consensus regions involved in S-adenosyl-l-methionine and metal binding (Vidgren et al., 1994; Ibrahim et al., 1998). CCoAOMTs also share significant amino acid similarity to catechol OMT from rat and human (Salminen et al., 1990; Bertocci et al., 1991).

After the finding that CCoAOMT expression was closely associated with lignification in zinnia, the next important issue was to determine whether this expression pattern was common in plants. Localization of CCoAOMT at the protein or mRNA level confirmed that CCoAOMT was expressed in lignifying tissues in both herbaceous plants including parsley, tobacco, tomato, soybean and alfalfa, and woody plants including forsythia and poplar (Ye, 1997; Inoue et al., 1998; Kersey et al., 1999; Chen et al., 2000; Zhong et al., 2001). CCoAOMT was found to be present in tracheal elements, xylary fibers, phloem fibers and xylem ray parenchyma cells. Studies of CCoAOMT activity and its gene expression in tobacco, quaking aspen and pine also showed a close correlation with lignification in xylem (Martz et al., 1998; Meng and Campbell, 1998; Li et al., 1999; Maury et al., 1999). These results further indicated that association of CCoAOMT with lignification was ubiquitous in both herbaceous and woody plants.

Isolation of CCoAOMT promoters from parsley (Grimmig and Matern, 1997) and poplar (Chen et al., 2000), and their functional characterization also proved CCoAOMT gene expression in actively lignifying tissues. The gene promoters directed the expression of GUS reporter gene in lignifying tissues, hence proving CCoAOMT’s active participation during lignification.

Ferrer et al. (2003) crystallized alfalfa 28KDa CCoAOMT protein, consisting of 247 amino acids that form a homodimer in solution (Inoue et al., 1998). The crystallographic structures obtained with CoA-linked substrates bound reveal the architecture of dimeric form, which was novel, as it lacked the N-terminal dimerization domain present in previously characterized plant OMTs (Zubieta et al., 2001, 2002). The dimerization interface involved mostly hydrophobic interactions. The first 20 amino acid residues were not observed in the electronic density. Each monomer consists of a single catalytic domain, which exhibits a core a/b Rossmann fold that provides the binding site for SAM/SAH (Rossmann et al., 1974). All structurally characterized SAM-dependent OMTs share this architectural motif.

They also suggested catalytic mechanism based on the metal-mediated deprotonation of the caffeoyl 3-hydroxyl group followed by transmethylation due to the juxtapositioning of a reactive phenolic oxyanion near the reactive methyl group of SAM. It was shown that the environment around the caffeoyl 3-hydroxyl moiety of the substrate was electropositive due to the positively charged sulfur of SAM and the Ca^{2+} ion, and the caffeoyl 3-hydroxyl exists as an oxyanion in order to balance charge in the active site region. Chelation to the metal ion further positions the 3-oxyanion moiety in close proximity to the reactive methyl group of SAM. This metal dependent catalytic mechanism is also postulated for mammalian catechol OMT and related methyltransferases (Vidgren et al., 1994; Ibrahim et al., 1998).

1.11 Why CCoAOMT from *Leucaena*?

This study is the first instance towards isolation and characterization of lignin biosynthetic pathway gene(s) for the development of transgenic *Leucaena* plants. This study is intended to lay down criteria for the development of the transgenic plants which would ultimately strive for: (a) reduced lignin content, (b) altered ratio of S/G lignin, (c) increased cellulose content and (d) enhanced growth. These criteria would be met by isolation, cloning and characterization of the selected and key lignin biosynthesis pathway genes *viz.* cinnamyl alcohol dehydrogenase (CAD), 4-coumarate-CoA-ligase (4CL), Caffeate O-methyltransferase (COMT or AldOMT), caffeoyl-CoA-3-O-methyltransferase (CCoAOMT), cinnamoyl-CoA-reductase (CCR), UDP glucose-glucosyl transferase (UDPG-GT) and coniferin-β-glucosidase (CBG). A lignin biosynthesis related gene promoter sequences will also be helpful in this regard. The above genes of the lignin biosynthesis pathway in the target plant species could then be down regulated by affecting transformations with the gene constructs in sense or antisense orientations under the control of suitable tissue specific promoters. It will, however, be beyond the scope of this thesis to cover all the genes listed above.

Pulp and paper industry will benefit by either reduced lignin synthesis or by alteration of the lignin components in a manner where in the sinapyl alcohol to guaiacyl alcohol ratios are in favour of sinapyl alcohol. The later of the two may be achieved by down regulating the Caffeoyl CoA 3-O-Methyl Transferase (CCoAOMT; EC 2.1.1.104), which would divert the flux towards the formation of sinapyl alcohol. Hence, the present study was aimed at the

isolation of CCoAOMT genomic and cDNA gene(s) clones from *Leucaena leucocephala* and molecular characterization *vis-à-vis* their pattern of expression and promoter analysis.

1.12 References

Abbott JC, Barakate A, Pincon G, Legrand M , Lapierre C, Mila I, Schuch W, Halpin C (2002) Simultaneous suppression of multiple genes by single transgenes. Down-regulation of three unrelated lignin biosynthetic genes in tobacco. Plant Physiol 128: 844–853

Adler E (1977) Lignin chemistry—past present and future. Wood Sci Technol 11: 169–218

Akhtar M, Scott GM, Lentz MJ, Horn E, Scaney RE, Kirk, TK (1998b) Commercialization of biopulping for mechanical pulping. In Biotechnology in the Pulp and Paper Industry, Vol. A: Oral Presentations A55–A58. Paice M, Saddler J. Eds, Canadian Pulp and Paper Association, Montreal.

Akhtar, M, Lentz MJ, Swaney RE, Scott GM, Horn E, Kirk TK (1998a) Biopulping: technology learned from nature that gives back to nature. In Competing and Cooperating for Stakeholder Value, (Proceedings of the 23rd Annual Summit) 149–157. Technology Transfer Society, Ed, T2S, Chicago, IL.

Allison MJ, Hammond AC, Jones RJ (1990) Detection of ruminal bacteria that degrade toxic dihydroxypyridine compounds produced from mimosine. Appl Environ Microbiol 56: 590-594

Atalla RH, Agarwal UP (1985) Raman microprobe evidence for lignin orientation in the cell walls of native woody tissue. Science 227: 636–638

Atanassova R, Favet N, Martz F, Chabbert B, Tollier MT, Monties B, Fritig B, Legrand M (1995) Altered lignin composition in transgenic tobacco expressing O-methyltransferase sequences in sense and antisense orientation. Plant J 8: 465–477

Bao W, O'Malley DM, Whetten R, Sederoff RR (1993) A laccase associated with lignification in loblolly pine xylem. Science 260: 672–74

Bate NJ, Orr J, Ni W, Meromi A, Nadler-Hassar T, Doerner PW, Dixon RA, Lamb CJ, Elkind Y (1994) Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate-

determining step in natural product synthesis. Proc Natl Acad Sci USA **91**: 7608–7612

Baucher M, Bernard-Vailhe MA, Chabbert B, Besle J.-M., Opsomer C, Van Montagu M, Botterman J (1999) Down-regulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (*Medicago sativa* L.) and the impact on lignin composition and digestibility. Plant Mol Biol 39: 437–447

Baucher M, Chabbert B, Pilate G, Van Doorsselaere J, Tollier M.-T, Petit-Conil M, Cornu, D, Monties B, Van Montagu M, Inze D, Jouanin L, Boerjan W (1996) Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar (*Populus tremula* x *P. alba*). Plant Physiol 112: 1479–1490

Baucher M, Halpin C, Petit-Conil M, Boerjan W (2003) Lignin: Genetic Engineering and Impact on Pulping. Crit Rev Biochem Mol Biol 38: 305-350

Baucher M, Monties B, Van Montagu M, Boerjan W (1998) Biosynthesis and genetic engineering of lignin. Crit Rev Plant Sci 17: 125–97

Bell-Lelong DA, Cusumano JC, Meyer K, Chapple C (1997). Cinnamate-4-hydroxylase expression in *Arabidopsis*. Plant Physiol 113: 729–738

Bernard-Vailhe MA, Cornu A, Robert D, Maillot M-P, Besle J-M (1996) Cell wall degradability of transgenic tobacco stems in relation to their chemical extraction and lignin quality. J Agric Food Chem 44: 1164–1169

Bertocci B, Miggiano V, Prada MD, Dembic Z, Lahm, H-W, Malherbe P (1991) Human catechol-O-methyltransferase: cloning and expression of the membrane-associated form.Proc Natl Acad Sci 88: 1416–1420

Blee K, Choi JW, O'Connell AP, Jupe SC, Schuch W, Lewis NG, Bolwell GP (2001) Antisense and sense expression of cDNA coding for CYP73A15, a class II cinnamate 4-hydroxylase, leads to a delayed and reduced production of lignin in tobacco. Phytochemistry 57: 1159–1166

Blount JW, Korth KL, Masoud SA, Rasmussen S, Lamb C, Dixon RA (2000) Altering expression of cinnamic acid 4-hydroxylase in transgenic plants provides evidence for a feedback loop at the entry point into the phenylpropanoid pathway. Plant Physiol 122: 107–116

Boerjan W, Ralph J, Baucher M (2003) Lignin Biosynthesis. Annu Rev Plant Biol 54: 519-546

Bolwell GP, Buti VS, Davies DR, Zimmerlin A (1995) The origin of the oxidative burst in plants. Free Radic Res 25: 517–532

Bomati EK, Noel JP (2005) Structural and kinetic basis for substrate selectivity in Populus tremuloides sinapyl alcohol dehydrogenase. Plant Cell 17: 1598–1611

Boniwell JM, Butt VS (1986) Flavin nucleotide-dependent 3- hydroxylation of 4-hydroxyphenylpropanoid carboxylic acids by particulate preparations from potato tubers. Z Naturforsch 41: 56–60

Boudet AM (2000) Lignins and lignification: selected issues. Plant Physiol Biochem 38: 81-96

Boudet AM, Goffner D, Grima-Pettenati J (1997) In Lignin and Lignification: Recent biochemical and biotechnological advances. C R Acad Sci Paris pp 319, 317-331

Boudet AM, Lapierre C, Grima-Pettenati J (1995) Biochemistry and molecular biology of lignification. New Phytol 129: 203–236

Bray, RA (1995) The Leucaena psyllid. In: Gutteridge, R.C. and Shelton, H.M. (Eds), forage tree legumes in tropical agriculture. CAB International, Wallingford Oxon, U.K. pp 283

Breen A, Singleton FL (1999) Fungi in lignocellulose breakdown and biopulping. Curr Opin Biotech 10: 252–258

Brewbaker JL, Sorenson CT (1990) Leucaena: New tree crops from interspecific hybrids. In: Janick. J and Simon J. (Eds). Advances in new crops Timber Press, Oregon. pp 238

Brown SA (1966) Lignins. Annu Rev Plant Physiol 17: 223–244

Brown SA (1969) Biochemistry of lignin formation. Bioscience 19: 115–121

Busam G, Junghanns K, Kneusel RE, Kassemeyer HH, Matern,U (1997) Characterization and expression of caffeoyl-coenzyme A 3-O-methyltransferase proposed for the induced resistance response of *Vitis vinifera* L. Plant Physiol 115: 1039-1048

Buxton DR, Redfearn DD (1997) Plant limitation to fiber digestion and utilization. J Nutr 127: S814-S818

Campbell MM, Sederoff RR (1996) Variation in lignin content and composition. Mechanisms of control and implications for the genetic improvement of plants. Plant Physiol 110: 3–13

Carpin S, Cr`evecoeur M, de Meyer M, Simon P, Greppin H, Penel C (2001) Identification of a Ca_2C-pectate binding site on an apoplastic peroxidase. Plant Cell 13: 511–520

Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. Plant J 3: 1– 30

Chabannes M, Barakate A, Lapierre C, Marita JM, Ralph J, Pean M, Danoun S, Halpin C, Grima-Pettenati J, Boudet AM (2001b) Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) in tobacco plants. Plant J 28: 257–270

Chabannes M, Ruel K, Yoshinaga A, Chabbert B, Jauneau A, Joseleau J-P, Boudet A-M (2001a) In situ analysis of lignins in transgenic tobacco reveals a differential impact of individual transformations on the spatial patterns of lignin deposition at the cellular and subcellular levels. Plant J 28: 271–82

Chabbert B, Monties B, Liu YT, Lagrimini M (1992) Lignin content and composition in transgenic tobacco plants with altered peroxidase activity. In Proceedings of the 5th International Conference on Biotechnology in the Pulp and Paper Industry. pp 481–485 Kuwahara M, Shimada M, Eds., Uni Publishers Co., Tokyo.

Chang HM, Sarkanen KV (1973) Species variation in lignin, effect of species on the rate of kraft delignification. Techn Assoc Pulp Pap Ind 56: 132–143.

Chapple CCS, Vogt T, Ellis BE, Somerville CR (1992) An Arabidopsis mutant defective in the general phenylpropanoid pathway. Plant Cell 4: 1413–1424.

Chen C, Meyermans H, Burggraeve B, De Rycke RM, Inoue K, Vleesschauwer VD, Steenackers M, Van Montagu MC, Engler GJ, Boerjan, WA (2000) Cell-specific and conditional expression of caffeoyl-coenzyme A-3-O-methyltransferase in poplar. Plant Physiol 123: 853-867

Chen F, Kota P, Blount JW, Dixon RA (2001) Chemical syntheses of caffeoyl and 5-OH coniferyl aldehydes and alcohols and determination of lignin O-methyltransferase activities in dicot and monocot species. Phytochemistry 58: 1035–1042

Christensen JH, Baucher M, O'Connell AP, Van Montagu M, Boerjan W (2000) Control of lignin biosynthesis. In Molecular Biology of Woody Plants, Volume 1, ed. SM Jain, SC Minocha, For. Sci., 64: 227–67. Dordrecht: Kluwer. pp 520 Fengel D and Wegener G, Eds., De Gruyter, Berlin

Christensen JH, Bauw G, Welinder KG, Montagu M, Boerjan W (1998) Purification and characterization of peroxidases correlated with lignification in poplar xylem. Plant Physiol. 118: 125–135

Christensen JH, Overney S, Rohde A, Ardiles Diaz W, Bauw G, Simon P, Van Montagu M, Boerjan W (2001b) The syringaldazine-oxidizing peroxidase PXP 3-4 from poplar xylem: cDNA isolation, characterization and expression. Plant Mol Biol 47:581–593

Christensen JH, Van Montagu M, Bauw G, Boerjan W (2001a) Xylem peroxidases: purification and altered expression. In Molecular Breeding of Woody Plants, (Progress in Biotechnology Series, Vol. 18), pp 171–176. Morohoshi N and Komamine A, Eds., Elsevier Science, Amsterdam.

Cramer CL, Edwards K, Dron M,. Liang X, Dildine SL, Bolwell GP, Dixon RA, Lamb CJ, Schuch W (1989) Phenylalanine ammonialyase gene organization and structure. Plant Mol Biol 12: 367–383

Croteau R, Kutchan TM, Lewis NG (2000) Natural products. In: Biochemistry and Molecular Biology of Plants. pp 1250–1318. Buchana, B. B., Gruissem,W., and Jones, R. L. Eds. American Society of Plant Biologist, Rockville, MD.

Daguma B, Kang BT, Okali DUU (1988) Factors affecting germination of Leucaena leucocephala (Lam.) de Wit. seed. Seed Science and Technology 16: 489-500

Davin LB, Lewis NG (2000) Dirigent proteins and dirigent sites explain the mystery of specificity of radial precursor coupling in lignan and lignin biosynthesis. Plant Physiol 123: 453–461

Davin LB, Lewis NG (2005) Lignin primary structures and dirigent sites. Curr Opin Biotechnol 16: 407–415

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

De Melis LE, Whiteman PH, Stevenson TW (1999) Isolation and characterization of a cDNA clone encoding cinnamyl alcohol dehydrogenase in *Eucalyptus globulus* Labill. Plant Sci 143: 173-182

Dean JFD, LaFayette PR, Rugh C, Tristram AH, Hoopes JT (1998) Laccases associated with lignifying vascular tissues. In Lignin and Lignan Biosynthesis (ACS Symp. Series), ed. Lewis NG, Sarkanen S, 697: 96–108. Washington, DC: Am. Chem. Soc. pp 436

Delmer DP, Amor Y (1995) Cellulose biosynthesis. Plant Cell 7: 987–1000

Dharmawardhana DP, Ellis BE, Carlson JE (1995) A β-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. Plant Physiol 107: 331–39

Dixon RA, Chen F, Guo D, Parvathi K (2001) The biosynthesis of monolignols: a "metabolic grid," or independent pathways to guaiacyl and syringyl units? Phytochemistry 57: 1069–1084

Donaldson LA (2001) Lignification and lignin topochemistry—an ultrastructural view. Phytochemistry 57: 859–873

Driouich A, Lain´e A-C, Vian B, Faye L (1992) Characterization and localization of laccase forms in stem and cell cultures of sycomore. Plant J 2: 13–24

Dwivedi UN, Campbell WH, Yu J, Datla RSS, Bugos RC, Chiang VL, Podila GK (1994) Modification of lignin biosynthesis in transgenic *Nicotiana* through expression of an antisense O-methyltransferase gene from *Populus*. Plant Mol Biol 26: 61–71

El Mansouri I, Mercado JA, Santiago-Domenech N, Pliego-Alfaro F, Valpuesta V, Quesada MA (1999) Biochemical and phenotypical characterization of transgenic tomato plants overexpressing a basic peroxidase. Physiol Plant 106: 355–362

El-Basyouni SZ, Neish AC (1966) Occurrence of metabolicallyactive bound forms of cinnamic acid and its phenolic derivatives in acetone powders of wheat and barley plants. Phytochemistry 5: 683–691

El-Basyouni SZ, Neish AC, Tower GHN (1964) The phenolic acids in wheat.III.Insoluble derivatives of phenolic cinnamic acids as natural intermediates in lignin biosynthesis. Phytochemistry 3: 627–639

Elfstrand M, Sitbon F, Lapierre C, Bottin A, von Arnold S (2002) Altered lignin structure and resistance to pathogens in spi 2-expressing tobacco plants. Planta 214: 708–716

Elkind Y, Edwards R, Mavandad M, Hedrick SA, Ribak O, Dixon RA, Lamb CJ (1990) Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene. Proc Natl Acad Sci 87: 9057–9061

Fengel D, Wegener, G (1984) Lignin. In Wood: Chemistry, Ultrastructure, Reactions pp 132–181

Fergus BJ, Goring DAI (1970) The location of guaiacyl and syringyl lignins in birch xylem tissue. Holzforschung 24: 113–17

Ferrer JL, Zubieta C, Dixon RA, Noel JP (2005) Crystal Structures of Alfalfa Caffeoyl Coenzyme A 3-O-Methyltransferase. Plant Physiol 137: 1009-1017

Ferrer MA, Pedre˜no MA, Mu˜noz R, Ros Barcel´o A (1990) Oxidation of coniferyl alcohol by cell wall peroxidases at the expense of indole-3-acetic acid and O2. A model for lignification of plant cell walls in the absence of H_2O_2. FEBS Lett 276: 127–130

Food and Agriculture Organization of the United Nations (2001). Pulp and Paper Capacities—Survey 2000– 2005. FAO, Rome.

Franke R, Hemm MR, Denault JW, Ruegger MO, Humphreys JM, Chapple C (2002b) Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of *Arabidopsis*. Plant J 30: 47–59

Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C (2002a) The *Arabidopsis* REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. Plant J 30: 33–45

Franke R, McMichael CM, Meyer K, Shirley AM, Cusumano JC, Chapple C (2000) Modified lignin in tobacco and poplar plants overexpressing the *Arabidopsis* gene encoding ferulate 5-hydroxylase. Plant J 22: 223–234

Freudenberg K, Neish AC, ed. (1968) Constitution and Biosynthesis of Lignin. Berlin: Springer-Verlag. pp 129

Gierer J (1985) Chemistry of delignification. Part 1: General concept and reaction during pulping. Wood Sci Technol 19: 289–312

Goffner D, Campbell MM, Campargue C, Clastre M, Borderies G, Boudet A, Boudet AM (1994) Purification and characterization of cinnamoyl-CoA:NADP oxidoreductase in *Eucalyptus gunnii*. Plant Physiol 106: 625–632

Goujon T, Ferret V, Mila I, Pollet B, Ruel K, Burlat V, Joseleau J-P, Barri`ere Y, Lapierre C, Jouanin L (2003) Down-regulation of the AtCCR1 gene in *Arabidopsis thaliana*: effects on phenotype, lignins and cell wall degradability. Planta 217: 218–228.

Grand C, Boudet A, Boudet AM (1983) Isoenzymes of hydroxycinnamate:CoA ligase from poplar stems: properties and tissue distribution. Planta 158: 225–229

Grand C, Parmentier P, Boudet A, Boudet AM (1985) Comparison of lignins and of enzymes involved in lignification in normal and brown midrib (bm3) mutant corn seedlings. Physiol V´eg 23: 905– 911.

Grima-Pettenati J, Goffner D (1999) Lignin genetic engineering revisited. Plant Sci 145: 51–65.

Grimmig B, Matern U (1997) Structure of the parsley caffeoyl-CoA O-methyltransferase gene, harbouring a novel elicitor responsive cis-acting element. Plant Mol Biol 33: 323-341

Guo D, Chen F, Inoue K, Blout JW, Dixon RA (2001) Downregulation of caffeic acid 3-O-methyltransferase and caffeoyl CoA 3-O-methyltransferase in transgenic alfalfa: impacts on lignin structure and implications for the biosynthesis of G and S lignin. Plant Cell 13: 73–88

Gupta H.K, Atreja PP (1999) Influence of feeding increasing levels of *Leucaena* leaf meal on the performance of milch goats and metabolism of mimosine and 3, 4- DHP. Anim Feed Sci and Technol 78: 159

Halpin C, Holt K, Chojecki J, Oliver D, Chabbert B, Monties B, Edwards K, Barakate A, Foxon GA (1998) Brown-midrib maize (bm1)—a mutation affecting the cinnamyl alcohol dehydrogenase gene. Plant J 14: 545–553

Halpin C, Knight ME, Foxon GA, Campbell MM, Boudet AM, Boon JJ, Chabbert B, Tollier M-T, and Schuch W (1994) Manipulation of lignin quality by downregulation of cinnamyl alcohol dehydrogenase. Plant J 6: 339–350

Hamberger B, Hahlbrock K (2004) The 4-coumarate:CoA ligase gene family in Arabidopsis thaliana comprises one rare, sinapate-activating and three commonly occurring isoenzymes. Proc Natl Acad Sci 101: 2209– 2214

Hammond AC, Allison MJ, Williams MJ (1989b) Persistence of DHP-degrading bacteria between growing seasons in subtropical Florida. Leucaena Research Report. 10: 66

Hammond AC, Allison MJ, Williams MJ, Prine, GM Bates DB (1989a) Prevention of Leucaena toxicosis of cattle in Florida by ruminal inoculation with 3, 4-DHP-degrading bacteria. Am J Vet Res 50: 2176

Harding SA, Leshkevich J, Chiang VL, Tsai C-J (2002) Differential substrate inhibition couples kinetically distinct 4-coumarate:coenzyme A ligases with spatially distinct metabolic roles in quaking aspen. Plant Physiol 128: 428-438

Hatfield R, Vermerris W (2001) Lignin formation in plants. The dilemma of linkage specificity. Plant Physiol 126: 1351–1357

Hatton D, Sablowski R, Yung M-H, Smith C, Schuch W, Bevan M (1995) Two classes of cis sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. Plant J 7: 859–876

He L, Terashima N (1991) Formation and structure of lignin in monocotyledons. IV. Deposition process and structural diversity f the lignin in the cell wall of sugarcane and rice plant studied by ultraviolet icroscopic spectroscopy. Holzforschung 45:191–98

Hegarty MP, Court RD, Thorne PM (1964) The determination of mimosine and 3, 4- DHP in biological material. Aust J Agric Res 15: 168

Hibino T, Takabe K, Kawazu T, Shibata D, Higuchi T (1995) Increase of cinnamaldehyde groups in lignin of transgenic tobacco plants carrying an antisense gene for cin9namyl alcohol dehydrogenase. Biosci Biotechnol Biochem 59: 929–931

Highley TL, Dashek WV (1998) Biotechnology in the study of brown-and white-rot decay. In Forest Products Biotechnology. pp 15–36 Bruce AM, Palfreyman JW, Eds., Taylor & Francis, London

Higuchi T (1997) Biochemistry and Molecular Biology of Wood (Springer Series in Wood Science). Springer, Berlin.

Higuchi, T (1990) Lignin biochemistry: biosynthesis and biodegradation. Wood Sci Technol 24: 23–63

Hosokawa M, Suzuki S, Umezawa T, Sato Y (2001) Progress of lignification mediated by intercellular transportation of monolignols during tracheary element differentiation of isolated Zinnia mesophyll cells. Plant Cell Physiol 42: 959– 68

Howles PA, Sewalt VJH, Paiva NL, Elkind Y, Bate NJ, Lamb C, Dixon R (1996) Over expression of L-phenylalanine ammonia-lyase in transgenic tobacco plants reveals control points for flux into phenylpropanoid biosynthesis. Plant Physiol 112: 1617–1624

Hu W-J, Harding SA, Lung J, Popko JL., Ralph J, Stokke DD, Tsai C-J, Chiang, VL (1999) Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. Nat Biotechnol 17: 808–812

Hughes CE (1998) *Leucaena*: A genetic resources handbook. Tropical Forestry Papers No. 37. Oxford Forestry Institute, Department of Plant Science. Oxford University, UK

Humphreys JM, Chapple C (2002) Rewriting the lignin roadmap. Curr Opin Plant Biol 5: 224–29

Humphreys JM, Hemm MR, Chapple, C (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci 96: 10045–10050

Ibrahim RK, Bruneau A, Bantignies B (1998) Plant O-methyltransferases: molecular analysis, common signature and classification. Plant Mol Biol 36: 1–10

Inoue K, Parvathi K, Dixon RA (2000) Substrate preferences of caffeic acid/5-hydroxyferulic acid 3-O-methyltransferases in developing stems of alfalfa (*Medicago sativa* L.). Arch Biochem Biophys 375: 175–182

Inoue K, Sewalt VJH, Murray BG, Ni W, Sturzer C, Dixon RA (1998) Developmental expression and substrate specificities of alfalfa caffeic acid 3-O-methyltransferase and caffeoyl coenzyme A 3-O-methyltransferase in relation to lignification. Plant Physiology 117: 761–770

Ishii T (1997) Structure and functions of feruloylated polysaccharides. Plant Sci 127: 111–127

Jones L, Ennos AR, Turner SR (2001) Cloning and characterization of irregular xylem4 (irx4): a severely lignin deficient mutant of Arabidopsis. Plant J 26: 205–216

Jones RJ (1979) The value of *Leucaena leucocephala* as feed for ruminants in the tropics World Animal review 32: 10

Jones RJ (1994) Management of anti-nutritive factors with special reference to Leucaena.In: Gutteridge RC, Shelton HM (Eds). Forage tree legumes in tropical agriculture. CAB International. Wallingford, UK. pp 216

Jones, HD (1984) Phenylalanine ammonia-lyase: Regulation of its induction, and its role in plant development. Phytochemistry 23: 1349–1359

Joseleau J-P, Ruel K (1997) Study of lignification by noninvasive techniques in growing maize internodes. An investigation by Fourier transform infrared cross-polarization-magic angle spinning 13C-nuclear magnetic resonance spectroscopy and immunocytochemical transmission electron microscopy. Plant Physiol 114: 1123–1133

Jouanin L, Goujon T, de Nada V, Martin M-T, Mila I, Vallet C, Pollet B, Yoshinaga A, Chabbert B, Petit-Conil M, Lapierre C (2000) Lignification in transgenic poplars with extremely reduced caffeic acid O-methyltransferase activity. Plant Physiol 123: 1363–1373

Kajita S, Hishiyama S, Tomimura Y, Katayama Y, Omori S (1997) Structural characterization of modified lignin in transgenic tobacco plants in which the activity of 4-coumarate: coenzyme A ligase is depressed. Plant Physiol 114: 871– 879

Kajita S, Katayama Y, Omori S (1996) Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate: coenzyme A ligase. Plant Cell Physiol 37: 957–965

KaoYY, Harding SA, Tsai CJ (2002) Differential expression of two distinct phenylalanine ammonia-lyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen. Plant Physiol 130: 796–807

Keller B, Templeton MD, Lamb CJ (1989) Specific localization of a plant cell wall glycine-rich protein in protoxylem cells of the vascular system. Proc Natl Acad Sci USA 86: 1529–1533

Kersey R, Inoue K, Schubert KR, Dixon RA (1999) Immunolocalization of two lignin O-methyltransferases in stems of alfalfa (*Medicago sativa* L.). Protoplasma 209: 46–57

Kim H, Ralph J, Lu F, Pilate G, Lepl´e JC (2002) Identification of the structure and origin of thioacidolysis marker compounds or cinnamyl alcohol dehydrogenase deficiency in angiosperms. J Biol Chem 277: 47412–47419

Kojima M, Takeuchi W (1989) Detection and characterization of p-coumaric acid hydroxylase in mung bean, *Vigna mungo*, seedlings. J Biochem 105: 265–270

Korth KL, Blount JW, Chen F, Rasmussen S, Lamb C, Dixon RA (2001) Changes in phenylpropanoid metabolites associated with homologydependent silencing of phenylalanine ammonialyase and its somatic reversion in tobacco. Physiol Plant 111: 137–143

Ku˘c J, Nelson OE, Flanagan P (1968) Degradation of abnormal lignins in the brown-midrib mutants and double mutants of maize. Phytochemistry 7: 1435–1436

Kuhnl T, Koch U, Heller W, Wellmann E (1989) Elicitor induced S-adenosyl-l-methionine:caffeoyl-CoA 3-O-methyltransferase from carrot cell suspension cultures. Plant Science 60: 21–25

Kumar A, Ellis BE (2001) The phenylalanine ammonia-lyase gene family in raspberry. Structure, expression, and evolution. Plant Physiol 127: 230– 239

Lacombe E, Hawkins S, Doorsselaere JV, Piquemal J, Goffner D, Poeydomenge O, Boudet, AM (1997) Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: Cloning, expression and phylogenetic relationships. Plant J 11: 429–441

Lagrimini LM (1991) Wound-induced deposition of polyphenols in transgenic plants over expressing peroxidase. Plant Physiol 96: 577–583

Lagrimini LM, Burkhart W, Moyer M, Rothstein S (1987) Molecular cloning of complementary DNA encoding the lignin-forming peroxidase from tobacco: molecular analysis and tissue-specific expression. Proc Natl Acad Sci USA 84: 7542–7546

Lagrimini LM, Gingas V, Finger F, Rothstein S, Liu T-TY (1997) Characterization of antisense transformed plants deficient in the tobacco anionic peroxidase. Plant Physiol 114: 1187–1196

Lagrimini LM, Vaughn J, Erb WA, Miller SA (1993) Peroxidase overproduction in tomato: wound-induced polyphenol deposition and disease resistance. Hort Science 28: 218–221

Lapierre C, Pollet B, MacKay JJ, Sederoff RR (2000) Lignin structure in a mutant pine deficient in cinnamyl alcohol dehydrogenase. J Agric Food Chem 48: 2326–2331

Lapierre C, Pollet B, Petit-Conil M, Toval G, Romero J, Pilate G, Leple J-C, Boerjan W, Ferret V, De Nadai V, Jouanin, L (1999) Structural alterations of lignins in transgenic poplars with depressed cinnamyl alcohol dehydrogenase or caffeic acid O-methyltransferase activity have opposite impact on the efficiency of industrial Kraft pulping. Plant Physiol 119: 153–163

Laura C, Joan R, Pere P (1999) Nucleotide Sequence of two cDNAs coding for Caffeoyl-coenzyme A O-Methyltransferase (CCoAOMT) and study of their expression in *Zea mays*. Plant Physiol 120: 1206-1206

Lee D, Meyer K, Chapple C, Douglas CJ (1997) Antisense suppression of 4- coumarate: coenzyme A ligase activity in *Arabidopsis* leads to altered lignin subunit composition. Plant Cell 9: 985– 998

Lewis NG (1999) A 20th century roller coaster ride: a short account of lignification. Curr Opin Plant Biol 2: 153–162

Lewis NG, Yamamoto E (1990) Lignin: Occurrence, biogenesis, and biodegradation. Annu Rev Plant Physiol Plant Mol Biol 41: 455–496

Leyva A, Liang X, Pintor-Toro JA, Dixon RA, Lamb CJ (1992) cis- Element combinations determine phenylalanine ammonia-lyase gene tissue specific expression patterns. Plant Cell 4: 263–271

Li L, Cheng X, Lu S, Nakatsubo T, Umezawa T, Chiang VL (2005) Clarification of cinnamoyl co-enzyme A reductase catalysis in monolignol biosynthesis of aspen. Plant Cell Physiol 46: 1073–1082

Li L, Cheng XF, Leshkevich J, Umezawa T, Harding SA, Chiang VL (2001) The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. Plant Cell 13: 1567–1585

Li L, Osakabe K, Joshi CP, Chiang VL (1999) Secondary xylem specific expression of caffeoyl-coenzyme A 3-O-methyltransferase plays an important role in the

methylation pathway associated with lignin biosynthesis in loblolly pine. Plant Mol Biol 40: 555–565

Li L, Popko JL, Umezawa T, Chiang VL (2000) 5-Hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms. J Biol Chem 275: 6537–6545

Li L, Popko JL, Zhang X.-H, Osakabe K, Tsai C-J, Joshi C P, Chiang VL (1997) A novel multifunctional O-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine. Proc Natl Acad Sci 94: 5461–5466

Li L, Zhou Y, Cheng X, Sun J, Marita JM, Ralph J, Chiang VL (2003) Combinatorial modification of multiple lignin traits in trees through multigene cotransformation. Proc Natl Acad Sci 100: 4939–4944

Lindermayr C, Fliegmann J, Ebel J (2003) Deletion of a single amino acid residue from different 4-coumarate-CoA ligases from soybean results in the generation of new substrate specificities. J Biol Chem 278: 2781–2786

Lois R, Dietrich A, Hahlbrock K, Schulz W (1989) A phenylalanine ammonia-lyase gene from parsley: Structure, regulation and identification of elicitor and light responsive cis-acting elements. EMBO 8: 1641–1648

Lu F, Ralph J (1999) Detection and determination of p-coumaroylated units in lignins. J Agric Food Chem. 47:1988– 92

Lu F, Ralph J (2002) Preliminary evidence for sinapyl acetate as a lignin monomer in kenaf. Chem Commun 1: 90–91

Luderitz T, Grisebach H (1981) Enzyme synthesis of lignin precursors. Comparison of cinnamoyl-CoA reductase and cinnamyl alcohol:NADP+ dehydrogenase from spruce (*Picea abies* L.) and soybean (*Glycine max* L.). Eur J Biochem 119: 115–124

MacKay JJ, O'Malley DM, Presnell T, Booker FL, Campbell MM, Whetten RW, Sederoff, RR (1997) Inheritance, gene expression, and lignin characterization in a mutant pine deficient in cinnamyl alcohol dehydrogenase. Proc Natl Acad Sci 94: 8255–8260

Mader M, Fussl, R (1982) Role of peroxidase in lignification of tobacco cells. II. Regulation by phenolic compounds. Plant Physiol 70: 1132–1134

Maher EA, Bate NJ, Ni W, Elkind Y, Dixon RA, Lamb CJ (1994) Increased disease susceptibility of transgenic tobacco plants with suppressed levels of preformed phenylpropanoid products. Proc Natl Acad Sci 91: 7802–7806

Marita JM, Ralph J, Hatfield RD, Chapple C (1999) NMR characterization of lignins in Arabidopsis altered in the activity of ferulate 5-hydroxylase. Proc Natl Acad Sci 96: 12328–12332

Marita JM, Ralph J, Hatfield RD, Guo D, Chen F, Dixon RA (2003) Structural and compositional modifications in lignin of transgenic alfalfa down-regulated in caffeic acid 3 O- methyltransferase and caffeoyl coenzyme A 3 O- methyltransferase. Phytochemistry 62: 53–65

Martz F, Maury S, Pincon G, Legrand M (1998) cDNA cloning, substrate specificity and expression study of tobacco caffeoyl-CoA 3-O-methyltransferase, a lignin biosynthetic enzyme. Plant Mol Biol 36: 427–437

Matern U, Wendorff H, Hamerski D, Pakusch AE, Kneusel RE (1988) Elicitor-induced phenylpropanoid synthesis in Apiaceae cell cultures.Bull. Liaison Group Polyphenols 14: 173–184

Maury S, Geoffroy P, Legrand M (1999) Tobacco O-methyltransferases involved in phenylpropanoid metabolism. The different caffeoyl-coenzyme A/5-hydroxyferuloyl-coenzyme A 3/5-O-methyltransferase and caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase classes have distinct substrate specificities and expression patterns. Plant Physiology 121: 215–223

McDougall GJ, Stewart D, Morrison IM (1994) Cell-wall-bound oxidases from tobacco (*Nicotiana tabacum*) xylem participate in lignin formation. Planta 194: 9–14

Mellerowicz EJ, Baucher M, Sundberg B, Boerjan W (2001) Unravelling cell wall formation in the woody dicot stem. Plant Mol Biol 47: 239– 247

Meng H, Campbell WH (1998) Substrate profiles and expression of caffeoyl coenzyme A and caffeic acid O-methyltransferases in secondary xylem of aspen during seasonal development. Plant Mol Biol 38: 513–520

Meyer K, Cusumano JC, Somerville C, Chapple CCS (1996) Ferulate-5-hydroxylase from Arabidopsis thaliana defines a new family of cytochrome P450-dependent monooxygenases. Proc Natl Acad Sci 93: 6869–6874

Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. Proc Natl Acad Sci 95: 6619– 6623

Meyermans H, Morreel K, Lapierre C, Pollet B, De Bruyn A, Busson R, Herdewijn P, Devreese B, Van Beeumen J, Marita JM, Ralph J, Chen C, Burggraeve B, Van Montagu M, Messens E, Boerjan W (2000) Modification in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A O-methyltransferase, an enzyme involved in lignin biosynthesis. J Biol Chem 275: 36899–36909

Moller SG, McPherson MJ (1998) Developmental expression and biochemical analysis of the Arabidopsis atao1 gene encoding an H_2O_2-generating diamine oxidase. Plant J 13: 781–791

Monties B (1998) Novel structures and properties of lignins in relation to their natural and induced variability in ecotypes, mutants and transgenic plants. Polymer Degrad Stabil 59: 53–64

Nair RB, Xia Q, Kartha CJ, Kurylo E, Hirji RN, Datla R, Selvaraj G (2002) *Arabidopsis* CYP98A3 mediating aromatic 3-hydroxylation. Developmental regulation of the gene, and expression in yeast. Plant Physiol 130: 210–20

NAS [National Academy of Sciences] (1984) *Leucaena*: promising forage and tree crop for the tropics. 2nd ed. Washington, DC: National Academy of Sciences. pp 100

Neish AC (1968) Monomeric intermediates in the biosynthesis of lignin. In: Freudenberg, K., Neish AC (Eds.), Constitution and Biosynthesis of Lignin.Springer-Verlag, New York pp1–43

Ni W, Paiva NL, Dixon RA (1994) Reduced lignin in transgenic plants containing a caffeic acid Omethyltransferase antisense gene. Transgenic Res 3: 120–126

Nimz HH, Tutschek R (1977) Holzforschung 31: 101–106

Norton BW, Poppi DP (1995) Composition and nutritional attributes of pasture legumes. In: D'Mello JPF, Devendra C (Eds). Tropical legumes in animal nutrition. CAB International. Wallingford. Oxon, UK. pp 23

NRC (National Research Council) (1984) Nutritional requirements of goats. National Academy Press. Washington DC.

O'Connell A, Holt K, Piquemal J, Grima-Pettenati J, Boudet A, Pollet B, Lapierre C, Petit-Conil M, Schuch W, Halpin C (2002) Improved paper pulp from plants with suppressed cinnamoyl- CoA reductase or cinnamyl alcohol dehydrogenase. Transgenic Res 11: 495–503

Ogawa K, Kanematsu S, Asada K (1997) Generation of superoxide anion and localization of CuZn-superoxide dismutase in the vascular tissue of spinach hypocotyls: their association with lignification. Plant Cell Physiol 38: 1118– 1126

Ohl S, Hedrick SA, Chory J, Lamb CJ (1990) Functional properties of a phenylalanine ammonia-lyase promoter from *Arabidopsis*. Plant Cell 2: 837–848

Onnerud H, Zhang L, Gellerstedt G, Henriksson G (2002) Polymerization of monolignols by redox shuttle-mediated enzymatic oxidation: a new model in lignin biosynthesis. I Plant Cell 14: 1953–62

Onysko KA (1993) Biological bleaching of chemical pulps: a review. Biotechnol Adv 11:179–198

Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, Carraway DT, Smeltzer RH, Joshi CP, Chiang VL (1999) Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. Proc Natl Acad Sci 96: 8955–8960

Osakabe Y, Ohtsubo Y, Kawai S, Katayama Y, Morohoshi N (1995) Structures and tissue-specific expression of genes for phenylalanine ammonialyase from a hybrid aspen. Plant Sci 105: 217–226

Ostergaard L, Teilum K, Mirza O, Mattsson O, Petersen M, Welinder KG, Mundy J, Gajhede M, Henriksen A (2000) *Arabidopsis* ATP A2 peroxidase. Expression and high-resolution structure of a plant peroxidase with implications for lignification. Plant Mol Biol 44: 231–243

Pakusch A-E, Kneusel RE, Matern U (1989) S-Adenosyl-lmethionine: trans-caffeoyl-coenzyme A 3-O-methyltransferase from elicitor-treated parsley cell suspension cultures. Arch Biochem Biophy 271: 488–494

Pakusch A-E, Matern U, Schiltz E (1991) Elicitor-inducible caffeoyl-coenzyme A 3-Omethyltransferase from Petroselinum crispum cell suspensions. Purification, partial sequence, and antigenicity. Plant Physiol 95: 137–143

Parrotta JA (1992) *Leucaena leucocephala* (Lam.) de Wit: leucaena, tantan. Res. Note SO-ITFSM-52. New Orleans: USDA Forest Service, Southern Forest Experiment Station. pp 8

Parham RA (1987) Structure, chemistry and physical properties of woody raw materials. In *Pulp and Paper Manufacture, Vol. 1: Properties of Fibrous Raw Materials and Their Preparation for Pulping*, 3rd ed. M.J. Kocurek, and F. Stevens, Eds., Tappi and Canadian Pulp and Paper Association Joint Textbook Committee of the Paper Industry, Montreal.

Parvathi K, Chen F, Guo D, Blount JW, Dixon RA (2001) Substrate preferences of Omethyltransferases in alfalfa suggest new pathways for 3-O-methylation of monolignols. Plant J 25:193–202

Petersen M, Strack D, Matern U (1999) Biosynthesis of phenylpropanoids and related compounds. In: Biochemistry of Plant Secondary Metabolism. Wink, M., Ed., Vol. 2, pp 151–222, Sheffield Academic Press, Sheffield, UK.

Petit-Conil M (1995) Principes de preparation de pates chimicothermomecaniques de bois resineux et feuillus. Ph. D. thesis, Institut National Polytechnique, Grenoble, France.

Pickett-Heaps JD (1968) Xylem wall deposition. Radioautographic investigations using lignin precursors. Protoplasma 65: 181–205

Pilate G, Guiney E, Holt K, Petit-Conil M, Lapierre C, Leple J-C, Pollet B, Mila I, Webster EA, Marstorp HG, Hopkins DW, Jouanin L, Boerjan W, Schuch W, Cornu D, Halpin C (2002) Field and pulping performances of transgenic trees with altered lignification. Nature Biotechnol 20: 607–612

Pincon G, Chabannes M, Lapierre C, Pollet B, Ruel K, Joseleau J-P, Boudet AM, Legrand M (2001b) Simultaneous down-regulation of caffeic/5- hydroxy ferulic acid-O-methyltransferase I and cinnamoyl-coenzyme A reductase in the progeny from a cross between tobacco lines homozygous for each transgene. Consequences for plant development and lignin synthesis. Plant Physiol 126: 145–155

Pincon G, Maury S, Hoffmann L, Geoffroy P, Lapierre C, Pollet B, Legrand M (2001a) Repression of O-methyltransferase genes in transgenic tobacco affects lignin synthesis and plant growth. Phytochemistry 57: 1167–1176

Piquemal J, Lapierre C, Myton K, O'Connell A, Schuch W, Grima-Pettenati J, Boudet A- M (1998) Down-regulation in cinnamoyl-CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. Plant J 13: 71–83

Pottinger AJ, Hughes CE (1995) A review of wood quality in *Leucaena*. In H.M. Shelton, C.M. Piggin & J.L. Brewbaker (Eds.) Leucaena opportunities and limitations. Proceedings of a workshop held in Bogor, Indonesia, January 1994. ACIAR Proceedings Canberra. 57: 98

Raes J, Rohde A, Christensen JH, Peer YV, Boerjan W (2003) Genome-wide characterization of the lignification toolbox in arabidopsis. Plant Physiol 133: 1051–1071

Ralph J, Grabber JH, Hatfield RD (1995) Lignin-ferulate crosslinks in grasses: active incorporation of ferulate polysaccharide esters into ryegrass lignins. Carbohydr Res 275: 167–78

Ralph J, Hatfield RD, Piquemal J, Yahiaoui N, Pean M, Lapierre C, Boudet AM (1998) NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignification enzymes cinnamyl alcohol dehydrogenase and cinnamoyl CoA reductase. Proc Natl Acad Sci 95: 12803–12808

Ralph J, Kim H, Peng J, Lu F (1999b) Arylpropane-1, 3-diols in lignins from normal and CAD-deficient pines. Org Lett 1: 323–26

Ralph J, Lapierre C, Lu F, Marita JM, Pilate G (2001) NMR evidence for benzodioxane structures resulting from incorporation of 5- hydroxyconiferyl alcohol into lignins of O-methyltransferase-deficient plants. J Agric Food Chem 49: 86–91

Ralph J, Lapierre C, Marita JM, Kim H, Lu F (2001) Elucidation of new structures in Lignins of CAD- and COMT-deficient plants by NMR. Phytochemistry 57: 993–1003

Ralph J, Lapierre C, Marita JM, Kim H, Lu F, Hatfield RD, Ralph S, Chapple C, Franke R, Hemm MR, Van Doorsselaere J, Sederoff RR, O'Malley DM, Scott JT, Mackay JJ, Yahiaoui N, Boudet A-M, Pean M, Pilate G, Jouanin L, Boerjan W (2001a) Elucidation of new structures in lignins of CAD-and COMT-deficient plants by NMR. Phytochemistry 57: 993–1003

Ralph J, Lundquist K, Brunow G, Lu F, Kim H, Schatz PF, Marita JM, Hatfield RD, Christensen JH, Boerjan W (2004) Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochemistry Rev 3: 29–60

Ralph J, MacKay JJ, Hatfield RD, O'Malley DM, Whetten RW, Sederoff RR (1997) Abnormal lignin in a loblolly pine mutant. Science 277: 235–39

Ralph J, Marita JM, Ralph SA, Hatfield RD, Lu F, Ede RM, Peng J, Quideau S, Helm RF, Grabber JH, Kim H, Jimenez-Monteon G, Zhang Y, Jung H-JG, Landucci LL, MacKay JJ, Sederoff RR, Chapple C, Boudet AM (1999a) Solution-state NMR of lignins. In Advances in Lignocellulosic Characterization. pp 55–108 Argyropoulos DS, Ed.,TAPPI Press, Atlanta, GA.

Ralph J, Peng J, Lu F, Hatfield RD (1999) Are lignins optically active? J Agric Food Chem 47: 2991–2996

Ralph, J (1996) An unusual lignin from kenaf. J Nat Prod 59: 341–342

Ranocha P, Chabannes M, Chamayou S, Danoun S, Jauneau A, Boudet A-M, Goffner D (2002) Laccase down-regulation causes alterations in phenolic metabolism and cell wall structure in poplar. Plant Physiol 129: 145–155

Ranocha P, Goffner D, Boudet AM (2000) Plant laccases: are they involved in lignification. In Cell and Molecular Biology of Wood Formation. pp 397–410 Savidge R, Barnett J, Napier R, Eds., BIOS Scientific Publications, Oxford.

Ranocha P, McDougall G, Hawkins S, Sterjiades R, Borderies G (1999) Biochemical characterization, molecular cloning and expression of laccases—a divergent gene family—in poplar. Eur J Biochem 259: 485–495

Reid ID, Paice MG (1994) Biological bleaching of kraft pulps by white-rot fungi and their enzymes. FEMS Microbiol Rev 13: 369–376

Ros-Barcel´o A, Pomar F, Ferrer MA, Mart´ınez P, Ballesta MC, Pedre˜no MA (2002) In situ characterization of a NO sensitive peroxidase in the lignifying xylem of *Zinnia elegans*. Physiol Plant 114: 33–40

Ros-Barcel´o A, Pomar F, L´opez-Serrano M, Mart´ınez P, Pedre˜no MA (2002) Developmental regulation of the H_2O_2- producing system and of a basic peroxidase

isoenzyme in the *Zinnia elegans* lignifying xylem. Plant Physiol Biochem 40: 325–332

Rossmann MG, Moras D, Olsen KW (1974) Chemical and biological evolution of nucleotide-binding protein. Nature 250: 194–199

Rouhi AM (2001) Only facts will end the lignin war. Chem Eng News 79: 52–56

Saka S, Goring DAI (1985) Localization of lignins in wood cell walls. In Biosynthesis and Biodegradation ofWood Components, ed. Higuchi T, pp 51–62. Orlando: Academic. pp 679

Salminen M, Lundstrom K, Tilgmann C, Savolainen R, Kalkkinen N, Ulmanen I, (1990) Molecular cloning and characterization of rat liver catechol-O-methyltransferase. Gene 93: 241–247

Samuels AL, Rensing K, Douglas CJ, Mansfield S, Dharmawardhana P, Ellis B (2002) Cellular machinery of wood production: differentiation of secondary xylem in *Pinus contorta* var. latifolia. Planta 216:72–82

Sarkanen KV, Ludwig CH (1971) Lignins: Occurrence, Formation, Structure, and Reactions. New York: Wiley- Intersci pp 916

Sarni F, Grand G, Boudet AM (1984) Purification and properties of cinnamoyl-CoA reductase and cinnamyl alcohol dehydrogenase from poplar stems. Eur J Biochem 139: 259–265

Schmitt D, Pakusch A-E, Matern U (1991) Molecular cloning, induction, and taxonomic distribution of caffeoyl-CoA 3-O-methyltransferase, an enzyme involved in disease resistance. J Biol Chem 266: 17416–17423

Schoch G, Goepfert S, Morant M, Hehn A, Meyer D, Ullmann P, Werck-Reichhart D (2001) CYP98A3 from *Arabidopsis thaliana* is a 3 - hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. J Biol Chem 276: 36566–36574

Sederoff RR, MacKay JJ, Ralph J, Hatfield RD (1999) Unexpected variation in lignin. Curr Opin Plant Biol 2: 145–152

Sewalt VJH, Ni W, Blount JW, Jung HG, Masoud SA, Howles PA, Lamb C, Dixon RA (1997) Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of L-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase. Plant Physiol 115: 41– 50

Shelton M (1995) *Leucaena* forage production and quality in South East Queensland. *Leucaena* News. 2: 6 SO-ITFSM- 52. New Orleans: USDA Forest Service, Southern Forest Experiment Station. pp 8

Sibout R, Eudes A, Pollet B, Goujon T, Mila I, Granier F, Seguin A, Lapierre C, Jouanin L (2003) Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenase in *Arabidopsis*. Isolation and characterization of the corresponding mutants. Plant Physiol 132: 848–860

Stafford HA, Dresler S (1972) 4-Hydroxycinnamic acid hydroxylase and polyphenolase activities in *Sorghum vulgare*. Plant Physiol 49: 590–595

Steeves C, F¨orster H, Pommer U, Savidge R (2001) Coniferyl alcohol metabolism in conifers. I. Glucosidic turnover of cinnamyl aldehydes by UDPG: coniferyl alcohol glucosyltransferase from pine cambium. Phytochemistry 57:1085–93

Sterjiades R, Dean JFD, Eriksson KEL (1992) Laccase from sycamore maple (*Acer pseudoplatanus*) polymerizes monolignols. Plant Physiol 99:1162–68

Stewart D, Yahiaoui N, McDougall GJ, Myton K, Marque C, Boudet AM, Haigh J (1997) Fourier-transform infrared and Raman spectroscopic evidence for the incorporation of cinnamaldehydes into the lignin of transgenic tobacco (*Nicotiana tabacum* L.) plants with reduced expression of cinnamyl alcohol dehydrogenase. Planta 201: 311–318

Takahama U, Oniki T (1994) Effects of ascorbate on the oxidation of derivatives of hydroxycinnamic acid and the mechanism of oxidation of sinapic acid by cell wall-bound peroxidases. Plant Cell Physiol 35: 593–600

Tangendjaja B, Willis RBH (1980) Analysis of mimosine and 3, 4-DHP by High Performance Liquid Chromatography. J Chromatog 202: 317

Terashima N, Fukushima K, He LF, Takabe K (1993) Comprehensive model of the lignified plant cell wall. In Forage Cell Wall Structure and Digestibility, ed. Jung HG, Buxton DR, Hatfield RD, Ralph D pp 247–70. Madison: ASACSSA-SSSA pp 794

Timell TE (1986) CompressionWood in Gymnosperms, Vol. 1: Bibliography, Historical Background, Determination, Structure, Chemistry, Topochemistry, Physical

Properties, Origin, and Formation of CompressionWood. Berlin: Springer-Verlag. pp 706

Tognolli M, Penel C, Greppin H, Simon P (2002) Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene 288: 129–138

Tsai C-J, Popko JL, Mielke MR, Hu W-J, Podila GK, Chiang VL (1998) Suppression of O-methyltransferase gene by homologous sense transgene in quaking aspen causes red-brownwood phenotypes. Plant Physiol 117: 101–112

Tsuji Y, Fukushima K (2004) Behavior of monolignol glucosides in angiosperms. J Agric Food Chem 52: 7651–7659

Vallette P, de Choudens C (1992) Le bois, la pate, le papier, 3rd Ed. Centre Technique de l'Industrie des Papiers, Cartons et Celluloses, Grenoble, France.

Van Doorsselaere J, Baucher M, Chognot E, Chabbert B, Tollier M-T, Petit-Conil M, Leple J-C, Pilate G, Cornu D, Monties B, Van Montagu M, Inze D, Boerjan W, Jouanin, L (1995) A novel lignin in poplar trees with a reduced caffeic acid/5-hydroxyferulic acid O methyltransferase activity. Plant J 8: 855–864

Vidgren J, Svensson LA, Liljas A (1994) Crystal structure of catechol O-methyltransferase. Nature 368: 354–358

Viikari L, Buchert J, Suurnakki A (1998) Enzymes in pulp bleaching. In Forest Products Biotechnology. pp. 83–97. Bruce AM, and Palfreyman JW, Eds., Taylor & Francis, London.

Viikari L, Kantelinen A, Sundquist J, Linko M (1994) Xylanases in bleaching: from an idea to the industry. FEMS Microbiol Rev 13: 335–350

Voo KS, Whetten RW, O'Malley DM, Sederoff RR (1995) 4- coumarate:coenzyme a ligase from loblolly pine xylem. Isolation, characterization, and complementary DNA cloning. Plant Physiol 108: 85–97

Wengenmayer H, Ebel J, Grisebach H (1976) Enzymic synthesis of lignin precursors Purification and properties of a cinnamoyl-CoA: NADPH reductase from cell suspension cultures of soybean (*Glycine max*). Eur J Biochem 65: 529–536

Whetten R, Sederoff R (1995) Lignin biosynthesis. Plant Cell 7: 1001–1013

Whetten RW, MacKay JJ, Sederoff RR (1998) Recent advances in understanding lignin biosynthesis. Annu Rev Plant Physiol Plant Mol Biol 49: 585–609

Wu J, Fukazawa K, Ohtani J (1992) Distribution of syringyl and guaiacyl lignins in hardwoods in relation to habitat and porosity. Holzforschung 46: 181-185

Yahiaoui N, Marque C, Myton KE, Negrel J, Boudet AM (1998) Impact of different levels of cinnamyl alcohol dehydrogenase down-regulation on lignins of transgenic tobacco plants. Planta 204: 8–15

Yahong L, Tsuji Y, Nishikubo N, Kajita S, Morohoshi N (2001) Analysis of transgenic poplar in which the expression of peroxidase gene is suppressed. In Molecular Breeding of Woody Plants, (Progress in Biotechnology Series,Vol. 18). pp. 195–204 Morohoshi N, Komamine A, Eds., Elsevier Science, Amsterdam.

Ye Z-H (1997) Association of caffeoyl coenzyme A 3-O-methyltransferase expression with lignifying tissues in several dicot plants. Plant Physiol 115: 1341–1350

Ye Z-H, Kneusel RE, Matern U, Varner JE (1994) An alternative methylation pathway in lignin biosynthesis in *Zinnia*. Plant Cell 6: 1427–1439

Ye Z-H, Varner JE (1995) Differential expression of two O-methyltransferases in lignin biosynthesis in *Zinnia elegans*. Plant Physiol 108: 459–467

Zhang XH, Chiang VL (1997) Molecular cloning of 4- coumarate:coenzyme A ligase in loblolly pine and the roles of this enzyme in the biosynthesis of lignin in compression wood. Plant Physiol 113: 65– 74

Zhong R, Morrison III WH, Himmelsbach DS, Poole II FL, Ye Z-H (2000) Essential role of caffeoyl coenzyme A O-methyltransferase in lignin biosynthesis in woody poplar plants. Plant Physiol 124: 563–577

Zhong R, Morrison WH III, Negrel J, Ye Z-H (1998) Dual methylation pathways in lignin biosynthesis. Plant Cell 10: 2033–2046

Zimmermann HH, Nimz, E. Seemuller (1985) Holzforschung 39: 45–49

Zubieta C, He X-Z, Dixon RA, Noel JP (2001) Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant O-methyltransferases. Nat Struct Biol 8: 271–279

Zubieta C, Kota P, Ferrer J-L, Dixon RA, Noel JP (2002) Structural basis for the modulation of lignin monomer methylation by caffeic acid/5-hydroxyferulic acid 3/5-O methyltransferase. Plant Cell 14: 1265–1277

2.1 Materials

Ampicillin, Kanamycin, Tetracyclin, Hygromycin, Rifampicin, Acetosyringone, Tris, IPTG, X-gal, SDS, BSA, EDTA and Ethidium bromide were purchased from Sigma-Aldrich, USA. Agarose, restriction enzymes, T4 DNA ligase, RNase A and lysozyme were obtained from GIBCO-BRL (USA), Promega (USA) and Amersham (UK). Taq DNA polymerase was obtained from Bangalore Genei (India). Plasmid vectors pBSKS+, pGEM-T Easy Vector, pGEM 3Z, pET30b(+) were purchased from Stratagene (USA), Promega (USA) and Novagen (USA) respectively. Megaprime labeling kit and Hybond-N^+ membrane were obtained from Amersham (UK). α-^{32}P-dATP and α-^{32}P –dCTP were obtained from Bhabha Atomic Research Centre (BARC), India. X-ray films were obtained from Konica (Japan) or Kodak (USA). Affi-gel 10 and Affi-gel 15 matrices was purchased from BioRad (USA). All other chemicals and solvents of analytical grade were purchased from HIMEDIA, Qualigens Fine Chemicals and E-Merck Laboratories, India. Pipette tips and micro centrifuge tubes were purchased from Axygen (USA).

Table 2.1 Bacterial strains and plasmids used in the study

Strain or Plasmid	**Important features (reference or source)**
Escherichia coli XL-1 Blue	recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ proAB lacIqZΔM15 Tn10 (Tet^r)] (Stratagene, USA)
E. coli BL 21(DE3)	F^- ompT gal dcm lon $hsdS_B(r_B^- m_B^-)$ λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) (Invitogen, USA)
A. tumefaciens GV2260	C58, Rif^r, pGV2260 (pTiB6S3ΔT-DNA), $Carb^r$, Octopine type
pGEM-T Easy Vector	Cloning vector (Promega, USA)
pCAMBIA 1300 MCS	Binary cloning plant transformation vector (CAMBIA)
pET30b(+)	Expression vector, Novagen (USA)

2.2 Bacterial culture conditions

E. coli was grown at 37°C with shaking at 200 rpm in Luria Bertani (LB) broth and maintained on LB plates with 1.5% agar (Sambrook et al., 1989). For plasmid DNA preparation recombinant *E. coli* was grown in LB media supplemented with appropriate antibiotic.

Luria Bertani Broth (LB)	**(g L^{-1})**
Bactotryptone	10
Yeast extract	5
NaCl	10
pH adjusted to 7.0	

YEP medium	**(g L^{-1})**
Peptone	10
Sucrose	1
Yeast extract	5
$MgSO_4.7H_2O$	0.5
pH adjusted to 7.0-7.2	

Terrific Broth (TB)	**(g L^{-1})**
YTG base 900mL	
Bactotryptone	12
Bacto Yeast extract	24
Glycerol	4mL
Salt 100mL	
KH_2PO_4 monobasic	2.31
K_2HPO_4 dibasic	12.54

Agar-Agar was added at 1.5% (w/v) wherever required.

2.3 Bacterial cells transformation

2.3.1 *E. coli* transformation and selection

LB medium (50mL) was inoculated with 1% of the overnight grown *E. coli* culture and allowed to grow till 0.5 O.D. at 600nm. The cells were harvested by centrifugation at 5,000 x g for 10 min at 4°C, suspended in 100mM ice-cold $CaCl_2$ and kept on ice for 30 min. Cells were centrifuged, the pellet suspended in 1mL of 100mM ice-cold $CaCl_2$ and stored as aliquots of 200µL at 4°C. The competent *E. coli* cells, thus formed, were transformed according to Sambrook et al. (1989). Briefly, DNA (~50ng in 10µL or less) was added to the competent *E. coli* cells, mixed and kept on ice for 30 min. The cells were then incubated at 42°C for 2 min. To each tube 800µL of LB broth was added and further incubated at 37°C for 1 h. About 100µL of the transformed competent cells was spread onto LB plates containing appropriate antibiotic, IPTG and X-gal as per need (Sambrook et al., 1989).

Solutions	*Stock*	*Final conc.*
1) IPTG stock solution	200mg mL^{-1} in sterile distilled water	40µg mL^{-1}
2) X-gal stock solution	20mg mL^{-1} in dimethylformamide	40µg mL^{-1}

2.3.2 *A. tumefaciens* transformation and selection

The method adopted for the preparation of *A. tumefaciens* competent cells was essentially as reported by An (1987). A single colony of *A. tumefaciens* (GV2260) was inoculated in 50mL LB broth containing Rifampicin 150µg mL^{-1} and incubated at 28ºC with shaking at 200 rpm till O.D. 0.5 at 600nm. Cells were centrifuged at 5000 x g for 10 min at 4ºC and washed twice with ice cold 150mM $CaCl_2$. The cells were pelleted and resuspended in 1mL of ice cold 20mM $CaCl_2$. Aliquots of 200µL were made.

For transformation 1µg DNA of the desired binary plasmid vector was added to an aliquot of the competent *A. tumefaciens* cells and incubated on ice for 30 min. The cells were then snap frozen in liquid nitrogen and allowed to thaw at 37ºC. After thawing 1mL LB medium was added and the tubes incubated at 28ºC for 2 h with gentle shaking. The cells were centrifuged at 4,000 x g for 5 min and 100µL supernatant was retained. The cells were resuspended in the 100µL supernatant and plated on LB agar medium with appropriate antibiotic(s). The plates were incubated at 28ºC for two days to allow the transformed colonies to grow for further processing.

2.4 The Nucleic Acids

2.4.1 Isolation of plasmid DNA from *E. coli* cells

Solutions

TEG Buffer (Soln. I): 25mM Tris-HCl (pH 8.0), 10mM EDTA (pH 8.0), 50mM Glucose.

Soln. II: 0.2N NaOH, 1% SDS (freshly prepared)

Soln. III: 3.0M Potassium acetate (pH 4.8)

Chloroform, absolute ethanol, 3.0M Sodium acetate, 70% ethanol, deionized water

The alkaline lysis method of Sambrook et al. (1989) was improvised upon so that 12-24 samples could be processed conveniently for plasmid DNA extraction within 3 h, with yields of 5-30µg per 1.5mL culture depending on the host strain and the plasmid vector. An important feature of this protocol was the use of PEG for purification, which resulted in precipitation of high quality super-coiled plasmid DNA free of contamination. The bacterial cultures were grown overnight (O/N) with shaking (200 rpm) at 37°C in LB broth, with appropriate antibiotic(s). About 1.5 to 3mL culture was centrifuged for 1 min at 4,000 x g to pellet the bacterial cells. The pellet was resuspended in 100µL of TEG buffer by vigorous pipetting, 200µL of Soln. II was added, mixed by inversion till the solution becomes clear and incubated on ice for 5 min. The cell lysate was neutralized by addition of 150µL of Soln. III, mixed well and incubated on ice for 5 min. The cell debris was removed by centrifuging for 5 min at 12,000 x g at room temperature. The supernatant was transferred to a clean tube, RNase A to a final concentration of 20µg mL^{-1} (Sambrook et al., 1989) was added and incubated at 37°C for 20 min. To the above solution 400µL of chloroform was added, mixed for 30 s and centrifuged for 5 min at 12,000 x g at 4°C. The upper aqueous layer was transferred to a clean tube, $1/10^{th}$ volume sodium acetate and one volume absolute ethanol was added with mixing and kept at -20°C for 1-2 h. The sample was centrifuged at 12,000 x g for 10 min at room temperature. The pellet was washed thrice with 70% ethanol and dried under vacuum. The dried pellet was dissolved in 40µL of deionized water and 40µL of PEG-NaCl solution (20% PEG 8000 in 2.5M NaCl) was added. The mixture was incubated on ice for 20 min and the plasmid DNA pelleted out by centrifugation at 12,000 x g for 15 min at 4°C. The supernatant was aspirated carefully, the pellet washed with 70% ethanol and air-dried. The dried pellet was resuspended in 20µL deionized water and stored at -70°C.

2.4.2 Isolation of plasmid DNA from *Agrobacterium tumefaciens*

Solutions

TEG Buffer (Soln. I): 25mM Tris-HCl (pH 8.0), 10mM EDTA (pH 8.0), 50mM Glucose.

Soln. II: 0.2N NaOH, 1% SDS (freshly prepared)

Soln. III: 3.0M Potassium acetate (pH 4.8)

Chloroform: Isoamyl alcohol (24:1), Isopropanol, 70% Ethanol, deionized water

Overnight grown culture of *A. tumefaciens* was centrifuged at 4,000 x g for 10 min. The cells were washed with 500µL of Sol I containing 10µL Triton X 100 and resuspended in 300µL of Soln I and 400µL of Soln II. The cells were mixed by inverting till the solution turns clear. This was followed by the addition of 400µL of Soln III. The suspension was mixed by inverting the tubes several times and centrifuged at 12,000 x g for 10 min. Clear supernatant was transferred to fresh 1.5mL tubes and 2µL of RNase A (10mg mL^{-1}) added. The tube was incubated at 37ºC for 30 min and then extracted with 400µL of Chloroform: isoamyl alcohol. For phase separation the tubes were centrifuged at 12,000 x g for 10 min. The upper aqueous phase was transferred to a 1.5mL tube and 600µL of isopropanol added. It was mixed thoroughly and the plasmid DNA pelleted by centrifugation at 12,000 x g for 10 min. The DNA pellet was washed with 70% ethanol, air dried, dissolved in 40µL of deionized sterile water and stored at -20ºC till further use.

2.4.3 Genomic DNA Isolation

Solutions

Extraction buffer: 100mM Tris-HCl and 20mM sodium EDTA adjust pH to 8.0 with HCl; add NaCl to 1.4M and 2.0% (w/v) CTAB (cetyltrimethylammonium bromide). Dissolve CTAB by heating to 60°C. Store at 37°C. Add ß-mercaptoethanol to 0.2 % before use.

Chloroform: isoamyl alcohol 24:1(v/v), 5M NaCl, RNAase A (10mg mL^{1}), 95% ethanol, 76% ethanol TE buffer: 10mM Tris-HCl and 1mM EDTA, adjust pH to 8.0 and autoclave.

Genomic DNA was isolated by using the protocol of Lodhi et al. (1994). Fresh young leaves were collected, frozen in liquid nitrogen and crushed to a fine powder. About 1g of ground tissue was extracted with 10mL extraction buffer. The slurry was poured into a clean

autoclaved 50mL centrifuge tube and 100mg insoluble polyvinylpolypyrrolidone (PVPP) added. The tube was inverted several times to thoroughly mix the slurry, incubated at 60ºC for 30 min and then allowed to cool down to room temperature. Twelve milliliters of chloroform: isoamylalcohol mix was added and the contents mixed by inverting the tube gently till an emulsion formed. The mix was then centrifuged at 6,000 x g for 15 min at room temperature. Supernatant was carefully collected in a fresh tube and chloroform: isoamylalcohol extraction step repeated. To the clear supernatant 0.5 volume of 5M NaCl was added and mixed gently. Next two volumes of cold (-20ºC) 95% ethanol was added and the sample kept at 4º C until DNA strands appeared. The tube was centrifuged at 3,000 x g for 3 min and then at 5,000 x g for next 3 min. The supernatant was poured off, the DNA pellet washed with cold (4ºC) 76% ethanol and air dried. DNA was dissolved in 300µL of TE buffer.

The DNA solution was treated with 1µL RNase A (10mg mL^{-1}) per 100µL DNA and incubated at 37ºC for 30 min. The sample was extracted with chloroform: isoamyl alcohol to remove RNAase A, reprecipitated and dissolved in TE buffer. Purity of DNA was checked spectrophotometrically by measuring the ratio of OD at 260/280nm. DNA was stored at 4ºC.

2.4.4 Restriction digestion of DNA

Plasmid and genomic DNA restriction digestion was set up as per recommendations manufacturer.

2.4.5 Extraction and purification of DNA from agarose gels

The restriction digested DNA or PCR amplified products were run on an agarose gel in 1X TAE buffer (see Section 2.5). The gel was stained with ethidium bromide (0.5µg mL^{-1}) and viewed using a hand held long wavelength UV illuminator. The fragment of interest were excised from the gel and weighed. A 50-200mg gel slice was transferred to a 1.5mL microcentrifuge tube and 0.5mL GEX buffer (Auprep™ GEL^{X} kit, Life Technologies, USA) added. The tube was incubated at 60°C for 5 to 10 min with intermittent mixing until the gel slice was completely dissolved. The gel mixture was cooled down to room temperature. Approximately 0.7mL dissolved gel mixture was loaded into the GEL^{X} column which was placed on the collection tube. The assembly was centrifuged at 12,000 x g for 2 min and the

flow-through was discarded. The procedure was repeated for the balance of the dissolved gel mixture. The column was washed with 0.5mL of WF buffer by centrifuging for 30-60 s. The flow-through was discarded and the column washed with 0.7mL of WS buffer. The flow-through was discarded and the column again centrifuged at 12,000 x g for 3 min to remove residual ethanol. The column was placed onto a new 1.5mL centrifuge tube. About 30-50µL of elution buffer or sterile distilled water was added onto the center of the column membrane. The column was allowed to stand for 1-2 min and then centrifuged at 12,000 x g to elute the DNA. The eluted DNA was stored at -20°C.

2.4.6 Total RNA Isolation

Solutions

Biozol (BIOWORLD, USA)

Chloroform and Isoamyl Alcohol (24:1), Isopropanol, 70% ethanol in DEPC treated deionized water, DEPC treated deionized water

RNase free environment was created and maintained as described by Blumberg (1987). All glass and plastic ware was DEPC (0.1% in water) treated overnight and autoclaved. The pestle and mortar were also DEPC treated and then baked at 300ºC for 6 h. All materials were dried in a vacuum oven.

Total RNA from different plant tissues was isolated using Biozol reagent. The plant tissue was collected, washed with DEPC treated water, frozen in liquid nitrogen and crushed to a fine powder. To 100mg of the fine powder 1mL Biozol reagent was added and mixed thoroughly using a vortimix. Chloroform: isoamyl alcohol (300µL) was added and mixed thoroughly using vortimix. The tubes were centrifuged at 4ºC at 13,000 x g for 15 min. The supernatant was transferred to 1.5mL tubes and the chloroform: isoamyl alcohol step repeated. The aqueous phase was transferred to 1.5mL tubes and half volume isopropanol added. It was mixed thoroughly and kept for RNA precipitation for 1 h. Total RNA was pelleted out by centrifugation at 13,000 x g for 15 min at 4ºC. The RNA pellet was washed with 70% ethanol twice and dried in a SpeedVac centrifugal concentrator. RNA pellet was dissolved in 40µL of DEPC treated water and stored at -80ºC in aliquots. Purity of RNA was confirmed by measuring OD at 260/280 nm and also by visualization on 1.5% TAE Agarose gel.

2.4.7 cDNA first strand synthesis by Reverse Transcription

Complementary DNA (cDNA) is synthesized from a mature mRNA template in a reaction catalyzed by the enzyme reverse transcriptase. The resulting molecule is a DNA RNA hybrid and the process is called as cDNA 1st strand synthesis. For DNA double strand synthesis this hybrid molecule is digested with RNase H (specific for degrading RNA strand in a DNA RNA hybrid), DNA second strand is synthesized using DNA polymerase I (Okayama and Berg, 1982; Kimmel and Berger, 1987).

In the present study cDNA first strand was synthesized using ImProm-II™ Reverse Transcription System (Promega, USA). The reactions were set up as per the manufacturers guidelines.

A brief reverse transcription reactions of up to 1µg of total RNA performed in 20µL reactions comprised of components of the ImProm-II. Reverse Transcription System. Experimental RNA was combined with the oligo$(dT)_{15}$ primer. The primer/template mix was isothermally denatured at 70°C for 5 min and snap chilled on ice. A reverse transcription reaction mix was assembled on ice to contain nuclease-free water, reaction buffer, reverse transcriptase, magnesium chloride, dNTPs and ribonuclease inhibitor RNasin®. As a final step, the template-primer combination was added to the reaction mix on ice. Following an initial annealing at 25°C for 5 min, the reaction was incubated at 42°C for up to 1 h. The cDNA synthesized was directly added to amplification reactions.

The first strand reaction was set up as follows:

Experimental RNA (1µg)	1µL
Primer [Oligo$(dT)_{15}$ or Random (10pmol)	1µL
Nuclease-Free Water	3µL
Final volume	5µL

The tubes were incubated at 70°C for 5 min and then chilled in ice-water for 5 min. Tubes were briefly spun in a microcentrifuge to collect the condensate and maintain the original volume. The tubes were kept closed and on ice until addition of the reverse transcription reaction mix.

The reverse transcription reaction mix was prepared by combining the following components of the ImProm-II. Reverse Transcription System in a sterile 1.5mL microcentrifuge tube on ice.

Nuclease-free water	6.5µL
ImProm-II. 5X Reaction Buffer	4.0µL
$MgCl_2$ (15mM)	2.0µL
dNTP Mix (7.5mM)	1.0µL
RNasin® Ribonuclease Inhibitor (40U μL^{-1})	0.5µL
ImProm-II Reverse Transcriptase	1.0µL
Final volume	15.0µL

An aliquot of 1.0µg total RNA and oligo $(dT)_{15}$ or Random hexamer primer (10pmol) mix was added to the above reaction for a final reaction volume of 20µL per tube. The tube was incubated at 25°C for 5 min for primer annealing and then at 42°C for 1 h for cDNA first strand synthesis. Reverse transcriptase was thermally inactivated by incubation at 70°C for 15 min prior to proceeding with PCR amplification (Sellner et al., 1992; Chumakov, 1994).

2.4.8 Polymerase Chain Reaction (PCR)

PCR is a powerful technique to amplify a desired nucleotide sequence using sequence specific primers. This amplification may be either of and from a single template or of a template from a mixture of templates (Saiki et al., 1985; Mullis and Faloona, 1987; Saiki et al., 1988; Mullis, 1990; Arnheim and Erlich, 1992). This technique has been successfully used for various purposes like fishing out of gene(s) from genomic DNA or from cDNA population (Todd et al., 1987), introducing restriction sites of interest in the amplified product for directional cloning (Scharf et al., 1986), creating sequence mismatch/ deletion/ addition resulting in mutant version of a gene or nucleotide sequence (Goodenow et al., 1989; Fuchroen et al., 1989; Mendelman et al., 1990), differentiating between two alleles (Kwok et al., 1990) etc.

In the present study applications of PCR were exploited for a few of the above specified applications. The PCR reaction mixture and cycling conditions used were as follows:

Reaction mixture

Sterile deionized water	6.2µL
Template (50ng μL^{-1})	1.0µL
Forward primer (6pmol)	1.0µL
Reverse primer (6pmol)	1.0µL
dNTPs (0.2mM)	4.0µL
10 x Buffer (Mg^{+2} 1.5mM)	1.5µL
Taq Polymerase (1U μL^{-1})	0.3µL
Total volume	15.0µL

PCR cycle conditions

1 cycle	95ºC	5 min
35 cycles	95ºC	1 min
	45-65ºC	30-45 s (annealing temperature was dependent on primer Tm)
	72ºC	1-1 min 30 s
1 cycle	72ºC	5 min
1 cycle	4ºC	hold

2.4.9 Rapid Amplification of cDNA ends (RACE)

Generally using reverse transcription either partial cDNA fragments (both 5' and 3' ends missing) or cDNA with full 5'end missing are amplified from total cDNA. If a partial cDNA sequence is known, unknown sequences to the 5' and 3' of the known sequence can be reverse transcribed from RNA, amplified by PCR using RACE (Frohmman et al., 1988). In the present study SMART ™ RACE cDNA Amplification Kit (BD Biosciences, Clontech, USA) was used. The reactions were set up as per the manufacturer's guidelines.

Briefly SMART technology provides a mechanism for generating full-length cDNA's in reverse transcription reactions (Zhu et al., 2001). This is done by the joint action of the SMART II™ A Oligonucleotide and the PowerScript™ Reverse Transcriptase (RT). PowerScript RT is a variant of MMLV RT, which upon reaching the end of a RNA template,

exhibits terminal transferase activity by adding 3–5 residues (predominantly dC) to the 3' end of the first-strand cDNA. The SMART oligo contains a terminal stretch of G residues that anneal to the dC-rich cDNA tail and serves as an extended template for RT. PowerScript RT switches templates from the mRNA molecule to the SMART oligo, generating a complete cDNA copy of the original RNA with the additional SMART sequence at the end. Following reverse transcription, the first-strand cDNA is used directly in 5'- and 3'-RACE PCR reactions. The only requirement for SMART RACE cDNA amplification is 23–28 nucleotides of sequence information in order to design gene-specific primers (GSPs) for the 5'- and 3'-RACE reactions.

Using SMART RACE Kit two separate cDNA populations, 5'-RACE cDNA and 3'-RACE cDNA are synthesized. The cDNA for 5'-RACE is synthesized using a modified lock-docking oligo(dT) primer and the SMART II A oligo as described above. The modified oligo(dT) primer, termed the 5'-RACE CDS Primer A (5'-CDS), has two degenerate nucleotide positions at the 3' end. These nucleotides position the primer at the start of the poly A+ tail and thus eliminate the 3' heterogeneity inherent with conventional oligo(dT) priming (Borson et al., 1994).

The 3'-RACE cDNA is synthesized using a traditional reverse transcription procedure, but with a special oligo(dT) primer. This 3'-RACE CDS Primer A (3'-CDS) primer includes the lock-docking nucleotide positions as in the 5'-CDS primer and also has a portion of the SMART sequence at its 5' end. By incorporating the SMART sequence into both the 5'- and 3'-RACE cDNA populations, both RACE PCR reactions can be primed using the Universal Primer A Mix (UPM), which recognizes the SMART sequence, in conjunction with distinct gene-specific primers.

Once RACE cDNAs are prepared, 5'- and 3'-RACE can be performed using gene-specific primers. All PCR reactions in the SMART RACE protocol are carried out using the Advantage® 2 Polymerase Mix. The Polymerase Mix is comprised of TITANIUM™ Taq DNA Polymerase—a nuclease-deficient N-terminal deletion of Taq DNA polymerase plus TaqStart® Antibody to provide automatic hot-start PCR (Kellogg et al., 1994) and a minor amount of a proofreading polymerase.

The reactions were set up for 5’ and 3’ RACE cDNA as below:

For preparation of 5'-RACE cDNA

RNA sample (1μg μL^{-1})	1μL
5'-CDS primer A (12μM)	1μL
SMART II A oligo (12μM)	1μL

For preparation of 3'-RACE cDNA

RNA sample (1μg μL^{-1})	1μL
3'-CDS primer A (12μM)	1μL

Sterile H_2O was added to a final volume of 5μL for each of the above reaction. Contents were mixed and the tubes centrifuged briefly. The tubes were incubated at 70°C for 2 min and cooled on ice for 2 min. The tubes were briefly centrifuged and to each reaction tubes following reagents were added:

5X First-Strand Buffer	2μL
DTT (20mM)	1μL
dNTP Mix (10mM)	1μL
PowerScript Reverse Transcriptase	1μL
Total volume	10μL

The contents of the tubes were mixed by gentle pipetting and tubes brief centrifuged. The tubes were incubated at 42ºC for 1.5 h in a hot-lid thermal cycler. The first strand reaction mixture was diluted to 100μL with Tricine- EDTA buffer and heated at 70ºC for 7 min. The above diluted first strands were used for 5’ and 3’ RACE. The following master mix was prepared according to the number of PCR reactions to be set up:

PCR-Grade Water	34.5μL
10X Advantage 2 PCR Buffer	5.0μL
dNTP Mix (10mM)	1.0μL
50X Advantage 2 Polymerase Mix	1.0μL
Total volume	41.5μL

The above master mix was used for 5' and 3' RACE PCR. The reaction was set up as follows:

5'RACE

5' RACE cDNA	2.5µL
UPM (10X)	5.0µL
GSP1 (10µM)	1.0µL
Master Mix	41.5µL
Total volume	50.0µL

3'RACE

3'RACE cDNA	2.5µL
UPM (10X)	5.0µL
GSP2 (10µM)	1.0µL
Master Mix	41.5µL
Total volume	50.0µL

Following PCR cycling conditions were used for amplifying 5' and 3' RACE products.

5 cycles:	94°C 30 s	72°C 3 min	
5 cycles:	94°C 30 s	70°C 30 s	72°C 3 min
27 cycles:	94°C 30 s	68°C 30 s	72°C 3 min

After PCR, 10µL of the reaction mix was loaded on 1.2% agrose gel in 1XTAE buffer and checked for amplification.

Following is the list of primers and the other reagents which were used for the RACE:

SMART II™ A Oligonucleotide 5'(d)AAGCAGTGGTATCAACGCAGAGTACGCG GG3'

3'-RACE CDS Primer A 5'(d)AAGCAGTGGTATCAACGCAGAGTAC $(T)_{30}$V N–3'

5'-RACE CDS Primer A (5'-CDS) 5'– $(T)_{25}$V N–3'

(N = A, C, G, or T; V = A, G, or C)

10X universal Primer A Mix (UPM)

Long :5' (d) CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT 3'

Short :5' (d) CTAATACGACTCACTATAGGGC 3'

Nested universal Primer A (NUP) 5' (d) AAGCAGTGGTATCAACGCAGAGT3'

5X First-Strand Buffer 250mM Tris-HCl (pH 8.3), 375mM KCl, 30mM $MgCl_2$

Dithiothreitol (DTT) 20mM

Tricine-EDTA Buffer 10mM Tricine-KOH (pH 8.5), 1.0mM EDTA

2.4.10 Genome Walking

Genome walking is a novel method for walking upstream (i.e., towards the promoters) or downstream in a genomic DNA clone from a known sequence such as a cDNA (Siebert et al., 1995). In the present study GenomeWalker Kit from (BD Biosciences, Clontech, USA) was used for the purpose. Reactions were set up as per the manufacturers guidelines. A brief introduction and the protocol are discussed here.

Briefly four "libraries" of uncloned, adaptor-ligated genomic DNA fragments are made, (these are not libraries in the conventional sense as the DNA fragments are not ligated into a vector which is then propagated in *E. coli.* However, like conventional libraries, GenomeWalker Libraries are a pool of specially prepared DNA fragments from which specific DNA fragments can be identified, isolated, and cloned). Construction of GenomeWalker Libraries begins with isolation of very clean genomic DNA that has a very high average molecular weight. The starting DNA must be of considerably higher quality than the minimum suitable for Southern blotting or conventional PCR. Four separate aliquots are then thoroughly digested with four different restriction enzymes that recognize specific 6-base site, leaving blunt ends. Following digestion, each pool of DNA fragments is ligated to the GenomeWalker Adaptor.

The GenomeWalker protocol consists of two PCR amplifications per library. The first or "primary" PCR amplification uses the outer adaptor primer (AP1) provided in the kit and an outer gene-specific primer (GSP1). The primary PCR mixture is then diluted and used as a template for a secondary or "nested" PCR amplification using the nested adaptor primer (AP2) and a nested gene-specific primer (GSP2). Each of the DNA fragments—which begin in known sequence at the 5' end of GSP2 and extend into the unknown adjacent genomic DNA—can then be cloned and further analyzed.

2.4.10.1 Digestion of Genomic DNA

For each library construction four blunt-end digestions of experimental genomic DNA were setup, one for each of the blunt-end restriction enzyme provided.

For each reaction, the following were combined in a separate 1.5mL tubes:

Genomic DNA (0.1µg μL^{-1})	25µL
Restriction enzyme (10 U μL^{-1})	8µL
Enzyme buffer (10X)	10µL
Deionized H_2O	57µL

Contents in the tubes were mixed gently and incubated at 37°C for 2 h. The reaction mix was vortexed at slow speed for 5–10 s and kept at 37°C overnight (16–18 h). From each reaction tube 5µL of the reaction mix was run on a 0.6% agarose/EtBr gel to determine the extent of restriction.

2.4.10.2 Purification of DNA

To each reaction tube, an equal volume (95µL) of phenol was added and vortexed at slow speed for 5–10 s. The aqueous and organic phases were separated by brief centrifugation at room temperature. The upper aqueous layer was transferred to a fresh 1.5mL tube and an equal volume (95µL) of chloroform added. Tube was vortexed slowly and centrifuged briefly at room temperature to separate the phases. The upper (aqueous) layer was transferred to a fresh 1.5ml tube and 1/10 volume (9.5µL) of 3M Sodium Acetate (pH 4.5) and 2 volumes (190µL) of ice cold 95% ethanol were added. The mix was vortexed slowly. Tubes were centrifuged at 13,000 x g for 15 min at 4°C. The pellet was washed with 100µL of ice cold 80% ethanol and air dried. The pellet was dissolved in 20µL of TE buffer, pH 7.5. From each reaction tube 1µL was run on a 0.6% agarose/EtBr gel to determine the approximate quantity of DNA after purification.

2.4.10.3 Ligation of Genomic DNA to GenomeWalker™ Adaptors

From each of the above restriction digests, 4µL of digested, purified DNA was transferred to a fresh 0.5-mL tube and following added:

GenomeWalker Adaptor (25µM)	1.9µL
10X Ligation Buffer	1.6µL
T4 DNA Ligase (6U μL^{-1})	0.5µL

The tubes were incubated at 16°C overnight. Reaction was stopped by incubating the tubes at 70°C for 5 min. To each tube 72µL of TE (10/1, pH 7.5) was added and mixed gently by vortexing.

2.4.10.4 Primary PCR

Sufficient primary PCR master mix was prepared by combining the following reagents in a 0.5mL tube:

per reaction	
Deionized H2O	40µL
10X Advantage 2 PCR Buffer	5µL
dNTP (10mM each)	1µL
AP1 (10µM)	1µL
Advantage 2 Polymerase Mix (50X)	1µL
Total volume	48µL

Contents of the tube were mixed well by vortexing and briefly spun. From the above master mix 48µL of the primary PCR master mix was transferred to four tubes and 1µL of GSP1 added to each tube. From each library 1µL was added to each of the above four tubes, mixed and briefly spun. Following two step thermal cycling parameters were used:

7 cycles: 94°C 25 s 72°C 3 min

32 cycles: 94°C 25 s 67°C 3 min

67°C for an additional 7 min after the final cycle.

The primary PCR products (5µL) were analyzed on a 1.5% agarose/EtBr gel. Five additional cycles can be performed in case no bands are seen.

2.4.10.5 Secondary PCR

The primary PCR product was diluted 1:50. A clean 0.5mL tube was used for each sample; 1µL of each primary PCR was added into 49µL of deionized H_2O.A secondary PCR master mix for all four reactions was prepared. Following reagents were combined in a 0.5mL tube:

per reaction

Deionized H2O	40µL
10X Advantage 2 PCR buffer	5µL
dNTP (10mM each)	1µL
AP2 (10µM)	1µL
Advantage 2 Polymerase Mix (50X)	1µL
Total volume	48µL

Contents of the tube were mixed well by vortexing and briefly spun. From the above master mix 48µL was transferred to four different tubes and 1µL of GSP2 added to each tube. Diluted primary PCR product (1µL) was added to the tubes, mixed and briefly spun. Following two step cycle parameters were used for thermal cycling:

5 cycles: 94°C 25 s 72°C 3 min

20 cycles: 94°C 25 s 67°C 3 min

67°C for an additional 7 min after the final cycle.

For visual analysis 5µL of the secondary PCR products was loaded on a 1.5% agarose/EtBr gel. Four additional cycles can be performed in case of no bands are seen. The unused portion of each secondary PCR was stored at 4°C until confirmation that the procedure was successful. The bands of interest were gel purified and cloned in pGEM T Easy Vector (Promega, USA) and sequenced for further analysis.

Enzymes, adapter, primers and the other reagents used for Genome walking are listed below:

Enzymes and buffers:

Dra I (10U μL^{-1}); 10X *Dra* I Restriction Buffer

EcoR V (10U μL^{-1}); 10X *EcoR* V Restriction Buffer

Pvu II (10U μL^{-1}) ; 10X *Pvu* II Restriction Buffer

Stu I (10U μL^{-1}); 10X *Stu* I Restriction Buffer

T4 DNA Ligase (6U μL^{-1}); 10X Ligation Buffer

Genome walker adapter

5'–GTAATACGACTCACTATAGGGCACGCGTGGTCGACGGCCCGGGCTGGT 3’

3'–H_2N-CCCGACCA-PO_4–5’

Adaptor Primer 1 (AP1) 5' (d)GTAATACGACTCACTATAGGGC 3'

Nested Adaptor Primer 2 (AP2) 5' (d) ACTATAGGGCACGCGTGGT3’

Phenol, Chloroform, 3M sodium acetate, 95% ethanol, 80% ethanol

TE buffer : 10mM Tris, 0.1mM EDTA (10/0.1, pH 7.5)

TE buffer : 10mM Tris, 1mM EDTA (10/1, pH 7.5)

0.5X TBE buffer or 1X TAE buffer

Advantage 2 Polymerase Mix (50X)

10X PCR reaction buffer

dNTP mix: 10mM each of dATP, dCTP, dGTP & dTTP. Store at –20°C.

PCR reaction tubes, Deionized H_2O

1 Kb ladder of DNA size markers

2.4.11 Quantitative Real Time PCR (QPCR)

Quantitative Real Time PCR is a powerful tool for gene expression analysis and was first demonstrated by Higuchi et al. (1992, 1993). QPCR quantitates the initial amount of the template most specifically, sensitively and reproducibly, and is a preferable alternative to other forms of quantitative reverse transcriptase PCR that detect the amount of final amplified product at the end-point (Freeman et al., 1999; Raeymaekers, 2000). QPCR monitors the fluorescence emitted during the reaction as an indicator of amplicon production during each PCR cycle (*i.e.*, in real time) as opposed to the endpoint detection (Higuchi et al., 1992, 1993). The QPCR system is based on the detection and quantitation of a fluorescent reporter (Lee et al., 1993; Livak et al., 1995). There are three main fluorescence-monitoring systems for DNA amplification (Wittwer et al., 1997a): (1) hydrolysis probes, (2) hybridizing probes and (3) DNA-binding agents (Wittwer et al., 1997b; van der Velden et al., 2003). Most commonly used are the hydrolysis probes, which include TaqMan probes (Heid et al., 1996), molecular beacons (Mhlanga and Malmberg, 2001; Vet et al., 2002; Abravaya et al., 2003; Tan et al., 2004; Vet and Marras, 2005) and scorpions (Saha et al., 2001; Solinas et al., 2001; Terry et al., 2002). They use the 5' exonuclease activity of Taq polymerase to measure the amount of target sequences in cDNA samples. The use of fluorescent probe technologies reduces the risk of sample contamination while maintaining convenience, speed and high throughput screening capabilities (Mitchell, 2001; Reischl et al., 2002).

In the present study the Brilliant® II QPCR Master Mix (Stratagene, USA) was used. This kit supports quantitative amplification and detection with multiplex capability, and shows consistent high performance with various fluorescent detection systems, including molecular beacons and TaqMan® probes. The kit supports PCR amplifications and detection of a variety of DNA targets, including genomic DNA, plasmid DNA, and cDNA. The Brilliant II QPCR master mix includes SureStart® *Taq* DNA polymerase, a modified version of *Taq2000*™ DNA polymerase with hot start capability. A passive reference dye (an optional reaction component) is provided so as to control the final dye concentration, increasing the flexibility of the reagents for use with different platforms.

2.4.11.1 Pre-protocol considerations

Magnesium Chloride Concentration

Magnesium chloride concentration in the PCR reaction mix affects the specificity of the PCR primers and probe hybridization (Eckert and Kunkel, 1990). The Brilliant® II QPCR Master Mix kit contains $MgCl_2$ at a concentration of 5.5mM (in the 1 x solution), which is suitable for most targets.

Probe Design

Probes are designed in a fashion so as to have a melting temperature 7–10°C higher than the annealing temperature of the primers. Lyophilized custom molecular beacon or TaqMan probes are constituted in 5mM Tris-HCl, pH 8.0 and 0.1mM EDTA (low TE buffer).

Fluorescence Detection

Fluorescence may be detected either in real-time or at the endpoint of cycling using a real-time spectrofluorometric thermal cycler. For endpoint analysis, PCR reactions can be run on any thermal cycler and then analyzed with a fluorescence plate reader that has been designed to accommodate PCR tubes and is optimized for the detection of fluorescent PCR reactions. Readings may be acquired before and after PCR for comparison.

Optimal Concentrations for Experimental Probes and Primers Probes

The optimal concentration of the experimental probe is determined empirically. The optimal concentration is the lowest concentration that results in the lowest Ct and an adequate fluorescence for a given target concentration. In present study TaqMan® probes were used.

TaqMan® Probes

The TaqMan probe concentration can be optimized by varying the final concentration from 100 to 500nM in increments of 100nM.

PCR Primers

The optimal concentration of the upstream and downstream PCR primers is determined empirically. The optimal concentration is the lowest concentration that results in the lowest Ct and an adequate fluorescence for a given target concentration. The primer concentration for use with TaqMan probes can be optimized by varying the primer concentration from 50 to 600nM. The best concentrations of the upstream and downstream primers are not always of equal molarity.

Reference Dye

A passive reference dye may be added to compensate for non-PCR related variations in fluorescence. Fluorescence from the passive reference dye does not change during the course of the PCR reaction but provides a stable baseline to which samples are normalized. In this way, the reference dye compensates for changes in fluorescence between wells caused by slight volume differences in reaction tubes. The excitation and emission wavelengths of the reference dye are 584nm and 612nm respectively.

Reference Dye Dilution

Reference dye is prepared fresh prior to setting up the reactions. All reaction tubes containing reference dye are protected from light. Initial dilutions of the reference dye are made in nuclease-free PCR-grade H_2O and used at a final concentration of 30 - 300nM depending upon the instrument configuration.

Data Acquisition with a Spectrofluorometric Thermal Cycler

Acquisition of real-time data generated by fluorogenic probes was as recommended by the instrument manufacturer. Data collection was either at the annealing step of each cycle (3-step cycling protocol) or the annealing/extension step (2-step cycling protocol).

2.4.11.2 Preparing the Reactions

The reference dye was diluted 1:500 (recommended in kit) using nuclease-free PCR-grade H_2O resulting in a final reference dye concentration of 30nM in the reaction mixture. Real time PCR model ABI PRISM 7700 was used in present study.

The experimental reactions were prepared by adding the following components in order:

Reagent Mixture

Nuclease-free PCR-grade H_2O to adjust the final volume to 25µL (including experimental DNA)

2 x master mix	2.5µL
experimental probe (optimized concentration)	xµL
upstream primer (optimized concentration)	xµL
downstream primer (optimized concentration)	xµL
diluted reference dye	0.375µL

The reactions were mixed without creating bubbles and xµL of experimental gDNA, cDNA or plasmid DNA was added to each experimental reaction. Reaction was mixed gently and briefly spun.

2.4.11.3 PCR Cycling Programs

The reactions were placed in the instrument and a 2 step program was run (recommended for TaqMan reactions).

Two-step Cycling Protocol

Cycles	Duration of cycle	Temperature
1	10 min	95°C
40	5-30 s	95°C
	1.0 min	60°C

Initial 10 min incubation was to fully activate the DNA polymerase. The temperature cycler was set to detect and report fluorescence during the annealing/extension step of each cycle.

2.5 Nucleic acids blotting/hybridization

2.5.1 Southern Blotting

Solutions

1X TAE:	0.04M Tris-Acetate (pH 8.0), 0.001M EDTA (pH 8.0)
20X SSC:	3M NaCl, 0.3M Sodium citrate (pH 7.0)
Depurination solution:	0.25N HCl
Denaturation solution:	1.5M NaCl, 5 M NaOH

Neutralization solution: 0.5M Tris-HCl (pH 7.4), 3M NaCl

Gel loading dye (6X): 0.25% Bromophenol blue in 40% (w/v) sucrose in water

For Southern hybridization (Southern, 1975) the DNA samples were electrophoresed on an agarose gel in 1X TAE buffer containing 0.5μg mL^{-1} ethidium bromide. The gel was rinsed with deionized water (DW) and placed in depurination solution for 15 min. The gel was rinsed with deionized water and immersed in denaturation solution for 30 min with gentle shaking. The gel was again rinsed with deionized water and transferred to neutralization solution for 45 min. The gel was next set up for capillary transfer of DNA to solid membrane support.

A tray was filled with the transfer buffer (20X SSC). A platform was made and covered with a wick, made from 2 sheets of Whatman 3MM filter paper saturated with transfer buffer and the gel was placed on it. It was surrounded with Saran Wrap to prevent the transfer buffer from being absorbed directly by the paper towels stacked above the membrane. A sheet of Hybond-N^+ membrane (Amersham, UK) of the exact gel size was wetted with deionized water followed by transfer buffer (20X SSC) and then placed on top of the gel. A glass rod was rolled over the membrane to remove any trapped air bubbles. One piece of Whatman 3MM paper wetted with 20X SSC was placed on the membrane followed by Whatman 3MM paper pre-wetted in 2X SSC. On this paper another dry Whatman 3MM paper was placed followed by a stack of absorbent paper towels. A glass plate and a ~0.5 kg weight were placed on the top of the paper towels. Transfer of DNA was allowed to proceed for 18 h. The membrane was marked for orientation, removed carefully and washed with 6X SSC. The membrane was air dried and baked for 2h at 80°C to immobilize DNA onto the nylon membrane. Hybridization and autoradiography were carried out as is described in the following section 2.8.

2.5.2 Slot Blot Hybridization

Solutions

20X SSC 3M NaCl, 0.3M Sodium citrate (pH 7.0)

NaOH 3M

For slot blot hybridization DNA or RNA samples were diluted according to experimental requirements. The DNA samples were denatured by adding 1/10th volume 3M NaOH and incubation at 65ºC for 10 min. To the denatured sample an equal volume of 6X SSC was added. Two layers of Whatman 3MM filter paper wetted with sterile deionized water and 6X SSC were placed in the Slot Blot apparatus followed by Hybond-N$^+$ membrane (Amersham, UK) treated in the manner as above. The Slot Blot unit was assembled and wells washed with 500µL of 6X SSC by applying vacuum. After washing, samples prepared earlier were applied in the wells and vacuum applied till whole sample volume passed through the well slit and wells appear dry. The unit was carefully disassembled and the membrane taken out. The membrane was air dried and then baked for 2 h at 80°C to immobilize DNA. Hybridization and autoradiography were carried out as described in the following section 2.8.

2.6 Random primer labeling

Random primer labeling of the DNA probes (Feinbeng and Vogelstein, 1983, 1984) was done using the Megaprime DNA labeling kit (Amersham, UK). Reaction (50µL) was set up as follows:

25ng DNA (used as probe)	5.0µL
Primer solution (Random hexanucleotides) (3.5 A_{260} U)	5.0µL

Above mixture was heated in a boiling water bath for 10 min and cooled to room temperature facilitating primer annealing to the DNA.

10X reaction buffer (500mM Tris-HCl, pH 8.0; 100mM $MgCl_2$; 10mM DTT; 0.5mg mL^{-1} acetylated BSA)	5.0µL
dCTP, dGTP, dTTP (0.5mM in 333mM Tris-HCl, pH 8.0; 33.3mM $MgCl_2$; 10mM β-Mercaptoethanol)	12.0µL (4.0µL each)
[α-^{32}P-dATP (Sp. activity 3000 Ci $mmol^{-1}$)	5.0µL
Sterile deionized water	16.0µL
Exonuclease free Klenow fragment (2 U μL^{-1})	2.0µL
Total volume	50.0µL

The reaction was carried out at 37°C for 45 min and stopped by incubation in a boiling water bath for 10 min and snap chilled on ice.

2.7 Hybridization

Solutions

20X SSC	3M NaCl, 0.3M Sodium citrate (pH 7.0)
Hybridization buffer	1% BSA; 1.0 mM EDTA, pH 8.0; 0.5M Sodium phosphate, pH 8.0; 7% SDS
Low stringency wash buffer	2 X SSC, 0.1% SDS
High stringency wash buffer	0.2 X SSC, 1% SDS

The blots made as in sections 2.5 and 2.6 above were prehybridized at 45°C in 30mL of hybridization buffer for 6-8 h in a hybridization incubator (Robin Scientific, USA). The buffer was decanted and fresh buffer added with the denatured radiolabelled probe. Hybridization was carried out at 50°C for 14-18 h. The solution was decanted and the membrane washed with low stringency buffer at 55°C for 15 min followed by a high stringency wash at 55°C for 15 min. The moist blot was wrapped in Saran wrap and exposed to X-ray film at –70°C in a cassette with intensifying screen.

2.8 Protein purification

2.8.1 Inclusion Bodies

Solutions

Lysis Buffer	50mM Tris, 50mM EDTA, 15% sucrose, Lysozyme 2mg mL^{-1}, pH 8.0
Crystal Wash I	0.5M NaCl and 2% Triton X100
Crystal Wash II	0.5M NaCl
Solubilization Buffer	50mM Na_2CO_3, 10mM DTT, 5mM PMSF, 0.02% Themersol

E. coli cells expressing recombinant protein were grown in Terrific Broth at 37ºC in a shaker (200rpm) for 48 h. Cells were induced by with IPTG (0.5mM) and allowed to grow for 6 more hours.

Cells were harvested by centrifuging at 6000 x g for 10 min, suspended in lysis buffer and incubated for 1h at 37ºC. The suspension was sonicated for 1 min with pulses of 5 s on and 5 s off at 80% amplitude. After sonication the cells were centrifuged at 20,000 x g for 15 min. Supernatant was discarded and the cell pellet washed thrice each with Crystal Wash I, Crystal Wash II and finally with deionized water. The pellet was dissolved in solubilization buffer overnight at 37ºC. Proteins dissolved in the solubilization buffer was separated from debris by centrifuging at 20,000 x g for 15 min. Protein sample was then analyzed on 10% SDS PAGE.

2.8.2 Ni-chelated affinity column

Polyhistidine-tags are often used for affinity purification of polyhistidine-tagged recombinant proteins that are expressed in *Escherichia coli* or other prokaryotic expression systems (Hengen, 1995). The recombinant protein, among several other bacterial proteins is loaded on affinity matrix column such as Ni-agarose. This affinity matrix contains bound metal ion nickel, to which the polyhistidine-tag binds with micromolar affinity. The matrix is then washed with buffer to remove unbound proteins. The washing efficiency may be improved by the addition of 20mM imidazole and histidine-tagged proteins are then usually eluted with 150-300mM imidazole. The purity and amount of protein is assessed by SDS-PAGE**.**
In present study B-PER® 6xHis Fusion Protein Purification Kit (Pierce, USA) was used.

Solutions and affinity column

B-PER® Bacterial Protein Extraction Reagent: 20mM Tris buffer (pH 7.5) and a proprietary additive

Wash Buffer 1: 35mM Tris, 150mM NaCl, 10mM imidazole, 5% glycerol, pH 7.2, and 0.5X concentration of proprietary B-PER® Reagent additive

Wash Buffer 2: 50mM Tris, 300 mM NaCl, 25mM imidazole, 10% glycerol, pH 6.8

Elution Buffer: 50mM Tris, 300mM NaCl, 200mM imidazole, 10% glycerol, pH 6.8

Nickel Chelated Columns: Contains nickel (Ni^{+2}) chelated iminodiaceticacid (IDA) covalently immobilized to 4% beaded agarose; stored in 0.02 % sodium azide

The column(s) and buffers were equilibrated to room temperature. A 250mL culture of *E. coli* expressing the recombinant protein (as described in section 2.8.1) (O.D.600 = 1.5-3.0) was used and cells pelleted by centrifugation. Supernatant was discarded. The cell pellet was suspended in 10mL of B-PER® Reagent to a homogenous cell suspension and kept at room temperature for 10 min with gentle shaking. Soluble proteins were separated from insolubles by centrifugation at 27,000 x g for 15 min. The supernatant was collected in a new tube.

Sodium azide storage solution from the nickel chelated column was drained from the gel bed. The gel bed was prepared by adding 10mL (2 × 5mL) of B-PER® Reagent and allowed to flow through the column. Up to 10mL of sample (2 × 5mL) was applied to the column and allowed to flow through the gel bed. The column was washed by adding at least 6mL (2 × 3mL) of Wash Buffer 1 and 9mL (3 × 3mL) of Wash Buffer 2. The 6xHis-tagged protein was eluted by adding 6mL (2 × 3 mL) of Elution Buffer. Protein elution was monitored by measuring the absorbance at 280 nm of collected fractions. The eluted protein was analyzed by SDS-PAGE.

After protein was completely eluted, the gel bed was washed with 2-5mL of Wash Buffer 1. Some solution was left above the gel bed, the column capped and stored upright at 4°C. Microbial growth during long-term storage can be prevented by using Wash Buffer 1 to which sodium azide has been added to 0.01-0.02%.

2.9 Sodium dodecylsulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE system is the widely used electrophoresis system for protein separations (Laemmli 1970). The resolution in a Laemmli gel is excellent because the treated peptides are stacked in a stacking gel before entering the separating gel.

Solutions

Monomer solution	29.2% acrylamide/0.8% bis in water. Stored at 4°C in the dark.
Running Gel Buffer	3.0M Tris-Cl (pH 8.8)
Stacking Gel Buffer	1.0M Tris-Cl (pH 6.8)
SDS	10% (w/v) in water
Initiator (APS)	10% Ammonium persulfate (make fresh)

Tank Buffer	0.025M Tris; 0.192M glycine; 0.1% SDS pH 8.3
Water saturated n-Butanol	5mL water in 50mL n-butanol
2X loading Buffer	0.125M Tris-Cl (pH 6.8); 4% SDS; 20% (v/v) glycerol; 0.1% (w/v) Bromophenol blue; 10% (v/v) 2-mercaptoethanol

Table 2.2 Composition of 10% Sodium dodecylsulfate - Polyacrylamide Gel

Solution	Seperating gel for 8mL	Stacking gel for 5mL
Gel Buffer	2.0mL	1.25mL
Monomer	2.66mL	0.65mL
Solutioin	3.33mL	3.05mL
Distilled water	50.0µL	25.00µL
APS	400.0µL	400.00µL
SDS	10.0µL	5.00µL

2.9.1 Preparation of the separating gel

A vertical slab gel (Hoefer Scientific, U.S.A.) was assembled using 1.0mm spacers. In a side-armed vacuum flask, 10% separating gel solution was made according to Table. 2.2 except for the addition of ammonium persulfate and TEMED. The solution was degassed by applying vacuum while stirring on a magnetic stirrer. TEMED and ammonium persulfate were added, and mixed gently without generating bubbles. The solution was pipetted into the gel cassette leaving 1.5cm from the top unfilled. The gel solution was overlaid with water-saturated n-butanol to remove trapped air bubbles and to layer evenly across the entire surface. When a sharp liquid-gel inter-surface was observed after the gel polymerization, the slab was tilted to pour off the overlay.

2.9.2 Preparation of the stacking gel

Stacking gel solution was prepared according to Table. 2.2 leaving out ammonium persulfate and TEMED. As in the separating gel, this solution was degassed. TEMED and ammonium persulfate were added, mixed and overlaid on the separating gel. A comb was inserted taking care not to trap air bubbles beneath the comb teeth. The gel was left to polymerize.

2.9.3 Preparation of the sample

Equal parts of the protein sample (as described in section 2.8 above) and the loading buffer were mixed in a microcentrifuge tube and kept in a boiling water bath for 90 s. The samples were centrifuged at 10,000 x g for 10 min, supernatant collected and used for electrophoresis.

2.9.4 Loading and running the gel

The comb was removed from the gel, the wells were flushed with distilled water and drained off completely. The wells were filled with tank buffer and the samples under laid using a syringe. Lower and upper buffer chambers were filled with tank buffer. Voltage was set between 70 – 80 volts. The run was stopped when the dye reached the bottom of the gel.

2.9.5 Silver staining of the gel

Solutions

Fixer	40% (v/v) methanol, 12% (v/v) acetic acid in deionized water
Thiosulfate solution	20mg of $Na_2S_2O_3$ in 100mL deionized H_2O
Silver Stain	200mg $AgNO_3$, 75µL HCHO in 100mL deionized H_2O
Developer	6g of Na_2CO_3, 50µL HCHO, 100µL $Na_2S_2O_3$ (4mg mL^{-1} stock) in 100mL deionized H_2O

After running the gel, it was transferred to the fixer solution for 1 h. This was followed by 2 x 20 min washes in 50% ethanol. The gel was transferred to thiosulfate solution for 1 min. and rinsed thrice with deionized H_2O (20 s each). The gel was silver stained (HCHO was added to the solution just before use) for 20 min with intermittent shaking in dark. The gel

was then rinsed thrice with deionized H_2O (20 s each) and transferred to the developer till the bands developed. The gel was washed with deionized H_2O and stored in the fixer.

2.10 Raising Polyclonal Antibody against CCoAOMT in Rabbit and antigen affinity purification

The purified CCoAOMT protein was used for raising polyclonal antibodies in rabbit (New Zealand White).

Solutions

Solutions and chemicals required

Affi-Gel 15 (BIORAD)	1mL
MOPS Buffer	100mM (pH 7.5)
Ethanolamine HCl	1M (pH 8.0)
Tris- HCl	1M (pH 8.0)
Glycine- HCl	100mM (pH 2.5)
Deionized water	

2.10.1 Affinity Column preparation for polyclonal antibody purification

Affi-Gel 10 and Affi-Gel 15 affinity supports are activated immunoaffinity supports that offer rapid, high efficiency coupling for all ligands with primary amino group, including proteins throughout the entire range of pIs and low molecular weight compounds such as peptides. Both Affi-Gel 10 and Affi-Gel 15 supports are N-hydroxysuccinimide esters of derivatized cross linked agarose gel bead support, and both couple to ligands spontaneously in aqueous and non aqueous solutions. The Affi-Gel 10 support contains a neutral 10 atom spacer and Affi-Gel 15 support contains a cationic charge in its 15 atom spacer arm. Affi-Gel 10 support couples proteins best at a pH near or below their isoelectric point, and Affi-Gel 15 support couples protein best near or above their isoelectric point. Therefore, when coupling at neutral pH (6.5-7.5), Affi-Gel 10 support is recommended for proteins with isoelectric points of 6.6 to 11(neutral and basic proteins), and the Affi-Gel 15 support is recommended for proteins with isoelectric points below 6.5 (acidic proteins). CCoAOMT protein has a calculated pI of 5.31.

Affi-Gel 15 matrix (stored at -20ºC) was allowed to come to 4ºC on ice. It was properly mixed to homogenous solution and 1mL of this matrix pipetted into a clean 30mL polypropylene tube. The gel was centrifuged at 3,500 x g for 5 min and supernatant was carefully pipetted out. The gel was given three washes with cold deionized water. After final wash water was removed as much as possible and 20mg purified protein (antigen) in 100mM MOPS (pH 8.0) was added to the gel. The tube was kept on a rocker in cold room for 12 h to allow antigen binding to the gel. Tube was then centrifuged at 3,500 x g for 5 min. Supernatant (containing unbound protein) was removed and the gel given two washes with cold MOPS buffer. Ethanol amine-HCl (pH 8.0) 100µL was added to the gel to block any active ester sites present. The tube was kept on a rocker in cold room for 4 h.

The gel was transferred to a column and washed with cold MOPS buffer till the O.D at 280nm measured zero. The column was also washed with 100mM Glycine - HCl (pH 2.5). The column wash was measured for O.D. at 280nm to ensure that no antigen was getting eluted. The column was again equilibrated with MOPS buffer and kept in cold room.

2.10.2 Antigen affinity purification of antibodies

Rabbit immune serum (stored at -70ºC), containing polyclonal antibodies against the antigen, was thawed overnight at 4ºC. The immune serum was centrifuged at 12,000 x g for 15 min at 4ºC and the supernatant diluted by adding an equal volume of MOPS buffer. The diluted immune serum was passed through 0.22µ filter and loaded on the Affi-Gel 15 column to which the antigen was bound and as prepared above. The immune serum collected after passing through the column was reloaded on to the column three times and O.D. at 280nm of the eluate measured every time. Reloading of immune serum was stopped when O.D. at 280nm became constant. The column was washed with MOPS buffer till the flow through O.D at 280nm measured zero. The polyclonal antibodies affinity bound to the column were eluted using 100mM Glycine-HCl (pH 2.5) and at least ten 1mL fractions were collected directly into 1.5mL tubes containing 200µL of 1M Tris-HCl (pH 8.0). Fractions were subject to SDS PAGE to determine the purity of eluted antibodies.

2.11 Histology and Immunocytolocalization

Solutions

1X PBS	10mM NaH_2PO_4-Na_2HPO_4 buffer, pH 7.2; 130mM NaCl, pH 7.2
0.5 X SSC	75mM NaCl; 7.5mM Na Citrate, pH 7.0
Color development buffer	100mM Tris, pH 9.5, 150mM NaCl, 50mM $MgCl_2$
	BCIP/NBT mix **:** 0.577mM BCIP, 0.122mM NBT (Merck, USA)
EDTA	10mM

Polyvinyl alcohol, Ethanol, Tertiary butanol, Paraffin, Xylene, Glycerol

Free hand transverse sections were fixed overnight under vacuum in freshly prepared cold 4% buffered formaldehyde (4% paraformaldehyde in 1X PBS). The sections were dehydrated by passages through increasing ethanol: water series (30%, 50%, 70%, 85%, 95% and 100% ethanol) for 30 min each. This was followed by passages through tertiary butanol: ethanol series (25:75, 50:50, 75:25, and 100:0). The sections were rehydrated by treating with 70% and 50% ethanol and 0.5X SSC for 2 min. The rehydrated sections were soaked in two changes of 1X PBS for 10 min each. Next, the sections were washed in 1X PBS containing 0.1% BSA for 5 min and subjected to 30 min of blocking with 10% BSA at room temperature in a humidified chamber. Post blocking washes included three washes of 15 min each with 1X PBS containing 0.1% BSA. Primary antibody incubation was carried out overnight in a humidified chamber at 4ºC using 75µL of diluted antibody (1µg mL^{-1}) in 1 X PBS containing 0.1% BSA. Negative controls included either the use of pre-immune serum or the omission of both antibody and pre-immune serum. Following the primary antibody incubation, the sections were washed thrice for 15 min each in 1X PBS containing 0.1% BSA. A secondary antibody, 0.2% Anti-rabbit-IgG-goat alkaline phosphate conjugate antibody (diluted in 1X PBS with 0.1% BSA), was added to the tissue sections at this stage and incubated at 37ºC in a humidified chamber for 2 h in dark. Post secondary antibody washes were carried out at room temperature using 1X PBS with 0.1% BSA. Color was developed in dark by addition of 60µL of color development reagent (20µL BCIP/NBT mix in color development buffer containing 10% polyvinyl alcohol) to the color development buffer (100mM Tris, pH 9.5, 150mM NaCl, 50mM $MgCl_2$) treated slides. The slides were placed in humidified (color development buffer) chamber at RT in dark. Upon color development, 10mM EDTA was used to stop the reaction, rinsed with water, air dried, cover

slip-mounted using glycerol and observed under microscope, and microscopic image captured.

2.12 Green Fluorescence Protein (GFP) visualization

Primary screening of putative transformed plants was done by visualizing green fluorescence at 395nm using NightSea GFP flash light (NightSea, USA). Transverse sections of leaf mid rib, shoot and root were dehydrated and rehydrated as described above (section 2.11) and visualized at 395nm under fluorescence microscope (Zeiss, Axioplan 2) and pictures captured.

2.13 Histochemical staining

Solutions

Phloroglucinol 2% in 95% Ethanol

Free hand transverse sections were prepared for histochemistry as described above. Phloroglucinol staining was done by covering pretreated slide with phloroglucinol for half an hour and then concentrated HCl was added (Harris et al., 1980; Speer et al., 1987). Lignified tissue takes red color. A cover slip mounted on the specimen using glycerol was sealed to the slide using nail enamel. The slides were viewed under a light microscope and pictures captured.

2.14 *Agrobacterium* mediated Tobacco transformation

Solutions

MS medium (Murashige and Skoog 1962)

1.	**Major Components**	**mM**
	NH_4NO_3	20.61
	KNO_3	18.75

	$CaCl_2.2H_2O$	2.99
	$MgSO_4.7H_2O$	1.5
	KH_2PO_4	1.24
2.	**Minor Components**	**mM**
	$MnSO_4$	0.147
	$ZnSO_4$	5.3×10^{-2}
	$CuSO_4$	1.56×10^{-4}
	$CoCl_2.6H_2O$	1.05×10^{-4}
	KI	4.99×10^{-3}
	H_3BO_4	0.1
	$Na_2Mo_4.2H_2O$	1.03×10^{-3}
3.	**Vitamins**	**mM**
	Myoionsitol	5.55×10^{-2}
	Nicotinic acid	4.06×10^{-3}
	Pyridoxine HCl	2.43×10^{-3}
	Thymine HCl	2.96×10^{-4}
	Glycine	2.66×10^{-2}
4.	**Iron**	**mM**
	$FeSO_4.7H_20$	0.1
	Na_2EDTA	0.1

5.	**BAP**	**1.776 mM** (Dissolve 20mg BAP in 1mL 1N NaOH and make up to 50mL in ethanol)
6.	**NAA**	**2.148 mM** (NAA-20mg $50mL^{-1}$ dH_2O)
7.	**SIM(1000mL)**	Major (40x)-25mL Minor (100x)-10mL Iron (100x)-10mL Vitamins (200x)-5mL BAP (4.4 µM))-2.5mL NAA (5.37 µM)-250µL Glucose-1.5% Sucrose-2.0% pH-5.6 to 5.8 Agar-0.8%
8.	**RSM(1000mL)**	Major(40x)-25mL Minor(100x)-10mL Iron (100x)-10mL Vitamins (200x)-5mL

		BAP(4.4µM)-2.5mL NAA(5.37µM)-250µL Glucose-1.5% Sucrose-2.0% pH-5.6 to 5.8 Acetosyringone-200µM $MgSO_4$-40mM
9.	**Selection medium**	SIM Hygromycin- 7.09µM Cefotaxime-200µM
10.	**RIM**	Major(40x)-25mL Minor(100x)-10mL Iron (100x)-10mL Vitamins (200x)-5mL NAA(5.37 µM)-250µL Glucose-1.5% Sucrose-2.0% pH-5.6 to 5.8 Agar-0.8%

SIM	Shoot Induction Medium
RSM	Resuspension Medium
RIM	Root Induction Medium

The tobacco regeneration and transformation protocol is modified from Horch et al. (1985). *Agrobacterium tumefaciens* strain GV2260 harbouring the binary plasmid vector to be transformed in tobacco was inoculated in 5mL YEP media containing rifampicin (250mg L^{-1}) and kanamycin (50mg L^{-1}). The culture was allowed to grow overnight in 28ºC with shaking at 200rpm. Next day 1mL inoculum from this tube was taken and added to 50mL YEP flask containing appropriate antibiotics. The culture was incubated at 28ºC for 24 hours with shaking at 200 rpm.

Fresh leaves were taken from axenic cultures of *Nicotiana tabacum* var. Anand 119 and ~ 6 mm diameter disc were punched out. The leaf discs were inoculated on MS medium agar plates containing 4.4 µM BAP and 5.37µM NAA (MSBN) and kept for 48 h in dark.

A. tumefaciens GV2260 cells grown for 24 h were centrifuged for 10 min at 5000 x g. Cells were suspended in 20mL of 50mM $MgSO_4$ and 200µM Acetosyrigone and allowed to stand for 10 min with intermittent shaking. The tobacco leaf discs from axenic tobacco cultures were co-cultivated in this suspension for 10 min with intermittent shaking. The leaf discs were then transferred to MSBN plates without any antibiotics. The adaxial side of the leaf disc was in contact with the medium. The plates were incubated in dark at 28°C. After two days the leaf discs were harvested, washed with liquid MSBN and dried on sterile filter paper. Up to 10 leaf discs per plate were transferred to MSBN agar plates containing Hygromycin (7.09µM) and Cefotaxime (200µM). The cultures were incubated at 28ºC with 18/6 hours photoperiod in diffused light (60-80 mE/m.s.) The leaf discs were subcultured every week till elongated shoots were excisable. Elongated shoots were excised and transferred to root induction medium. The rooted plants were hardened and further analyzed.

2.15 References

Abravaya K, Huff J, Marshall R (2003) Molecular beacons as diagnostic tools: technology and applications. Clin Chem Lab Med 41: 468-474

An G (1987) Binary ti vectors for plant transformation and promoter analysis. Meth Enzymol 153: 292-305

Arnheim N, Erlich H (1992) Polymerase Chain Reaction Strategy. Annu Rev Biochem 61: 131-156

Blumberg DD (1987) Creating a ribonuclease-free environment. Meth. Enzymol 152: 20-24

Borson, ND, Sato WL, Drewes LR (1992) A lock-docking oligo(dT) primer for 5' and 3' RACE PCR. PCR Methods Appl 2: 144–148

Chumakov KM (1994) Reverse transcriptase can inhibit PCR and stimulate primer dimer formation. PCR Methods Appl 4: 62-64

Eckert KT, Kunkel TA (1990) High fidelity DNA synthesis by the Thermus aquaticus DNA polymerase. Nucleic Acids Res 18: 3739- 3744

Feinbeng AP, Vogelstein B (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 132: 6-13

Feinbeng AP, Vogelstein B (1984) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Addendum. Anal Biochem 137: 266-267

Freeman WM, Walker SJ, Vrana KE (1999) Quantitative RT-PCR: pitfalls and potential. Biotechniques 26: 112-122, 124-125.

Fucharoen S, Fucharoen G, Fucharoen P, Fukumaki Y (1989) A novel ochre mutation in the beta-thalassemia gene of a Thai. Identification by direct cloning of the entire beta-globin gene amplified using polymerase chain reactions. J Biol Chem 264: 7780-7783

Goodenow M, Huet T, Saurin W, Kwok S, Sninsky J, Wain-Hobson S (1989) HIV-1 isolates are rapidly evolving quasi species: evidence for viral mixtures and preferred nucleotide substitutions. J Acquired Immunol Defic Syndr 2: 344-352

Harris PJ, Kelderman MR, Kendon MF and McKenzie RJ (1997) Monosaccharide compositions of unlignified cell walls of monocotyledons in relation to the occurrence of wall-bound ferulic acid. Biochem Syst Ecol 25: 167-179

Heid CA, Stevens J, Livak KJ, Williams PM (1996) Real time quantitative PCR. Genome Res 6: 986-994

Hengen P (1995) Purification of His-Tag fusion proteins from *Escherichia coli*. Trends Biochem Sci 20: 285-286

Higuchi R, Dollinger G, Walsh PS, and Griffith R (1992) Simultaneous amplification and detection of specific DNA sequences. Biotechnology 10: 413–417

Higuchi R, Fockler C, Dollinger G, Watson R (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. Biotechnology 11: 1026–1030

Horch R, Fry J, Hoffmann N, Eichholtz P, Rogers R, Fratey T (1985) A simple method for transferring genes into plants. Science 227: 1229–1231

Kellogg DE, Rybalkin I, Chen S, Mukhamedova N, Vlasik T, Siebert P, Chenchik A (1994) TaqStart Antibody: Hot start PCR facilitated by a neutralizing monoclonal antibody directed against Taq DNA polymerase. BioTechniques 16: 1134–1137

Kimmel AR, Berger SL (1987) Preparation of cDNA and the generation of cDNA libraries: Overview. Meth Enzymol 152: 307-316

Kwok S, Kellogg DE, McKinney N, Spasic D, Goda L (1990) Effects of primer-template mismatches on the polymerase chain reaction: Human immunodeficiency virus type 1 model studies. Nucleic Acids Res 18: 999-1005

Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685

Lee LG, Connell CR, Bloch W (1993) Allelic discrimination by nick-translation PCR with fluorogenic probes. Nucleic Acids Res 21: 3761-3766

Livak KJ, Flood SJ, Marmaro J, Giusti W, Deetz K (1995) Oligonucleotides with fluorescent dyes at opposite ends provide a quenched probe system useful for detecting PCR product and nucleic acid hybridization. PCR Methods Appl 4: 357-362

Lodhi MA, Guang-Ning Ye, Norman FW, Bruce IR (1994) A simple and efficient method for DNA extraction from grapevine cultivars, *Vitis* species and *Ampelopsis*. Plant Mol Biol Reporter 12: 6-13

Mendelman LV, Boosalis MS, Petruska J, Goodman MF (1989) Nearest neighbor influences on DNA polymerase insertion fidelity. J Biol Chem 264:14415-14423

Mhlanga MM, Malmberg L (2001) Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR. Methods 25: 463-471.

Mitchell P (2001) Microfluidics– downsizing large-scale biology. Nat Biotechnol 19: 717–721

Mullis KB (1990) The unusual origin of the polymerase chain reaction. Sci Am 262: 56–61

Mullis KB, Faloona FA (1987) Specific synthesis of DNA in vitro via a polymerase catalysed reaction. Methods Enzymol 255: 335-350.

Murashige T, Skoog F (1962) A revised medium for rapid growth and bioassays with tobacco tissue cultures. Physiol Plant 15: 473-497

Okayama H, Berg P (1982) High-efficiency cloning of full-length cDNA. Mol Cell Biol 2: 161- 170

Raeymaekers L (2000) Basic principles of quantitative PCR. Mol Biotechnol 15: 115-122

Reischl U, Wittwer CT, Cockerill F (2002) Rapid Cycle Real-time PCR: Methods and Applications; Microbiology and Food Analysis. New York: Springer-Verlag

Saha BK, Tian B, Bucy RP (2001) Quantitation of HIV-1 by real-time PCR with a unique fluorogenic probe. J Virol Methods 93: 33-42

Saiki R, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, Arnheim N (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science 230: 1350-1354

Saiki RK, Gelfand DH, Stoffel S, Scharf S, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487-491

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning; A Laboratory Manual, 2nd ed., New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Scharf SJ, Horn GT, Erlich HA (1986) Direct cloning and sequence analysis of enzymatically amplified genomic sequences. Science 233: 1076-1078

Sellner LN, Coelen RJ, MacKenzie JS (1992) Reverse transcriptase inhibits *Taq* polymerase activity. Nucleic Acids Res 20: 1487-1490

Siebert PD, Chenchik A, Kellogg DE, Lukyanov KA, Lukyanov SA (1995) An improved method for walking in uncloned genomic DNA. Nucleic Acids Res 23: 1087–1088

Solinas A, Brown LJ, McKeen C, Mellor JM, Nicol J, Thelwell N, Brown T (2001) Duplex Scorpion primers in SNP analysis and FRET applications. Nucleic Acids Res 29 : E96

Speer EO (1987) A method of retaining phloroglucinol proof of lignin. Stain Technol 62: 279–280

Tan W, Wang K, Drake TJ (2004) Molecular beacons. Curr Opin Chem Biol 8: 547-553

Terry CF, Shanahan DJ, Ballam LD, Harris N, McDowell DG, Parkes HC (2002) Real-time detection of genetically modified soya using Lightcycler and ABI 7700 platforms with TaqMan, Scorpion, and SYBR Green I chemistries. J AOAC Int 85: 938-944

Todd JA, Bell JI, McDevitt HO (1987) HLA-DQ βgene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. Nature 329: 599-604

van der Velden VH, Hochhaus A, Cazzaniga G, Szczepanski T, Gabert J, van Dongen JJ (2003) Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. Leukemia 17: 1013-1034

Vet JA, Marras SA (2005) Design and optimization of molecular beacon real-time polymerase chain reaction assays. Methods Mol Biol 288: 273-290

Vet JA, Van der Rijt BJ, Blom HJ (2002) Molecular beacons: colorful analysis of nucleic acids. Expert Rev Mol Diagn 2: 77-86

Wittwer CT, Herrmann MG, Moss AA, Rasmussen RP (1997) Continuous fluorescence monitoring of rapid cycle DNA amplification. Biotechniques 22: 130-131, 134-138

Wittwer CT, Ririe KM, Andrew RV, David DA, Gundry RA, Balis UJ (1997) The LightCycler: a microvolume multisample fluorimeter with rapid temperature control. Biotechniques 22: 176-81

Zhu YY, Machleder EM, Chenchik A, Li R & Siebert PM (2001) Reverse transcriptase template switching: A SMART™ approach for full-length cDNA library construction. Biotechniques 30: 892–897

3.1 Introduction

S-adenosyl-L-methionine methyltransferases (EC 2.1.1. -) are key enzymes in the phenylpropanoid, flavanoid and many other metabolic pathways in plants. The enzymes Caffeate 3-O methyltransferase (COMT; EC 2.1.1.68) and Caffeoyl CoA 3-O methyltransferase (CCoAOMT; EC 2.1.1.104) control the degree of methoxylation at the C3 and C5 positions of the aromatic ring of caffeic and 5-hydroxyferulic acids, and caffeoyl and 5-hydroxyferuloyl CoAs respectively, resulting in *p*- hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units of lignin (Higuchi, 1990; Ralph et al., 1998; Boerjan et al., 2003). COMT was long considered as the only methylating enzyme involved in lignification (Lewis and Yanamoto, 1990). Later evidences, however, suggest O-methylation of lignin precursors may also occur at the level of hydroxycinnamoyl-CoA-ester and specific O-methyltransferase *i.e* CCoAOMT converts caffeoyl CoA into feruloyl CoA and 5-hydroxyferuloyl CoA into sinapyl CoA (Martz et al*.,* 1998).

To date many highly conserved CCoAOMT cDNAs (with respect to their deduced amino acid sequences) and a few genomic DNA clones have been isolated from angiosperm and gymnosperm species. For example, it has been isolated from zinnia (Ye et al., 1994), aspen (Meng and Campbell, 1995), grape vine (Busam et al., 1997), tobacco (Busam et al., 1997; Martz et al., 1998), alfalfa (Inoue et al., 1998), eucalyptus (De Melis et al., 1998), pine (Li et al., 1999), maize (Laura et al., 1999) etc. Different strategies have been employed to isolate these gene clones. These included PCR and reverse transcriptase PCR (Li et al., 1999), screening of genomic (Grimming and Matern, 1997) and cDNA libraries (Li et al., 1997; Martz et al., 1998).

In the present study the genomic and the cDNA gene clones of CCoAOMT were isolated from *Leucaena leucocephala.* This leguminous tree is used extensively as raw source in India by the pulp and paper industry.

3.2 Materials and Methods

3.2.1 Genomic DNA extraction and PCR

Genomic DNA was extracted from young and disease free leaves of *L. leucocephala* (Chapter 2; section 2.4.3). PCR reactions were set as described earlier (Chapter 2; section 2.4.8).

3.2.2 Slot Blot, Southern hybridization and random primer labeling

Slot blot, Southern blot and random primer labeling were done as described earlier (Chapter 2; sections 2.5.1, 2.5.2, 2.6 and 2.7 respectively).

3.2.3 Bioinformatic analysis

Nucleotide and amino acid sequence analysis was done using software **pDRAW 32** and online bioinformatics analysis facility available at **www.justbio.com**, **www.expasy.org** and **www.ncbi.nlm.nih.gov**.

3.2.4 Genome walking, cDNA synthesis, 5' and 3' RACE

Genome walking**,** cDNA synthesis, 5' and 3' RACE have been described earlier (Chapter 2; sections 2.4.10, 2.4.7 and 2.4.9 respectively).

3.3 Results and Discussion

3.3.1 Establishing the presence of the CCoAOMT gene in *L. leucocephala*

The presence of Caffeoyl CoA 3-O methyltransferase (CCoAOMT) gene in *Leucaena leucocephala* was established by slot blot analysis and Southern hybridization. Genomic DNA of good integration was isolated from the leaves of *L. leucocephala* (Fig. 3.1)

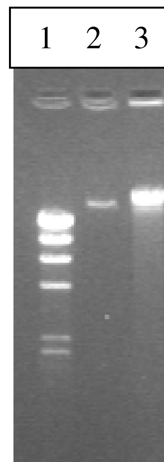


Fig. 3.1: Genomic DNA of *L.leucocephala*: λ DNA *Hind* III digest (lane 1), undigested λ phage DNA (lane 2), *L. leucocephala* genomic DNA (lane 3).

3.3.2 Slot Blot Hybridization

The genome size of *L. leucocephala* is 1.18 x 10^9 bp and 1.2 pg of the genomic DNA correspond to a single representation of the genome. *L. leucocephala* genomic DNA, 1.2µg representing the genome one million times was spotted on Hybond N^+ membrane. The CCoAOMT cDNA gene clone from alfalfa (996 bp; NCBI GenBank Accession number U20736, a kind gift from Dr. R. A. Dixon, Oklahoma, USA) was also spotted. The 1.1pg DNA, representing one million copies of the gene clone was spotted in triplicate to represent these gene copies once (1.1pg), twice (2.2pg) and three (3.3pg) times. The blot was probed with the radiolabelled alfalfa CCoAOMT gene. Positive signal obtained under high stringency hybridization conditions was indicative of the presence of the CCoAOMT gene in *L. leucocephala.* Based on signal intensity it was also inferred that in *L. leucocephala* CCoAOMT belonged to a gene family represented possibly by three members (Fig 3.2).

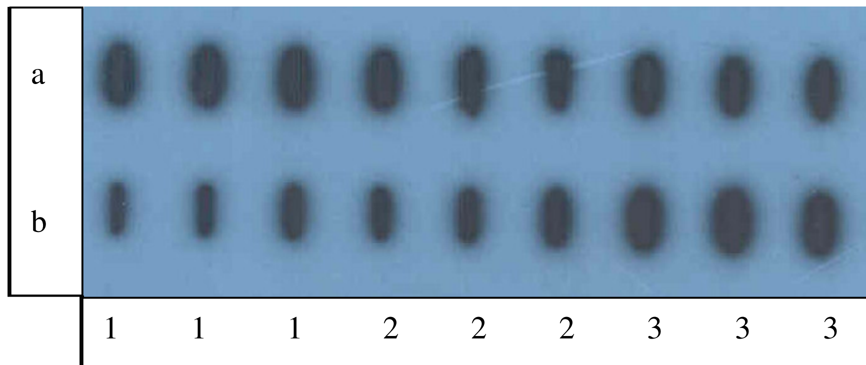


Fig. 3.2: Slot Blot of *L. leucocephala* genomic DNA hybridized with radiolabelled alfalfa CCoAOMT gene (U20736): (a) signals from 10^6 genomic DNA representations of *L. leucocephala*, (b) signals from 10^6 x 1, 2 and 3 copies of alfalfa CCoAOMT gene clone (each blotted in triplicate).

3.3.3 Southern Hybridization

To further validate the results from slot blot experiment and to understand the distribution of the CCoAOMT gene in the *L leucocephala* genome, Southern hybridization was performed. A 40µg aliquot of *L. leucocephala* genomic DNA was restriction digested with *BamH* I, *EcoR* I, *Hind* III and *Sac* I and electrophoresed on 0.7% agarose gel in 1XTAE buffer (Fig.

3.3 a). Southern hybridization with alfalfa CCoAOMT gene (U20736) revealed that from each of the restriction digested DNA samples multiple bands hybridized to the alfalfa CCoAOMT gene clone (Fig. 3.3 b). The hybridization signals from very low to very high molecular weight DNA fragments indicated possible multiple gene copies in the *L. leucocephala* genome. Low molecular weight signals from *Hind* III and *Sac* I digests suggested the presence of these restriction sites within the gene(s). As with the CCoAOMT gene family in *L. leucocephala* multiple CCoAOMT gene copies have been reported earlier: two copies in poplar (Chen et al., 2000), four copies in tobacco (Martz et al., 1998), two copies in maize (Civardi et al., 1999) etc.

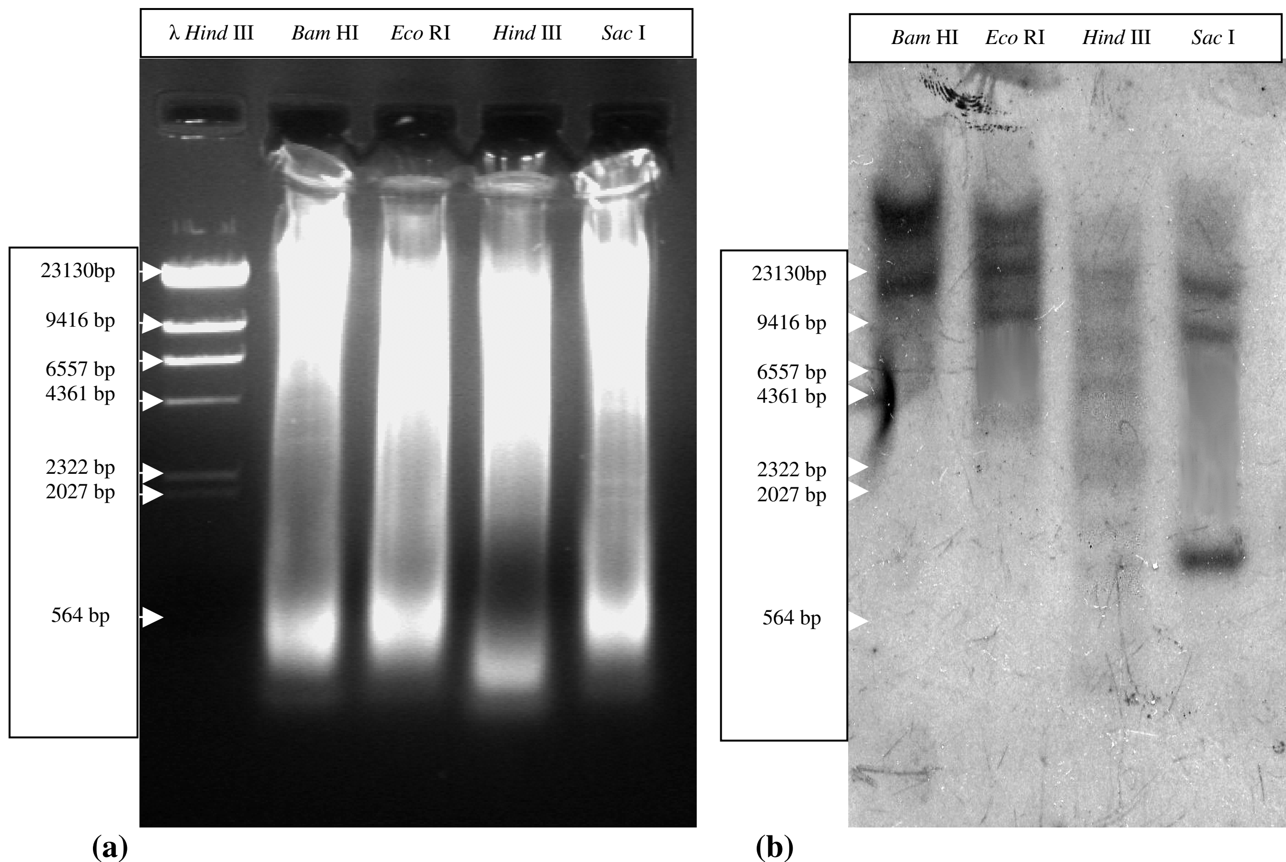


Fig 3.3: (a) 0.7% agarose gel showing *L. leucocephala* genomic DNA digestion with *BamH* I, *EcoR* I, *Hind* III and *Sac* I. λ phage DNA *Hind* III digest used as DNA size marker; (b) Southern hybridization of (a) with alfalfa CCoAOMT gene.

3.3.4 Isolation of the genomic Caffeoyl CoA 3-O methyltransferase (CCoAOMT) gene

A PCR based approach was adopted to isolate the genomic clone(s) of CCoAOMT. Selected nucleotide sequences of CCoAOMT genes available from NCBI GenBank database were aligned using multiple sequence alignment program Clustal W 1.8 (www.justbio.com). Nine sequences with high degree of sequence similarity were selected for the design and synthesis of sets of three forward and three reverse primers (Fig. 3.4).

Multiple sequence alignment for forward primers

```
AJ130841   -------------------------------------------------------------------
AF240466   ---CAAAG-CCAGGCAGGAAGGCACCAGGAAGTTGGCCACAAGAGCCTTTTGCAAAGTGATGC---
AF327458   ---CAAAG-CCAGGCAGGAAGGCACCAGGAAGTTGGCCACAAGAGCCTTTTGCAAAGTGATGC---
AY057554   ---CAGTC-TCAGAATCTTCGACATCAAGAAGTTGGTCACAAGAGTCTCTTACAGAGCGATGA---
AF022775   ---CAGAG-AACGGAA-TTAAACACCAAGAGGTTGGCCACAAAAGCCTTTTGCAAAGTGATGC---
U62736     ---TGGAG---AAAATGGAAGACATCAAGAAGTTGGACACAAGAGTCTTTTGCAAAGTGATGC---
U38612     -------------AATGGAAGACATCAAGAAGTTGGACACAAGAGTCTTTTGCAAAGTGATGC---
U62734     ---TGGAG---AAAATGGAAGACATCAAGAAGTTGGACACAAGAGTCTCTTGCAAAGTGATGC---
U13151     ---TGAAACTCAACCTGCTAAACACCAAGAAGTTGGTCACAAAAGCCTCCTTCAAAGTGATGC---
                                   * ** ** ***** ***** *
```

Forward primer set

CCo F1: 5′(d) ACCAGGAAGTTGGCCACAAGA 3′

CCo F2: 5′(d) ATCAAGAAGTTGGTCACAAGA 3′

CCo F3: 5′(d) ACCAAGAGGTTGGCCACAAAA 3′

Multiple sequence alignment for design of reverse primers

```
AJ130841   ---GATTGAAATTTGCATGCTTCCTGTTGGTGATGGCATCACTCTCTGCCGTCGGATCCAATGA
AF240466   ---GATTGAAATTTGCATGCTTCCTGTTGGTGATGGCATCACTCTCTGCCGTCGGATCCAATGA
AF327458   ---GATTGAAATTTGCATGCTTCCTGTTGGTGATGGCATCACTCTCTGCCGTCGGATCCAATGA
AY057554   ---GATCGAGATCTGTATGCTCCCTGTTGGTGATGGAATCACTATCTGCCGTCGGATCAGTTGA
AF022775   ---GATTGAGATTTGCATGCTACCCGTTGGTGATGGCATTACCTTGTGCCGCCGCATCACCTGA
U62736     ---AATTGAAATTTGTCAGCTTCCCGTTGGTGATGGCATCACCCTCTGCCGCCGCATCAGTTAA
U38612     ---AATCGAAATTTGCCAGCTACCTGTTGGTGACGGCATCACCCTTTGCCGCCGCATTAGTTAA
U62734     ---AATTGAAATCTGTCAGCTTCCCGTTGGCGATGGCATCACCCTTTGCCGACGCATTAGTTAA
U13151     ---AGTGGAGATCTGTCAGCTTCCGGTCGGTGATGGAATCACTTTGTGTCGCCGCATAAGCTAA
                                                       * ** ** ** **    * *
```

Reverse primer set

```
CCo R1: 5′(d) TCATTGGATCCGACGGCAGA 3′
CCo R2: 5′(d) TTAACTGATGCGGCGGCAGA 3′
CCo R3: 5′(d) TTAACTAATGCGTCGGCAAA 3′
```

Fig. 3.4: Clustal W 1.8 multiple sequence alignment of the CCoAOMT gene sequences: *Populus balsamifera* subsp. trichocarpa (AJ130841), *Populus tomentosa* (AF240466), *Populus alba* x *Populus glandulosa* (AF327458), *Arabidopsis thaliana* (AY057554), *Nicotiana tabacum* (AF022775), *Nicotiana tabacum* (U62736), *Nicotiana tabacum* (U38612), *Nicotiana tabacum* (U62734), *Zinnia elegans* (U13151). The region selected for designing primers are in bold letters. Only the 5' and the 3' regions of the gene coding sequences are shown.

PCR reactions were set up with *L. leucocephala* genomic DNA using all possible combinations of the forward and the reverse primers.

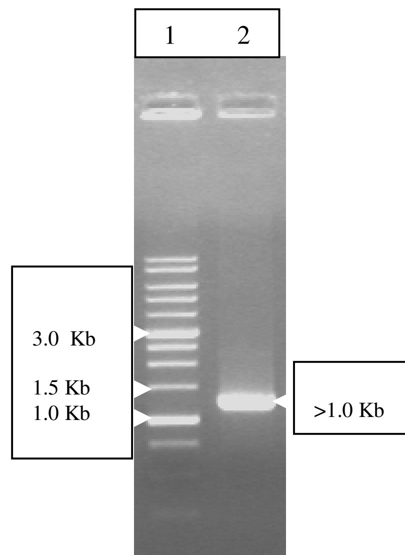
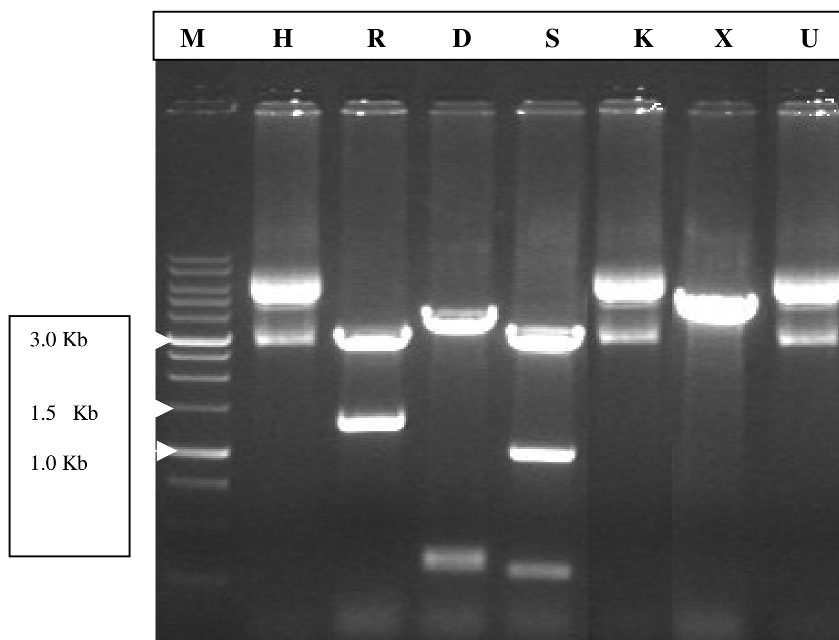
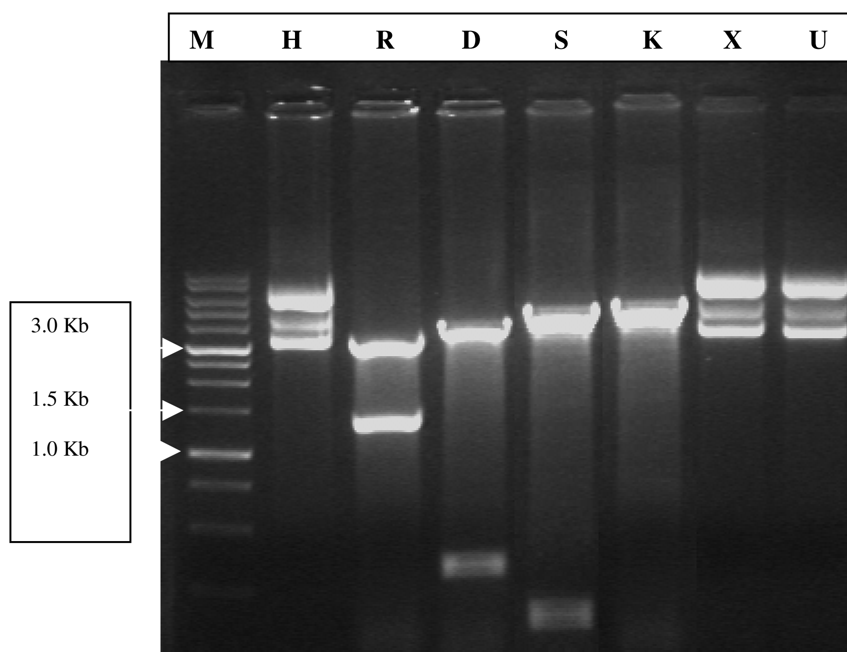


Fig. 3.5: PCR with *L. leucocephala* genomic DNA template and primers Cco F2 and Cco R3. DNA size marker (lane 1), PCR amplified product (lane 2).

A DNA fragment of > 1.0 Kb was amplified (Fig. 3.5) with forward primer **Cco F2 [5' (d) ATCAAGAAGTTGGTCACAAGA3']** and reverse primer **Cco R3 [5'(d)TTAACTAATGCGTCGGCAAA3'].** The amplicon was gel eluted and cloned in pGEM-T Easy Vector (Promega, USA). Eight recombinant plasmid DNAs were restriction digested with *BamH* I, *EcoR* I, *Hind* III, *Sac* I, *Kpn* I and *Xba* I. The digestion patterns revealed presence of two type of inserts (Fig. 3.6 a, b). In one recombinant a ~1Kb and an ~300 bp fragments were released, whereas in the other only the ~300 bp *Sac* I fragment was released (this enzyme also cuts in the MCS of the pGEM-T Easy vector). This suggests that two different amplicons have been cloned which were of same molecular weight. The two clones were designated as gCcoF2R31 and gCcoF2R32. No restriction site was noticed for *EcoR* I and *BamH* I. In both the clones two *Hind* III restriction sites result in the release of an ~300bp fragment. *Xba* I cuts once in the gCcoF2R31 but not in the gCcoF2R32 clone. *Kpn* I cuts once in the gCcoF2R32 but not in the gCcoF2R31 clone.



(a)



(b)

Fig. 3.6: Restriction digestion of gCcoF2R31 (a) and gCcoF2R32 (b). DNA size marker (M), restriction digestion with enzymes *BamH* I (H), *EcoR* I (R), *Hind* III (D), *Sac* I (S), *Kpn* I (K) and *Xba* I (X), uncut plasmid (U).

The **gCcoF2R31** and **gCcoF2R32** clones were sequenced bi-directionally and each clone was of 1258 bp (Fig. 3.7 a, b). The Integrated System of Michael Zhang's (1997) MZEF and Thanaraj's (2000, 2003) SpliceProximalCheck was used to identify the intron / exon junctions in the two clones. The **gCcoF2R31** and **gCcoF2R32** nucleotide sequences show presence of five putative exons and four putative introns. In the **gCcoF2R31** clone the five putative exons span nucleotide positions 1 - 50, 162 - 241, 483 - 627, 729 - 860 and 965 - 1258 (bold letters in Fig. 3.7 a). The putative introns spanned nucleotide positions 51 - 161, 242 - 482, 628 - 728 and 861 - 964 (lower case italics in Fig. 3.7 a). In the gCcoF2R32 clone the five putative exons spanned nucleotide positions 1 - 50, 177 - 256, 427 - 591, 708 - 839 and 965 - 1258 (bold letters in Fig. 3.7 b). The putative introns in this sequence spanned nucleotide positions 51 - 176, 257 - 426, 592 - 707 and 840 - 964 (lower case italics in Fig. 3.7 b). The first exon, however, was incomplete in both the sequences.

gCcoF2R31

CcoF2

1 ATCAAGAAGT TGGTCACAAG AGCCTTTTGC AGAGTGATGC TCTCTACCAG *gtactcctat*
61 *ttctctttga tcacgccaca catcaatccc acaaccacaa ccacaaccac tttcttcttc*
121 *ttcttcttct ggtctaaagg atcatatgat tttgtctgca g*TACATTCTA GAGACCAGTG
181 TCTACCCCAG AGAACCTGAA CCCATGAAGG AGCTCAGAGA AATCACAGCC AAACACCCTT
241 G*gtaagcttc ttcttaggac tcaaacatta tgaagttttc agcaaattac atcataatgc*
301 *tgtcgtgccc tatgtccgta atgcttaaga ttcagatcct ctgtttttgt gcttgctttg*
361 *agtttgtcct tcatccagac aaaaaatttg gagactttaa acaaaataga ttaggataag*
421 *aaaaattgcg tgtgtgatta tgggaagatg gtccattact gattgatgga ggaaaacaac*
481 *ag*GAACATCA TGACAACCTC AGCAGACGAG GGGCAATTCT TGAACATGCT CCTTAAGCTT
541 ATCAATGCTA AGAATACCAT GGAGATTGGT GTTTACACTG GCTACTCCCT GCTTGCCACT
601 GCTCTGGCTC TCCCTGAAGA TGGAAAG*gta agggccaatc aacaacagag tccaatgttc*
661 *tgaatgtttc catgttactt tatcaagatc atcgaaacat atctgagatt tggtcggaat*
721 *tgttgcag*AT ACTGGCCATG GACATTAACA GAGAAAACTA CGAGTTGGGT CTTCCGGTAA
781 TTCAGAAAGC TGGTGTTGCC CACAAAATTG AGTTCAGAGA GGGCCCTGCT CTCCCTGTTC
841 TTGACGAACT CGTTAAAGAT *gtaagctttt tctatctcca tcttatcatt acaactttct*
901 *ggtcttatct ccattataaa catcaaattg ataagagagg tatgattgtg ggggtggttt*
961 *acag*GAAAAG AATCATGGGA GCTATGATTT CATATTCGTG GATGCTGACA AGGACAACTA
1021 CTTGAATTAT CATAAGAGGT TGATCGATTT GGTGAAGGTA GGAGGAGTGA TCGGGTACGA
1081 TAACACCTTG TGGAACGGAT CTGTAGTGGC ACCACCAGAT GCTCCTCTGA GGAAGTACGT
1141 GAGGTATTAC AGGGACTTTG TGCTGGAGCT CAGCAAGGCT TTGGCTGTGG ACCCTAGGAT
1201 CGAGATCTGC ATGCTTCCTA TTGGTGATGG CATTACCCTT TGCCGACGCA TTAGTTAA

(a) CcoR3

gCcoF2R32

CcoF2

1 ATCAAGAAGT TGGTCACAAG AGCCTTTTGC AGAGTGATGC TCTCTACCAG *gtaccccaat*
61 *ttctctttta tcacaccaca catcatcaat cccacaaccg caaccagttt cttcttcttc*
121 *ttcttcttct tcctgttctt gtcctggtct aaaggatcat gtgattttgt atgcag*TACA
181 TTCTGGAGAC TAGTGTCTAC CCCAGAGAAC CTGAACCCAT GAAGGAGATC AGAGGAATCA
241 CAGCCAAACA CCCTTG*gtaa agcttctttt taggacgaca tagtgaaatt ttctgcaaat*
301 *cttttttttt ttgcttgctt tcggtttgtc cttcctcaag acgaaaaatc gagaaacttt*
361 *aaacaaaatc gatttggata cgaaagattg tccgtgtgtg attatgagtg gatgatccat*
421 *tactgattga aggaagaaaa caacag*GAAC ATCATGACAA CCTCAGCAGA CGAGGGGCAA
481 TTCTTGAACA TGCTCCTTAA GCTTATTAAT GCTAAGAACA CCATGGAGAT TGGTGTCTAC
541 ACTGGCTACT CCCTGCTTGC AACTGCTCTG GCTCTCCCTG AAGATGGAAA G*gtaacgacc*
601 *aatcaacaac agagtccaat gttctgaatg attcaatgtt acttttatca agatcctcca*
661 *aacatgtctg aaatgttttc gtattctttg gttgggaatt ggcgtag*ATA CTGGCCATGG
721 ACATTAACAG AGAAAACTAC GAATTGGGTC TTCCGGTAAT TCAGAAAGCT GGTGTTGCCC
781 ACAAAATTGA GTTCAAAGAG GGCCCTGCTC TTCCTGTTCT TGACGAACTC GTTAAAGAT*g*

```
841  taagcttttc tatctccttc ttgtcttcac cacttctttt tggtcttttc tccattatca
901  ctagttttta ggccaaacga acatcaaatt gacaagagtc gtatgtttgt ggggtgaata
961  gcagGAAAAG AATCATGGGA GCTATGATTT CATATTCGTG GATGCTGACA AGGACAACTA
1021 TTTGAATTAT CACAAGAGGT TGATCGATTT GGTGAAGGTA GGAGGAGTGA TAGGGTACGA
1081 TAATACCTTG TGGAATGGGT CTGTGGTGGC ACCACCAGAT GCTCCTCTGA GGAAGTACGT
1141 GAGGTATTAC CGGGACTTTG TGTTGGAGCT CAACAAGGCT TTGGCTGTGG ACCCTCGGAT
1201 TGAGATCTGC ATGCTTCCTA TTGGTGATGG CATTACCCTT TGCCGACGCA TTAGTTAA
                                          <-------------------------
                                                         CcoR3
```

(b)

Fig. 3.7: Nucleotide sequence of: (a) gCcoF2R31 and (b) gCcoF2R32. The sequence letters in shaded lower case italics are introns. Sequence letters in bold capitals are exons. Forward and reverse primers are indicated by arrows.

The nucleotide sequences of gCcoF2R31 and gCcoF2R32 genomic clones were searched for sequence similarity in NCBI GenBank database and they showed sequence similarity with other reported plant CCoAOMT genomic clones. The above two nucleotide sequences showed 72% and 79% sequence similarity with *Vitis vinifera* (AM478701) and *Petroselinum crispum* (Z54183) genomic clones.

Restriction analysis of the two genomic DNA clones gCcoF2R31 and gCcoF2R32 was done using bioinformatic software pDRAW32 (Fig. 3.8 a, b). Analysis was limited to enzyme site cutting not more than four times in the sequence. The sequence analysis data was in conformity with the restriction analysis of the amplicons as described above.

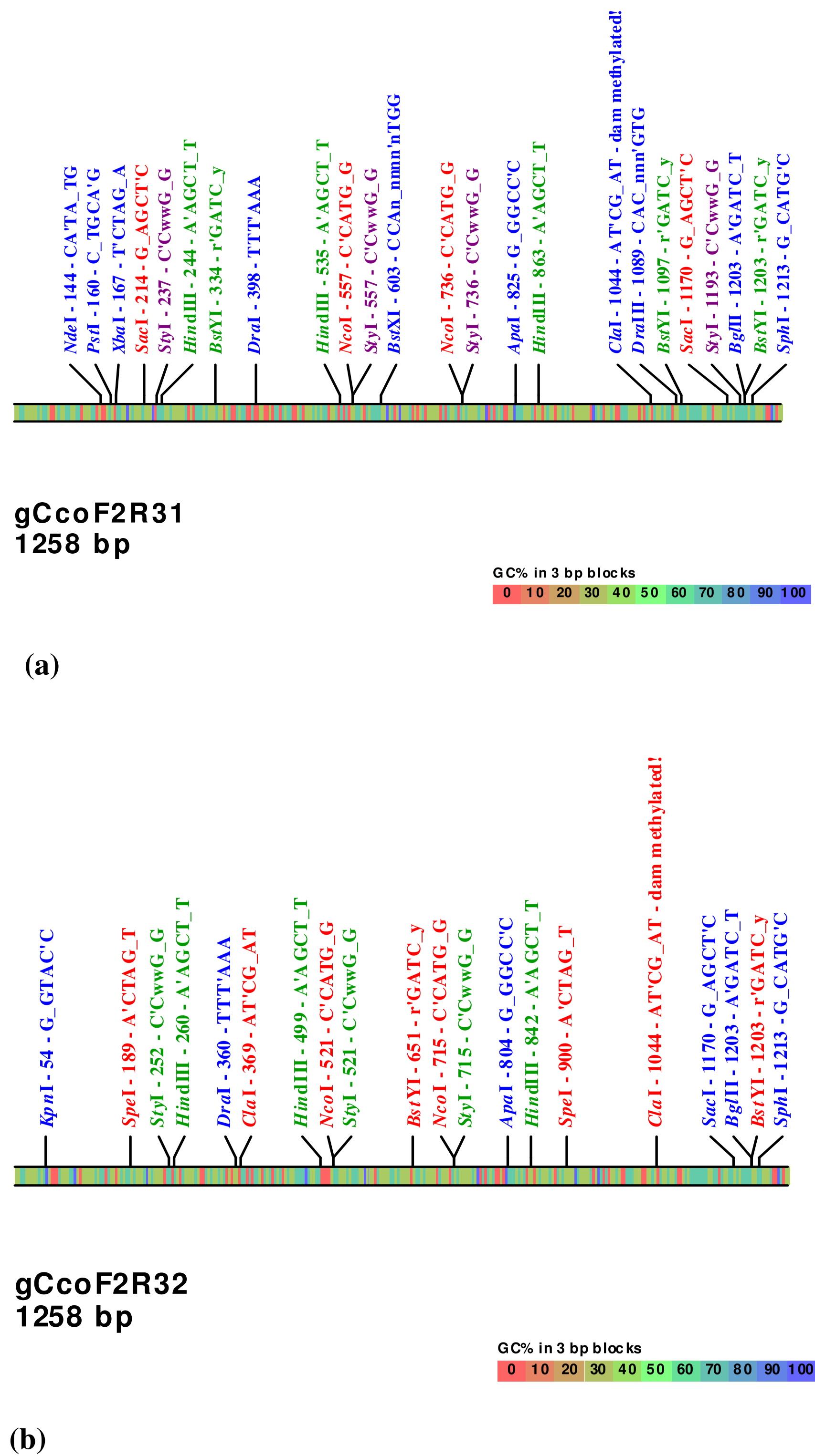


Fig. 3.8: Restriction analysis of genomic DNA CCoAOMT clone gCcoF2R31 (a) and gCcoF2R31 (b) using pDRAW32.

3.3.5 Genome Walking

From the preceding it was obvious that the amplified genomic DNA clones of Caffeoyl CoA 3-O methyltransferase (CCoAOMT) from *L. leucocephala* were partial sequences. The 5' end of the two gene clones were missing. In order to fish out the full-length genes, genome walking was resorted to. The nucleotide sequence spanning the two 5' exons of g**CcoF2R31** and g**CcoF2R32** genes were aligned to design primers for 5' upstream. Similarly the last exon at the 3' end of gCcoF2R31 and gCcoF2R32 were aligned to design primers for 3' downstream Genome walking. Genome walking was performed using Universal Genome Walking Kit (BD Biosciences, Clontech, USA) as described earlier (Chapter 2; section 2.4.10).

3.3.5.1 5' upstream walking

Two reverse primers, **N1: 5' (d) CTGGTAGAGAGCATCACTCTGCAAAAGGC 3'** from the first incomplete exon and **N2: 5'(d) TGGGTTCAGGTTCTCTGGGGTAG ACACT 3'** from the second exon were designed (Fig. 3.13) and synthesized.

```
      N1
gCcoF2R31     ATCAAGAAGTTGGTCACAAGAGCCTTTTGCAGAGTGATGCTCTCTACCAGgtactcctat
gCcoF2R32     ATCAAGAAGTTGGTCACAAGAGCCTTTTGCAGAGTGATGCTCTCTACCAGgtaccccaat
              ****************************************************** ** **

gCcoF2R31     ttctctttgatcacgccacacatca---atcccacaaccacaaccacaaccactttcttc
gCcoF2R32     ttctcttttatcacaccacacatcatcaatcccacaaccgcaaccagtttcttcttcttc
              ******** ***** **********   *********** ******    *   ******

gCcoF2R31     ttcttcttcttc------------tggtctaaaggatcatatgattttgtctgcagTACA
gCcoF2R32     ttcttcttcttcctgttcttgtcctggtctaaaggatcatgtgattttgtatgcagTACA
              ************            **************** ********* *********

                                       N2
gCcoF2R31     TTCTAGAGACCAGTGTCTACCCCAGAGAACCTGAACCCATGAAGGAGCTCAGAGAAATCA
gCcoF2R32     TTCTGGAGACTAGTGTCTACCCCAGAGAACCTGAACCCATGAAGGAGATCAGAGGAATCA
              **** ***** ************************************ ****** *****

gCcoF2R31     CAGCCAAACACCCTTG
gCcoF2R32     CAGCCAAACACCCTTG
              ****************
```

Fig. 3.13: Nucleotide sequence alignment of gCcoF2R31 and gCcoF2R32 (partial first exon, first intron and complete second intron) showing regions from where primers for 5' upstream genome walking were designed.

The **N2** reverse primer and **AP1** forward primer from the Genome walking kit were used to perform primary PCR (Fig. 3.14 a) using the *Dra* I, *EcoR* V, *Pvu* II and *Stu* I libraries

(Chapter 2; section 2.4.10) constructed from *L. leucocephala* gDNA. Multiple bands were amplified from all the four libraries.

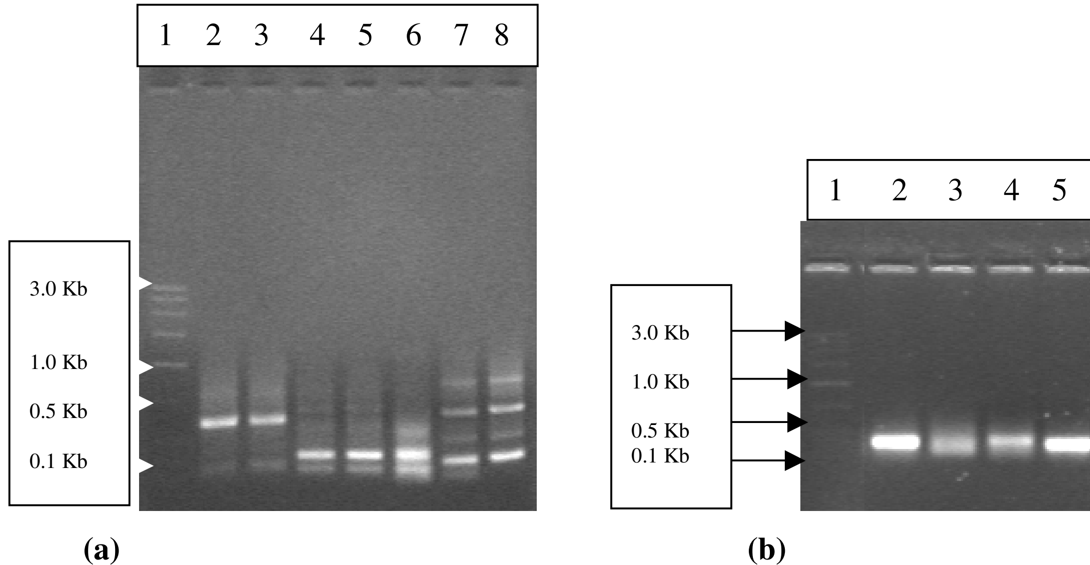


Fig. 3.14: (a) Primary PCR with *Dra* I (lanes 2, 3), *EcoR* V (lanes 4, 5), *Pvu* II (lane 6) and *Stu* I libraries (lanes 7, 8) of *L.leucocephala*. DNA size marker (lane 1); (b) Secondary PCR. DNA size marker (lane 1), amplification from primary PCR products of *Dra* I, *EcoR* V, *Pvu* II and *Stu* I libraries (lanes 2 – 5) respectively.

Secondary PCR was done as per guidelines of the Genome walker kit manufacturer. **N1** was used as the reverse primer and **AP2** from the Genome walking kit as the forward primer. Amplicons of approximately 400 bp were amplified (Fig. 3.14 b). All amplicons from the secondary PCR were gel purified, cloned in pGEM-T Easy Vector (Promega) and six recombinant clones from each of the library were sequenced. Two clones, designated **gMN5'UTR1** and **gMN5'UTR2** of 397 bp and 403 bp respectively were identified. Both clones showed the presence of the initiation codon **ATG** and the reverse primer **N1.**

The **gMN5'UTR1 and gMN5'UTR2** nucleotide sequences were aligned. The **gMN5'UTR1** started 313bp upstream of the translation initiation codon (ATG). The **gMN5'UTR2** started 319bp upstream of the ATG. The non coding sequences upstream of the ATG showed 93%

sequence similarity among themselves. While the nucleotide sequence from ATG and down stream of it showed 99% sequence similarity (Fig. 3.16).

```
gMN5'UTR1 CGACGGCCCGGGCTGGTAAAGGCGTGACTTATTTATCTATATAAAGAATTATACAAGTGG
gMN5'UTR2 CGATGGCCCGGGCTGGTAAAGGCATGACTTATTTATCTATACAGATAATTATACGAGTGG
          *** ******************* ***************** * * ******** *****
gMN5'UTR1 ATTGGGGTGGGGGTCGAATATCGCTGGGCCCACGTTCATGCCGTACAAGAACCTTACCAA
gMN5'UTR2 ATTGGGGTGGTGGTCGAATATTGCTGG-CCCACGTTCACGCCGTACAAGAACCTCACCAA
          ********** ********** ***** ********** *************** *****
gMN5'UTR1 CTGCACCCGGTCGGCAGCCGGTTCAACCAACCACTGCTTCTTTTGCTTCTCCCATGCCGG
gMN5'UTR2 CCACACCCGGTCGGCAGCCGGCTCAAC--ACCACTGCTGCTTTTGCTTCTCCCATGCCGG
          *  ****************** *****  ********* *********************
gMN5'UTR1 TTCAAACCGGACATTTCCCCCTCCTTATATATCATACAAGCCGCATTCGCCATGTTTCCC
gMN5'UTR2 TTCAAACCGGACAGTTCCCCCTCCCTATAAATCATACAAGCCGCCTTCGCCATGTTT---
          ************* ********** **** ************** ************
gMN5'UTR1 AAGGCTCTGGAAGAAAGACTCAAAAACAGAGCAGAAGAAAAGAAGAGAGAGGTCGAGGAA
gMN5'UTR2 AAGGCTCTGGAAGAAAGACTCGAAAACAGAGCAGAAGAAAAGAAGAGAGAGTTCAAGGAA
          ********************* ***************************** ** *****

                                 →
gMN5'UTR1 ACAAAAA-GGATAAGAAGTAATGGCGGATCAAAATCAAAGCGAAGCAGGAAGGCACCAAG
gMN5'UTR2 ACAAAAAAGGATAAGAAGT-ATGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAG
          ******* *********** *********** ****************************
gMN5'UTR1 AGGTTGGCCACAAGAGCCTTTTGCAGAGTGATGCTCTCTACCAG
gMN5'UTR2 AGGTTGGCCACAAGAGCCTTTTGCAGAGTGATGCTCTCTACCAG
          ********************************************
```

Fig. 3.16: Nucleotide sequence alignment of the gMN5'UTR1 and gMN5'UTR2.

3.3.5.2 3' downstream walking

For isolation of the 3'UTRs, the last exon of the g**CcoF2R31** and g**CcoF2R32** gene sequences were aligned (Fig. 3.16) and two forward primers, **C1: 5' (d) AGATGCTCCTC TGAGGAAGTACGTGAGGT 3'** and **C2: 5' (d) ATGCTTCCTATTGGTGATGGC ATTACCC 3',** were designed and used for 3' downstream walking.

```
gCcoF2R31     GAAAAGAATCATGGGAGCTATGATTTCATATTCGTGGATGCTGACAAGGACAACTACTTG
gCcoF2R32     GAAAAGAATCATGGGAGCTATGATTTCATATTCGTGGATGCTGACAAGGACAACTATTTG
              ******************************************************** ***

gCcoF2R31     AATTATCATAAGAGGTTGATCGATTTGGTGAAGGTAGGAGGAGTGATCGGGTACGATAAC
gCcoF2R32     AATTATCACAAGAGGTTGATCGATTTGGTGAAGGTAGGAGGAGTGATAGGGTACGATAAT
              ******** ************************************** ***********

                                                              C1
                                                 ─────────────────────►
gCcoF2R31     ACCTTGTGGAACGGATCTGTAGTGGCACCACCAGATGCTCCTCTGAGGAAGTACGTGAGG
gCcoF2R32     ACCTTGTGGAATGGGTCTGTGGTGGCACCACCAGATGCTCCTCTGAGGAAGTACGTGAGG
              *********** ** ***** ***************************************

gCcoF2R31     TATTACAGGGACTTTGTGCTGGAGCTCAGCAAGGCTTTGGCTGTGGACCCTAGGATCGAG
gCcoF2R32     TATTACCGGGACTTTGTGTTGGAGCTCAACAAGGCTTTGGCTGTGGACCCTCGGATTGAG
              ****** *********** ********* ********************** **** ***
```

```
                              C2
                    ——————————————————►
gCcoF2R31     ATCTGCATGCTTCCTATTGGTGATGGCATTACCCTTTGCCGACGCATTAGTTAA
gCcoF2R32     ATCTGCATGCTTCCTATTGGTGATGGCATTACCCTTTGCCGACGCATTAGTTAA
              ******************************************************
```

Fig. 3.16: Nucleotide sequence alignment of the last exon of gCcoF2R31 and gCcoF2R32 showing the regions where from the primers C1 and C2 were designed for 3' downstream Genome walking.

The primers **C1** was used as the forward primer and the **AP1** primer from the Genome walking kit was used as the reverse primer to perform primary PCR using *Dra* I, *EcoR* V, *Pvu* II and *Stu* I libraries (Chapter 2; section 2.4.10) of *L. leucocephala* as template. The forward primer C2 and the reverse primer **AP2** (from Genome walking kit) were used for secondary PCR. The primary PCR products were used as template for the secondary PCR. Multiple bands were amplified both in the primary and the secondary PCR (Fig. 3.17 a, b).

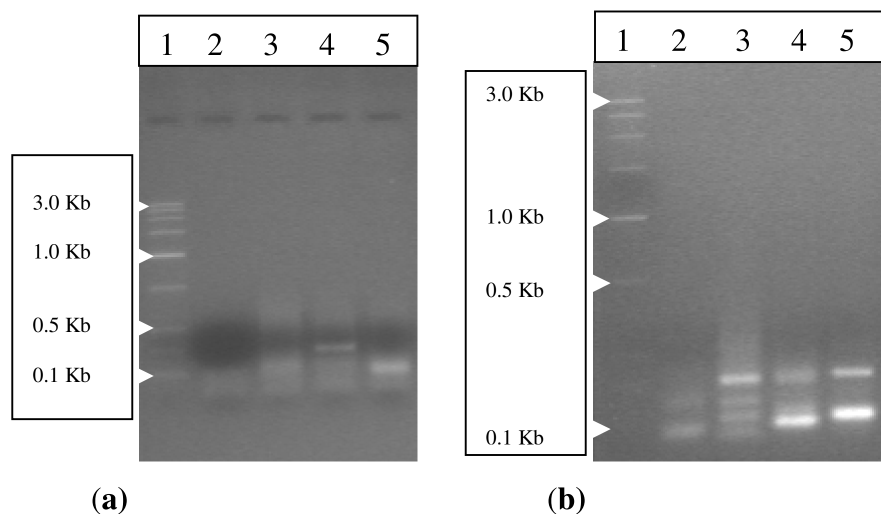


Fig. 3.17: (a) Primary PCR with *L. leucocephala Dra* I (lane 2), *EcoR* V (lane 3) , *Pvu* II (lane 4) and *Stu* I (lane 5) libraries. DNA size marker (lane 1); (b) secondary PCR using diluted primary PCR product of *Dra* I (lane 2), *EcoR* V (lane 3), *Pvu* II (lane 4) and *Stu* I (lane 5) libraries as template. DNA size marker (lane 1)

Secondary PCR amplicons from all the four libraries were cloned in pGEM-T Easy Vector (Promega, USA) and sequenced. The nucleotide sequence of a 313bp amplicons, **gMC3'UTR**, from the *Stu* I library was from the 3' coding region of the CCoAOMT gene. It included the translation termination codon and the putative poly A site (Fig. 3.18). Only one type of 3' downstream fragment was fished out. The **gMC3'UTR** nucleotide sequence revealed that the reverse primer, CcoR3, used earlier for amplification had mismatches and the stop codon for the gene was TGA instead of TAA.

gMC3'UTR:

```
  1 ATGCTTCCTA TTGGTGATGG CATTACCCTC TGCCGTAGGA TCAGCTGATT TCTCTGGACC
 61 TTCCTGCACC GCGAAGGGCT AAAAATATTT CTATCTGTAT TTTTTTCCAA ACCATTCTGA
121 GCATGGGGCT TAATTGGGAG TGTTTTATAT TTTCATATTC TTGCTTTTTC ATATTATATT
181 GAATCCAGTA ATCATAAAAC CCTCCTTGAG TTTATTGTGA AGGCTGACTT GTGTCCTTAG
241 TGGGAATTGA TTCCTAATAC TTCAACAGAA GATTATTATT AAATAATTTA TCGTTTACCA
301 GCCCGGGCCG TCG
```

Fig. 3.18: Nucleotide sequence of gMC3'UTR. Stop codon is in bold italics. Putative polyadenylation signal (shaded) is at nucleotides 173 –178.

3.3.6 Full length CCoAOMT genomic clone

The 5'upstream genome walking of the CCoAOMT partial gene sequence as discussed in section **3.3.5** above resulted in the isolation of two nucleotide sequence which spanned the first exon including the 5'UTR upstream of the ATG. Also the two sequences showed marked nucleotide dissimilarities in their UTR's. In the 3' downstream genome walking experiment a single nucleotide sequence was isolated. This sequence went beyond the putative polyadenylation site. Information from these nucleotide sequences were used for the design of gene specific primers for amplification of the CCoAOMT genes from translation start codon to translation stop codon. Forward primer **MF4: 5' (d) ATGGCGGATCAG AATCAAAGCGAA G 3'** starting from the translation initiation codon and a reverse primer **MR4: 5' (d) TCAGCTGATCCTACGGCAGAGAGTGAT 3'** upto the translation termination codon were designed and synthesized. PCR was performed and a DNA fragment of around 1.2 Kb amplified (Fig. 3.19).

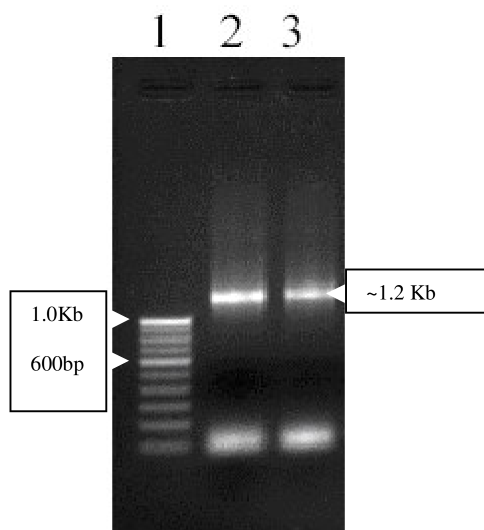


Fig. 3.19: PCR amplification from genomic DNA of *L. leucocephala with* MF4 and MR4 primers. DNA size marker (lane 1) and ~1.2 Kb amplified product (lanes 2, 3).

The ~1.2 Kb amplicon was cloned into pGEM-T Easy Vector (Promega, USA) and the recombinant plasmids were restriction digested with *Sac* I. As expected two clones showing different *Sac* I restriction patterns (Fig. 3.19) were observed. These were designated as **gMF4MR41** and **gMF4MR42,** and sequenced.

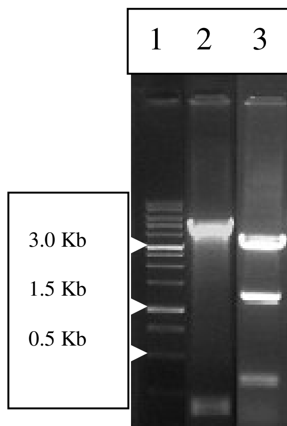


Fig. 3.20: Restriction digestion of gMF4MR41 and gMF4MR42 gene clones. DNA size marker (lane 1), *Sac* I restriction digestion of gMF4MR41 (lane 2) and gMF4MR42 (lane 3).

Sequencing data analysis of **gMF4MR41** and **gMF4MR42** showed that the clones were two different CCoAOMT genes, each of 1292 bp. The two sequences upon alignment showed 92.4% sequence similarity (Fig. 3.21). The calculated GC content (www.justbio.com) of the two genes was 41.67 % and 43.72 % respectively The Integrated System of Michael Zhang's MZEF and Thanaraj's SpliceProximalCheck was used to identify the intron / exon junctions in the two clones. The gMF4MR41 and gMF4MR42 nucleotide sequences show presence of putative five exons and four introns. In the gMF4MR41 gene clone the five exons spanned the nucleotide positions 1 - 84, 196 - 275, 517 - 661, 763 - 894 and 998 - 1292 (bold letters in Fig. 3.21). The putative introns spanned nucleotide positions 85 - 195, 276 - 516, 662 - 762 and 895 - 998 (shaded lower case italics in Fig. 3.21). The gMF4MR42 gene clone revealed five putative exons spanning nucleotide positions 1 - 84, 211 - 290, 481 - 625, 742 - 873 and 998 - 1292 (bold letters in Fig. 3.21). Introns in this sequence spanned nucleotide positions 51 - 210, 291 - 480, 626 - 741 and 874 – 998 (shaded lower case italics in Fig. 3.21). Both the CCoAOMT genomic DNA clones, gMF4MR41 and gMF4MR42, have been submitted to NCBI GenBank database with designations **CCoAOMT1** and **CCoAOMT2** and the sequence data is available under accession numbers **DQ517929** and **DQ517930** respectively**.**

```
                MF4
CCoAOMT1 ATGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAGAGGTTGGCCACAAGAGCCTT
CCoAOMT2 ATGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAGAGGTTGGCCACAAGAGCCTT
         ************************************************************

CCoAOMT1 TTGCAGAGTGATGCTCTCTACCAGgtactcctatttctctttgatcacgccacacatca-
CCoAOMT2 TTGCAGAGTGATGCTCTCTACCAGgtaccccaatttctcttttatcacaccacacatcat
         **************************** ** ********** ***** ********** 

CCoAOMT1 --atcccacaaccacaaccacaaccactttcttcttcttcttcttc------------tg
CCoAOMT2 caatcccacaaccgcaaccagtttcttcttcttcttcttcttcttcctgttcttgtcctg
           *********** ******    *   ******************            **

CCoAOMT1 gtctaaaggatcatatgattttgtctgcagTACATTCTAGAGACCAGTGTCTACCCCAGA
CCoAOMT2 gtctaaaggatcatgtgattttgtatgcagTACATTCTGGAGACTAGTGTCTACCCCAGA
         ************** ********* ************* ***** ***************

CCoAOMT1 GAACCTGAACCCATGAAGGAGCTCAGAGAAATCACAGCCAAACACCCTTGgtaa-gcttc
CCoAOMT2 GAACCTGAACCCATGAAGGAGATCAGAGAAATCACAGCCAAACACCCTTGgtaaagcttc
         ********************* ******************************** *****

CCoAOMT1 ttcttaggactcaaacattatgaagttttcagcaaattacatcataatgctgtcgtgccc
CCoAOMT2 tttttaggac---gacatagtgaaattttctgcaaat-----------------------
         ** *******    **** ****  ***** ******
```

```
CCoAOMT1 tatgtccgtaatgcttaagattcagatcctctgtttttgtgcttgctttgagtttgtcct
CCoAOMT2 ----------------------------ctttttttttttgcttgctttcggtttgtcct
                                     ** * ***** **********  *********
CCoAOMT1 tcatccagacaaaaaatttggagactttaaacaaaatagattaggataagaaaaattg--
CCoAOMT2 tcctcaagacgaaaaatcgagaaactttaaacaaaatcgatttggatacgaaagattgtc
         ** ** **** ******   ** ************** **** ***** **** ****
CCoAOMT1 cgtgtgtgattatgggaagatggtccattactgattgatggaggaaaacaacagGAACAT
CCoAOMT2 cgtgtgtgattatgagtggatgatccattactgattgaaggaagaaaacaacagGAACAT
         ************** *  **** *************** *** *****************
CCoAOMT1 CATGACAACCTCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGCTTATCAATGC
CCoAOMT2 CATGACAACCTCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGCTTATTAATGC
         ****************************************************** *****
CCoAOMT1 TAAGAATACCATGGAGATTGGTGTTTACACTGGCTACTCCCTGCTTGCCACTGCTCTGGC
CCoAOMT2 TAAGAACACCATGGAGATTGGTGTCTACACTGGCTACTCCCTGCTTGCAACTGCTCTGGC
         ****** ***************** *********************** ***********
CCoAOMT1 TCTCCCTGAAGATGGAAAGgtaagggccaatcaacaacagagtccaatgttctgaatgtt
CCoAOMT2 TCTCCCTGAAGATGGAAAGgtaacgaccaatcaacaacagagtccaatgttctgaatgat
         *********************** * ******************************** *
CCoAOMT1 tccatgttacttt-atcaagatcatcgaaacatatctgagat-------------ttggt
CCoAOMT2 tcaatgttacttttatcaagatcctccaaacatgtctgaaatgttttcgtattctttggt
         ** ********** ********* ** ****** ***** **             *****
CCoAOMT1 cgg-aattgttgcagATACTGGCCATGGACATTAACAGAGAAAACTACGAGTTGGGTCTT
CCoAOMT2 tgggaattggcgtagATACTGGCCATGGACATTAACAGAGAAAACTACGAATTGGGTCTT
         **  *****  * ************************************* *********
CCoAOMT1 CCGGTAATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGTTCAGAGAGGGCCCTGCTCTC
CCoAOMT2 CCGGTAATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGTTCAAAGAGGGCCCTGCTCTT
          ******************************************* ***************
CCoAOMT1 CCTGTTCTTGACGAACTCGTTAAAGATgtaagctttttctatctccatcttatcattaca
CCoAOMT2 CCTGTTCTTGACGAACTCGTTAAAGATgtaagctttt-ctatctccttcttgtcttcacc
         ************************************* ******** **** ** * **
CCoAOMT1 actt---tctggtcttatctccattat--------------------aaacatcaaattg
CCoAOMT2 acttctttttggtcttttctccattatcactagtttttaggccaaacgaacatcaaattg
         ****   * ******* **********                     ************
CCoAOMT1 ataagagaggtatgattgtgggggtggtttacagGAAAAGAATCATGGGAGCTATGATTT
CCoAOMT2 acaagagtcgtatgtttgtgggg-tgaatagcagGAAAAGAATCATGGGAGCTATGATTT
         * ***** ***** ********  **  *  *****************************
CCoAOMT1 CATATTCGTGGATGCTGACAAGGACAACTACTTGAATTATCATAAGAGGTTGATCGATTT
CCoAOMT2 CATATTCGTGGATGCTGACAAGGACAACTATTTGAATTATCACAAGAGGTTGATCGATTT
         ****************************** *********** *****************
CCoAOMT1 GGTGAAGGTAGGAGGAGTGATCGGGTACGATAACACCTTGTGGAACGGATCTGTAGTGGC
CCoAOMT2 GGTGAAGGTAGGAGGAGTGATAGGGTACGATAATACCTTGTGGAATGGGTCTGTGGTGGC
         ********************* *********** *********** ** ***** *****
```

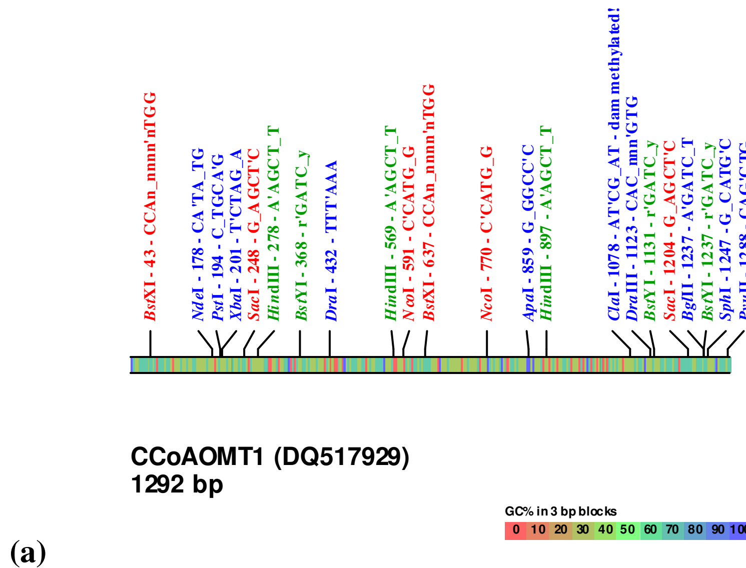
```
CCoAOMT1 ACCACCAGATGCTCCTCTGAGGAAGTACGTGAGGTATTACAGGGACTTTGTGCTGGAGCT
CCoAOMT2 ACCACCAGATGCTCCTCTGAGGAAGTACGTGAGGTATTACCGGGACTTTGTGTTGGAGCT
         ****************************************  ***********  *******

CCoAOMT1 CAACAAGGCTTTGGCTGTGGACCCTAGGATCGAGATCTGCATGCTTCCAGTTGGTGATGG
CCoAOMT2 CAACAAGGCTTTGGCTGTGGACCCTCGGATTGAGATCTGCATGCTTCCAGTTGGTGATGG
         ************************* **** *****************************

                    MR4
         <---------------------------
CCoAOMT1 CATCACTCTCTGCCGTAGGATCAGCTGA
CCoAOMT2 CATCACTCTCTGCCGTAGGATCAGCTGA
         ****************************
```

Fig. 3.21: Nucleotide sequence alignment of *L. leucocephala* CCoAOMT1 (DQ517929) and (DQ517930) CCoAOMT2 genomic gene clones using Clustal W 1.8. The sequences in shaded lower case italics are introns. Sequence letters in capitals are exons. The forward MF4 and the reverse MR4 primers are indicated by arrow.

Restriction analysis of **CCoAOMT1** and **CCoAOMT2** was done using bioinformatic software **pDRAW32** (Fig. 3.22 A and B). Analysis was limited to enzyme site cutting not more than four times in the sequence.



(a)

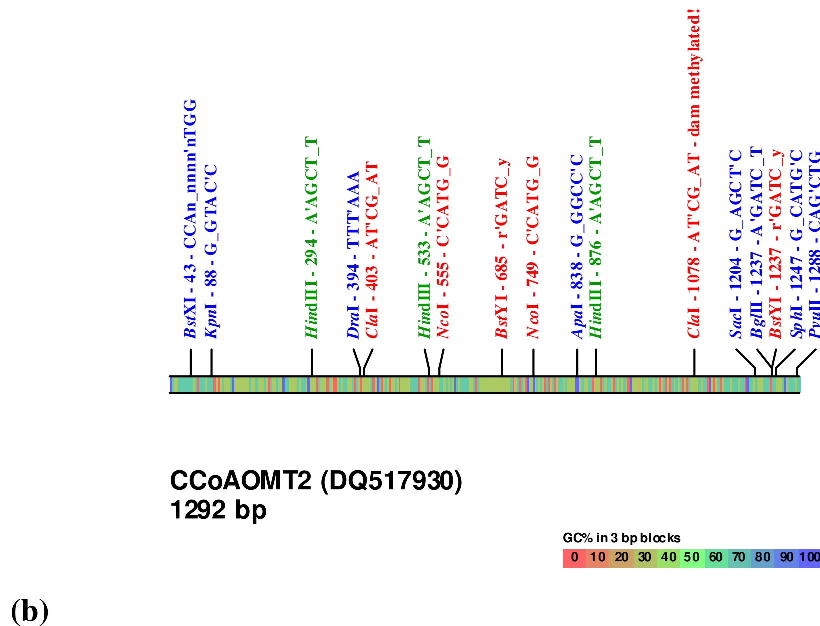


Fig. 3.22: Restriction analysis of (a) CCoAOMT1 (DQ517929) and (b) CCoAOMT2 (DQ517930) using pDRAW 32.

The introns as identified by the Michael Zhang's Exon finder, were excised from the **CCoAOMT1 and CCoAOMT2** genomic gene clones and the exons joined to deduce the coding sequences. Both the deduced coding sequences were of 735 bp and show 97% nucleotide sequence similarity upon alignment (Fig. 3.23).

```
CCoAOMT1 ATGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAGAGGTTGGCCACAAGAGCCTT
CCoAOMT2 ATGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAGAGGTTGGCCACAAGAGCCTT
         ************************************************************
CCoAOMT1 TTGCAGAGTGATGCTCTCTACCAGTACATTCTAGAGACCAGTGTCTACCCCAGAGAACCT
CCoAOMT2 TTGCAGAGTGATGCTCTCTACCAGTACATTCTGGAGACTAGTGTCTACCCCAGAGAACCT
         ******************************** ***** *********************
CCoAOMT1 GAACCCATGAAGGAGCTCAGAGAAATCACAGCCAAACACCCTTGGAACATCATGACAACC
CCoAOMT2 GAACCCATGAAGGAGATCAGAGAAATCACAGCCAAACACCCTTGGAACATCATGACAACC
         *************** ********************************************
CCoAOMT1 TCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGCTTATCAATGCTAAGAATACC
CCoAOMT2 TCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGCTTATTAATGCTAAGAACACC
         ******************************************** *********** ***
CCoAOMT1 ATGGAGATTGGTGTTTACACTGGCTACTCCCTGCTTGCCACTGCTCTGGCTCTCCCTGAA
CCoAOMT2 ATGGAGATTGGTGTCTACACTGGCTACTCCCTGCTTGCAACTGCTCTGGCTCTCCCTGAA
         ************** ************************ ********************
CCoAOMT1 GATGGAAAGATACTGGCCATGGACATTAACAGAGAAAACTACGAGTTGGGTCTTCCGGTA
CCoAOMT2 GATGGAAAGATACTGGCCATGGACATTAACAGAGAAAACTACGAATTGGGTCTTCCGGTA
         ******************************************** ***************
CCoAOMT1 ATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGTTCAGAGAGGGCCCTGCTCTCCCTGTT
CCoAOMT2 ATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGTTCAAAGAGGGCCCTGCTCTTCCTGTT
         ************************************* *************** ******
CCoAOMT1 CTTGACGAACTCGTTAAAGATGAAAAGAATCATGGGAGCTATGATTTCATATTCGTGGAT
CCoAOMT2 CTTGACGAACTCGTTAAAGATGAAAAGAATCATGGGAGCTATGATTTCATATTCGTGGAT
         ************************************************************
CCoAOMT1 GCTGACAAGGACAACTACTTGAATTATCATAAGAGGTTGATCGATTTGGTGAAGGTAGGA
CCoAOMT2 GCTGACAAGGACAACTATTTGAATTATCACAAGAGGTTGATCGATTTGGTGAAGGTAGGA
         ***************** *********** ******************************
CCoAOMT1 GGAGTGATCGGGTACGATAACACCTTGTGGAACGGATCTGTAGTGGCACCACCAGATGCT
CCoAOMT2 GGAGTGATAGGGTACGATAATACCTTGTGGAATGGGTCTGTGGTGGCACCACCAGATGCT
         ******** *********** *********** ** ***** ******************
CCoAOMT1 CCTCTGAGGAAGTACGTGAGGTATTACAGGGACTTTGTGCTGGAGCTCAACAAGGCTTTG
CCoAOMT2 CCTCTGAGGAAGTACGTGAGGTATTACCGGGACTTTGTGTTGGAGCTCAACAAGGCTTTG
         *************************** *********** ********************
CCoAOMT1 GCTGTGGACCCTAGGATCGAGATCTGCATGCTTCCAGTTGGTGATGGCATCACTCTCTGC
CCoAOMT2 GCTGTGGACCCTCGGATTGAGATCTGCATGCTTCCAGTTGGTGATGGCATCACTCTCTGC
         ************ **** ******************************************
CCoAOMT1 CGTAGGATCAGCTGA
CCoAOMT2 CGTAGGATCAGCTGA
         ***************
```

Fig. 3.23: Nucleotide sequence alignment of the deduced coding sequences of CCoAOMT1 and CCoAOMT2 genes using Clustal W 1.8.

3.3.7 Which 5’ UTR belongs to which CCoAOMT gene?

The 5’ upstream genome walking of CCoAOMT1 and CCoAOMT2 genes resulted in isolation of two 5’ UTR’s sequences (gMN5’UTR1 and gMN5’UTR2; see section 3.3.5.1 above). These 5’UTR’s needed to be assigned to their respective genes. The nucleotide sequences of gMN5’UTR1 and the gMN5’UTR2 were aligned to identify regions from where the forward primers would be designed. Primers were designed in a manner so that the 3’ nucleotide of the primers was different and would, hence, selectively amplify CCoAOMT1 and CCoAOMT2 gene sequences (Fig. 3.24).

```
gMN5'UTR1   CGACGGCCCGGGctggtaaaggcgtgacttatttatctatatAAAGAATTATACAAGTGG
gMN5'UTR2   CGATGGCCCGGGctggtaaaggcatgacttatttatctatacAGATAATTATACGAGTGG
            *** ******************* ***************** * * ******** *****

gMN5'UTR1   ATTGGGGTGGGGGTCGAATATCGCTGGGCCCACGTTCATGCCGTACAAGAACCTTACCAA
gMN5'UTR2   ATTGGGGTGGTGGTCGAATATTGCTGG-CCCACGTTCACGCCGTACAAGAACCTCACCAA
            ********** ********** ***** ********** *************** *****

gMN5'UTR1   CTGCACCCGGTCGGCAGCCGGTTCAACCAACCACTGCTTCTTTTGCTTCTCCCATGCCGG
gMN5'UTR2   CCACACCCGGTCGGCAGCCGGCTCAAC--ACCACTGCTGCTTTTGCTTCTCCCATGCCGG
            *  ****************** *****  ********* *********************

gMN5'UTR1   TTCAAACCGGACATTTCCCCCTCCTTATATATCATACAAGCCGCATTCGCCATGTTTCCC
gMN5'UTR2   TTCAAACCGGACAGTTCCCCCTCCCTATAAATCATACAAGCCGCCTTCGCCATGTTT---
            ************* ********** **** ************** ************

gMN5'UTR1   AAGGCTCTGGAAGAAAGACTCAAAAACAGAGCAGAAGAAAAGAAGAGAGAGGTCGAGGAA
gMN5'UTR2   AAGGCTCTGGAAGAAAGACTCGAAAACAGAGCAGAAGAAAAGAAGAGAGAGTTCAAGGAA
            ********************* ***************************** ** *****

gMN5'UTR1   ACAAAAA-GGATAAGAAGTAATGGCGGATCAAAATCAAAGCGAAGCAGGAAGGCACCAAG
gMN5'UTR2   ACAAAAAAGGATAAGAAGT-ATGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAG
            ******* *********** *********** ****************************

gMN5'UTR1   AGGTTGGCCACAAGAGCCTTTTGCAGAGTGATGCTCTCTACCAG
gMN5'UTR2   AGGTTGGCCACAAGAGCCTTTTGCAGAGTGATGCTCTCTACCAG
            ********************************************
```

Fig. 3.24: Nucleotide sequence alignment of gMN5’UTR1 and gMN5’UTR2. The regions used for designing the forward primers for selective amplification of CCoAOMT1 and CCoAOMT2 genes are shaded and in lower case italics.

Forward primers designed from **gMN5’UTR1 and gMN5’UTR2:**

MPC1: **5’ (d) CTGGTAAAGGCGTGACTTATTTATCTATAT 3’**

MPC2: **5’ (d) CTGGTAAAGGCATGACTTATTTATCTATAC 3’**

PCR was performed using the MPC1 and MPC2 as forward primers, the previously designed MR4 reverse primer (section 3.4 above) and genomic DNA of *L. leucocephala* as template. Approximately 1.5 Kb DNA fragment were amplified (Fig. 3.25). All the amplicons were cloned in pGEM-T Easy Vector and sequenced. The sequencing data revealed that the gMN5'UTR1 was the 5' UTR of CCoAOMT2 gene and gMN5'UTR2 the 5' UTR of the CCoAOMT1 gene.

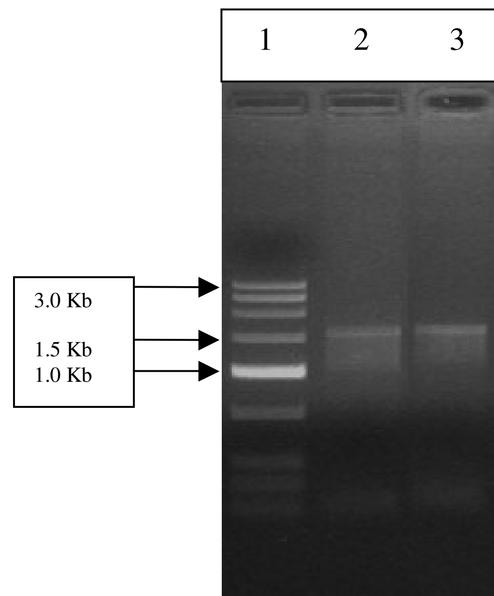


Fig. 3.25: PCR amplification using genomic DNA of *L. leucocephala* as template. MPC1 and MPC2 were used as forward primers and MR4 as the reverse primer (lanes 2, 3 respectively).

The nucleotide sequences of the CCoAOMT1 and CCoAOMT2 genes with their 5' UTRs are shown (Fig. 3.26 a, b) and their nucleotide sequence alignment in Fig. 3.36 c. The nucleotide sequence alignment of the two CCoAOMT gene with each other shows 92% sequence similarity. The CCoAOMT1 and CCoAOMT2 genomic clones when aligned with the genomic DNA clones of *Vitis vinifera* (AM478701) and *Petroselinum crispum* (Z54183) showed 72% and 79% sequence similarity respectively (Fig. 3.27)

1 CGACGGCCCG GGCTGGTAAA GGCGTGACTT ATTTATCTAT ATAAAGAATT ATACAAGTGG

61 ATTGGGGTGG GGGTCGAATA TCGCTGGGCC CACGTTCATG CCGTACAAGA ACCTTACCAA

121 CTGCACCCGG TCGGCAGCCG GTTCAACCAA CCACTGCTTC TTTTGCTTCT CCCATGCCGG

181 TTCAAACCGG ACATTTCCCC CTCCTTATAT ATCATACAAG CCGCATTCGC CATGTTTCCC

241 AAGGCTCTGG AAGAAAGACT CAAAAACAGA GCAGAAGAAA AGAAGAGAGA GGTCGAGGAA

*301 ACAAAAAGGA TAAGAAGTA***A TGGCGGATCA GAATCAAAGC GAAGCAGGAA GGCACCAAGA**

361 GGTTGGCCAC AAGAGCCTTT TGCAGAGTGA TGCTCTCTAC CAGGTACTCC TATTTCTCTT

421 TGATCACGCC ACACATCAAT CCCACAACCA CAACCACAAC CACTTTCTTC TTCTTCTTCT

481 TCTGGTCTAA AGGATCATAT GATTTTGTCT GCAGTACATT CTAGAGACCA GTGTCTACCC

541 CAGAGAACCT GAACCCATGA AGGAGCTCAG AGAAATCACA GCCAAACACC CTTGGTAAGC

601 TTCTTCTTAG GACTCAAACA TTATGAAGTT TTCAGCAAAT TACATCATAA TGCTGTCGTG

661 CCCTATGTCC GTAATGCTTA AGATTCAGAT CCTCTGTTTT TGTGCTTGCT TTGAGTTTGT

721 CCTTCATCCA GACAAAAAAT TTGGAGACTT TAAACAAAAT AGATTAGGAT AAGAAAAATT

781 GCGTGTGTGA TTATGGGAAG ATGGTCCATT ACTGATTGAT GGAGGAAAAC AACAGGAACA

841 TCATGACAAC CTCAGCAGAC GAGGGGCAAT TCTTGAACAT GCTCCTTAAG CTTATCAATG

901 CTAAGAATAC CATGGAGATT GGTGTTTACA CTGGCTACTC CCTGCTTGCC ACTGCTCTGG

961 CTCTCCCTGA AGATGGAAAG GTAAGGGCCA ATCAACAACA GAGTCCAATG TTCTGAATGT

1021 TTCCATGTTA CTTTATCAAG ATCATCGAAA CATATCTGAG ATTTGGTCGG AATTGTTGCA

1081 GATACTGGCC ATGGACATTA ACAGAGAAAA CTACGAGTTG GGTCTTCCGG TAATTCAGAA

1141 AGCTGGTGTT GCCCACAAAA TTGAGTTCAG AGAGGGCCCT GCTCTCCCTG TTCTTGACGA

1201 ACTCGTTAAA GATGTAAGCT TTTTCTATCT CCATCTTATC ATTACAACTT TCTGGTCTTA

1261 TCTCCATTAT AAACATCAAA TTGATAAGAG AGGTATGATT GTGGGGGTGG TTTACAGGAA

1321 AAGAATCATG GGAGCTATGA TTTCATATTC GTGGATGCTG ACAAGGACAA CTACTTGAAT

1381 TATCATAAGA GGTTGATCGA TTTGGTGAAG GTAGGAGGAG TGATCGGGTA CGATAACACC

1441 TTGTGGAACG GATCTGTAGT GGCACCACCA GATGCTCCTC TGAGGAAGTA CGTGAGGTAT

1501 TACAGGGACT TTGTGCTGGA GCTCAGCAAG GCTTTGGCTG TGGACCCTAG GATCGAGATC

1561 TGCATGCTTC CAGTTGGTGA TGGCATCACT CTCTGCCGTA GGATCAGCTG A

(a)

1 CGATGGCCCG GGCTGGTAAA GGCATGACTT ATTTATCTAT ACAGATAATT ATACGAGTGG

61 ATTGGGGTGG TGGTCGAATA TTGCTGGCCC ACGTTCACGC CGTACAAGAA CCTCACCAAC

121 CACACCCGGT CGGCAGCCGG CTCAACACCA CTGCTGCTTT TGCTTCTCCC ATGCCGGTTC

181 AAACCGGACA GTTCCCCCTC CCTATAAATC ATACAAGCCG CCTTCGCCAT GTTTAAGGCT

241 CTGGAAGAAA GACTCGAAAA CAGAGCAGAA GAAAAGAAGA GAGAGTTCAA GGAAACAAAA

*301 AAGGATAAGA AGT***ATGGCGG ATCAGAATCA AAGCGAAGCA GGAAGGCACC AAGAGGTTGG**

361 CCACAAGAGC CTTTTGCAGA GTGATGCTCT CTACCAGGTA CCCCAATTTC TCTTTTATCA

421 CACCACACAT CATCAATCCC ACAACCGCAA CCAGTTTCTT CTTCTTCTTC TTCTTCTTCC

481 TGTTCTTGTC CTGGTCTAAA GGATCATGTG ATTTTGTATG CAGTACATTC TGGAGACTAG

```
 541 TGTCTACCCC AGAGAACCTG AACCCATGAA GGAGATCAGA GAAATCACAG CCAAACACCC
 601 TTGGTAAAGC TTCTTTTTAG GACGACATAG TGAAATTTTC TGCAAATCTT TTTTTTTTTG
 661 CTTGCTTTCG GTTTGTCCTT CCTCAAGACG AAAAATCGAG AAACTTTAAA CAAAATCGAT
 721 TTGGATACGA AAGATTGTCC GTGTGTGATT ATGAGTGGAT GATCCATTAC TGATTGAAGG
 781 AAGAAAACAA CAGGAACATC ATGACAACCT CAGCAGACGA GGGGCAATTC TTGAACATGC
 841 TCCTTAAGCT TATTAATGCT AAGAACACCA TGGAGATTGG TGTCTACACT GGCTACTCCC
 901 TGCTTGCAAC TGCTCTGGCT CTCCCTGAAG ATGGAAAGGT AACGACCAAT CAACAACAGA
 961 GTCCAATGTT CTGAATGATT CAATGTTACT TTTATCAAGA TCCTCCAAAC ATGTCTGAAA
1021 TGTTTTCGTA TTCTTTGGTT GGGAATTGGC GTAGATACTG GCCATGGACA TTAACAGAGA
1081 AAACTACGAA TTGGGTCTTC CGGTAATTCA GAAAGCTGGT GTTGCCCACA AAATTGAGTT
1141 CAAAGAGGGC CCTGCTCTTC CTGTTCTTGA CGAACTCGTT AAAGATGTAA GCTTTTCTAT
1201 CTCCTTCTTG TCTTCACCAC TTCTTTTTGG TCTTTTCTCC ATTATCACTA GTTTTTAGGC
1261 CAAACGAACA TCAAATTGAC AAGAGTCGTA TGTTTGTGGG GTGAATAGCA GGAAAAGAAT
1321 CATGGGAGCT ATGATTTCAT ATTCGTGGAT GCTGACAAGG ACAACTATTT GAATTATCAC
1381 AAGAGGTTGA TCGATTTGGT GAAGGTAGGA GGAGTGATAG GGTACGATAA TACCTTGTGG
1441 AATGGGTCTG TGGTGGCACC ACCAGATGCT CCTCTGAGGA AGTACGTGAG GTATTACCGG
1501 GACTTTGTGT TGGAGCTCAA CAAGGCTTTG GCTGTGGACC CTCGGATTGA GATCTGCATG
1561 CTTCCAGTTG GTGATGGCAT CACTCTCTGC CGTAGGATCA GCTGA
```

(b)

```
CCoAOMT1   CGACGGCCCGGGCTGGTAAAGGCGTGACTTATTTATCTATATAAAGAATTATACAAGTGG
CCoAOMT2   CGATGGCCCGGGCTGGTAAAGGCATGACTTATTTATCTATACAGATAATTATACGAGTGG
           *** ******************* ***************** * * ******** *****

CCoAOMT1   ATTGGGGTGGGGGTCGAATATCGCTGGGCCCACGTTCATGCCGTACAAGAACCTTACCAA
CCoAOMT2   ATTGGGGTGGTGGTCGAATATTGCTGG-CCCACGTTCACGCCGTACAAGAACCTCACCAA
           ********** ********** ***** ********** *************** *****

CCoAOMT1   CTGCACCCGGTCGGCAGCCGGTTCAACCAACCACTGCTTCTTTTGCTTCTCCCATGCCGG
CCoAOMT2   CCACACCCGGTCGGCAGCCGGCTCAAC--ACCACTGCTGCTTTTGCTTCTCCCATGCCGG
           *  ****************** *****  ********* *********************

CCoAOMT1   TTCAAACCGGACATTTCCCCCTCCTTATATATCATACAAGCCGCATTCGCCATGTTTCCC
CCoAOMT2   TTCAAACCGGACAGTTCCCCCTCCCTATAAATCATACAAGCCGCCTTCGCCATGTTT---
           ************* ********** **** ************** ************

CCoAOMT1   AAGGCTCTGGAAGAAAGACTCAAAAACAGAGCAGAAGAAAAGAAGAGAGAGGTCGAGGAA
CCoAOMT2   AAGGCTCTGGAAGAAAGACTCGAAAACAGAGCAGAAGAAAAGAAGAGAGAGTTCAAGGAA
           ********************* ***************************** ** *****
```

```
CCoAOMT1   ACAAAAA-GGATAAGAAGTAATGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAG
CCoAOMT2   ACAAAAAAGGATAAGAAGTA-TGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAG
           ******* ************ ***************************************

CCoAOMT1   AGGTTGGCCACAAGAGCCTTTTGCAGAGTGATGCTCTCTACCAGGTACTCCTATTTCTCT
CCoAOMT2   AGGTTGGCCACAAGAGCCTTTTGCAGAGTGATGCTCTCTACCAGGTACCCCAATTTCTCT
           ************************************************ ** ********

CCoAOMT1   TTGATCACGCCACACATCA---ATCCCACAACCACAACCACAACCACTTTCTTCTTCTTC
CCoAOMT2   TTTATCACACCACACATCATCAATCCCACAACCGCAACCAGTTTCTTCTTCTTCTTCTTC
           ** ***** **********   *********** ******    *   ************

CCoAOMT1   TTCTTC------------TGGTCTAAAGGATCATATGATTTTGTCTGCAGTACATTCTAG
CCoAOMT2   TTCTTCCTGTTCTTGTCCTGGTCTAAAGGATCATGTGATTTTGTATGCAGTACATTCTGG
           ******            **************** ********* ************* *

CCoAOMT1   AGACCAGTGTCTACCCCAGAGAACCTGAACCCATGAAGGAGCTCAGAGAAATCACAGCCA
CCoAOMT2   AGACTAGTGTCTACCCCAGAGAACCTGAACCCATGAAGGAGATCAGAGAAATCACAGCCA
           **** ************************************ ******************

CCoAOMT1   AACACCCTTGGTAA-GCTTCTTCTTAGGACTCAAACATTATGAAGTTTTCAGCAAATTAC
CCoAOMT2   AACACCCTTGGTAAAGCTTCTTTTTAGGAC---GACATAGTGAAATTTTCTGCAAAT---
           ************** ******* *******    ****  **** ***** ******

CCoAOMT1   ATCATAATGCTGTCGTGCCCTATGTCCGTAATGCTTAAGATTCAGATCCTCTGTTTTTGT
CCoAOMT2   ------------------------------------------------CTTTTTTTTTTT
                                                           ** * ***** *

CCoAOMT1   GCTTGCTTTGAGTTTGTCCTTCATCCAGACAAAAAATTTGGAGACTTTAAACAAAATAGA
CCoAOMT2   GCTTGCTTTCGGTTTGTCCTTCCTCAAGACGAAAAATCGAGAAACTTTAAACAAAATCGA
           *********  *********** ** **** ******   ** ************** **

CCoAOMT1   TTAGGATAAGAAAAATTG--CGTGTGTGATTATGGGAAGATGGTCCATTACTGATTGATG
CCoAOMT2   TTTGGATACGAAAGATTGTCCGTGTGTGATTATGAGTGGATGATCCATTACTGATTGAAG
           ** ***** **** ****  ************** *  **** *************** *

CCoAOMT1   GAGGAAAACAACAGGAACATCATGACAACCTCAGCAGACGAGGGGCAATTCTTGAACATG
CCoAOMT2   GAAGAAAACAACAGGAACATCATGACAACCTCAGCAGACGAGGGGCAATTCTTGAACATG
           ** *********************************************************

CCoAOMT1   CTCCTTAAGCTTATCAATGCTAAGAATACCATGGAGATTGGTGTTTACACTGGCTACTCC
CCoAOMT2   CTCCTTAAGCTTATTAATGCTAAGAACACCATGGAGATTGGTGTCTACACTGGCTACTCC
           ************** *********** ***************** ***************

CCoAOMT1   CTGCTTGCCACTGCTCTGGCTCTCCCTGAAGATGGAAAGGTAAGGGCCAATCAACAACAG
CCoAOMT2   CTGCTTGCAACTGCTCTGGCTCTCCCTGAAGATGGAAAGGTAACGACCAATCAACAACAG
           ******** ********************************** * **************

CCoAOMT1   AGTCCAATGTTCTGAATGTTTCCATGTTACTTT-ATCAAGATCATCGAAACATATCTGAG
CCoAOMT2   AGTCCAATGTTCTGAATGATTCAATGTTACTTTTATCAAGATCCTCCAAACATGTCTGAA
           ****************** *** ********** ********* ** ****** *****
```

```
CCoAOMT1   AT-------------TTGGTCGG-AATTGTTGCAGATACTGGCCATGGACATTAACAGAG
CCoAOMT2   ATGTTTTCGTATTCTTTGGTTGGGAATTGGCGTAGATACTGGCCATGGACATTAACAGAG
           **             ***** ** *****  * ***************************

CCoAOMT1   AAAACTACGAGTTGGGTCTTCCGGTAATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGT
CCoAOMT2   AAAACTACGAATTGGGTCTTCCGGTAATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGT
           ********** *************************************************

CCoAOMT1   TCAGAGAGGGCCCTGCTCTCCCTGTTCTTGACGAACTCGTTAAAGATGTAAGCTTTTTCT
CCoAOMT2   TCAAAGAGGGCCCTGCTCTTCCTGTTCTTGACGAACTCGTTAAAGATGTAAGCTTTT-CT
           *** *************** ************************************* **

CCoAOMT1   ATCTCCATCTTATCATTACAACTT---TCTGGTCTTATCTCCATTAT-------------
CCoAOMT2   ATCTCCTTCTTGTCTTCACCACTTCTTTTGGTCTTTTCTCCATTATCACTAGTTTTTAG
           ****** **** ** * ** ****   * ******* **********

CCoAOMT1   -------AAACATCAAATTGATAAGAGAGGTATGATTGTGGGGGTGGTTTACAGGAAAAG
CCoAOMT2   GCCAAACGAACATCAAATTGACAAGAGTCGTATGTTTGTGGGG-TGAATAGCAGGAAAAG
           ************* *****  ***** ******** **   *  *********

CCoAOMT1   AATCATGGGAGCTATGATTTCATATTCGTGGATGCTGACAAGGACAACTACTTGAATTAT
CCoAOMT2   AATCATGGGAGCTATGATTTCATATTCGTGGATGCTGACAAGGACAACTATTTGAATTAT
           ************************************************** *********

CCoAOMT1   CATAAGAGGTTGATCGATTTGGTGAAGGTAGGAGGAGTGATCGGGTACGATAACACCTTG
CCoAOMT2   CACAAGAGGTTGATCGATTTGGTGAAGGTAGGAGGAGTGATAGGGTACGATAATACCTTG
           ** ************************************** *********** ******

CCoAOMT1   TGGAACGGATCTGTAGTGGCACCACCAGATGCTCCTCTGAGGAAGTACGTGAGGTATTAC
CCoAOMT2   TGGAATGGGTCTGTGGTGGCACCACCAGATGCTCCTCTGAGGAAGTACGTGAGGTATTAC
           ***** ** ***** *********************************************

CCoAOMT1  AGGGACTTTGTGCTGGAGCTCAGCAAGGCTTTGGCTGTGGACCCTAGGATCGAGATCTGC
CCoAOMT2  CGGGACTTTGTGTTGGAGCTCAACAAGGCTTTGGCTGTGGACCCTCGGATTGAGATCTGC
          *********** ********* ********************** **** *********

CCoAOMT1  ATGCTTCCAGTTGGTGATGGCATCACTCTCTGCCGTAGGATCAGCTGA
CCoAOMT2  ATGCTTCCAGTTGGTGATGGCATCACTCTCTGCCGTAGGATCAGCTGA
          ************************************************
```

(c)

Fig. 3.26: Nucleotide sequence of CCoAOMT1 (a), CCoAOMT2 (b) genomic clones and their nucleotide sequence alignment (c) using Clustal W 1.8 (c).

```
CCoAOMT1  ------------------------------CGACGGCCCGGGCTGGTAAAGGCGTGACTT
CCoAOMT2  ------------------------------CGATGGCCCGGGCTGGTAAAGGCATGACTT
AM478701  -------TTCGACCCATCTAGCCCTTGGATCAATATCATTGGTCGGATGACAGCTCGTGA
Z54183    GCCCTCCACTTCACCAACTATTATTCCAAACAACATGCATGAACACTTCACAGGTAA-AA
                                        * *       *        *    *
```

```
CCoAOMT1  ATTTATCTATATAAAGAATTATACAAGTGGATTGGGGTGGGGGTCGAATATCGCTGGGCC
CCoAOMT2  ATTTATCTATACAGATAATTATACGAGTGGATTGGGGTGGTGGTCGAATATTGCTGG-CC
AM478701  CGTCATCACCACCGAATAGTA-ACCAATTGATCTGCATAGGCGCATACTTCAATTGTGGT
Z54183    CCCAATCTCCACCGTTGATCG-TCTAGCTCATCATATACCAAGTTTATACCAAACGTGGG
              ***   *      *     * *    **          *   *        *

CCoAOMT1  CACGTTCATGCCGTACAAGAACCTTACCAACTGCACCCGGTCGGCAGCCGGTTCAACCAA
CCoAOMT2  CACGTTCACGCCGTACAAGAACCTCACCAACCACACCCGGTCGGCAGCCGGCTCAAC--A
AM478701  TTGGACGGGGAGACAGTAGCAATCCACCAACC---CCCGACCGGGTAGTGGGCCGTC---
Z54183    ACCCATCAAGATCAATCTCCATCTCACCAACCG---CAGACCGGGTTG-GTGAAAACGGT
                   *    *     *    ******     * *  ***     *      *

CCoAOMT1  CCACTGCTTCTTTTGCTTCTCCCATGCCGGTTCAAACCGGACATTTCCCCCTCCTTATAT
CCoAOMT2  CCACTGCTGCTTTTGCTTCTCCCATGCCGGTTCAAACCGGACAGTTCCCCCTCCCTATAA
AM478701  CAACCACTCCATAAGCGTAATTC-CACCGGTTCGAACCGGAC---CCAGCCCCATTATAA
Z54183    CTA--GCCACCTTTACATAAAAACAACCGGTTCGAACCCTTT-TTTCCTTTCCTTTATAT
          * *   *  * *   * *        ******* ****        *     *  ****

CCoAOMT1  ATCATACAAGCCGCATTCGCCATGTTTCCCAAGGCTCTGGAAGAAAGACTCAAAAACAGA
CCoAOMT2  ATCATACAAGCCGCCTTCGCCATGTTT---AAGGCTCTGGAAGAAAGACTCGAAAACAGA
AM478701  ACACTGCCATTTCCTCCTCCTAGACTC----AGGTTCACGGTTAA----TCGGCGACAAA
Z54183    ACCCTTTCATTCTTTCATTTCTTCTCA---GAAGCTCCG-----T-----CAACTACAAA
          *   *   *                      * * **             *    *** *

CCoAOMT1  GCAGAAGAAAAGAAGAGAGAGGTCGAGGAAACAAAAA-GGATAAGAAGTAATGGCGGATC
CCoAOMT2  GCAGAAGAAAAGAAGAGAGAGTTCAAGGAAACAAAAAAGGATAAGAAGTA-TGGCGGATC
AM478701  GTTTCCGGCGAATACTAGAAGG--AAAGGAACAGAAGATCTCTAGAAATGGCCACGAACC
Z54183    CAACTAAAAGAAGAAGAAAAGCT-TAAAAATTAAAAACAATTCCAGATGGCTTCTAATG
                    * *      **    *   *  * **         *  *      *  *

CCoAOMT1  AGAATCAAAGCGAAGCAGGAAGGCACCAAGAGGTTGGCCACAAGAGCCTTTTGCAGAGTG
CCoAOMT2  AGAATCAAAGCGAAGCAGGAAGGCACCAAGAGGTTGGCCACAAGAGCCTTTTGCAGAGTG
AM478701  A-----A--G--AAGCTGGGAGGCACCAGGAGGTTGGCCACAAGAGCCTTTTGCAGAGTG
Z54183    G---------TGAATCT---AAACATTCAGAAGTTGGGCACAAGAGTCTTTTGCAGAGTG
                      ** *    *  **    ** ***** ******** *************

CCoAOMT1  ATGCTCTCTACCAGGTACTCCTATTTCTCTTTGATCACGCCACACATCA---ATCCCACA
CCoAOMT2  ATGCTCTCTACCAGGTACCCCAATTTCTCTTTTATCACACCACACATCATCAATCCCACA
AM478701  ATGCTCTTTATCAGGTA----AGCTCTTCATT--CCCCTCGAGGCTTCAT--------CA
Z54183    ATGCTCTTTATCAGGTAC---ACATTTTTTTTGTCATCAAAATGATCAAT-GCTTTTTT
          ******* ** ******       *  *  **   *     *    ***

CCoAOMT1  ACCACAACCACAACCACTTTCTTCTTCTTCTTCTTC------------TGGTCTAAAGGA
CCoAOMT2  ACCGCAACCAGTTTCTTCTTCTTCTTCTTCTTCTTCCTGTTCTTGTCCTGGTCTAAAGGA
AM478701  GTCTTCGTCCTTTTAATCTGTATTTTCTCGTTCTCT-------------GATTTT-----
Z54183   GTTAAGTTTCATACATTCTTGATTATCTTGACCAACTAA--------CTAATTCAATG--
                           *   *  ***    *                  *

CCoAOMT1 TCATATGATTTTGTCTGCAGTACATTCTAGAGACCAGTGTCTACCCCAGAGAACCTGAAC
CCoAOMT2 TCATGTGATTTTGTATGCAGTACATTCTGGAGACTAGTGTCTACCCCAGAGAACCTGAAC
AM478701 TCTTGTTGTTGT---TGCAGTATATACTTGAAACCAGTGTGTACCCAAGAGAGCCTGAAT
Z54183   TTTTCTGTTTTT-----CAGTATATACTTGAAACAAGTGTGTACCCAAGAGAACCAGAGG
         *  * *  ** *     ***** ** ** ** ** ***** ***** ***** ** **
```

```
CCoAOMT1 CCATGAAGGAGCTCAGAGAAATCACAGCCAAACACCCTTGGTAA----GCTTCTTCTTAG
CCoAOMT2 CCATGAAGGAGATCAGAGAAATCACAGCCAAACACCCTTGGTAAA---GCTTCTTTTTAG
AM478701 CCATGAAGGAGCTCAGAGAGTTGACTGCCCAGCATCCATGGTGA----GCTTTTCTT--G
Z54183   CAATGAAAGAGCTTAGAGAAGTCACCGCAAAGCATCCATGGTTAGTTTACTAATCCTAAG
         * ***** *** * *****  * ** **  * ** ** **** *     **  *  *  *

CCoAOMT1 GACTCAAA--CATTATGAAGTTTTCAGCAAATTACATCATAATGCTGTCGTGCCCTATGT
CCoAOMT2 GAC---GA--CATAGTGAAATTTTCTGCAAAT----------------------------
AM478701 ATT------------CGGA--TTCT--CAAAT----------------------------
Z54183   TTCTAAGTTTCATTTTAGTGTTCGCATTTAAACATAATAAAACACACACATTACGGCCCC
                              *       **

CCoAOMT1 CCGTAATGCTTAAGATTCAGATCCTCTGTTTTTGTGCTTGCTTTGAGTTT-GTCCTTCAT
CCoAOMT2 ------------------------CTTTTTTTTTTGCTTGCTTTCGGTTT-GTCCTTCCT
AM478701 ----------------------------------------CTTTGAAT---ACCCGTTGC
Z54183   GCTCGGAAGTTATATTTTTAAGAATAAGATGT-ACGCAGACTTTACGTCTAGTCCTGAGA
                                                 ****   *     **

CCoAOMT1 CCAGACAAA------------AAATTTGGAGACTTTAAACAAAATAGATTAGGATAAGAA
CCoAOMT2 CAAGACGAA------------AAATCGAGAAACTTTAAACAAAATCGATTTGGATACGAA
AM478701 CTA---------------------T---------TTGA----------------------
Z54183   GTAGATAAGTTGTTTCTTGTAGATCTCGGTTTTTAAATTTATCCCAAAAATCTTTAGAA
           *                     *         *  *

CCoAOMT1 AAATTG--CGTGTGTGATTATGGGAAGATGGTCCATTACTGATTGATGGAGGAAAACA--
CCoAOMT2 AGATTGTCCGTGTGTGATTATGAGTGGATGATCCATTACTGATTGAAGGAAGAAAACA--
AM478701 ----------TGCTTGTTTGTA-----A-----CGTTAGTG-TTGGTTTGAAAACGAA--
Z54183   GTCTTGAAACATTTTTAGTGTGTGTAATTTATAATTTAGCTAATGGTAGAATCTCAAATT
                       *   * *              ***     **            *

CCoAOMT1 ACAGGAACATCATGACAACCTCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGC
CCoAOMT2 ACAGGAACATCATGACAACCTCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGC
AM478701 ACAGGAACATCATGACTACGTCTGCTGATGAAGGGCAGTTCTTGAACATGCTTCTCAAGC
Z54183   ACAGGAATCTGATGACAACATCAGCTGATGAAGGGCAGTTCTTGAACATGCTTTTGAAGC
         *******  * ***** ** ** ** ** ** ***** **************  * ****

CCoAOMT1 TTATCAATGCTAAGAATACCATGGAGATTGGTGTTTACACTGGCTACTCCCTGCTTGCCA
CCoAOMT2 TTATTAATGCTAAGAACACCATGGAGATTGGTGTCTACACTGGCTACTCCCTGCTTGCAA
AM478701 TCATCAATGCCAAGAACACCATGGAGATAGGCGTCTACACTGGCTACTCTCTTCTGGCCA
Z54183   TCATCAATGCCAAAAACACCATGGAGATTGGTGTTTATACTGGTTATTCTCTCCTTGCCA
         * ** ***** ** ** *********** ** ** ** ***** ** ** ** ** ** *

CCoAOMT1 CTGCTCTGGCTCTCCCTGAAGATGGAAAGGTAAGGGCCAATCAACA---ACAGAGTCCAA
CCoAOMT2 CTGCTCTGGCTCTCCCTGAAGATGGAAAGGTAACGACCAATCAACA---ACAGAGTCCAA
AM478701 CAGCCCTTGCTCTCCCCGATGACGGAAAGGTCGGA---AAATGGTC---ATAGA--TCAA
Z54183   CTGCCCTGGCTCTTCCAGATGATGGAAAGGTGTGTTAATATCTATATGTATAGGA-TTTA
         * ** ** ***** ** ** ** ********        *         * **      *

CCoAOMT1 TGTTCTGAATGTTTCCATGTTACTTT-ATCAAG--ATCATCGA---AACATATCTGAGAT
CCoAOMT2 TGTTCTGAATGATTCAATGTTACTTTTATCAAG--ATCCTCCA---AACATGTCTGAAAT
AM478701 CACAGT-AATCATCCAAG--CACGTATAAAAAA--ATGTT--------A--TGGCTGAAAA
Z54183   TTTATTCTTTTCTTAAATTCTACTGCAAGTAATTTATATTTAGTTTAGAGCATTAAAAT
              *   *  *   *    **    *  **   **  *                * *
```

```
CCoAOMT1 -------------TTGGTCGG-AATTGTT---GCAGATACTGGCCATGGACATTAACAGA
CCoAOMT2 GTTTTCGTATTCTTTGGTTGGGAATTGGC---GTAGATACTGGCCATGGACATTAACAGA
AM478701 AA----------ATTGGTTCTTGATTGTT---GCAGATCCTGGCTATGGACATCAACAAA
Z54183   GACATGGATTGTGTGAATTAATCATTGTTTTTGCAGATTTTGGCGATGGATATCAACAGA
                      *   *     ****     * ****  **** ***** ** **** *

CCoAOMT1 GAAAACTACGAGTTGGGTCTTCCGGTAATTCAGAAAGCTGGTGTTGCCCACAAAATTGAG
CCoAOMT2 GAAAACTACGAATTGGGTCTTCCGGTAATTCAGAAAGCTGGTGTTGCCCACAAAATTGAG
AM478701 GAAAATTACGAGCTGGGTCTGCCAGTAATTCAAAAGGCAGGGGTTGCCCACAAGATTGAC
Z54183   GAAAACTATGAAATTGGATTACCCATTATTGAAAAAGCTGGAGTTGGTCACAAAATTGAC
         ***** ** **  * **  * **  * *** * ** ** ** ****  ***** *****

CCoAOMT1 TTCAGAGAGGGCCCTGCTCTCCCTGTTCTTGACGAACTCGTTAAAGATGTAAGCTTTTT-
CCoAOMT2 TTCAAAGAGGGCCCTGCTCTTCCTGTTCTTGACGAACTCGTTAAAGATGTAAGCTTTT--
AM478701 TTCAAAGAAGGCCCTGCTTTGCCTGTTCTTGATCAGATGATCGAAGATGTAAGCGTCAC-
Z54183   TTCAGAGAAGGCCCTGCTTTGCCTGTTCTTGATCATATGCTTGAAGATGTATGCATCATT
         **** *** ********* * ***********  *  *  *  ******** ** *

CCoAOMT1 ---CTATCTCCATCTTATCATTACAACT------------------TTCTGGT-----CT
CCoAOMT2 ---CTATCTCCTTCTTGTCTTCACCACTTCT---------------TTTTGGT-----CT
AM478701 ---AAAAACCCATCATCCAAACATTCTA------------------TTTAAAT-----CC
Z54183   CTCCAACCCCCTGCACACAGTTACACACACTAAAATGCTGCAATTTTTCAAATATTTATC
              *   **  *        *                       **    *

CCoAOMT1 TATCTCC---ATTAT-----------------------AAACATCAAATTGATAAGA--G
CCoAOMT2 TTTCTCC---ATTATCACTAGTTTTTAGGCCAA---ACGAACATCAAATTGACAAGA--G
AM478701 AAACATG---ATTGA-----------------------AACAAGCTAATTGAC-------
Z54183   AAATACGTGTATCAAATTATGATATAAGACCTAGTCATGTATATTTAATTACTCCTACTA
                   **                              *   ****

CCoAOMT1 AGGTATGATTGTGGGGGTGGTTTACAGGAAAAGAATCATGGGAGCTATGATTTCATATTC
CCoAOMT2 TCGTATGTTTGTGGGG-TGAATAGCAGGAAAAGAATCATGGGAGCTATGATTTCATATTC
AM478701 ---TTTGTTTCTGTGA-TGAAAAACAGGGCAAGTATCACGGGTCGTTCGACTTCATATTC
Z54183   ATATTCTTTTGTGCAATGAAATTTCAGGGAAAGTATCATGGAACATTTGATTTTGTATTT
            *    ** **           ****  *** **** **    *  ** **  ****

CCoAOMT1 GTGGATGCTGACAAGGACAACTACTTGAATTATCATAAGAGGTTGATCGATTTGGTGAAG
CCoAOMT2 GTGGATGCTGACAAGGACAACTATTTGAATTATCACAAGAGGTTGATCGATTTGGTGAAG
AM478701 GTGGACGCAGACAAGGACAATTATCTGAACTACCACAAGAGATTGATCGATTTGGTGAAG
Z54183   GTTGATGCTGACAAGGATAACTATATCAACTACCACAAGAGATTAATTGATTTAGTAAAA
         ** ** ** ******** ** **  * ** ** ** ***** ** ** ***** ** **

CCoAOMT1 GTAGGAGGAGTGATCGGGTACGATAACACCTTGTGGAACGGATCTGTAGTGGCACCACCA
CCoAOMT2 GTAGGAGGAGTGATAGGGTACGATAATACCTTGTGGAATGGGTCTGTGGTGGCACCACCA
AM478701 GTGGGGGGAATCATCGGCTACGACAACACCCTCTGGAACGGGTCGGTGGTGGCGCCGCCC
Z54183   ATCGGAGGACTTATCGGCTACGACAACACCCTATGGAATGGTTCTGTGGCTCAGCCAGCT
          * ** *** * ** ** ***** ** *** * ***** ** ** ** *     **  *

CCoAOMT1 GATGCTCCTCTGAGGAAGTACGTGAGGTATTACAGGGACTTTGTGCTGGAGCTCAGCAAG
CCoAOMT2 GATGCTCCTCTGAGGAAGTACGTGAGGTATTACCGGGACTTTGTGTTGGAGCTCAACAAG
AM478701 GATGCTCCGCTGCGGAAGTACGTGAGGTACTACAGAGACTTCGTGTTGGAGCTGAACAAG
Z54183   GATGCTCCAATGAGAAAGTATGTAAGGTACTACAGAGACTTTGTGATTGAGCTTAACAAA
         ********  ** * ***** ** ***** *** * ***** *** * ***** * ***
```

```
CCoAOMT1 GCTTTGGCTGTGGACCCTAGGATCGAGATCTGCATGCTTCCAGTTGGTGATGGCATCACT
CCoAOMT2 GCTTTGGCTGTGGACCCTCGGATTGAGATCTGCATGCTTCCAGTTGGTGATGGCATCACT
AM478701 GCTCTTGCTGCTGACCCAAGAATCGAGATCTGTATGCTTCCGGTTGGTGACGGGATCACC
Z54183   GCTCTGGCCGCTGATCCCAGGATTGAGATCTGTATGCTTCCTGTTGGTGATGGAGTTACC
         *** * ** *  ** **  * ** ******** ******** ******** **  * **

CCoAOMT1 CTCTGCCGTAGGATCAGCTGA
CCoAOMT2 CTCTGCCGTAGGATCAGCTGA
AM478701 CTTTGCCGTCGGCTAAGCTGA
Z54183   CTGTGCCGTCGTATCAGCTGA
         ** ****** *  * ******
```

Figure 3.27: Nucleotide sequence alignment of CCoAOMT1 and CCoAOMT2 genomic clone with genomic CCoAOMT clones of *Vitis vinifera* (AM478701) and *Petroselinum crispum* (Z54183).

3.3.8 CCoAOMT1 and CCoAOMT2 cDNA clones

Total RNA was extracted from xylem tissue of *L. leucocephala* to obtain a rich population of lignin biosynthetic pathway gene transcripts. RNA was extracted using BIOZOL reagent (Chapter 2, section 2.4.6).

cDNA first strand was synthesized using the ImProm-II Reverse Transcription System (Promega, USA), brief protocol is described in Chapter 2, section 2.4.7. PCR was setup with the cDNA first strand as template. MF4 forward primer and MR4 reverse primer were used (see section 3.3.6 above). A fragment of ~750 bp was amplified. (Fig. 3.28)

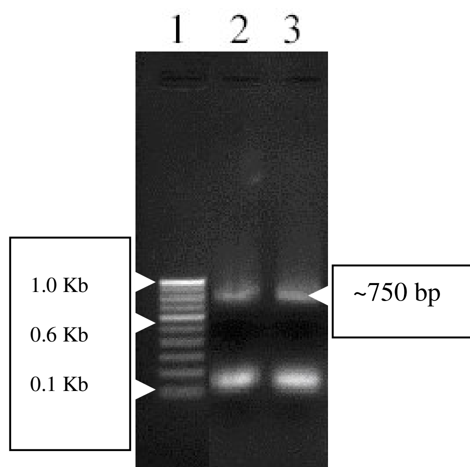


Fig. 3.28: PCR using forward primer MF4 with reverse primer MR4 and cDNA first strand as template (lane 2 and 3). DNA size marker (lane 1).

The amplicon was cloned in pGEM-T Easy Vector. The recombinant clones were restriction digested with *Sac* I since the digestion by this enzyme could discriminate between CCoAOMT1 and CCoAOMT2. As in the case of genomic clones of CCoAOMT1 and CCoAOMT2, two cDNA clones were identified. These were isolated and sequenced. The two sequences upon alignment showed 97% sequence similarity. Both sequences were of 735bp in length (Fig. 3.29). The deduced amino acid sequences showed 99% sequence similarity (Fig. 3.30). The cDNA clones have been submitted to NCBI GenBank database with designations **CCoAOMT1** and **CCoAOMT2** and their accession numbers are: **DQ431233** and **DQ431234** respectively. The nucleotide sequence of the cDNA clones **CCoAOMT1** and **CCoAOMT2** showed absolute similarity to the deduced coding sequences from the genomic DNA clones **CCoAOMT1** and **CCoAOMT2** respectively.

The deduced amino acid sequences of CCoAOMT1 and CCoAOMT2 genes show the presence of three domains, proposed to be the SAM- binding domain of O-Methyltransferases (Joshi and Chiang, 1998; Li et al., 1997 and 1999). SAM- binding domain I (LIDLVKVGGVI), domain II (VAPPDAPLRKYV) and domain III (ALAVDPRIEI) were at amino acid residues 173-183, 195-206 and 219-228 respectively (Fig. 3.30).

```
CCoAOMT1 ATGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAGAGGTTGGCCACAAGAGCCTT
CCoAOMT2 ATGGCGGATCAGAATCAAAGCGAAGCAGGAAGGCACCAAGAGGTTGGCCACAAGAGCCTT
         ************************************************************
CCoAOMT1 TTGCAGAGTGATGCTCTCTACCAGTACATTCTAGAGACCAGTGTCTACCCCAGAGAACCT
CCoAOMT2 TTGCAGAGTGATGCTCTCTACCAGTACATTCTGGAGACTAGTGTCTACCCCAGAGAACCT
         ******************************** ***** *********************
CCoAOMT1 GAACCCATGAAGGAGCTCAGAGAAATCACAGCCAAACACCCTTGGAACATCATGACAACC
CCoAOMT2 GAACCCATGAAGGAGATCAGAGAAATCACAGCCAAACACCCTTGGAACATCATGACAACC
         *************** ********************************************
CCoAOMT1 TCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGCTTATCAATGCTAAGAATACC
CCoAOMT2 TCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGCTTATTAATGCTAAGAACACC
         ******************************************** *********** ***
CCoAOMT1 ATGGAGATTGGTGTTTACACTGGCTACTCCCTGCTTGCCACTGCTCTGGCTCTCCCTGAA
CCoAOMT2 ATGGAGATTGGTGTCTACACTGGCTACTCCCTGCTTGCAACTGCTCTGGCTCTCCCTGAA
         ************** *********************** *********************
CCoAOMT1 GATGGAAAGATACTGGCCATGGACATTAACAGAGAAAACTACGAGTTGGGTCTTCCGGTA
CCoAOMT2 GATGGAAAGATACTGGCCATGGACATTAACAGAGAAAACTACGAATTGGGTCTTCCGGTA
         ******************************************** ***************
```

```
CCoAOMT1 ATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGTTCAGAGAGGGCCCTGCTCTCCCTGTT
CCoAOMT2 ATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGTTCAAAGAGGGCCCTGCTCTTCCTGTT
         ************************************* *************** ******
CCoAOMT1 CTTGACGAACTCGTTAAAGATGAAAAGAATCATGGGAGCTATGATTTCATATTCGTGGAT
CCoAOMT2 CTTGACGAACTCGTTAAAGATGAAAAGAATCATGGGAGCTATGATTTCATATTCGTGGAT
         ************************************************************
CCoAOMT1 GCTGACAAGGACAACTACTTGAATTATCATAAGAGGTTGATCGATTTGGTGAAGGTAGGA
CCoAOMT2 GCTGACAAGGACAACTATTTGAATTATCACAAGAGGTTGATCGATTTGGTGAAGGTAGGA
         ***************** *********** ******************************
CCoAOMT1 GGAGTGATCGGGTACGATAACACCTTGTGGAACGGATCTGTAGTGGCACCACCAGATGCT
CCoAOMT2 GGAGTGATAGGGTACGATAATACCTTGTGGAATGGGTCTGTGGTGGCACCACCAGATGCT
         ******** *********** *********** ** ***** ******************
CCoAOMT1 CCTCTGAGGAAGTACGTGAGGTATTACAGGGACTTTGTGCTGGAGCTCAACAAGGCTTTG
CCoAOMT2 CCTCTGAGGAAGTACGTGAGGTATTACCGGGACTTTGTGTTGGAGCTCAACAAGGCTTTG
         *************************** *********** ********************
CCoAOMT1 GCTGTGGACCCTAGGATCGAGATCTGCATGCTTCCAGTTGGTGATGGCATCACTCTCTGC
CCoAOMT2 GCTGTGGACCCTCGGATTGAGATCTGCATGCTTCCAGTTGGTGATGGCATCACTCTCTGC
         ************ **** ******************************************
CCoAOMT1 CGTAGGATCAGCTGA
CCoAOMT2 CGTAGGATCAGCTGA
         ***************
```

Fig. 3.29: Nucleotide sequence alignment of CCoAOMT1 (DQ431233) and CCoAOMT2 (DQ431234) cDNA gene clones using Clustal W 1.8.

```
CCoAOMT1 MADQNQSEAGRHQEVGHKSLLQSDALYQYILETSVYPREPEPMKELREITAKHPWNIMTT
CCoAOMT2 MADQNQSEAGRHQEVGHKSLLQSDALYQYILETSVYPREPEPMKEIREITAKHPWNIMTT
         *********************************************:**************
CCoAOMT1 SADEGQFLNMLLKLINAKNTMEIGVYTGYSLLATALALPEDGKILAMDINRENYELGLPV
CCoAOMT2 SADEGQFLNMLLKLINAKNTMEIGVYTGYSLLATALALPEDGKILAMDINRENYELGLPV
         ************************************************************
CCoAOMT1 IQKAGVAHKIEFREGPALPVLDELVKDEKNHGSYDFIFVDADKDNYLNYHKRLIDLVKVG
CCoAOMT2 IQKAGVAHKIEFKEGPALPVLDELVKDEKNHGSYDFIFVDADKDNYLNYHKRLIDLVKVG
         ************:***********************************************
```

```
CCoAOMT1 GVIGYDNTLWNGSVVAPPDAPLRKYVRYYRDFVLELNKALAVDPRIEICMLPVGDGITLC
CCoAOMT2 GVIGYDNTLWNGSVVAPPDAPLRKYVRYYRDFVLELNKALAVDPRIEICMLPVGDGITLC
         ************************************************************
CCoAOMT1 RRIS
CCoAOMT2 RRIS
         ****
```

Fig. 3.30: Deduced amino acid sequence alignment of CCoAOMT1 (DQ431233) and CCoAOMT2 (DQ431234) cDNA genes using Clustal W 1.8. The highly divergent N terminal amino acids vis a vis other plant CCoAOMT is shaded at the N terminus of the amino acid sequence. The three SAM binding domains are also shaded.

The methyltransf_3 domain is a feature of the O-methyltransferase family. The family includes catechol O-methyltransferase, caffeoyl-CoA O-methyltransferase and the bacterial O-methyltransferases that may be involved in antibiotic production (Marchler-Bauer et al., 2004). The deduced amino acid sequences of the CCoAOMT 1 and CCoAOMT 2 were searched in the NCBI database for the presence of the conserved methyltransf_3 domain. The 213 amino acids domain present in all the CCoAOMT proteins reported from different plants was located here too (Fig. 3.31).

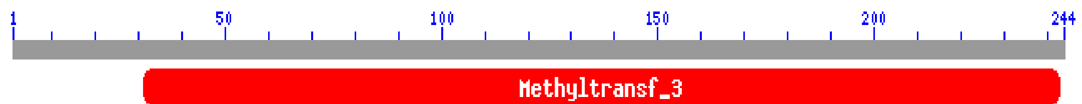


Fig. 3.31: Image showing Methyltransf_3 conserved domain present in CCoAOMT1 and CCoAOMT2.

Restriction analysis of CCoAOMT 1 and CCoAOMT 2 was done using pDRAW32 (Fig. 3.32 a, b). Analysis was limited to enzyme site cutting not more than four times in the sequence.

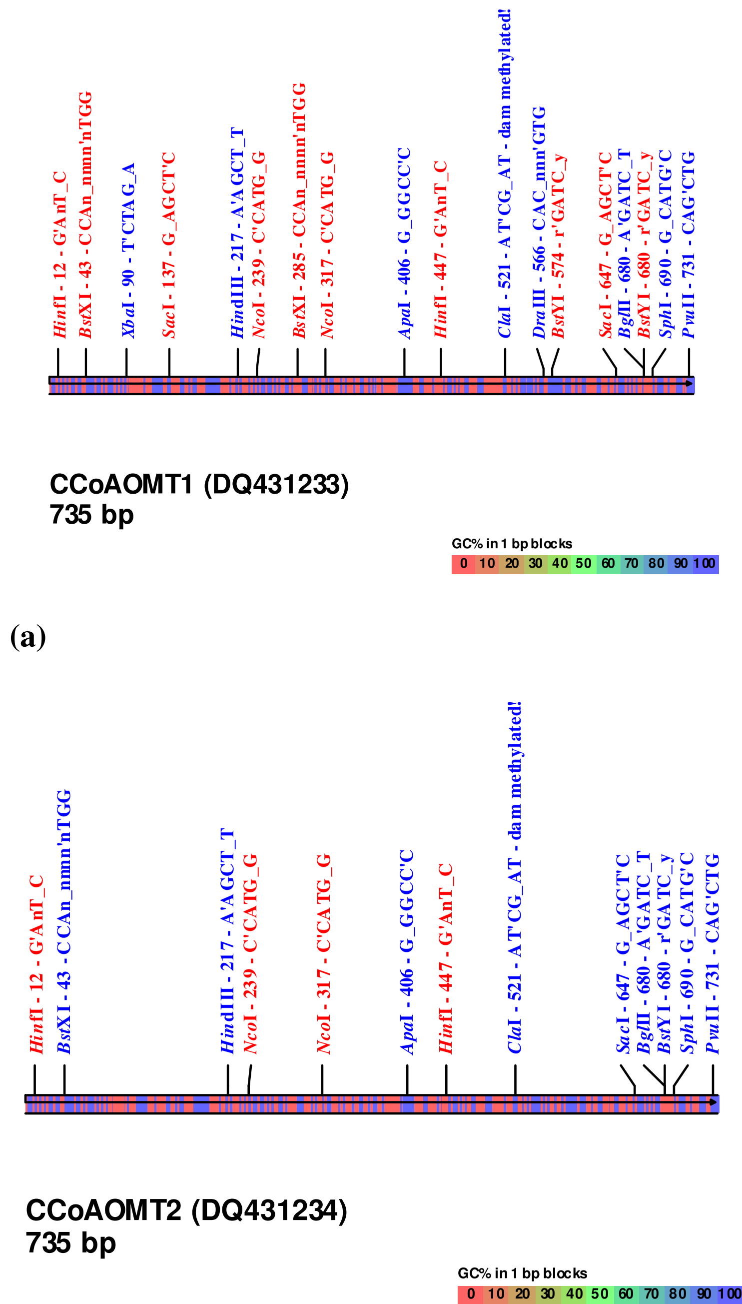


Fig. 3.32: (a) and (b) - Restriction analysis of CCoAOMT1 (DQ431233) and CCoAOMT2 (DQ431234) respectively using pDRAW 32.

The CCoAOMT1 and the CCoAOMT2 cDNA gene sequences and their respective deduced amino acids sequences were searched for sequence similarity in the GenBank data base

(NCBI) using BLAST. The nucleotide sequence alignment of the CCoAOMT1 and CCoAOMT2 gene sequences with representative member of the dicots *e.g. Nicotiana tabacum* U62735, the monocots *e.g Zea mays* AJ242981 and the gymnosperms e.g *Pinus taeda* AF036095 (Fig. 3.33) showed 79%, 75% and 76% sequence similarity respectively.

```
DQ431233     -------------------------------------------------ATGGCGGATCAG
DQ431234     -------------------------------------------------ATGGCGGATCAG
U62735       -------------------------------------------------------ATGGC
AF036095     ---------------ATGGCAAGCACAAGTGTTGCTGCAGCAGAGGTGAAGGCTCAGACA
AJ242981     ATGGCCACCACGGCGACCGAGGCGACCAAGACGACTGCACCGGCGCAGGAGCAGCAGGCC
                                                                        *

DQ431233     AAT--CAA--AGCGAAGCAGG--AAGG--------CACCAAGAGGTTGGCCACAAGAGCC
DQ431234     AAT--CAA--AGCGAAGCAGG--AAGG--------CACCAAGAGGTTGGCCACAAGAGCC
U62735       AAC--CAAT-GGAGAAAATGG--AAGA--------CATCAAGAAGTTGGACACAAGAGTC
AF036095     ACC--CAAGCAGAGGAGCCGGTTAAGGTTGTCCGCCATCAAGAAGTGGGACACAAAAGTC
AJ242981     AACGGCAACGGCAACGGCGAGCAGAAGACG--CGCCACTCCGAGGTCGGCCACAAGAGCC
             *    ***            *   *          **    ** ** ** ***** ** *

DQ431233     TTTTGCAGAGTGATGCTCTCTACCAGTACATTCTAGAGACCAGTGTCTACCCCAGAGAAC
DQ431234     TTTTGCAGAGTGATGCTCTCTACCAGTACATTCTGGAGACTAGTGTCTACCCCAGAGAAC
U62735       TTTTGCAAAGTGATGCCCTTTACCAGTACATTCTTGAAACAAGCGTGTATCCAAGAGAGC
AF036095     TCTTGCAGAGCGATGCCCTCTATCAGTATATATTGGAAACGAGTGTGTACCCTCGAGAGC
AJ242981     TGCTCAAGAGCGACGACCTCTACCAGTACATCCTGGACACGAGCGTGTACCCGCGGGAGC
             *  * ** ** *  ** ** ***** **  * ** ** ** ** ** **  * ** *

DQ431233     CTGAACCCATGAAGGAGCTCAGAGAAATCACAGCCAAACACCCTTGGAACATCATGACAA
DQ431234     CTGAACCCATGAAGGAGATCAGAGAAATCACAGCCAAACACCCTTGGAACATCATGACAA
U62735       CTGAACCCATGAAAGAGCTAAGAGAGATTACTGCAAAACACCCCTGGAACATCATGACCA
AF036095     CTGAGCCAATGAAGGAGCTCCCGAGAGTGACTGCCAAGCATCCCTGGAACCTCATGACTA
AJ242981     CGGAGAGCATGAAGGAGCTGCGCGAGATCACCGCCAAGCACCCATGGAACCTGATGACCA
             * **    ***** *** *        * ** ** ** ** ** ****** * ***** *

DQ431233     CCTCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGCTTATCAATGCTAAGAATA
DQ431234     CCTCAGCAGACGAGGGGCAATTCTTGAACATGCTCCTTAAGCTTATTAATGCTAAGAACA
U62735       CCTCTGCTGATGAAGGGCAATTTTTGAGCATGCTTCTTAAACTCATCAATGCCAAAAACA
AF036095     CTTCTGCCGATGAGGGTCAATTTCTGGGCCTCCTGCTGAAGCTCATTAACGCCAAGAACA
AJ242981     CCTCCGCCGACGAGGGCCAGTTCCTCAACATGCTCATCAAGCTCATCGGCGCCAAGAAGA
             * ** ** ** ** ** ** **  *   * * **  * ** ** **    ** ** ** *

DQ431233     CCATGGAGATTGGTGTTTACACTGGCTACTCCCTGCTTGCCACTGCTCTGGCTCTCCCTG
DQ431234     CCATGGAGATTGGTGTCTACACTGGCTACTCCCTGCTTGCAACTGCTCTGGCTCTCCCTG
U62735       CAATGGAGATTGGTGTTTTTACTGGTTACTCTCTTCTTGCTACTGCCATGGCTCTTCCTG
AF036095     CCATGGAGATTGGGGTGTACACTGGTTACTCGCTTCTCAGCACAGCCCTTGCATTGCCCG
AJ242981     CCATGGAGATCGGCGTCTACACCGGCTACTCGCTCCTCGCCACCGCGCTCGCACTCCCGG
             * ******** ** ** *  ** ** ***** ** **    ** **  * **  * ** *

DQ431233     AAGATGGAAAGATACTGGCCATGGACATTAACAGAGAAAACTACGAGTTGGGTCTTCCGG
DQ431234     AAGATGGAAAGATACTGGCCATGGACATTAACAGAGAAAACTACGAATTGGGTCTTCCGG
U62735       ATGATGGCAAGATTTTAGCCATGGATATCAACCGGGACAACTACGAGATTGGTCTTCCAG
AF036095     ATGATGGAAAGATTCTAGCCATGGACATCAACAGAGAGAACTATGATATCGGATTGCCTA
AJ242981     AGGACGGCACGATCTTGGCCATGGACATCAACCGCGAGAACTACGAGCTAGGCCTTCCCT
             * ** ** * ***  * ******** ** *** * ** ***** **  * **  * **
```

```
DQ431233      TAATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGTTCAGAGAGGGCCCTGCTCTCCCTG
DQ431234      TAATTCAGAAAGCTGGTGTTGCCCACAAAATTGAGTTCAAAGAGGGCCCTGCTCTTCCTG
U62735        TAATTGAAAAGGCTGGACTAGCTCACAAAATTGAATTCAAAGAAGGCCCTGCACTTCCCG
AF036095      TTATTGAGAAAGCAGGAGTTGCCCACAAGATTGACTTCAGAGAGGGCCCTGCTCTGCCAG
AJ242981      GCATCAACAAGGCCGGCGTGGGCCACAAGATCGACTTCCGCGAGGGCCCCGCGCTCCCCG
              **   * ** ** **   * *  ***** ** ** ***    ** ***** ** ** ** *

DQ431233      TTCTTGACGAACTCGTTAAAGATGAAAAGAATCATGGGAGCTATGATTTCATATTCGTGG
DQ431234      TTCTTGACGAACTCGTTAAAGATGAAAAGAATCATGGGAGCTATGATTTCATATTCGTGG
U62735        TTCTTGATCAAATGATTGAAGACGGCAAATACCATGGATCATATGACTTCATATTTGTGG
AF036095      TTCTGGACGAACTGCTTAAGAATGAGGACATGCATGGATCGTTCGACTTTGTGTTCGTGG
AJ242981      TCCTGGACGACCTCGTGGCGGACAAGGAGCAGCACGGGTCGTTCGACTTCGCCTTCGTGG
              * ** **  *  *  *      *      *     ** **     *  ** **    ** ****

DQ431233      ATGCTGACAAGGACAACTACTTGAATTATCATAAGAGGTTGATCGATTTGGTGAAGGTAG
DQ431234      ATGCTGACAAGGACAACTATTTGAATTATCACAAGAGGTTGATCGATTTGGTGAAGGTAG
U62735        ACGCTGACAAAGACAATTACTTGAACTATCACAAGAGATTAATCGACTTGGTCAAAGTTG
AF036095      ATCGGGACAAAGACAACTATCTAAACTACCACAAGCGTCTGATCGATCTGGTGAAGGTTG
AJ242981      ACGCCGACAAGGACAACTACCTCAACTACCACGAGCGGCTCCTGAAGCTGGTGAGGCCCG
              *    ***** ***** **  * ** ** **  ** *  *  *  *  **** *      *

DQ431233      GAGGAGTGATCGGGTACGATAACACCTTGTGGAACGGATCTGTAGTGGCACCACCAGATG
DQ431234      GAGGAGTGATAGGGTACGATAATACCTTGTGGAATGGGTCTGTGGTGGCACCACCAGATG
U62735        GGGGACTAATTGGATATGACAACACCCTATGGAATGGATCAGTGGTTGCACCACCAGATG
AF036095      GAGGTCTGATTGCATATGACAACACCCTGTGGAACGGATCTGTGGTGGCTCCACCCGATG
AJ242981      GCGGCCTCATCGGCTACGACAACACGCTGTGGAACGGCTCCGTCGTGCTCCCCGACGACG
              * **  * ** *  ** ** ** **  * ***** ** ** ** **     **     ** *

DQ431233      CTCCTCTGAGGAAGTACGTGAGGTATTACAGGGACTTTGTGCTGGAGCTCAGCAAGGCTT
DQ431234      CTCCTCTGAGGAAGTACGTGAGGTATTACCGGGACTTTGTGTTGGAGCTCAACAAGGCTT
U62735        CACCCCTTAGGAAATATGTTAGGTATTATAGGGATTTCGTATTGGAACTCAACAAGGCTT
AF036095      CTCCCCTGAGGAAATATGTGAGATATTACAGAGATTTCGTGATGGAGCTAAACAAGGCCC
AJ242981      CGCCCATGCGCAAGTACATCCGCTTCTACCGCGACTTCGTCCTCGCCCTCAACAGCGCGC
              * **  *  * ** **  *  * *  **  * ** ** **  * *  ** * **  **

DQ431233      TGGCTGTGGACCCTAGGATCGAGATCTGCATGCTTCCAGTTGGTGATGGCATCACTCTCT
DQ431234      TGGCTGTGGACCCTCGGATTGAGATCTGCATGCTTCCAGTTGGTGATGGCATCACTCTCT
U62735        TGGCTGCTGATTCAAGAATTGAAATCTGTCAGCTTCCCGTTGGGGATGGCATCACTCTCT
AF036095      TTGCTGTCGATCCCCGCATTGAGATCAGCCAAATCCCAGTGCTGGACGGCGTCACCCTTT
AJ242981      TCGCCGCCGACGACCGCGTCGAGATCTGCCAGCTCCCCGTCGGCGACGGCGTCACGCTCT
              * ** *  **      *  * ** *** *      * ** **     ** *** **** ** *

DQ431233      GCCGTAGGATCAGCTGA
DQ431234      GCCGTAGGATCAGCTGA
U62735        GCCGCCGCATCAGTTAA
AF036095      GCAGGCGTGTCTATTGA
AJ242981      GCCGCCGCGTCAAGTGA
              ** *  *  **   * *
```

Fig. 3.33: Nucleotide sequence alignment of CCoAOMT1 (DQ431233) and CCoAOMT2 (DQ431234) with *Nicotiana tabacum* (U62735), *Zea mays (*AJ242981) and *Pinus taeda (*AF036095) using Clustal W 1.8.

The deduced amino acid sequences of CCoAOMT1 and CCoAOMT2 were aligned with CCoAOMT amino acid sequences of *Nicotiana tabacum (*U62735), *Zea mays (*AJ242981) and *Pinus taeda (*AF036095). Respective sequence similarity was 90%, 82% and 86% (Fig. 3.34).

```
DQ431233    -----------MADQNQSEA---------GRHQEVGHKSLLQSDALYQYILETSVYPREP
DQ431234    -----------MADQNQSEA---------GRHQEVGHKSLLQSDALYQYILETSVYPREP
U62735      -----------MATN--GEN---------GRHQEVGHKSLLQSDALYQYILETSVYPREP
AF036095    MASTSVAAAEVKAQTTQAEEP-----VKVVRHQEVGHKSLLQSDALYQYILETSVYPREP
AJ242981    MATTATEATKTTAPAQEQQANGNGNGEQKTRHSEVGHKSLLKSDDLYQYILDTSVYPREP
                        *     :           **.********:** ******:********

DQ431233    EPMKELREITAKHPWNIMTTSADEGQFLNMLLKLINAKNTMEIGVYTGYSLLATALALPE
DQ431234    EPMKEIREITAKHPWNIMTTSADEGQFLNMLLKLINAKNTMEIGVYTGYSLLATALALPE
U62735      EPMKELREITAKHPWNIMTTSADEGQFLSMLLKLINAKNTMEIGVFTGYSLLATAMALPD
AF036095    EPMKELPRVTAKHPWNLMTTSADEGQFLGLLLKLINAKNTMEIGVYTGYSLLSTALALPD
AJ242981    ESMKELREITAKHPWNLMTTSADEGQFLNMLIKLIGAKKTMEIGVYTGYSLLATALALPE
            *.***: .:*******:***********.:*:***.**:******:******:**:***:

DQ431233    DGKILAMDINRENYELGLPVIQKAGVAHKIEFREGPALPVLDELVKDEKNHGSYDFIFVD
DQ431234    DGKILAMDINRENYELGLPVIQKAGVAHKIEFKEGPALPVLDELVKDEKNHGSYDFIFVD
U62735      DGKILAMDINRDNYEIGLPVIEKAGLAHKIEFKEGPALPVLDQMIEDGKYHGSYDFIFVD
AF036095    DGKILAMDINRENYDIGLPIIEKAGVAHKIDFREGPALPVLDELLKNEDMHGSFDFVFVD
AJ242981    DGTILAMDINRENYELGLPCINKAGVGHKIDFREGPALPVLDDLVADKEQHGSFDFAFVD
            **.********:**::*** *:***:.***:*:*********::: : . ***:** ***

DQ431233    ADKDNYLNYHKRLIDLVKVGGVIGYDNTLWNGSVVAPPDAPLRKYVRYYRDFVLELNKAL
DQ431234    ADKDNYLNYHKRLIDLVKVGGVIGYDNTLWNGSVVAPPDAPLRKYVRYYRDFVLELNKAL
U62735      ADKDNYLNYHKRLIDLVKVGGLIGYDNTLWNGSVVAPPDAPLRKYVRYYRDFVLELNKAL
AF036095    RDKDNYLNYHKRLIDLVKVGGLIAYDNTLWNGSVVAPPDAPLRKYVRYYRDFVMELNKAL
AJ242981    ADKDNYLNYHERLLKLVRPGGLIGYDNTLWNGSVVLPDDAPMRKYIRFYRDFVLALNSAL
             *********:**:.**: **:*.*********** * ***:***:*:*****: **.**

DQ431233    AVDPRIEICMLPVGDGITLCRRIS
DQ431234    AVDPRIEICMLPVGDGITLCRRIS
U62735      AADSRIEICQLPVGDGITLCRRIS
AF036095    AVDPRIEISQIPVLDGVTLCRRVY
AJ242981    AADDRVEICQLPVGDGVTLCRRVK
            *.* *:**. :** **:*****:
```

Fig. 3.34: Deduced amino acid sequence alignment of CCoAOMT1 (DQ431233) and CCoAOMT2 (DQ431234) with CCoAOMT amino acid sequences of *Nicotiana tabacum* (U62735), *Zea mays* (AJ242981) and *Pinus taeda* (AF036095) using Clustal W 1.8.

The GC% and codon usage of the two CCoAOMT cDNA clones was calculated using online software (**www.justbio.com**). The GC content of CCoAOMT1 and CCoAOMT2 was 47.52 % and 46.70 % respectively. The codon usage for CCoAOMT1 and CCoAOMT2 is given in Table 3.1 (a, b) and it is expressed as % of total codons. Standard genetic codes were used for the purpose.

Table 3.1 (a): Codon usage of CCoAOMT1

Phe	UUU	1	0.40	Ser	UCU	1	0.67	Tyr	UAU	3	0.46	Cys	**UGU**	**0**	**0.00**
	UUC	4	1.60		UCC	1	0.67		UAC	10	1.54		UGC	2	2.00
Leu	**UUA**	**0**	**0.00**		UCA	1	0.67	Ter	**UAA**	**0**	**0.00**	Ter	UGA	1	3.00
	UUG	8	1.66		**UCG**	**0**	**0.00**		**UAG**	**0**	**0.00**	Trp	UGG	2	1.00
Leu	CUU	7	1.45	Pro	CCU	7	2.15	His	CAU	2	0.67	Arg	CGU	1	0.50
	CUC	8	1.66		CCC	2	0.62		CAC	4	1.33		**CGC**	**0**	**0.00**
	CUA	1	0.21		CCA	3	0.92	Gln	CAA	3	0.86		**CGA**	**0**	**0.00**
	CUG	5	1.03		CCG	1	0.31		CAG	4	1.14		**CGG**	**0**	**0.00**
Ile	AUU	5	0.94	Thr	ACU	3	1.33	Asn	AAU	5	0.77	Ser	AGU	2	1.33
	AUC	9	1.69		ACC	4	1.78		AAC	8	1.23		AGC	4	2.67
	AUA	2	0.38		ACA	2	0.89	Lys	AAA	4	0.53	Arg	AGA	4	2.00
Met	AUG	7	1.00		**ACG**	**0**	**0.00**		AAG	11	1.47		AGG	7	3.50
Val	GUU	6	1.41	Ala	GCU	10	2.22	Asp	GAU	10	1.18	Gly	GGU	4	1.00
	GUC	1	0.24		GCC	4	0.89		GAC	7	0.82		GGC	4	1.00
	GUA	3	0.71		GCA	3	0.67	Glu	GAA	8	0.89		GGA	5	1.25
	GUG	7	1.65		GCG	1	0.22		GAG	10	1.11		GGG	3	0.75

Table 3.1 (b): Codon usage of CCoAOMT2

Phe	UUU	1	0.40	Ser	UCU	1	0.67	Tyr	UAU	4	0.62	Cys	**UGU**	**0**	**0.00**
	UUC	4	1.60		UCC	1	0.67		UAC	9	1.38		UGC	2	2.00
Leu	**UUA**	**0**	**0.00**		UCA	1	0.67	Ter	**UAA**	**0**	**0.00**	Ter	UGA	1	3.00
	UUG	9	1.93		**UCG**	**0**	**0.00**		**UAG**	**0**	**0.00**	Trp	UGG	2	1.00
Leu	CUU	8	1.71	Pro	CCU	7	2.15	His	CAU	1	0.33	Arg	CGU	1	0.55
	CUC	6	1.29		CCC	2	0.62		CAC	5	1.67		**CGC**	**0**	**0.00**
	CUA	**0**	**0.00**		CCA	3	0.92	Gln	CAA	3	0.86		**CGA**	**0**	**0.00**
	CUG	5	1.07		CCG	1	0.31		CAG	4	1.14		CGG	2	1.09
Ile	AUU	7	1.24	Thr	ACU	4	1.78	Asn	AAU	6	0.92	Ser	AGU	2	1.33
	AUC	7	1.24		ACC	3	1.33		AAC	7	1.08		AGC	4	2.67
	AUA	3	0.53		ACA	2	0.89	Lys	AAA	5	0.62	Arg	AGA	3	1.64
Met	AUG	7	1.00		**ACG**	**0**	**0.00**		AAG	11	1.38		AGG	5	2.73
Val	GUU	5	1.18	Ala	GCU	10	2.22	Asp	GAU	10	1.18	Gly	GGU	4	1.00
	GUC	2	0.47		GCC	3	0.67		GAC	7	0.82		GGC	4	1.00
	GUA	2	0.47		GCA	4	0.89	Glu	GAA	9	1.00		GGA	4	1.00
	GUG	8	1.88		GCG	1	0.22		GAG	9	1.00		GGG	4	1.00

3.3.9 5' and 3' RACE

The CCoAOMT1 and CCoAOMT2 cDNA gene clones were without their 5' and 3' UTR's. Hence, 5'and 3' RACE was performed to isolate the 5' and the 3' UTR's. The BD Clontech Smart RACE kit was used for the purpose. A brief protocol is described in Chapter 2, section 2.4.9.

cDNA first strand was synthesized from RNA isolated from xylem tissue of *L. leucocephala*. The primer MF4 was used as forward primer and MR4 was used as reverse primer for 3' and 5' RACE.

Approximately 850 bp and 1000 bp PCR amplicons was amplified for 5' and 3'RACE respectively (Fig. 3.35). These amplicons were cloned in pGEM-T Easy Vector and sequenced. The nucleotide sequences of the CCoAOMT1 and CCoAOMT2 cDNA gene clones along with their 5' and 3' UTRs are shown in Fig. 3.36 a and b. The CCoAOMT1 and CCoAOMT2 have 3' UTRs of 189 bp and 176 bp respectively. The CCoAOMT1 and CCoAOMT2 5' UTRs are of 57 bp each. Putative polyadenylation signal in CCoAOMT1 and CCoAOMT2 are at position 930-935 and 917-922 respectively.

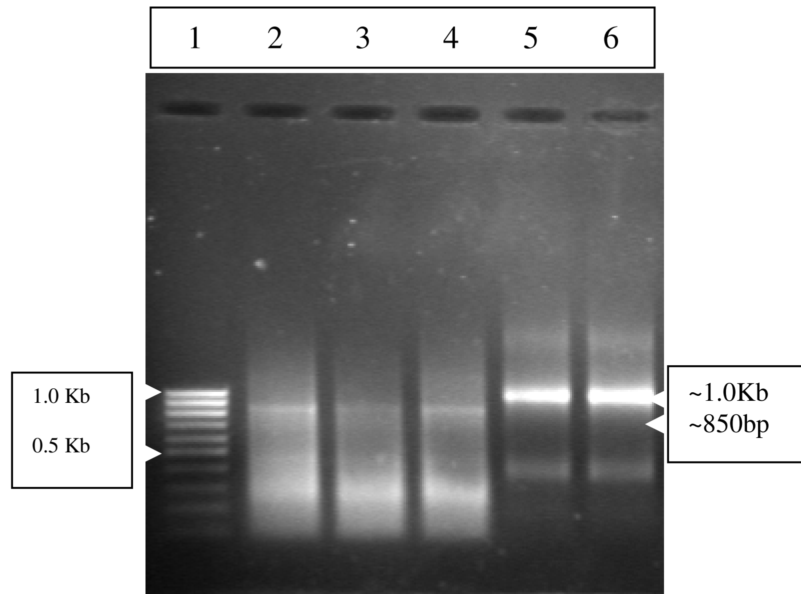


Fig. 3.35: 5' and 3' RACE using primers MR4 and MF4 respectively. DNA size marker (lane 1), 5' RACE product (lanes 2, 3 and 4) and 3'RACE product (lanes 5 and 6).

A nucleotide sequence comparison of the 3' RACE products of CCoAOMT1 and CCoAOMT2 with the CCoAOMT 3' Genome walking product nucleotide sequence (gMC3'UTR; see section 3.3.5.2 above) was done. The 3' Genome walking product shows absolute sequence similarity with the 3' RACE product of CCoAOMT2 gene. Thus concluding that the 3' Genome walking product was from CCoAOMT2 gene.

CCoAOMT1

```
  1 aaaacagagc agaagaaaag aagagagagg tcgaggaaac aaaaaggata agaagtaATG
 61 GCGGATCAGA ATCAAAGCGA AGCAGGAAGG CACCAAGAGG TTGGCCACAA GAGCCTTTTG
121 CAGAGTGATG CTCTCTACCA GTACATTCTA GAGACCAGTG TCTACCCCAG AGAACCTGAA
181 CCCATGAAGG AGCTCAGAGA AATCACAGCC AAACACCCTT GGAACATCAT GACAACCTCA
241 GCAGACGAGG GGCAATTCTT GAACATGCTC CTTAAGCTTA TCAATGCTAA GAATACCATG
301 GAGATTGGTG TTTACACTGG CTACTCCCTG CTTGCCACTG CTCTGGCTCT CCCTGAAGAT
361 GGAAAGATAC TGGCCATGGA CATTAACAGA GAAAACTACG AGTTGGGTCT TCCGGTAATT
421 CAGAAAGCTG GTGTTGCCCA CAAAATTGAG TTCAGAGAGG GCCCTGCTCT CCCTGTTCTT
481 GACGAACTCG TTAAAGATGA AAAGAATCAT GGGAGCTATG ATTTCATATT CGTGGATGCT
541 GACAAGGACA ACTACTTGAA TTATCATAAG AGGTTGATCG ATTTGGTGAA GGTAGGAGGA
601 GTGATCGGGT ACGATAACAC CTTGTGGAAC GGATCTGTAG TGGCACCACC AGATGCTCCT
661 CTGAGGAAGT ACGTGAGGTA TTACAGGGAC TTTGTGCTGG AGCTCAACAA GGCTTTGGCT
721 GTGGACCCTA GGATCGAGAT CTGCATGCTT CCAGTTGGTG ATGGCATCAC TCTCTGCCGT
781 AGGATCAGCT GAtttctttg taccctcttg caccacgtca tttccctcta caaaaatatt
841 tctgtctgta tttttttct ccaaattatt ccgagtgcat ggtgcttaat tgggagtgtt
901 ttatattttt catattcttg ccttttcata ttatatagaa tccagtaatc ataaaaaaaa
961 aaaaaaaaaa aaaaaaaaaa
```

(a)

CCoAOMT2

```
  1 aaaacagagc agaagaaaag aagagagagt tcaaggaaac aaaaaaggat aagaagtATG
 61 GCGGATCAGA ATCAAAGCGA AGCAGGAAGG CACCAAGAGG TTGGCCACAA GAGCCTTTTG
121 CAGAGTGATG CTCTCTACCA GTACATTCTG GAGACTAGTG TCTACCCCAG AGAACCTGAA
181 CCCATGAAGG AGATCAGAGA AATCACAGCC AAACACCCTT GGAACATCAT GACAACCTCA
241 GCAGACGAGG GGCAATTCTT GAACATGCTC CTTAAGCTTA TTAATGCTAA GAACACCATG
301 GAGATTGGTG TCTACACTGG CTACTCCCTG CTTGCAACTG CTCTGGCTCT CCCTGAAGAT
361 GGAAAGATAC TGGCCATGGA CATTAACAGA GAAAACTACG AATTGGGTCT TCCGGTAATT
421 CAGAAAGCTG GTGTTGCCCA CAAAATTGAG TTCAAAGAGG GCCCTGCTCT TCCTGTTCTT
```

481 GACGAACTCG TTAAAGATGA AAAGAATCAT GGGAGCTATG ATTTCATATT CGTGGATGCT

541 GACAAGGACA ACTATTTGAA TTATCACAAG AGGTTGATCG ATTTGGTGAA GGTAGGAGGA

601 GTGATAGGGT ACGATAATAC CTTGTGGAAT GGGTCTGTGG TGGCACCACC AGATGCTCCT

661 CTGAGGAAGT ACGTGAGGTA TTACCGGGAC TTTGTGTTGG AGCTCAACAA GGCTTTGGCT

721 GTGGACCCTC GGATTGAGAT CTGCATGCTT CCAGTTGGTG ATGGCATCAC TCTCTGCCGT

781 AGGATCAGCT GA*tttctctg gaccttcctg caccgcgaag ggctaaaaat atttctatct*

841 gtattttttt ccaaaccatt ctgagcatgg ggcttaattg ggagtgtttt atattttcat

901 attcttgctt tttcatatta tattgaatcc agtaatcata aaaaaaaaaa aaaaaaaaaa

961 aaaaaaaa

(b)

Fig. 3.36: Nucleotide sequence of (a) CCoAOMT1 and (b) CCoAOMT2 with their respective 5' and 3' UTRs. Both the 5' and 3' UTRs are in lower case italics' Start and stop codons are in shaded bold capitals. Polyadenylation signal is indicated in shaded small case.

3.3.10 Distance Tree results

The nucleotide sequence and deduced amino acid sequences of CCoAOMT1 and CCoAOMT2 cDNA genes were used as query in BLAST search (NCBI) and matching sequences from different members of plant groups were selected for constructing Distance Tree (Fast minimum evolution method using maximum sequence difference of 0.75). Resultant trees clearly groups the CCoAOMT1 and CCoAOMT2 genes sequences from *L. leucocephala* with those of the other reported Fabaceae family members (Fig. 3.37 a, b; 3.38 a, b).

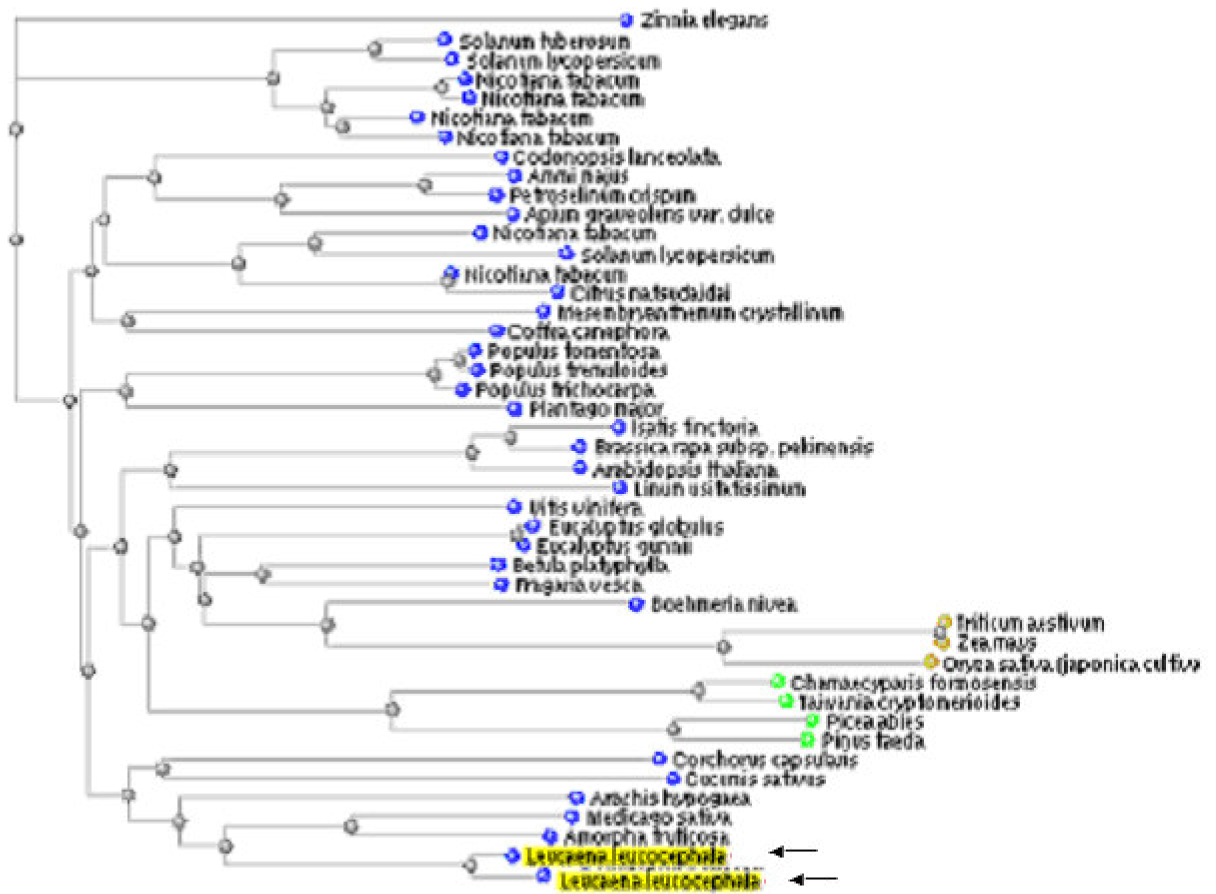


Fig. 3.37 a: Distance tree (Sequence label-Taxonomic name) for CCoAOMT1 and CCoAOMT2 nucleotide sequences. *L. leucocephala* position is indicated by arrows.

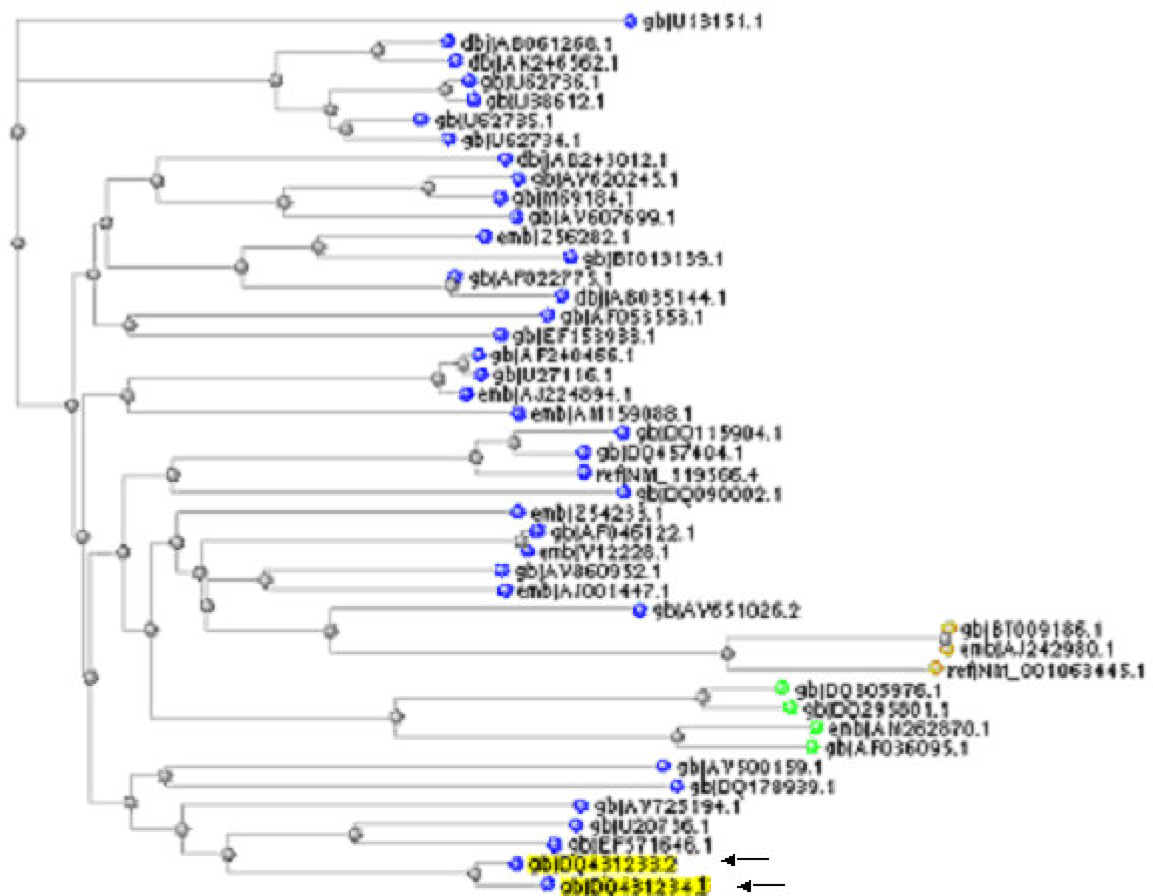


Fig. 3.37 b: Distance tree (Sequence label – sequence ID) for CCoAOMT1 and CCoAOMT2 nucleotide sequences. *L. leucocephala* position is indicated by arrows.

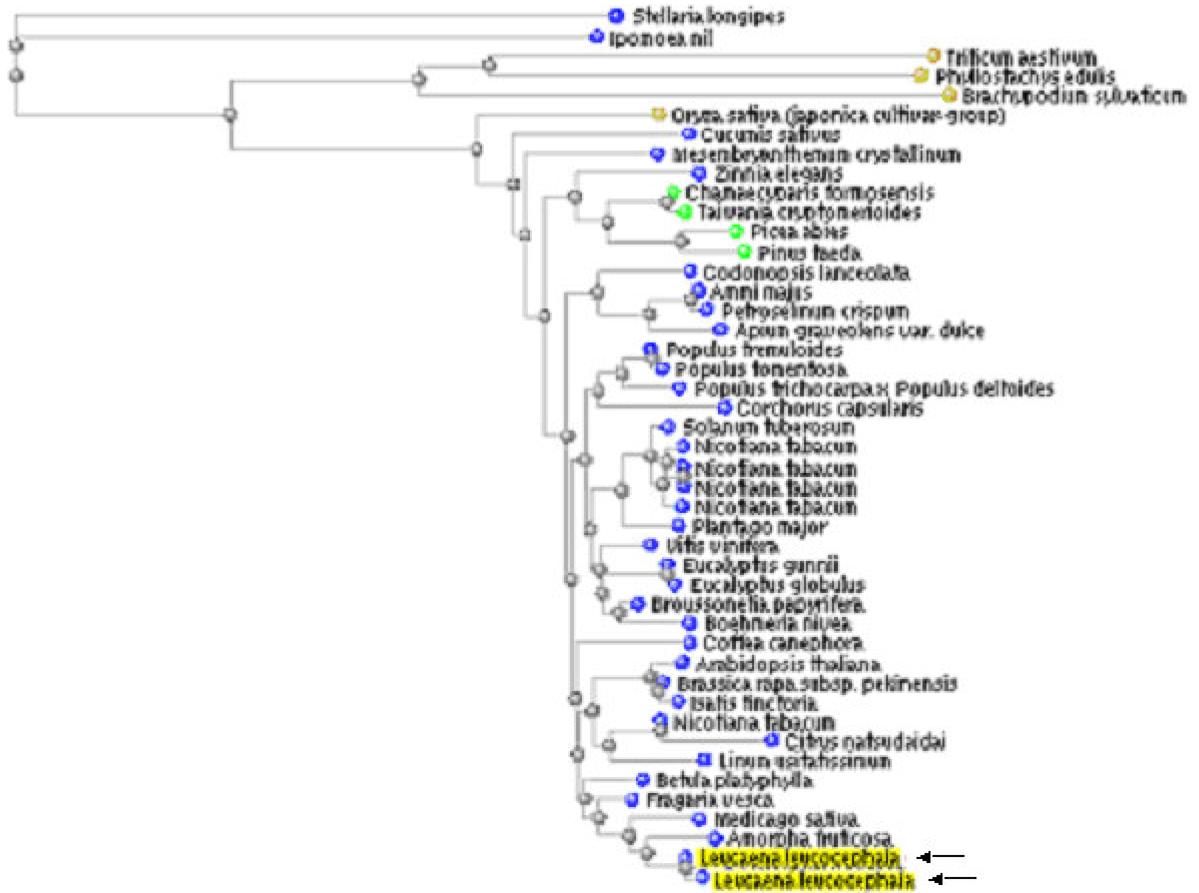


Fig. 3.38 a: Distance tree (Sequence label – Taxonomic name) for CCoAOMT1 and CCoAOMT2 amino acid sequences. *L. leucocephala* position is indicated by arrows.

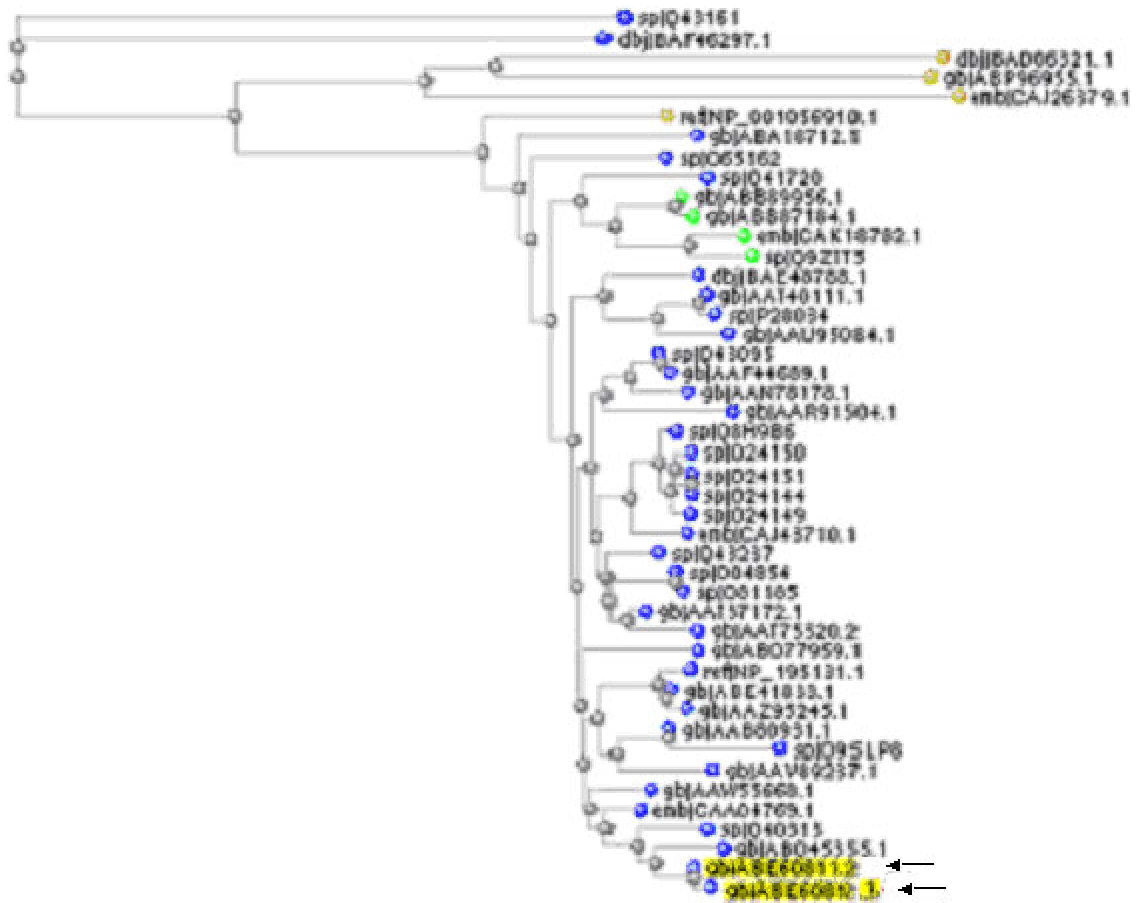


Fig. 3.38 b: Distance tree (Sequence label – sequence ID) for CCoAOMT1 and CCoAOMT2 amino acid sequences. *L. leucocephala* position is highlighted.

3.4 Conclusions

- CCoAOMT in *L. leucocephala* is gene family of possibly 3 members.
- CCoAOMT1 and CCoAOMT2 two genomic gene clones, 1292 bp each, isolated. Their NCBI GenBank database accession numbers are DQ517929 and DQ517930 respectively. The genes show 92.7% nucleotide sequence similarity with each other. Nucleotide sequence similarity with CCoAOMT genes from other plants is between 70-90%.
- CCoAOMT1 and CCoAOMT2 genomic clones comprised of five exons and four introns. Deduced coding sequence of both the genomic CCoAOMT1 and CCoAOMT2 is of 735 nucleotides.
- Two 5' upstream (313 and 319 bp) and one 3' downstream (265 bp) nucleotide sequence of CCoAOMT1 and CCoAOMT2 isolated by Genome walking.
- The 3' downstream UTR is from CCoAOMT2 gene.
- Two cDNA clones of CCoAOMT1 and CCoAOMT2, 735 bp each, isolated. Their NCBI GenBank database accession numbers are DQ431233 and DQ431234. They show 97% nucleotide and 99% deduced amino acid sequence similarity with each other. Nucleotide sequence similarity with CCoAOMT cDNA gene clones from other plants was 70-90%.
- Deduced amino acid sequences of CCoAOMT1 and CCoAOMT2 genes show presence of SAM binding domain I (LIDLVKVGGVI), domain II (VAPPDAPLRKYV) and domain III (ALAVDPRIEI).
- Two 57 bp long 5'UTR associated with both the CCoAOMT1 and CCoAOMT2 cDNA genes isolated using 5' RACE.
- Two 3'UTRs of 189 bp and 176 bp associated with the two CCoAOMT1 and CCoAOMT2 cDNA genes respectively isolated by 3' RACE.
- Distance tree results group the CCoAOMT1 and CCoAOMT2 genes from *L. leucocephala* with CCoAOMTs of other Fabaceae members.

3.5 References

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54**:** 519–546.

Busam G, Junghanns K., Kneusel RE, Kassemeyer HH, Matern,U (1997) Characterization and expression of caffeoyl-coenzyme A 3-O-methyltransferase proposed for the induced resistance response of *Vitis vinifera* L. Plant Physiol 115: 1039-1048

Busam G, Grimmig B, Kneusel RE, Matern U (1997) The electronic Plant Gene Register. Plant Physiol 113: 1003-1005

Chen C, Meyermans H, Burggraeve B, De Rycke RM, Inoue K, Vleesschauwer VD, Steenackers M, Van Montagu MC, Engler GJ, Boerjan, WA (2000) Cell-specific and conditional expression of caffeoyl-coenzyme A-3-O-methyltransferase in poplar. Plant Physiol. 123: 853-867

De Melis LE, Whiteman PH, Stevenson TW (1999) Isolation and characterization of a cDNA clone encoding cinnamyl alcohol dehydrogenase in *Eucalyptus globulus* Labill. Plant Sci 143: 173-182

Grimmig B, Matern U (1997) Structure of the parsley caffeoyl-CoA *O*-methyltransferase gene, harbouring a novel elicitor responsive *cis*-acting element. Plant Mol Biol 33: 323-341

Higuchi, T. (1990) Lignin biochemistry: biosynthesis and biodegradation. Wood Sci Technol 24: 23–63

Inoue K, Sewalt VJH, Balance GM, Ni W, Sturzer C, Dixon RA (1998) Developmental expression and substrate specificities of alfalfa caffeic acid 3-O-methyltransferase and caffeoyl coenzyme A 3-O-methyltransferase in relation to lignification. Plant Physiol. 117: 761–770

Joshi CP, Chiang VL (1998) Conserved sequence motifs in plant S-adenosyl-L-methionine-depedent methyltransferases. Plant Mol Biol 37: 663-674

Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY, Bryant SH (2002) CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res 30: 281–283

Martz F, Maury S, Pincon G, Legrand M (1998) cDNA cloning, substrate specificity and expression study of tobacco caffeoyl-CoA 3-O-methyltransferase, a lignin biosynthetic enzyme. Plant Mol Biol 36: 427–437

Meng H, Campbell WH (1998) Substrate profiles and expression of caffeoyl coenzyme A and caffeic acid O-methyltransferases in secondary xylem of aspen during seasonal development. Plant Mol Biol 38: 513–520

Laura C, Joan R, Pere P (1999) Nucleotide Sequence of two cDNAs coding for Caffeoyl-coenzyme A O-Methyltransferase (CCoAOMT) and study of their expression in Zea mays. Plant Physiol 120: 1206-1206

Lewis NG, Yamamoto E (1990) Lignin: Occurrence, biogenesis and biodegradation. Annu Rev Plant Physiol Plant Mol Biol 41: 455-496

Li L, Osakabe Y, Joshi CP, Chiang VL (1999) Secondary xylem-specific expression of caffeoyl-coenzyme A 3-O-methyltransferase plays an important role in the methylation pathway associated with lignin biosynthesis in loblolly pine. Plant Mol Biol 40: 555-565

Li L, Popko JL, Zhang XH, Osakabe K, Tsai CJ, Joshi CP, Chiang VL (1997) A novel multifunctional O-methyltransferase implicated in a dual methylation pathway associated with lignin biosynthesis in loblolly pine. Proc Natl Acad Sci 94: 5461-5466

Ralph J, Hatfield RD, Piquemal J, Yahiaoui N, Pean M, Lapierre C, Boudet AM (1998) NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignification enzymes cinnamyl alcohol dehydrogenase and cinnamoyl CoA reductase. *Proc Natl Acad Sci* 95: 12803–12808

Thanaraj TA, Stamm S (2003) Prediction and statistical analysis of alternatively spliced exons. Prog Mol Subcell Biol 31: 1–31

Thanaraj TA, Robinson AJ (2000) Prediction of exact boundaries of exons. Brief Bioinform 4: 343-356

Ye ZH, Kneusel RE, Matern U, Varner JE (1994) An alternative methylation pathway in lignin biosynthesis in Zinnia. Plant Cell 6: 1427-1439

Zhang MQ (1997) Identification of protein coding regions in the human genome by quadratic discriminant analysis. *Proc Natl Acad Sci* 94: 565-568

4.1 Introduction

CCoAOMT gene expression was for the first time reported in *Zinnia* and found to be closely associated with lignification (Ye et al., 1994). After this finding, the emphasis shifted towards determining whether this expression pattern was common to other plants. Localization of CCoAOMT at the protein and mRNA levels confirmed CCoAOMT gene expression in lignifying tissues in herbaceous plants like tobacco, tomato, soybean (Ye, 1997), alfalfa (Kersey et al., 1999; Inoue et al., 1998), and in woody plants like forsythia (Ye, 1997) and poplar (Chen et al., 2000; Zhong et al., 2000). CCoAOMT gene expression and localization was shown to be essentially restricted to the tracheary elements, xylary fibres, phloem fibers and xylem ray parenchyma cells. In tobacco after tobacco mosaic virus infection, two members of the CCoAOMT gene family are constitutively expressed in various plant organs and tissue, whereas two other are preferentially expressed in flower organs (Martz et al., 1998). Three distinct classes of CCoAOMT have been characterized in tobacco and levels of expression varied with the stages of stem development and in accordance with their involvement in the synthesis of guaiacyl units of lignin (Maury et al., 1999). In quaking aspen seasonal expression of CCoAOMT gene was analyzed in developing secondary xylem. CCoAOMT activity and expression was maximum during the middle of the growing season (Meng and Campbell, 1998). CCoAOMT gene expression and activity has also been studied in pine. Promoter of pine CCoAOMT localized GUS activity to secondary xylem suggesting the expression of CCoAOMT gene in secondary xylem and a close correlation with xylem lignification (Li et al., 1999).

4.2 Material and Methods

4.2.1 Cloning CCoAOMT cDNA in Expression vector

The pET-30b(+) vector (Novagen, USA) was used for expression of CCoAOMT gene in *E.coli* BL 21 (DE3). This vector carries a N-terminal His•Tag®/thrombin/S•Tag™/ enterokinase configuration plus an optional C-terminal His•Tag sequence. This vector has unique restriction sites for cloning ease (Fig. 4.1). The cloned gene is expressed under the T7 RNA polymerase promoter.

Fig. 4.1: Vector diagram of expression vector pET 30 a,b and c(+).The maps for pET-30b(+) and pET-30c(+) are the same as pET-30a(+) (shown) with the following exceptions: pET-30b(+) is a 5421 bp plasmid; subtract 1 bp from each site beyond *Bam*H I at 198. pET-30c(+) is a 5423 bp plasmid; add 1 bp to each site beyond *Bam*H I at 198. The sequence of pET 30b(+) is highlighted in yellow.

The CCoAOMT1 gene specific, MF4 forward and the MR4 reverse primer sequences (Chapter 3; section 3.3.6) were modified to include *Nde* I restriction site at the 5'end of the

MF4 forward primer and *Xho* I site at the 5' end of the MR4 reverse primer. The TGA stop codon was deleted from the MR4 reverse primer. This ensured that the **His•Tag** from the cloning vector was in frame at the C-terminus of the expressed protein. The modified primers were designated as **PetF [5' (d) CAT ATG GCG GAT CAG AAT CAA AGC GAA G 3']** and **PetR [5' (d)CTC GAG GCT GAT CCT ACG GCA GAG AGT GAT 3']** were used to reamplify the CCoAOMT1 cDNA gene clone. The reamplified gene was cloned in the *Nde* I and *Xho* I sites of the pET-30b(+) vector under the control of the T7 promoter.

4.2.2 Protein expression and purification

Protein was expressed in the host *E. coli* BL21 (DE3) cell line and purified as described earlier (Chapter 2; section 2.8). Purity of protein was checked on 10% SDS PAGE (Chapter 2; section 2.9).

4.2.3 Primary and Secondary Antibodies

Purified CCoAOMT protein was used for raising rabbit immune-serum. Primary Polyclonal antibodies were purified from the immune-serum. Secondary antibodies *i.e* anti rabbit goat IgG conjugated with alkaline phosphatase was from Merck, USA.

4.2.4 Affinity purification of CCoAOMT IgG from rabbit immune serum

The CCoAOMT protein specific polyclonal IgG were antigen affinity purified from rabbit immune-serum using Affi-gel 15 (BIORAD) (Chapter 2; section 2.10.2)

4.2.5 Histology, histochemical staining and immunocytolocalization

Transverse sections of root, shoot, leaf rachis and inflorescence of 0, 5, 10 and 15 days old seedlings as well as from one and two season old plant were used for histology, immunocytolocalization and histochemical staining (Chapter 2; section 2.11 and 2.13).

4.2.6 Semi- and absolute quantification of CCoAOMT1 and CCoAOMT2 gene expression

Total RNA isolated from root, shoot, leaf and inflorescence of 0, 5, 10 and 15 day old seedlings as well as of one and two season old plant (Chapter 2; section 2.4.6) was used for

cDNA first strand synthesis. The cDNA first strands were used as template for semi - and absolute quantification of CCoAOMT transcripts as described earlier (Chapter 2; section 2.4.11).

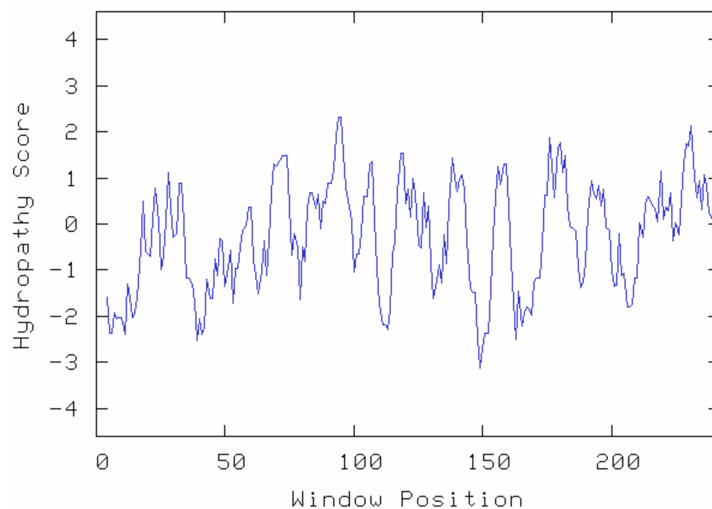
4.3 Results and Discussion

4.3.1 Expression of CCoAOMT gene in *E.coli*

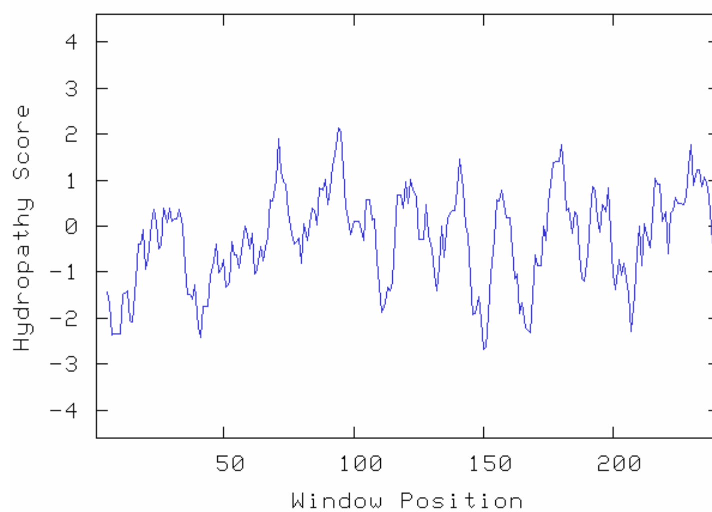
The CCoAOMT1 gene clone in pET-30b(+) vector was expressed in *E.coli* BL21 (DE3) cell line. In view of the fact that the deduced amino acid sequence of CCoAOMT1 and CCoAOMT2 genes is 99% similar and differ only at the amino acids L 46 I and R 133 K (see Chapter 3; Fig. 3.30), and the hydropathy plots (Kite and Dolittle, 1982; Fig. 4.2 a, b) for the two proteins was highly similar and superimposable, protein expressed from only one gene was used for raising antibodies. The hydropathy plots also show similar strong negative peaks indicating possible exposed surface regions of the protein at almost similar amino acid positions (Hopp and Woods, 1981). Because of the similarity of the hydropathy plots and the deduced amino acid sequence of CCoAOMT1 and CCoAOMT2, it was assumed that selective antibodies for the two proteins cannot be raised.

The pET-30b(+) expression vector provides the choice of having the **His•Tag** either at the N or the C terminus of the expressed protein. The CCoAOMT protein is reported to form homodimer in solution (Inoue et al., 1998) and during this homodimerization the N - terminus of the protein is not exposed. Since protein was purified from inclusion bodies and the use of Ni – chelated affinity column was intended, this information was taken into account and the **His•Tag** was provided at the exposed C-terminus of the protein.

The CCoAOMT1 gene reamplified using PetF and PetR primers was cloned in pGEM T Easy vector and sequenced. The gene insert was released with *Nde* I and *Xho* I restriction digestion, purified and cloned in the *Nde* I and *Xho* I sited of the pET-30b(+) vector. The pET-30b(+) vector harboring CCoAOMT1 gene was designated as pETCT7 (Fig. 4.3). The vector pETCT7 was then transformed into *E. coli* BL21 (DE3) cell line.



(a)



(b)

Fig. 4.2: Kite and Dolittle plots for the CCoAOMT 1 (a) and CCoAOMT 2 (b). In both the plots window size is 9.

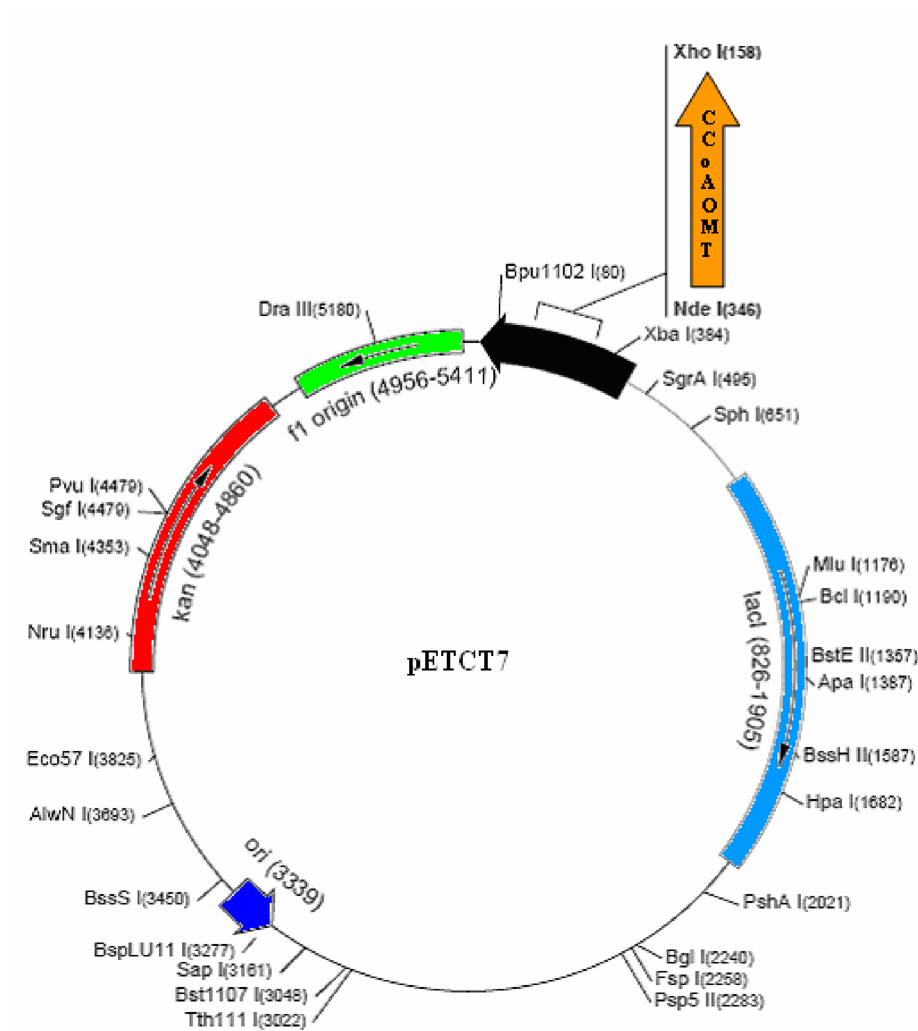


Fig. 4.3: Vector map of pETCT7

The deduced amino acids sequence of the CCoAOMT 1 gene sequence was used to calculate its approximate molecular mass, percent amino acid composition and theoretical pI. The calculated approximate molecular mass of CCoAOMT 1 was 27.5KDa with pI of 5.31. These calculations compare well with other CCoAOMT proteins reported from Zinnia, tobacco, tomato, soybean (Ye, 1997), alfalfa (Inoue et al, 1998.), poplar (Zhong et al., 2000). Percent amino acid composition is shown in table 4.1.

Table 4.1: Percentage amino acid composition of CCoAOMT 1

Amino Acid Comp.		
Aa	no	Mol%
A	18	7.38%
C	2	0.82%
D	17	6.97%
E	18	7.38%
F	5	2.05%
G	16	6.56%
H	6	2.46%
I	16	6.56%
K	15	6.15%
L	29	11.89%
M	7	2.87%
N	13	5.33%
P	13	5.33%
Q	7	2.87%
R	12	4.92%
S	9	3.69%
T	9	3.69%
V	17	6.97%
W	2	0.82%
Y	13	5.33%
Total:	244	100.00%

The CCoAOMT protein with **His•Tag** at its C-terminus was purified from *E.coli* culture lysate and inclusion bodies by using Ni- chelated column (Pierce, USA). The protein purity was checked by 10% SDS PAGE. The Ni- column purified CCoAOMT protein was homogenous and free of other contaminating protein. Total denatured protein profile, before and after induction, from cell lysate of *E.coli* BL21 (DE3) harbouring pETCT7 plasmid was compared. No appreciable difference was observed in the level of protein expression pre- and post –IPTG induction. The Ni-chelated affinity column purified CCoAOMT protein was seen to be of ~ 29KDa molecular mass (Fig. 4.4).

Cloning of alfalfa and tobacco CCoAOMT cDNA in expression vectors, its expression in *E coli* BL21 (DE3) cells and purification through metal ion (Ni^{+2} and Co^{+2}) chelated affinity columns and agarose-glutathione matrix have been reported. The purified protein was then used for raising polyclonal antibodies (Gowri et al., 1991; Inoue et al., 1998; Maury et al., 1999) and crystal formation (Ferrer et al., 2005).

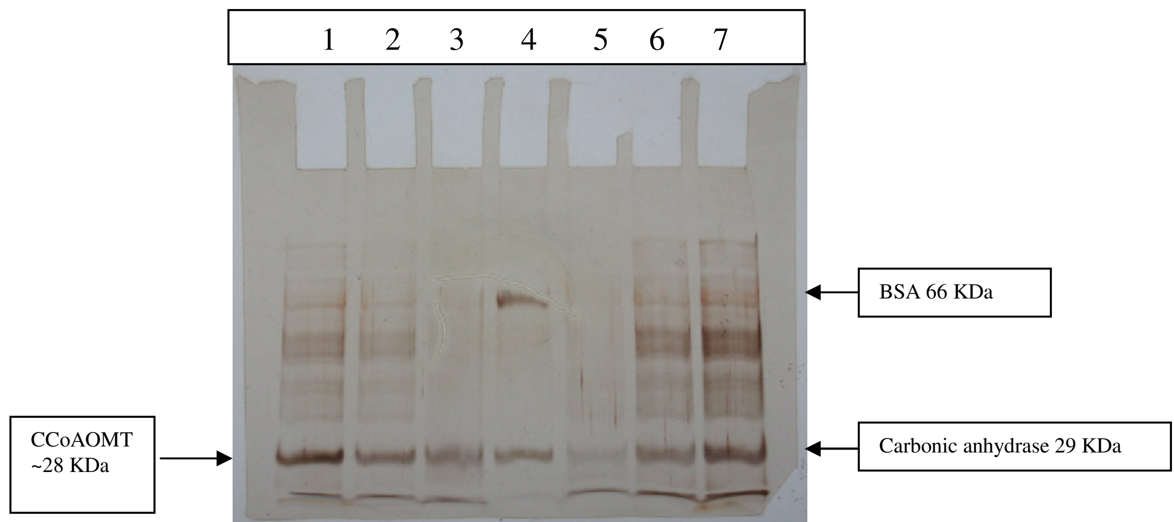


Fig. 4.4: SDS PAGE (10%) of *E. coli* BL21 (pETCT7) induced lysate (lanes 1, 7), uninduced lysate ((lane 2, 6), protein from inclusion bodies ~29Kda (lane 3), protein size markers BSA (66 KDa) and Carbonic anhydrase (29KDa) (lane 4), Ni column purified protein (lane 5).

4.3.2 Purification of polyclonal antibodies

CCoAOMT protein purified from recombinant *E. coli* inclusion bodies (since the yield was better than from lysate) was used to raise rabbit immune serum. The immune serum was centrifuged at 12,000 x g for 15 min and the clear supernatant heated for 1h at 55ºC to deactivate the compliment system. Heat denatured immune serum was again centrifuged and the supernatant transferred to fresh tubes. EDTA was added to a final concentration of 10mM and Themersol to a final concentration of 0.02%. The immune serum was aliquoted, frozen in liquid nitrogen and stored at -70ºC till further use.

The CCoAOMT specific polyclonal IgG was antigen affinity purified using Affi-Gel 15 (BioRad). The IgG was checked for purity on 10% SDS PAGE. As is obvious from Fig. 4.5, antigen affinity purified IgG show characteristic 44KDa heavy chain and the 22KDa light chain peptides. The recoved IgG was also free of other contaminating serum proteins.

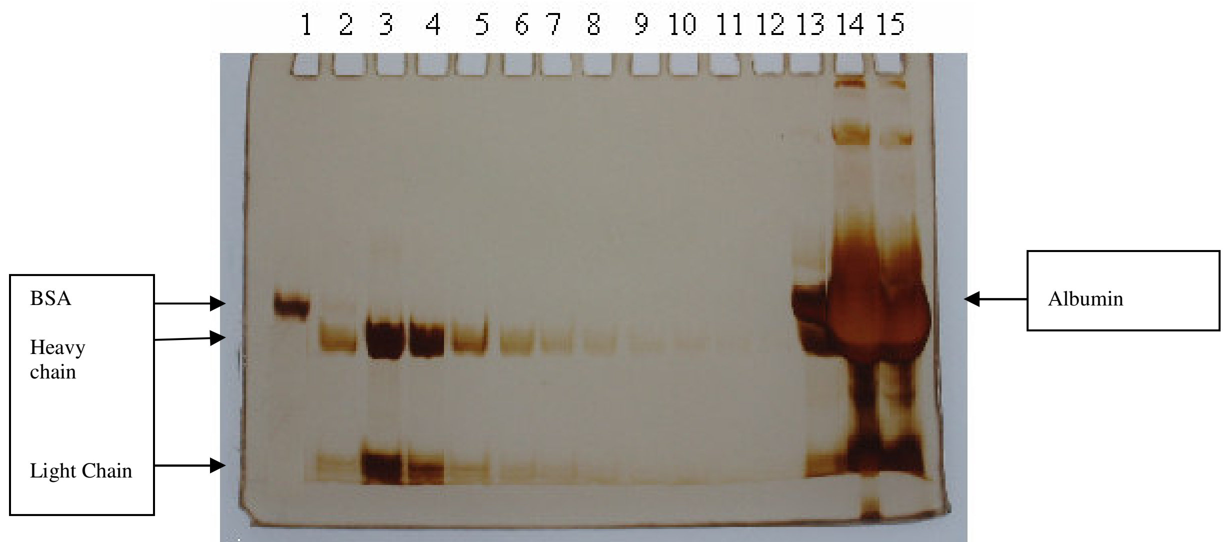


Fig. 4.5: 10% SDS PAGE of the affinity purified CCoAOMT specific IgG. BSA (lane 1), eluted fractions of affinity purified IgG (lane 2- 12), crude serum (lane 13) and serum flow through (lane 14-15).

4.3.3 Histology

Leucaena leucocephala seeds were scarified using concentrated H_2SO_4. The seeds were washed thoroughly with running tap water and imbibed for 12 h (in dark). Imbibed seeds were transferred to pots (1:1 coco peat and sand mixture) for germination. Transverse sections from the shoot, root, rachis and inflorescence were prepared. Seedlings at 0 day (when the seedling fully emerged from seed coat), 5, 10 and 15 days old, and one and two seasons old plants were used for analysis.

Anatomy of *L. leucocephala* resembles the anatomy of a typical dicot plant (Esau, 1977). The sequential and progressive development of vascular tissue was evident in the transverse sections of different plant parts of 0, 5, 10 and 15 days old seedlings, and in one and two season old plants. Phloroglucinol-HCl staining of the transverse section stained lignified xylem, secondary xylem and phloem fiber in brown red (Figure 4.6 a, b; Left panel). Development of vascular tissue, particularly xylem tissue, was visible in phloroglucinol-HCl stained transverse sections at different growth stage of seedlings. Increase in number of differentiating xylem cells as well as their stain intensity indicated progression of lignification. In older tissue well developed and lignified secondary growth and development of fibers was observed. When the sections were visualized under polarized light (Fig. 4.2 a, b; Central panel), the xylem tissue which was bright reddish yellow in 0 day plant, progressively changed to bright and deep orange red in 5 and 10 days old seedlings and then to light pinkish red in 15 days old seedlings, and one season and two season old plants. The phloem fibers, however, were of bright reddish yellow initially up to 15 day but appeared orange red and pink red in I and II season old plant. Since, polarizing light produces different hues depending on the chemical bond configurations the change in color may be as a consequence of the change in cell wall chemical composition with age. As the plant matures, change in the color of lignified tissue might also be as a consequence of change in ratio or the type of lignin.

4.3.4 Immunocytolocalization

In the present study CCoAOMT was immunolocalized in the transverse sections of root, shoot, rachis and inflorescence of 0, 5, 10, 15 days old seedling as well as from one and two season old plant. The affinity purified CCoAOMT specific IgG and anti rabbit goat IgG

conjugated with alkaline phophatase were used as primary and secondary antibodies respectively.

Results revealed CCoAOMT enzyme expression in differentiating xylem tissue and phloem fibers (Figure 4.6 a, b; Right panel). The CCoAOMT presence was detected as early as in 0 day old seedling. In transverse sections of 0 day seedling blue black precipitate was visible even in the potential phloem fiber regions. This observation was confirmed when these areas differentiated into fiber as seedlings grew older. Comparison of the sections where CCoAOMT was immunolocalized, with phloroglucinol-HCl stained sections confirmed the presence of CCoAOMT at the locations where lignin was stained. Hence, suggesting that CCoAOMT activity was confined to cell undergoing lignification.

No CCoAOMT protein was immunolocalized in cortex and phloem tissue (except in phloem fibers) showing absence of CCoAOMT activity in these regions. No CCoAOMT was detected in pith cells either.

As in the present study, CCoAOMT protein immunolocalized in xylem and phloem fiber and its active participation in lignin biosynthesis has been reported in many plants like Zinnia (Ye and Varner, 1995; Ye et al., 1997), alfalfa (Kersey et al., 1999), poplar (Chen et al., 2000; Zhong et al., 2000; Ye et al., 2001), tobacco and forsythia (Ye et al., 2001).

4.3.5 Semiquantitative and Quantitative Real Time PCR

The immunocytolocalization of CCoAOMT in different parts of *L. leucocephala* plants at different growth stages showed that CCoAOMT activity was localized to xylem tissue and phloem fibers. This study, however, could not elucidate or discriminate the involvement of CCoAOMT 1 and CCoAOMT 2 in lignification during different stages of plant development. The differential expression of CCoAOMT1 and CCoAOMT2 genes during lignification process was assayed by semiquantitative and quantitative real time PCR (QPCR). TaqMan probe based chemistry was used during the study.

The nucleotide sequence alignment of the CCoAOMT1 and CCoAOMT2 cDNA clones reveals many a positions where the nucleotide sequences differed. One of these differences in nucleotide sequence was used to design and synthesize gene specific primers to selectively amplify CCoAOMT1 and/or CCoAOMT2 transcripts from a cDNA population (Figs. 4.7 and 4.8 a, b). TaqMan probes were also designed and synthesized.

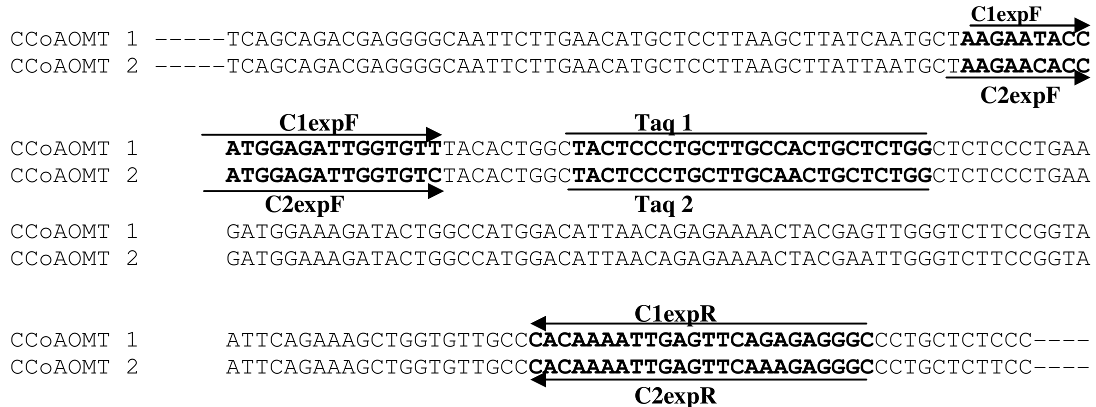


Fig. 4.7: Nucleotide sequence alignment (nucleotide position 232 to 416) of CCoAOMT 1 (DQ431233) and CCoAOMT 2 (DQ431234) genes. Nucleotide sequences highlighted in bold are regions from where forward primer, probe and reverse primers were designed.

C1expF 5′ (d) AAGAATACCATGGAGATTGGTGTT 3′

C2expF 5′ (d) AAGAACACCATGGAGATTGGTGTC 3′

C1expR 5′ (d) GCCCTCTCTGAACTCAATTTTGTG 3′

C2expR 5′ (d) GCCCTCTTTGAACTCAATTTTGTG 3′

(a)

Taq1 – 5′-/6-FAM/TACTCCCTGCTTGCCACTGCTCTGG/BHQ_2/3′

Taq2 – 5′-/Cy5/TACTCCCTGCTTGCAACTGCTCTGG/BHQ_1/3′

(b)

Fig. 4.8: (a) C1expF and C1expR are gene specific forward and reverse primer sequences for CCoAOMT 1. C2expF and C2expR are gene specific forward and reverse primer sequences for CCoAOMT 2. (b) Taq1 and Taq2 are TaqMan probes for the CCoAOMT 1 and CCoAOMT 2 genes respectively.

Total RNA was isolated from shoot, root, leaf and inflorescence of 0, 5, 10, 15 day old seedling, and one and two season old *L. leucocephala* plants. An aliquot of total RNA normalized for uniform amplification of the 5.8S rRNA was used for synthesis of cDNA first strand which was used as template for semiquantitative PCR.

The primers C1expF and C1expR, C2expF and C2expR were used for, semiquantitative PCR of CCoAOMT1 and CCoAOMT2 transcripts respectively. The amplification obtained from cDNAs of different plant tissues at different growth stages is shown in Fig. 4.9. The results revealed that both CCoAOMT 1 and CCoAOMT 2 were expressed simultaneously, but the levels of expression differed with the tissue type and plant age.

Low gene expression of CCoAOMT1 and CCoAOMT2 was observed in 0 day shoots. The expression increased to its maximum level by day 5. There after the gene expression declined by days 10 and 15. Low expression was seen in one season old plants also, however, the CCoAOMT2 gene expression was marginally higher than CCoAOMT1. In the two season old plants the gene expression levels increased again and here too the CCoAOMT2 gene expression was substantially higher than the expression level of CCoAOMT1.

In the roots of 0 day old seedlings both CCoAOMT1 and CCoAOMT2 genes were expressed, CCoAOMT2 relatively more than CCoAOMT1. In 5 day old seedling roots the expression levels of both the genes declined. The gene expression of CCoAOMT1 and CCoAOMT2 increased again by days 10 and 15. The expression level at day 10 in the roots was slightly higher for the CCoAOMT2 as compared to CCoAOMT1.This trend, however, reversed by day 15, when CCoAOMT1 expression was higher than CCoAOMT2. Negligible expression was seen in roots from one season and two season old plants.

Expression of CCoAOMT1 and CCoAOMT2 gene in the leaf tissue was seen in the leaves from 15 day old seedlings. Here too the expression level of CCoAOMT2 gene seemed to be marginally better than that of CCoAOMT1. No gene expression was seen in leaves of one season and two season old plants. No gene expression was seen in the inflorescence.

From the above study it was also obvious that both the CCoAOMT genes expressed or shut down in tandem.

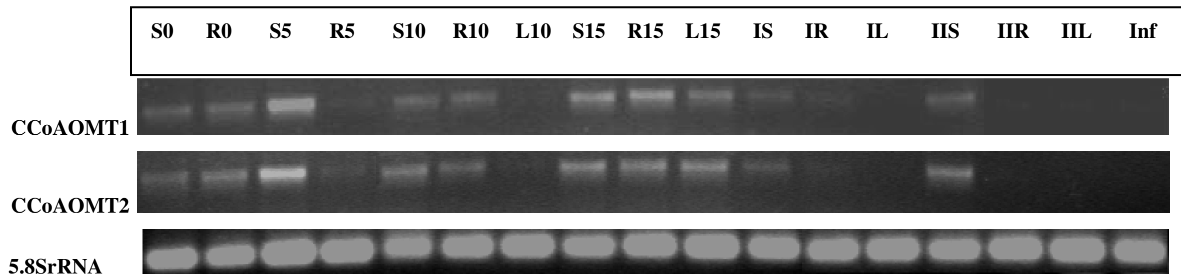


Fig. 4.9: Semiquantitative PCR for CCoAOMT 1 and CCoAOMT 2. S, R, L and Inf stands for shoot, root, leaf rachis and inflorescence respectively. The numbers 0, 5, 10, 15 represents age of seedling in days. I and II represents season one and two. 5.8S rRNA was used as internal control.

The CCoAOMT 1 and CCoAOMT 2 specific sets of primers with TaqMan probes were used for quantitative real time PCR. Absolute quantification method was used for the purpose. In this method a standard curve was prepared using cDNA clone of the gene to be quantified. Serial dilutions were used for real time PCR reactions and a Ct value for each dilution plotted against diluted gene clone quantity. Once Ct values for experiments were obtained, the expression levels of the target gene in the sample were determined using the standard curve.

The cDNA clones of CCoAOMT 1 (DQ431233) and CCoAOMT 2 (DQ431234) were used for standard curve preparation for absolute quantification of the two gene transcripts. CCoAOMT 1 and CCoAOMT 2 concentrations used were in nanogram (ng) quantities. The cDNA first strand synthesized using total RNA quantities, earlier normalized for uniform amplification of the 5.8S rRNA isolated from shoot, root, leaf and inflorescence of 0, 5, 10, 15 day old seedling, and one and two season old *L. leucocephala* plants was used as template and reactions were carried out in ABI PRISM 7700 instrument.

The Ct values obtained for the two genes and their respective quantity determined using standard curve is given in table 4.2. The relative expression of the two genes in terms of ng μL^{-1} is shown in Fig. 4.10.

Table 4.2: Ct values for the two CCoAOMTs and their respective quantities in ng μL^{-1}.

S.No	Tissue type and Plant age	CCoAOMT 1		CCoAOMT 2	
		Ct value	Quantity (ng μL^{-1})	Ct value	Quantity (ng μL^{-1})
1	Shoot 0 day	30.60	6.15 E-4	32.32	1.50 E-3
2	Shoot 5 day	26.26	1.38 E-2	29.81	8.50 E-3
3	Shoot 10 day	29.24	1.64 E-3	31.30	3.10 E-3
4	Shoot 15 day	29.33	1.52 E-3	32.57	1.31 E-3
5	Shoot I season	31.97	2.29 E-4	33.94	5.15 E-4
6	Shoot II season	28.96	1.98 E-3	30.39	5.80 E-3
7	Root 0 day	28.00	3.97 E-3	31.46	2.70 E-3
8	Root 5 day	30.50	6.55 E-4	33.72	5.90 E-4
9	Root 10 day	29.83	1.07 E-3	32.02	1.90 E-3
10	Root 15 day	29.16	1.72 E-3	32.36	1.50 E-3
11	Root I season	33.72	6.55 E-5	36.24	2.07 E-4
12	Root II season	33.11	1.01 E-4	36.00	1.26 E-4
13	Leaf Rachis 10 day	32.46	1.62 E-4	35.72	1.52 E-4
14	Leaf Rachis 15 day	27.98	4.00 E-3	30.94	3.98 E-3
15	Leaf Rachis I season	36.19	1.11 E-5	38.70	1.99 E-5
16	Leaf Rachis II season	36.15	1.14 E-5	38.07	3.05 E-5
17	Inflorescence	36.20	1.10 E-5	38.21	2.79 E-5

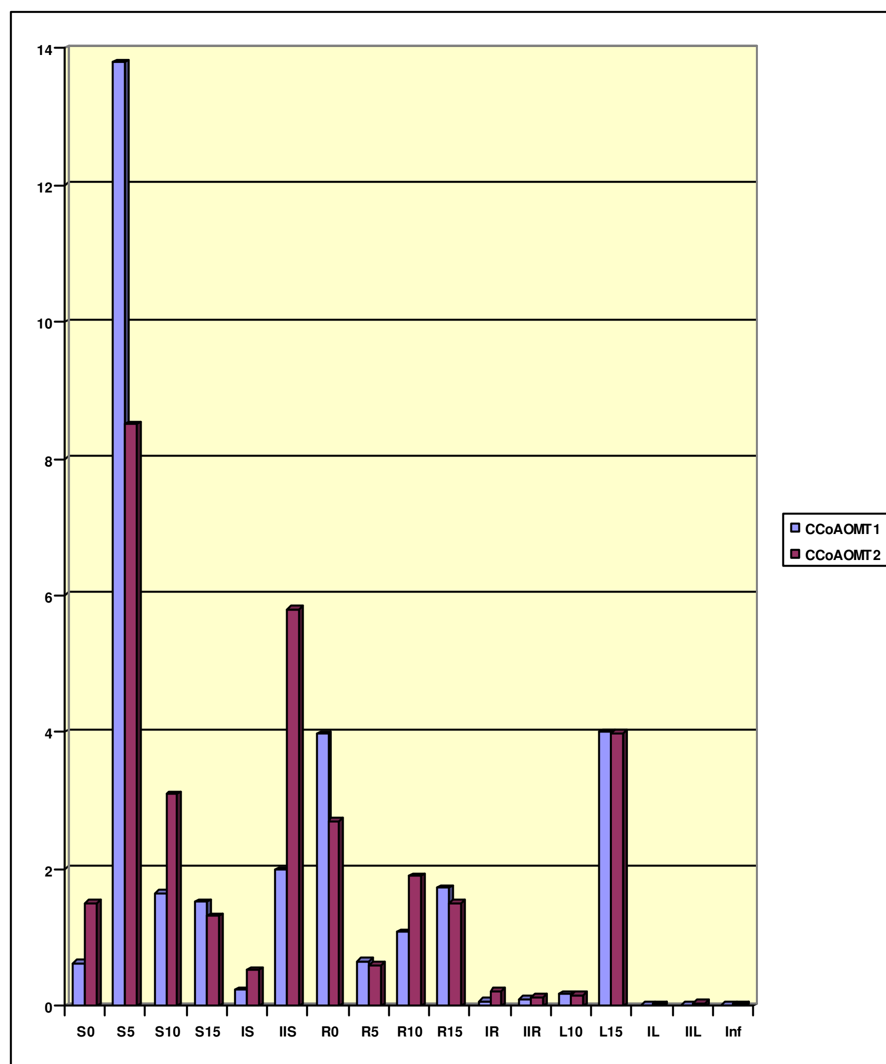


Fig. 4.10: Relative expression of CCoAOMT 1 and CCoAOMT 2 in terms of ng µL^{-1} in different plant tissues of different age plants. On X axis is tissue type with age of seedling or plant. On Y axis is quantity in ng µL^{-1}. S, R, L and Inf stands for shoot, root and leaf rachis respectively. The numbers 0, 5, 10, 15 represents age of seedling in days. I and II represents season one and two respectively.

QPCR results provided answers to the differential expression of CCoAOMT1 and CCoAOMT2 genes. These experiments also fine tuned and refined the semiquantitative PCR results. The QPCR results revealed that both CCoAOMT1 and CCoAOMT2 gene are expressed throughout the developmental stages of *L. leucocephala*, though the levels of expression differ greatly depending on tissue type and the seedling age. Highest abundance

of the CCoAOMT1 and CCoAOMT2 transcripts was observed in shoots of 5 day old seedling. However, the relative abundance of CCoAOMT1 transcript was more than the CCoAOMT2 transcripts. It was also apparent that with advance in the age of the plants, generally the CCoAOMT2 transcript levels are higher than or equal to the CCoAOMT1 transcript levels. In the root tissues the CCoAOMT1 transcript is of more abundance than the CCoAOMT2 transcript in 0 day old seedlings. The abundance of CCoAOMT1 and CCoAOMT2 is almost negligible in roots from I season and II season old plants. High and equal abundance of CCoAOMT1 and CCoAOMT2 transcript was seen in 15 days old leaves. The abundance of transcripts were negligible in leaves of I season and II season old plants and in inflorescence.

In all the lignifying tissues of different growth stages, the two CCoAOMTs were expressing in tandem. However, no correlation could be drawn between the extent of lignification and CCoAOMT gene expression levels as colour intensity of immunolocalized CCoAOMT was similar in lignifying tissue of different growth stages though QPCR results revealed different levels of transcripts.

4.4 Conclusions

- CCoAOMT gene expressed in *E. coli* BL21 (DE3) and protein purified from inclusion bodies using Ni-chelated affinity column.
- Polyclonal antibodies were raised against purified CCoAOMT protein in rabbit.
- CCoAOMT specific polyclonal IgG purified using Affi-gel 15 affinity matrix.
- Transverse sections of different plant parts of different age stained with phloroglucinol-HCL show increase in number of differentiating xylem cells as well as their stain intensity indicating progression of lignification with age.
- Visualization under polarized light of phloroglucinol-HCL stained tissues showed different colour in different plant parts of different ages, suggesting chemical compositional changes in lignin with tissue age.
- CCoAOMT immunolocalized in xylem and fibers suggesting its presence at the sites of extensive lignification.
- Semiquantitative and QPCR results showed that both the CCoAOMT1 and CCoAOMT2 genes were expressed in tandem.

4.5 References

Chen C, Meyermans H, Burggraeve B, De Rycke RM, Inoue K, Vleesschauwer VD, Steenackers M, Van Montagu MC, Engler GJ, Boerjan WA (2000) Cell-specific and conditional expression of caffeoyl-coenzyme A-3-O-methyltransferase in poplar. Plant Physiol 123: 853–867

Esau K (1977) Anatomy of Seed Plants, 2nd ed. John Wiley and Sons, New York.

Ferrer JL, Zubieta C, Dixon RA, Noel JP (2005) Crystal Structures of Alfalfa Caffeoyl Coenzyme A 3-*O*-Methyltransferase. Plant Physiol 137: 1009-1017

Gowri G, Bugos RC, Campbell WH, Maxwell CA, Dixon RA (1991) Stress responses in alfalfa (*Medicago sativa* L.) Molecular cloning and expression of S-adenosyl-L-methionine: caffeic acid 3-O-methyltransferase, a key enzyme of lignin biosynthesis. Plant Physiol 97: 7-14

Hopps T, Woods KR (1981) Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci 78: 3824-3828

Inoue K, Sewalt VJH, Balance GM, Ni W, Sturzer C, Dixon RA (1998) Developmental expression and substrate specificities of alfalfa caffeic acid 3-O-methyltransferase and caffeoyl coenzyme A 3-O-methyltransferase in relation to lignification. Plant Physiol 117: 761–770

Kersey R, Inoue K, Schubert KR, Dixon RA (1999) Immunolocalization of two lignin O-methyltransferases in stems of alfalfa (*Medicago sativa* L.). Protoplasma 209: 46–57

Kyte J, Doolittle R (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105-132

Li L, Osakabe K, Joshi CP, Chiang VL (1999) Secondary xylem specific expression of caffeoyl-coenzyme A 3-O-methyltransferase plays an important role in the methylation pathway associated with lignin biosynthesis in loblolly pine. Plant Mol Biol 40: 555–565

Martz F, Maury S, Pincon G, Legrand M (1998) cDNA cloning, substrate specificity and expression study of tobacco caffeoyl-CoA 3-O-methyltransferase, a lignin biosynthetic enzyme. Plant Mol Biol 36: 427–437

Maury S, Geoffroy P, Legrand M (1999) Tobacco O-methyltransferases involved in phenylpropanoid metabolism. The different caffeoyl-coenzyme A/5-hydroxyferuloyl-

coenzyme A 3/5-O-methyltransferase and caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase classes have distinct substrate specificities and expression patterns. Plant Physiol 121: 215–223

Meng H, Campbell WH (1998) Substrate profiles and expression of caffeoyl coenzyme A and caffeic acid O-methyltransferases in secondary xylem of aspen during seasonal development. Plant Mol Biol 38, 513–520

Ye Z-H (1997) Association of Caffeoyl coenzyme A 3-O-Methyltransferase Expression with Lignifying Tissues in Several Dicot Plants. Plant Physiol 115: 1341-1350

Ye Z-H, Kneusel RE, Matern U, Varner JE (1994) An alternative methylation pathway in lignin biosynthesis in Zinnia. Plant Cell 6: 1427–1439

Ye Z-H, Varner JE (1995) Differential Expression of Two-Methyltransferases in Lignin Biosynthesis in *Zinnia elegans*. Plant Physiol 108: 459-467

Ye Z-H, Zhong R, Morrison WH III, Himmelsbach DS (2001) Caffeoyl coenzyme A O-methyltransferase and lignin biosynthesis. Phytochemistry 57: 1177-1185

Zhong R, Morrison WH III, Himmelsbach DS, Poole FL II, Ye Z-H (2000) Essential role of caffeoyl coenzyme A *O*methyltransferase in lignin biosynthesis in woody poplar plants. Plant Physiol 124: 563–77

5.1 Introduction

Promoter analysis has been the main tool for studying gene regulation. Many phenylpropanoid pathway genes have been cloned and the activities of their promoters analyzed. Promoter analysis has been done for Phenyl ammonia-lyase (PAL; Bevan et al., 1989; Ohl et al., 1990; Hauffe et al., 1991; Levya et al., 1992), cinnamic acid 4-hydroxylase (C4H; Bell-Lelong et al., 1997), hydroxycinnamate: CoA ligase (4CL; Douglas et al., 1991), caffeic acid O-methyltransferase (COMT; Capellades et al., 1996), caffeoyl CoA O-methyltransferase (CCoAOMT; Grimmig and Matern, 1997; Chen et al., 2000) and hydroxycinnamyl alcohol dehydrogenase (Feuillet et al., 1995). All these gene promoters directed the expression of GUS reporter gene in lignifying tissues. Deletion analyses in the PAL and 4CL gene promoters, identified the cis-acting AC elements as being critical for tissue-specific expression (Hauffe et al., 1991; Levya et al., 1992). These elements are also conserved in the promoters of other phenylpropanoid pathway genes. The Myb protein binds to the AC elements and turns on the expression of a reporter gene driven by the PAL promoter (Sablowski et al., 1994, 1995). Over expression of the two *Antirrhinum Myb* genes in transgenic tobacco plants down-regulated the expression of certain genes in the phenylpropanoid biosynthetic pathway, resulting in a reduction in lignin accumulation (Tamagnone et al., 1998). The common *cis*-elements in the promoter regions of a number of known PAL genes, which were also found in 4CL gene, have been reported as box P (box 2: C/TTT/CC/TA/CA/CCA/CAA/CCC/AC/AC), box A (C/ACGTCT/C), and box L (box 1: T/CCT/CC/TACCTACC). These elements play important roles in the regulation of *PAL* promoter activity (Cramer et al., 1989; Lois et al., 1989; Logemann et al., 1995).

Other cis-elements, such as the FP56 identified in the 4CL promoter, have also been shown to be important for regulation of gene expression in addition to the conserved AC elements (Neustaedter et al., 1999).

In ~5Kb CCoAOMT promoter fragment isolated from parsley (Grimmig and Matern, 1997), presence of three putative cis-regulatory elements, boxes P, A and L, were shown. These elements represent motifs recurring in the promoters of several genes of stress inducible phenylpropanoid pathway. The transient expression assays with sets of 5'- truncated promoter –GUS fusion showed that significant promoter activity was retained in the 354 bp

fragment. Presence of a novel cis-regulatory element (E box) was also shown by *in vitro* DNase I footprinting.

Chen et al.(2000) reported the activity of two CCoAOMT promoters from poplar in regulating gene expression in xylem and differentially in phloem. In xylem, expression was preferentially observed in vessels and contact rays, whereas no detectable expression was found in storage rays and fibers. Increased expression levels were reported after fungal infection, wounding and bending. Upon bending and leaning of stem, it was reported that the cell specific expression pattern was lost and both genes were expressed in all cell types of xylem.

In the present study the partial promoters of CCoAOMT1 and CCoAOMT2 genes from *L. leucocephala* were analyzed for presence of regulatory elements, using online software MatInspector 2.2. The two promoters were also analyzed for their ability to drive the expression of the reporter green fluorescent protein gene (GFP).

5.2 Material and Methods

5.2.1 Promoter analysis

The two partial CCoAOMT1 and CCoAOMT2 gene promoter sequences, designated **ProC1** and **ProC2** respectively, were analyzed for presence of regulatory elements using online web based software MatInspector 2.2 available at www.genomatix.de.

5.2.2 Vector construction

Two primer sets were designed to reamplify the promoters **ProC1** and **ProC2**. The forward and the reverse primers were modified to introduce *EcoR* I and *BamH* I restriction site and the modified primer sets were designated as **ProC1F: 5' (d) GAA TTC CGA CGG CCC GGG CTG GTA A 3'** and **ProC1R: 5' (d) GGA TCC TAC TTC TTA TCC TTT TTG TTT CCT CG 3'** for promoter ProC1 , and **ProC2F: 5' (d) GAA TTC CGA TGG CCC GGG CTG GTA A 3'** and **ProC2R: 5' (d) GGA TCC ACT TCT TAT CCT TTT TTG TTT CCT TG 3'** for promoter ProC2. The modified promoters were reamplified, cloned in pGEM T Easy vector and sequenced for validation. The modified promoter fragments were separately cloned in *EcoR* I and *BamH* I digested pJ4GFP-XB (a gift from Dr Anjan Banerjee, Iowa, USA; Fig. 5.1 a) and upstream of the GFP gene. The two plasmid constructs

were designated pProG1 and pProG2 (Fig. 5.1 b, c). The pProG1 and pProG2 plasmids were digested with *EcoR* I and *Hind* III to release the ProC1/ProC2 promoter-GFP-NOS 3’ cassette. The released cassette was cloned in *EcoR* I and *Hind* III digested pCAMBIA 1300 MCS (Fig. 5.2). The two resulting plasmid were designated as pPC1G and pPC2G respectively.

5.2.3 *Agrobacterium tumefaciens* transformation

The **pPC1G** and **pPC2G** vectors harboring **ProC1 or ProC2 – GFP – NOS 3’** cassette were transformed independently into *A. tumefaciens* GV2260 (Chapter 2; section 2.3.2).

5.2.4 *Agrobacterium* mediated transformation of tobacco

Tobacco plants were transformed independently using the above *A. tumefaciens* cultures, harboring the **pPC1G** or the **pPC2G** vectors (Chapter 2; section 2.14).

5.2.5 Genomic DNA extraction and Polymerase Chain Reaction

Genomic DNA was extracted from plant leaves and PCR reactions set up as described earlier (Chapter 2; section 2.4.3 and 2.4.8).

5.2.6 Histology and Fluorescent Microscopy

Transverse sections of root, stem and leaf midrib were cut, dehydrated and rehydrated (Chapter 2; section 2.11). The dechlorophylled putative transformed plants and transverse sections of leaf mid rib, shoot and root were visualized for green fluorescence and pictures captured (Chapter 2; section 2.12).

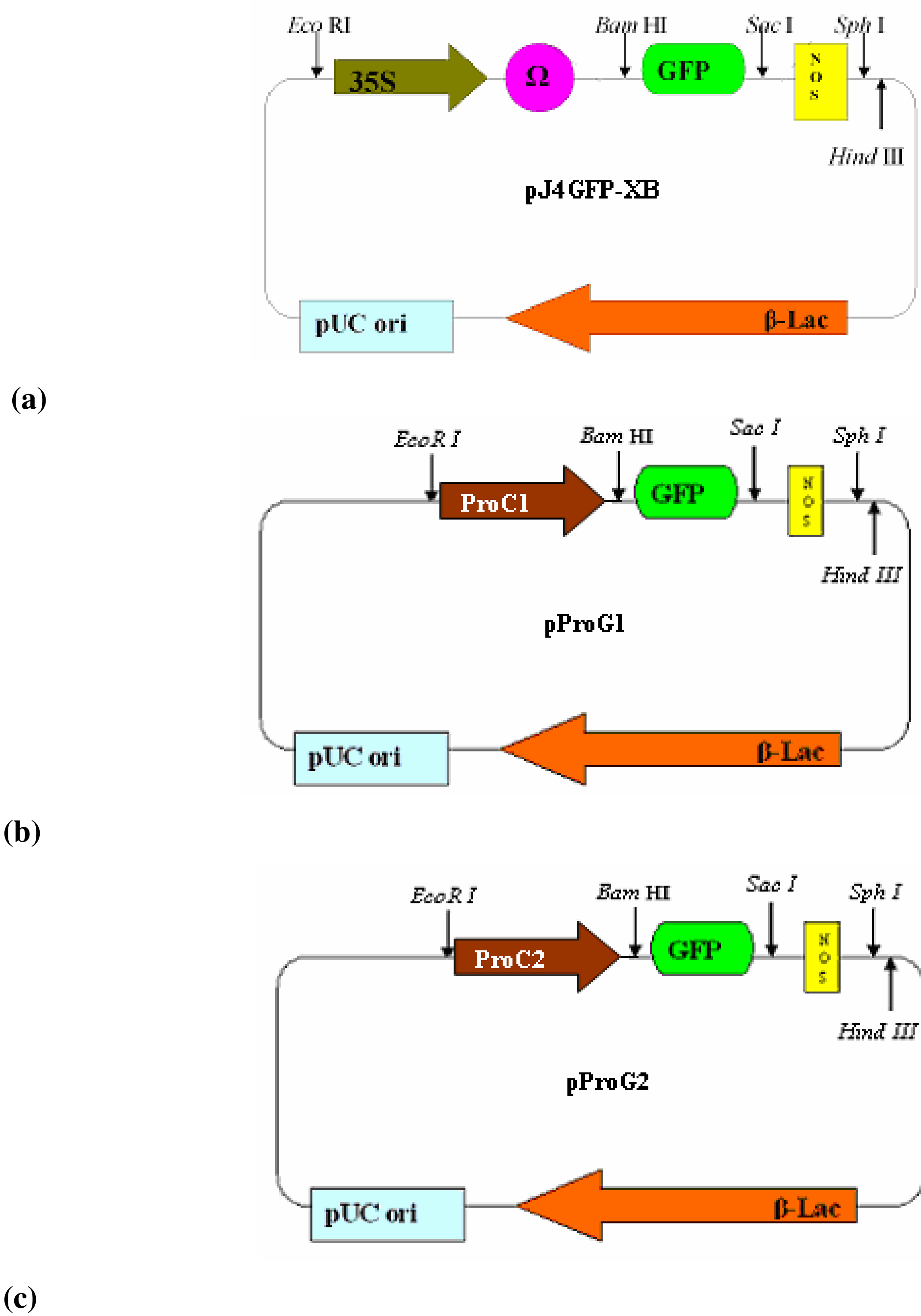
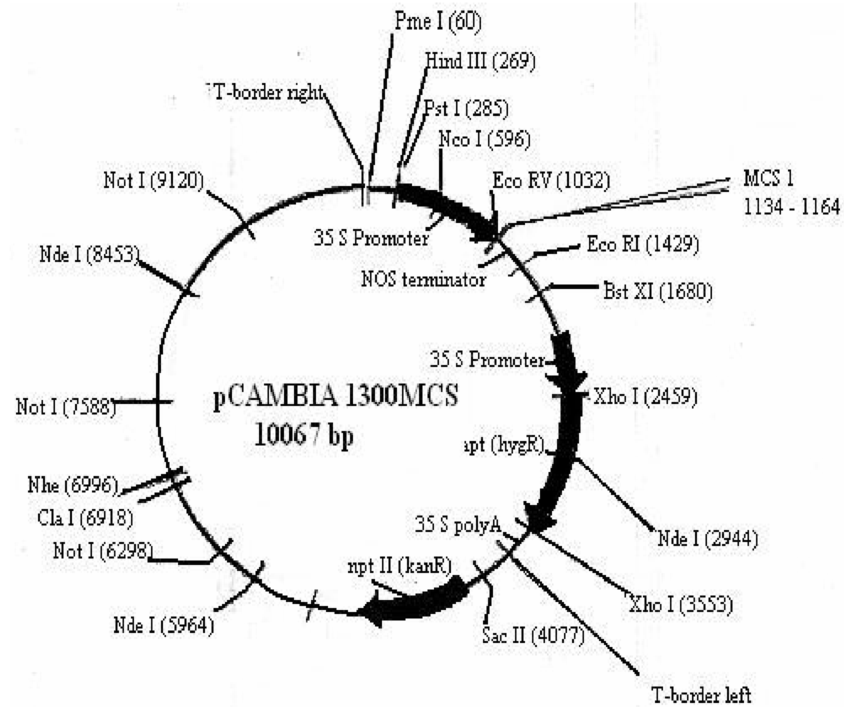
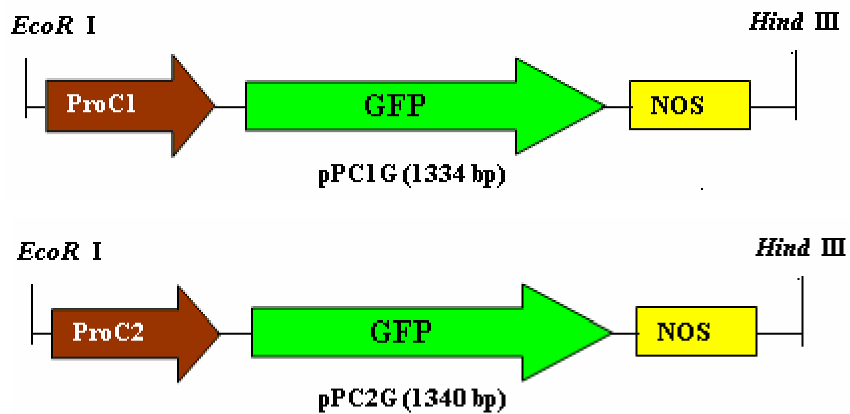


Fig. 5.1: Vector maps: (a) pJ4GFP- XB, (b) pProG1 and (c) pProG2.



(a)



(b)

Fig. 5.2: (a) Vector map of pCAMBIA 1300MCS; (b) ProC1 promoter-GFP-NOS 3' (pPC1G) and ProC2-GFP-NOS 3' (pPC2G) cassettes.

5.3 Results and Discussion

5.3.1 Promoter analysis using MatInspector version 2.2

The nucleotide sequences of the promoters **ProC1** and **ProC2** were analyzed using online software MatInspector 2.2. The analysis revealed presence of different motifs, which have been earlier reported in different plant promoters and play role in gene regulation (Table 5.1 a, b).

Fourteen and twelve gene regulation related motifs respectively were identified in ProC1 and ProC2. Out of these, eight motifs were of common occurrence in both the promoters. In promoter ProC1 six motifs were identified which were absent in promoter ProC2. Similarly three motifs were seen in ProC2, which were absent in ProC1.

Table 5.1: MatInspector 2.2 analysis of ProC1 (a) and ProC2 (b) showing presence of different motifs involved in gene regulation. Highlighted in bold letters are motifs unique to each of the promoters.

S. No	IUPAC	Position	Sequence Capitals: core sequence
1	P$FAM267/P$TAAAGSTKST1	15 - 21	GGTAAAG
2	P$FAM010/P$WBOXHVISO1	23 - 37	CATGACTTATTTATC
3	**P$FAM012/P$IBOXCORE**	**44 - 50**	**GATAATT**
4	P$FAM272/P$SV40COREENHAN	57 - 64	GTGGATTG
5	**P$FAM004/P$PALBOXLPC**	**111 - 121**	**CCTCACCAACC**
6	P$FAM003/P$MYBPLANT	114 - 124	CACCAACCACA
7	P$FAM171/P$BOXLCOREDCPAL	115 - 121	ACCAACC
8	**P$FAM263/P$DPBFCOREDCDC3**	**122 - 128**	**ACACCCG**
9	**P$FAM270/P$RAV1AAT**	**143 - 147**	**CAACA**
10	**P$FAM241/P$TATABOX2**	**203 - 209**	**TATAAAT**
11	**P$FAM061/P$AGCBOXNPGLB**	**216 - 222**	**AGCCGCC**
12	P$FAM305/P$ANAERO1CONSENSUS	293 - 299	AAACAAA
13	P$FAM014/P$MYBST1	302 - 308	AGGATAA
14	P$FAM012/P$IBOXCORENT	304 - 310	GATAAGA

(a)

S. No	IUPAC	Position	Sequence capitals: core sequence
1	**P$FAM107/P$CGACGOSAMY3**	**1 - 5**	**CGACG**
2	P$FAM267/P$TAAAGSTKST1	15 - 21	GGTAAAG
3	P$FAM010/P$WBOXHVISO1	23 - 37	CGTGACTTATTTATC
4	**P$FAM243/P$TATABOX4**	**38 - 44**	**TATATAA**
5	P$FAM267/P$TAAAGSTKST1	40 - 46	TATAAAG
6	P$FAM272/P$SV40COREENHAN	57 - 64	GTGGATTG
7	**P$FAM302/P$SITEIIATCYTC**	**85 - 95**	**TGGGCCCACGT**
8	P$FAM003/P$MYBPLANT	145 - 155	AACCAACCACT
9	P$FAM171/P$BOXLCOREDCPAL	146 - 152	ACCAACC
10	P$FAM305/P$ANAERO1CONSENSUS	299 - 305	AAACAAA
11	P$FAM014/P$MYBST1	307 - 313	AGGATAA
12	P$FAM012/P$IBOXCORENT	309 - 315	GATAAGA

(b)

5.3.1.1 Common Regulatory Motifs

The eight common motifs present in the ProC1 and ProC2 are described below:

IUPAC Name: P$TAAAGSTKST1

Position in Promoter ProC1: 15 - 21 bp.

Position in Promoter ProC2: 15 - 21 and 40 – 46 bp.

This motif belongs to family P$FAM267 and its IUPAC ambiguity code is TAAAG. This TAAAG motif is found in promoter of *Solanum tuberosum* KST1 gene, which encodes a K+ influx channel. It is the target site for trans-acting StDof1 (DNA binding with one finger) protein controlling guard cell-specific gene expression (Plesch et al., 2001).

IUPAC Name: P$WBOXHVISO1

Position in Promoter ProC1: 23 – 37 bp.

Position in Promoter ProC2: 23 – 37 bp.

This motif belongs to family P$FAM010 and its IUPAC ambiguity code is TGACT. A novel WRKY transcription factor SUSIBA2, involved in sugar signaling in barley by binding to the

sugar-responsive elements of the iso1 promoter, bind to W-box element in barley iso1 (encoding isoamylase1) promoter (Sun et al., 2003).

IUPAC Name: P$SV40COREENHAN

Position in Promoter ProC1: 57 – 64 bp.

Position in Promoter ProC2: 57 – 64 bp.

This motif belongs to family P$FAM272 and its IUPAC ambiguity code is GTGGWWHG. This sequence shows similarity to SV40 core enhancer sequence found in rbcs genes (Weiher et al., 1983; Green et al., 1987; Donald and Cashmore, 1990).

IUPAC Name: P$MYBPLANT

Position in Promoter ProC1: 114 – 124 bp.

Position in Promoter ProC2: 145 – 155 bp.

This motif belongs to family P$FAM003 and its IUPAC ambiguity code is MACCWAMC. This element is plant MYB binding site and consensus sequence is related to box P in promoters of phenylpropanoid biosynthetic genes such as PAL, CHS, CHI, DFR, CL, Bz1. The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum majus* have been shown to regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco (Tamagnone et al., 1998; Sablowski et al., 1994).

IUPAC Name: P$BOXLCOREDCPAL

Position in Promoter ProC1: 115 – 121 bp.

Position in Promoter ProC2: 146 – 152 bp.

This motif belongs to family P$FAM171 and its IUPAC ambiguity code is ACCWWCC. This element is consensus of the putative core sequence of box –L- like sequence in carrot (*Daucus carota*) PAL1 promoter region. *In vitro* it has been shown that DCMYB1 binds to these sequences (Maeda et al., 2005).

IUPAC Name: P$ANAERO1CONSENSUS

Position in Promoter ProC1: 293 – 299 bp.

Position in Promoter ProC2: 299 – 305 bp.

This motif belongs to family: P$FAM305 and its IUPAC ambiguity code is AAACAAA. It is one of 16 motifs found in silico in promoters of 13 anaerobic genes involved in the fermentative pathway (anaerobic set 1) (Mohanty et al., 2005).

IUPAC Name: P$MYBST1

Position in Promoter ProC1: 302 – 308 bp.

Position in Promoter ProC2: 307 – 313 bp.

This motif belongs to family P$FAM014 and its IUPAC ambiguity code is GGATA. It is core motif of MybSt1 (a potato MYB homolog) binding site. The MybSt1 cDNA clone was isolated by using CaMV 35S promoter domain A as a probe. The Myb motif of the MybSt1 protein is distinct from the other plant Myb DNA binding domain described so far (Baranowskij et al., 1994).

IUPAC Name: P$IBOXCORENT

Position in Promoter ProC1: 304 – 310 bp.

Position in Promoter ProC2: 309 – 315 bp.

This motif belongs to family P$FAM012 and its IUPAC ambiguity code is GATAAGR. It is "I-box core motif" in the CAMs (conserved DNA modular arrays) and associated with light-responsive promoter regions (Martinez et al., 2002).

5.3.1.2 Different Regulatory Motifs

The motifs which are present only in individual ProC1 and ProC2 promoter are as follows:

Promoter ProC1

IUPAC Name: P$IBOXCORE

Position: 44 – 50 bp

This motif belongs to family P$FAM012 and its IUPAC ambiguity code is GATAA. I-box is conserved sequence found upstream of light-regulated genes of both monocots and dicots (Terzaghi and Cashmore, 1995).

IUPAC Name: P$PALBOXLPC

Position: 111 - 121

This motif belongs to family P$FAM004 and its IUPAC ambiguity code is YCYYACCWACC. This element is ‘Box L’ which is one of three putative cis-acting elements (boxes P, A, and L) of phenylalanine ammonia lyase genes in parsley (*Petroselinum crispum*). None of these elements (boxes P, A, and L) alone, or the promoter region containing all of them together, conferred elicitor or light responsiveness. These elements appear to be necessary but not sufficient for elicitor or light-mediated PAL gene activation (Logemann et al., 1995).

IUPAC Name: P$DPBFCOREDCDC3

Position: 122 - 128

This motif belongs to family P$FAM263 and its IUPAC ambiguity code is ACACNNG. It is a novel class of bZIP (basic leucine zipper) transcription factors, DPBF-1 and 2 (Dc3 promoter-binding factor-1 and 2), binding core sequence found in the carrot Dc3 gene promoter. Dc3 expression is normally embryo-specific, and also can be induced by ABA. The *Arabidopsis* abscisic acid response gene ABI5 encodes a bZIP transcription DE factor. The abi5 mutant has pleiotropic defects in ABA response. The ABI5 DE regulates a subset of late embryogenesis-abundant genes. GIA1 (growth-insensitivity to ABA) is identical to ABI5 (Kim et al., 1997; Finkelstein et al., 2000; Lopez-Molina and Chua, 2000).

IUPAC Name: P$RAV1AAT

Position: 143 – 147 bp.

This motif belongs to family P$FAM270 and its IUPAC ambiguity code is CAACA. It is binding consensus sequence of *Arabidopsis* transcription factor, RAV1. RAV1 specifically binds to DNA with bipartite sequence motifs of RAV1-A (CAACA) and RAV1-B (CACCTG). RAV1 protein contain AP2-like and B3-like domains which recognize the CAACA and CACCTG motifs, respectively. The expression level of RAV1 was relatively high in rosette leaves and roots (Kagaya et al., 1999).

IUPAC Name: P$TATABOX2

Position: 203 – 209 bp.

This motif belongs to family P$FAM241 and its IUPAC ambiguity code is TATAAAT. It is commonly known as "TATA box". This TATA box is found in the 5' upstream region of pea legA gene (Shirsat et al., 1989), sporamin A of sweet potato and in beta-phaseolin promoter (Grace et al., 2004). The sequence and spacing of TATA box elements are critical for accurate initiation (Grace et al., 2004).

IUPAC Name: P$AGCBOXNPGLB

Position: 216 – 222 bp.

This motif belongs to family P$FAM061 and its IUPAC ambiguity code is AGCCGCC. This element, "AGC box", is repeated twice in a 61 bp enhancer element in tobacco class I beta-1,3-glucanase (GLB) gene. "GCC-box" which is a binding sequence of *Arabidopsis* AtERFs (ethylene response factors). AtERF1, 2 and 5 functions as activators of GCC box-dependent transcription whereas AtERF3 and 4 acts as repressors. AtERF proteins are stress signal-response factors and are found in the promoter of tobacco Osmotin-like protein (OLP) genes. It is EREBP2 (ethylene-responsive element binding proteins) binding site and is conserved in most PR-protein genes (Hart et al., 1993; Sato et al., 1996; Fujimoto et al., 2000; Ohme et al., 2000).

Promoter ProC2

IUPAC Name: P$CGACGOSAMY3

Position: 1 – 5 bp.

This motif belongs to family P$FAM107 and its IUPAC ambiguity code is CGACG. "CGACG element" has been reported from the GC-rich regions of the rice Amy3D and Amy3E amylase genes, but not in Amy3E gene (Hwang et al., 1998). This element may function as a coupling element for the G box element.

IUPAC Name: P$TATABOX4

Position: 38 – 44 bp.

This motif belongs to family P$FAM243 and its IUPAC ambiguity code is TATATAA. It is commonly known as "TATA box". This TATA box is found in the 5'upstream region of sweet potato sporamin A gene and in beta-phaseolin promoter (Grace et al., 2004). Sequence and spacing of TATA box elements are critical for accurate initiation (Grace et al., 2004).

IUPAC Name: P$SITEIIATCYTC

Position: 85 – 95 bp.

This motif belongs to family P$FAM302 and its IUPAC ambiguity code is TGGGCY. This element, "Site II element", is found in the promoter regions of cytochrome genes (Cytc-1, Cytc-2) in *Arabidopsis* and is located between -147 and -156 from the translational start sites (Welchen et al., 2005).

5.3.2 Promoter ProC1 and ProC2 directed GFP gene expression in Tobacco

The two isolated promoters ProC1 (313 bp) and ProC2 (319 bp) were relatively small compared to earlier reported promoters sequences for different phenylpropanoid pathway genes (Bevan et al., 1989; Ohl et al., 1990; Hauffe et al., 1991; Capellades et al., 1996; Mizutani et al., 1997; Bell-Lelong et al., 1997). Grimmig and Matern (1997) reported ~5.0 Kb promoter from parsley CCoAOMT gene. Two CCoAOMT promoters of 2.0 Kb and 1.4 Kb have been reported from poplar (Chen et al., 2000). The two promoters ProC1 and ProC2 were used to construct vectors pPC1G and pPC2G where the GFP gene was placed under their control (section 5.2.2 above).

5.3.2.1 Tobacco transformation

Tobacco (*N. tabaccum* var. Anand 119) leaf discs were transformed separately with *A. tumefaciens* cultures harboring the pPC1G or the pPC2G vectors. Shoots regenerated *in vitro* under selection pressure from the cut surface of the leaf disc after 2 weeks (Fig. 5.3). The sets of putative transformants were designated as CProG1 and CProG2 depending on the promoter used.

The regenerants were allowed to grow for 12 weeks and then shifted to root induction medium. Roots were initiated within 2 weeks of shifting (Fig. 5.4). The transformed plants were hardened and transferred to pots (Fig. 5.5 a, b).

5.3.2.2 GFP visualization in transformed tobacco plants

The two transformation events, CProG1 and CProG2 were visualized at 395nm using NightSea GFP flash light (NightSea, USA). The dechlorophylled putative transgenic tobacco plants, (regenerated *in vitro*) showed green fluorescence, when illuminated at 395nm (Fig. 5.6 a, b). No fluorescence was seen in untransformed control tobacco plants (Fig. 5 c). This was first indication of the integration and expression of the GFP gene under the control of both the promoters ProC1 and ProC2. As compared to other reported CCoAOMT promoters, the ProC1 and ProC2 were small and are possibly partial promoter sequences. Grimmig and Matern (1997) have shown that a 5’ truncated 354 bp fragment of CCoAOMT promoter from parsley was sufficient to drive the expression of GUS gene. The truncated ProC1 and ProC2 also are sufficient to drive GFP gene expression respectively.

5.3.2.3 Integration of promoters ProC1, ProC2 and GFP in tobacco genome

Genomic DNA isolated from untransformed control tobacco and the two transformation events CProG1 and CProG2 was used as template for PCR based amplification of the GFP gene using gene specific forward primer **GF: 5’ (d) ATG GTG AGC AAG GGC GAG GAG CTG TTC A 3’** and the reverse primer **GR: 5’ (d) TTA GAT CTC TTG TAC AGCTCG TCC ATG CCG TG 3’.** A fragment of ~700bp was amplified from the genomic DNA templates of CProG1 and CProG2. There was no amplification from the untransformed tobacco plant (Figure 5.7 a). The amplicons were isolated, cloned in pGEM-T Easy vector and sequenced. The nucleotide sequence of the amplicons was confirmed to be that of the GFP gene. PCR amplifications from the genomic DNA of the transformation events CProG1 and CProG2 with forward primer Pro1F and Pro2F designed from ProC1 and ProC2 promoter sequences respectively and the GFP gene reverse primer GR, amplified an ~1Kb band (Fig. 5.7 b). The amplicons were cloned in pGEM-T Easy vector and sequenced. The sequence data matched the promoter plus the GFP gene sequences. These results indicate the

integration of the CCoAOMT promoters ProC1 and ProC2 together with the GFP gene in the tobacco genome.

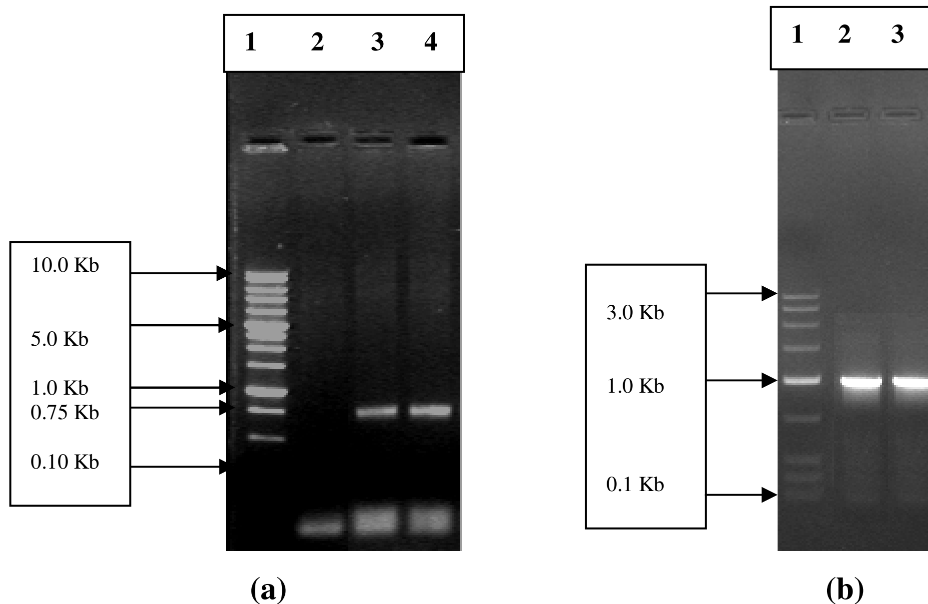


Fig. 5.7: (a) PCR amplification of GFP from genomic DNA of control plant (lane 2), CProG1 (lane 3), CProG2 (lane 4). DNA size marker (lane 1). (b) PCR amplification of Promoter plus GFP from genomic DNA of control plant (lane 2), CProG1 (lane 3), CProG2 (lane 4). DNA size marker (lane 1).

5.3.2.4 Visualization of GFP in transverse sections of leaf midrib, stem, and root of CProG1 and CProG2 transgenic tobacco plants

Transverse sections of leaf midrib, stem and root of transgenic tobacco plants, CProG1 and CProG2, were cut. The sections were dehydrated and rehydrated by passing them through alcohol: water series. The sections were mounted in glycerol, visualized at 395nm under fluorescence microscope and pictures captured. In the transverse sections of leaf mid rib and stem, green fluorescence was observed, however, the fluorescence was limited to the xylem tissue. Green fluorescence was also observed in occasional epidermal cells (Fig. 5.8 a-d).

The pattern of GFP gene expression in the roots was, however, different. Fluorescence was not limited to the xylem tissue. Fluorescence was also seen on the inner side of the epiblema cell layer and in the pericycle ring (Fig. 5.8 e, f). This observation is in agreement with the involvement of phenylpropanoid pathway genes during suberin biosynthesis. The aromatic moiety of suberin is synthesized via the general phenylpropanoid pathway with its key enzyme PAL (Kolattukudy, 1981). There was no apparent difference in the level or pattern of GFP expression as driven by the ProC1 and ProC2 promoter as was obvious by the extent of fluorescence of GFP in the two transformation events.

The xylem specific expression of GFP gene under control of the two promoters, ProC1 and ProC2, proved that the two truncated promoters were sufficient to drive the expression of a gene (GFP gene in the present case). Similar observation was made by Grimmig and Matern (1997) using truncated CCoAOMT promoter. Moreover, it was evident from the expression pattern of the GFP gene, that the gene regulation elements directing xylem specific gene expression were present in truncated promoters. These results are in confirmation with the CCoAOMT immunocytolocalization results where CCoAOMT protein was shown to be localized in the xylem tissue (Chapter 4; section 4.3.4). Since both ProC1 and ProC2 promoters from CCoAOMT1 and CCoAOMT2 drive GFP gene expression in the xylem tissue, the observation from QPCR that the CCoAOMT1 and CCoAOMT2 are expressed in tandem is substantiated here. However, the stomata guard cell specific expression of GFP was not observed though the TAAAG motif, responsible for it, was present in both the CPro1 and CPro2 promoters. Green fluorescence observed in tissue other than xylem especially in roots, suggests presence of other unidentified elements regulating CCoAOMT1 and CCoAOMT2 expression in roots.

The above results from promoter ProC1 and ProC2 were similar to earlier findings (Grimmig and Matern, 1997 and Chen et al., 2000) of CCoAOMT being selectively expressed in xylem and fibers. Negligible expression in phloem was also in agreement.

5.4 Conclusions

- Analysis of the two CCoAOMT1 (ProC1) and CCoAOMT2 (ProC2) promoter nucleotide sequences revealed presence of different cis-regulatory elements involved in CCoAOMT gene regulation.

- Tobacco plants transformed with cassettes where promoters ProC1 and ProC2 drive GFP gene were recovered. GFP was visualized in xylem and fibers in leaf mid rib, shoot and root.
- In root GFP was visualized in tissues other than xylem suggesting involvement of other regulatory elements in the two promoters.

5.5 References

Baranowskij N, Frohberg C, Prat S, Willmitzer L (1994) A novel DNA binding protein with homology to Myb oncoproteins containing only one repeat can function as a transcriptional activator. EMBO J 13: 5383-5392

Bell-Lelong DA, Cusumano JC, Meyer K, Chapple C (1997) Cinnamate-4-hydroxylase expression in Arabidopsis: regulation in response to development and the environment. Plant Physiol 113: 729-738

Bevan M, Shufflebottom D, Edwards K, Jefferson R, Schuch W (1989) Tissue- and cell-specific activity of a phenylalanine ammonia-lyase promoter in transgenic plants. EMBO J 8: 1899-1906

Capellades M, Torres MA, Bastisch I, Stiefel V, Vignols F, Bruce WB, Peterson D, Puigdoménech P, Rigau J (1996) The maize caffeic acid *O*-methyltransferase gene promoter is active in transgenic tobacco and maize plant tissues. Plant Mol Biol 31: 307-322

Chen C, Meyermans H, Burggraeve B, De Rycke RM, Inoue K, De Vleesschauwer V, Steenackers M, Van Montagu MC, Engler GJ, Boerjan WA (2000) Cell-specific and conditional expression of caffeoyl-CoA O-methyltransferase in poplar. Plant Physiol 123: 853–867

Cramer CL, Edwards K, Dron M, Liang X, Dildine SL, Bolwell GP,Dixon RA, Lamb CJ, Schuch W (1989) Phenylalanine ammonia lyase gene organization and structure. Plant Mol Biol 12:367–383

Donald RGK, Cashmore AR (1990) Mutation of either G box or I box sequences profoundly affects expression from the Arabidopsis rbcs-1A promoter. EMBO J 9: 1717-1726

Douglas CJ, Hauffe KD, Ites-Morales M-E, Ellard M, Paszkowski U, Hahlbrock K, Dangl JL (1991) Exonic sequences are required for elicitor and light activation of a plant

defense gene, but promoter sequences are sufficient for tissue specific expression. EMBO J 10: 1767–1775

Feuillet C, Lauvergeat V, Deswarte C, Pilate G, Boudet A, Grima-Pettenati J (1995) Tissue- and cell-specific expression of a cinnamyl alcohol dehydrogenase promoter in transgenic poplar plants. Plant Mol Biol 27: 651–667

Finkelstein RR, Lynch TJ (2000) The *Arabidopsis* abscisic acid response gene ABI5 encodes a basic leucine zipper transcription factor. Plant Cell 12: 599-609

Fujimoto SY, Ohta M, Usui A, Shinshi H, Ohme-Takagi M (2000) *Arabidopsis* ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. Plant Cell 12: 393-404

Grace ML, Chandrasekharan MB, Hall TC, Crowe AJ (2004) Sequence and spacing of TATA box elements are critical for accurate initiation from the beta-phaseolin promoter. J Biol Chem 279: 8102-8110

Green PJ, Kay SA, Chua N-H (1987) Sequence-specific interactions of a pea nuclear factor with light-responsive elements upstream of the rbcs3A gene. EMBO J 6: 2543-2549

Grimmig B, Matern U (1997) Structure of the parsley caffeoyl-CoA *O*-methyltransferase gene, harbouring a novel elicitor responsive *cis*-acting element. Plant Mol Biol 33: 323-341

Hart CM, Nagy F, Meins Jr F (1993) A 61 bp enhancer element of the tobacco beta-1, 3-glucanase B gene interacts with one or more regulated nuclear proteins. Plant Mol Biol 21: 121-131

Hauffe KD, Paszkowski U, Schulze-Lefert P, Hahlbrock K, Dangl JL, Douglas CJ (1991) A parsley 4CL-1 promoter fragment specifies complex expression patterns in transgenic tobacco. Plant Cell 3: 435-443

Hwang YS, Karrer EE, Thomas BR, Chen L, Rodriguez RL (1998) Three cis-lements required for rice alpha-amylase Amy3D expression during sugar starvation. Plant Mol Biol 36: 331-341

Kagaya Y, Ohmiya K, Hattori T (1999) RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. Nucleic Acids Res 27: 470-478

Kim SY, Chung HJ, Thomas TL (1997) Isolation of a novel class of bZIP transcription factors that interact with ABA-responsive and embryo-specification elements in the Dc3 promoter using a modified yeast one-hybrid system. Plant J 11: 1237-1251

Kolattukudy PE (1981) Structure, biosynthesis, and biodegradation of cutin and suberin. Ann Rev Plant Physiol 32: 539–567

Levya A, Liang X, Pintor-Toro JA, Dixon RA, Lamb CJ (1992) cis-Element combinations determine phenylalanine ammonia-lyase gene tissue-specific expression patterns. Plant Cell 4: 263–271

Logemann E, Parniske M, Hahlbrock K (1995) Modes of expression and common structural features of the complete phenylalanine ammonia-lyase gene family in parsley. Proc Natl Acad Sci USA 92: 5905-5909

Lois R, Dietrich A, Hahlbrock K, Schulz W (1989) A phenylalanine ammonia-lyase gene from parsley: structure, regulation and identification of elicitor and light responsive cis-acting elements. EMBO J 8: 1641–1648

Lopez-Molina L, Chua NH (2000) A null mutation in a bZIP factor confers ABA-insensitivity in *Arabidopsis thaliana*. Plant Cell Physiol 41: 541-547

Maeda K, Kimura S, Demura T, Takeda J, Ozeki Y (2005) DcMYB1 acts as a transcriptional activator of the carrot phenylalanine ammonia-lyase gene (DcPAL1) in response to elicitor treatment, UV-B irradiation and the dilution effect. Plant Mol Biol 59: 739-752

Martinez-Hernandez A, Lopez-Ochoa L, Arguello-Astorga G, Herrera-Estrella L (2002) Functional properties and regulatory complexity of a minimal RBCS light-responsive unit activated by phytochrome, cryptochrome, and plastid signals. Plant Physiol 128: 1223-1233

Mizutani M, Ohta D, Sato R (1997) Isolation of a cDNA and a Genomic Clone Encoding Cinnamate 4-Hydroxylase from Arabidopsis and Its Expression Manner in Planta. Plant Physiol 113: 755-763

Mohanty B, Krishnan SP, Swarup S, Bajic VB (2005) Detection and preliminary analysis of motifs in promoters of anaerobically induced genes of different plant species. Ann Bot 96: 669-681

Neustaedter DA, Lee SP, Douglas CJ (1999) A novel parsley *4CL1* cis-element is required for developmentally regulated expression and protein-DNA complex formation. Plant J 18: 77–88

Ohl S, Hedrick SA, Chory J, Lamb CJ (1990) Functional Properties of a Phenylalanine Ammonia-Lyase Promoter from *Arabidopsis.* Plant Cell 2: 837-848

Ohme-Takagi M, Suzuki K, Shinshi H (2000) Regulation of Ethylene-Induced Transcription of Defense Genes. Plant Cell Physiol 41: 1187-1192

Plesch G, Ehrhardt T, Mueller-Roeber B (2001) Involvement of TAAAG elements suggests a role for Dof transcription factors in guard cell-specific gene expression. Plant J 28: 455-464

Sablowski RWM, Baulcombe DC, Bevan M (1995) Expression of a flower-specific Myb protein in leaf cells using a viral vector causes ectopic activation of a target promoter. Proc Natl Acad Sci 92: 6901–6905

Sablowski RWM, Moyano E, Culianez-Macia FA, Schuch W, Martin C, Bevan M (1994) A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. EMBO J 13: 128-137

Sato F, Kitajima S, Koyama T, Yamada Y (1996) Ethylene-induced gene expression of osmotin-like protein, a neutral isoform of tobacco PR-5, is mediated by the AGCCGCC cis-sequence. Plant Cell Physiol 37: 249-255

Shirsat A, Wilford N, Croy R, Boulter D (1989) Sequences responsible for the tissue specific promoter activity of a pea legumin gene in tobacco. Mol Gen Genet 215: 326-331

Sun C, Palmqvist S, Olsson H, Boren M, Ahlandsberg S, Jansson C (2003) A novel WRKY transcription factor, SUSIBA2, participates in sugar signaling in barley by binding to the sugar-responsive elements of the iso1 promoter. Plant Cell 15: 2076-2092

Tamagnone L, Merida A, Parr A, Mackay S, Culianez-Macia FA, Roberts K, Martin C (1998) The AmMYB308 and AmMYB330 transcription factors from antirrhinum regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. Plant Cell 10: 135-154

Terzaghi WB, Cashmore AR (1995) Light-regulated transcription. Annu Rev Plant Physiol Plant Mol Biol 46: 445-474

Weiher H, Konig M, Gruss P (1983) Multiple point mutations affecting the simian virus 40 enhancer. Science 219: 626-631

Welchen E, Gonzalez DH (2005) Differential expression of the *Arabidopsis* cytochrome c genes Cytc-1 and Cytc-2. Evidence for the involvement of TCP-domain protein-binding elements in anther and meristem-specific expression of the Cytc-1 gene. Plant Physiol 139: 88-100

Yukawa Y, Sugita M, Choisne N, Small I, Sugiura M (2000) The TATA motif, the CAA motif and the poly(T) transcription termination motif are all important for transcription re-initiation on plant tRNA genes. Plant J 22: 439-447

Summary

The present study was aimed at understanding lignin biosynthesis pathway in *Leucaena leucocephala*. Choice of *L. leucocephala* was based on the fact that this hard wood tree is used exclusively in India for pulp and paper production with contribution of around 25%. No study has been done so far in this regard anywhere. CCoAOMT being a key enzyme responsible for diversion of flux towards formation of G lignin was chosen as target gene as it may be helpful for development of transgenic *L. leucocephala* plants with desired characters suitable for Indian pulp and paper industry.

The findings of present study are as follows CCoAOMT in *L. leucocephala* is gene family of possibly 3 members. Two of the possible three, CCoAOMT1 and CCoAOMT2 genomic gene clones, 1292 bp each were isolated. Their NCBI GenBank database accession numbers are DQ517929 and DQ517930 respectively. The genes show 92.7% nucleotide sequence similarity with each other. Nucleotide sequence similarity with CCoAOMT genes from other plants is between 70-90%.

CCoAOMT1 and CCoAOMT2 genomic clones comprised of five exons and four introns. Deduced coding sequence of both the genomic CCoAOMT1 and CCoAOMT2 is of 735 nucleotides. Two 5' upstream (313 and 319 bp) and one 3' downstream (265 bp) nucleotide sequence of CCoAOMT1 and CCoAOMT2 were isolated by Genome walking. The isolated 3' downstream UTR is from CCoAOMT2 gene.

Two cDNA clones of CCoAOMT1 and CCoAOMT2, 735 bp each, were isolated. Their NCBI GenBank database accession numbers are DQ431233 and DQ431234. They show 97% nucleotide and 99% deduced amino acid sequence similarity with each other. Nucleotide sequence similarity with CCoAOMT cDNA gene clones from other plants was 70-90%. Deduced amino acid sequences of CCoAOMT1 and CCoAOMT2 genes show presence of SAM binding domain I (LIDLVKVGGVI), domain II (VAPPDAPLRKYV) and domain III (ALAVDPRIEI) which are also present in other plant CCoAOMTs. Two 57 bp long 5'UTR associated with both the CCoAOMT1 and CCoAOMT2 cDNA genes were isolated using 5' RACE. Similarly two 3'UTRs of 189 bp and 176 bp associated with the two CCoAOMT1 and CCoAOMT2 cDNA genes respectively were isolated by 3' RACE.

Distance tree results group the CCoAOMT1 and CCoAOMT2 genes from *L. leucocephala* with CCoAOMTs of other Fabaceae members.

The CCoAOMT gene was expressed in *E. coli* BL21 (DE3) and protein purified from inclusion bodies using Ni-chelated affinity column. Polyclonal antibodies were raised against purified CCoAOMT protein in rabbit. CCoAOMT specific polyclonal IgG were purified using Affi-gel 15 affinity matrix.

Transverse sections of different plant parts of different age stained with phloroglucinol-HCL show increase in number of differentiating xylem cells as well as their stain intensity indicated progression of lignification with age. When these transverse sections were visualization under polarized light, phloroglucinol-HCL stained tissues showed different colour in different plant parts of different ages, suggesting chemical compositional changes in lignin with tissue age.

CCoAOMT was immunolocalized in xylem and fibers suggesting its presence at the sites of extensive lignification. The semiquantitative and QPCR results showed that both the CCoAOMT1 and CCoAOMT2 genes were expressed in tandem in all tissues. The genes are, however, differentially expressed with the tissue age.

Analysis of the two CCoAOMT1 (ProC1) and CCoAOMT2 (ProC2) promoter nucleotide sequences was done using MatInspector 2.2 revealed presence of different cis-regulatory elements involved in CCoAOMT gene regulation. Tobacco plants transformed with cassettes where promoters ProC1 and ProC2 drive GFP gene were recovered. GFP was visualized in xylem and fibers in leaf mid rib, shoot and root. In root GFP was visualized in tissues other than xylem suggesting involvement of other regulatory elements in the two promoters. The above GFP expression and visualization results confirmed that the two promoters were able to drive GFP expression in the aerial plant organs in a tissue specific fashion. The gene expression in the roots was across all tissue types. However, the GFP expression was higher in epiblema and pericycle cells as compared to the other tissue.

RESEARCH PAPERS/ ABSTRACTS PUBLISHED/ ACCEPTED FOR PUBLICATION/ COMMUNICATED FOR PUBLICATION

1. **M. Arha**, S. K. Gupta, N. M. Shaik, S. Srivastava, A. K. Yadav, P. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal **(2006)** Metabolic engineering of *Leucaena leucocephala* for eco-friendly paper and pulp industry. At proceedings of National Science Day held at NCL, February'2006.
2. **M. Arha**, S. K. Gupta, N. M. Shaik, S. Srivastava, A. K. Yadav, P. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal **(2006)** An approach for genetic modification of *leucaena leucocephala* for eco-friendly pulp and paper production. At "International Symposium on Frontiers of Geneti Engineering and Biotechnology: Retrospect and Prospect", at Osmania University, Hyderabad, Jan., 2006.
3. O. U. Abhilash, **M. Arha**, S. K. Gupta, N. M. Shaik, S. Srivastava, A. K. Yadav, P. Kulkarni, B. M. Khan and S. K. Rawal **(2006)** Cell wall bound peroxidase from *Leucaena leucocephala-* A candidate enzyme involved in lignin polymerisation. At the "International Symposium on Frontiers of Geneti Engineering and Biotechnology: Retrospect and Prospect", at Osmania University, Hyderabad, Jan., 2006.
4. **M. Arha**, S. K. Gupta, N. M. Shaik, S. Srivastava, A. K. Yadav, P. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal **(2005)** Isolation, cloning and characterization of Caffeoyl CoA 3-O- methyl transferase (CCoAOMT) from *Leucaena leucocephala*. At the "National Symposium on Plant Biotechnology: New Frontiers", at CIMAP, Lucknow, Nov., 2005
5. **M. Arha**, S. K. Gupta, N. M. Shaik, S. Srivastava, A. K. Yadav, P. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal **(2005)** High frequency regeneration and transformation of *Leucaena leucocephala.* At the "National Symposium on Plant Biotechnology: New Frontiers", at CIMAP, Lucknow, Nov., 2005
6. N. M. Shaik, **M. Arha**, S. K. Gupta, S. Srivastava, A. K. Yadava, P. Kulkarni, O. U.Abhilash, B. M. Khan and S. K. Rawal **(2005)**. Multiple shoot regeneration from cotyledonary nodes of *Leucaena leuc*ocephala. ICAR. At National Symposium on "Biotechnological Interventions for Improvement of Horticultural Crops: Issues and Strategies. Kerala Agriculture University, Thrissur, Kerala 10-12 January, 2005

7. **M. Arha**, S. K. Gupta, N. M. Shaik, S. Srivastava, A. K. Yadav, P. Kulkarni, O. U. Abhilash, B. M. Khan and S. K. Rawal **(2004).** Assessment of *Leucaena* Biodiversity Using Molecular Marker at proceedings of National Science Day Feb 2004 held at National Chemical Laboratory.
8. T.V. N. Ramachander, D. Rohini Devi, **Manish Arha** and S. K. Rawal **(2003)**. A 4.3 kb g-DNA fragment from *Streptomyces aureofaciens* supports PHB synthesis in recombinant *E. coli.* At "Proceedings of the 5th National Symposium on Biochemical Engineering and Biotechnology". BE19. 2003.
9. D.Rohini, T.V.N. Ramachander, **Manish Arha** and S.K. Rawal **(2003)**. Low cost production of Polyhydroxybutyrate by *Bacillus* sp. using Molasses as carbon source. At "Proceedings of the 5th National Symposium on Biochemical Engineering and Biotechnology". BE13. 2003.
10. N. M. Shaik, M. Arha, S. K. Gupta, S. Srivastava, A. K. Yadav, P. S. Kulkarni, O.U. Abhilash, V.L. Sirisha, S. Prashant, D. Randheer, P.B. Kavi Kishor B. M. Khan and S. K. Rawal (2006). Multiple shoot regeneration from cotyledonary nodes of *Leucaena leucocephala.* **(Communicated)**
11. V.L. Sirisha, S. Prashant, D. Randheer, N. M. Shaik, **M. Arha**, S. K. Gupta, S. Srivastava, A. K. Yadav, P. S. Kulkarni, O.U. Abhilash, B. M. Khan, S. K. Rawal and P.B. Kavi Kishor **(2006)** Direct shoot organogenesis and plant regeneration from hypocotyl explants in selected genotypes of *Leucaena leucocephala* – a leguminous pulpwood tree species. **(Communicated).**
12. P. Nataraj Sekhar, **Manish Arha**, S. K. Gupta, Sameer Srivastava, A. K. Yadav, Ranadheer Kumar, V.L. Sirisha, S. Prashant, Bashir Khan, P.B. Kavi Kishor and S.K. Rawal **(2007)** Homology modeling, molecular dynamics simulation and docking studies of caffeoyl–CoA-*O*- methyl transferase (CCoAOMT 1 and 2) isoforms in a leguminous tree subabul (*Leucaena leucocephala*). **(Communicated)**

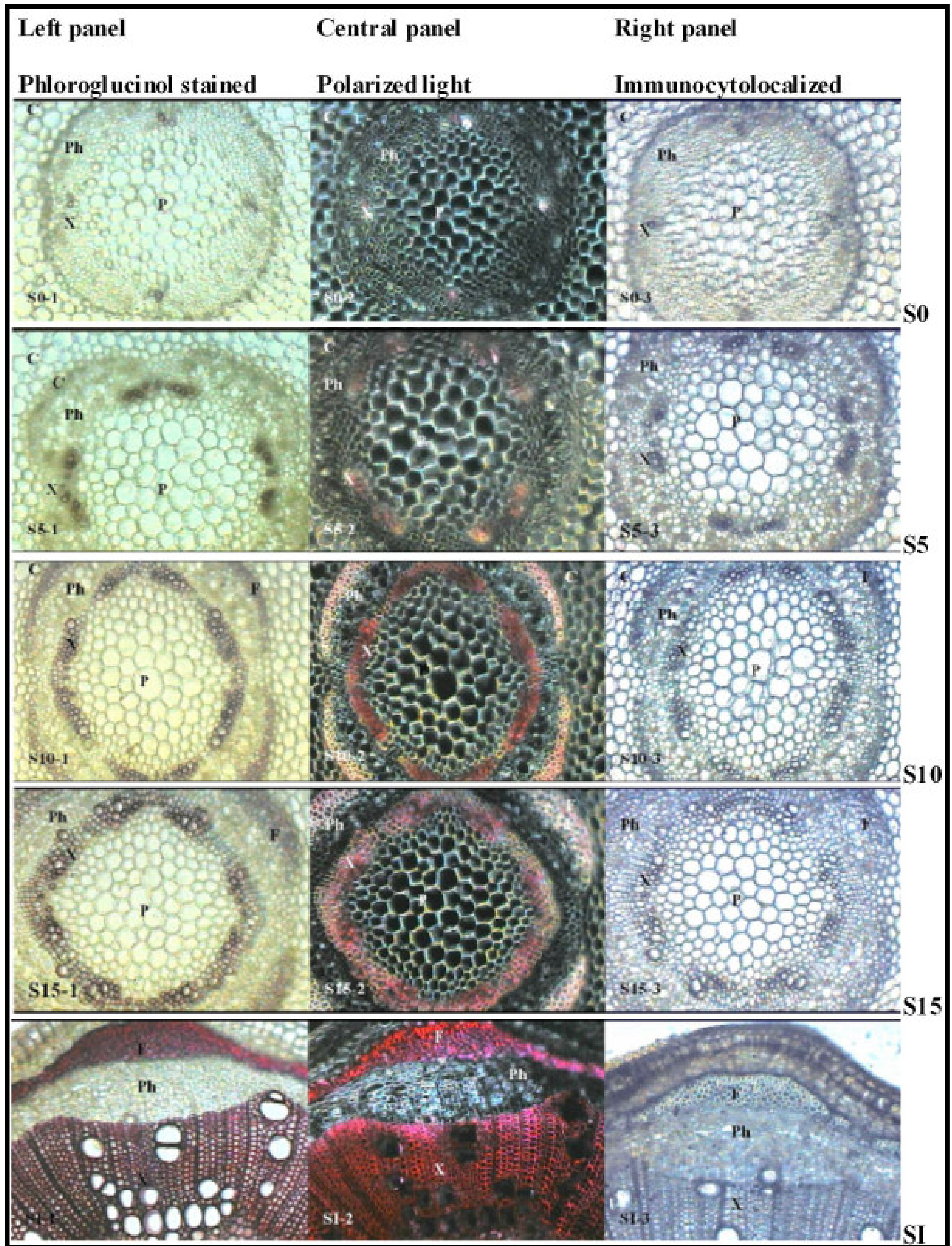


Fig 4.6 a **Contd…**

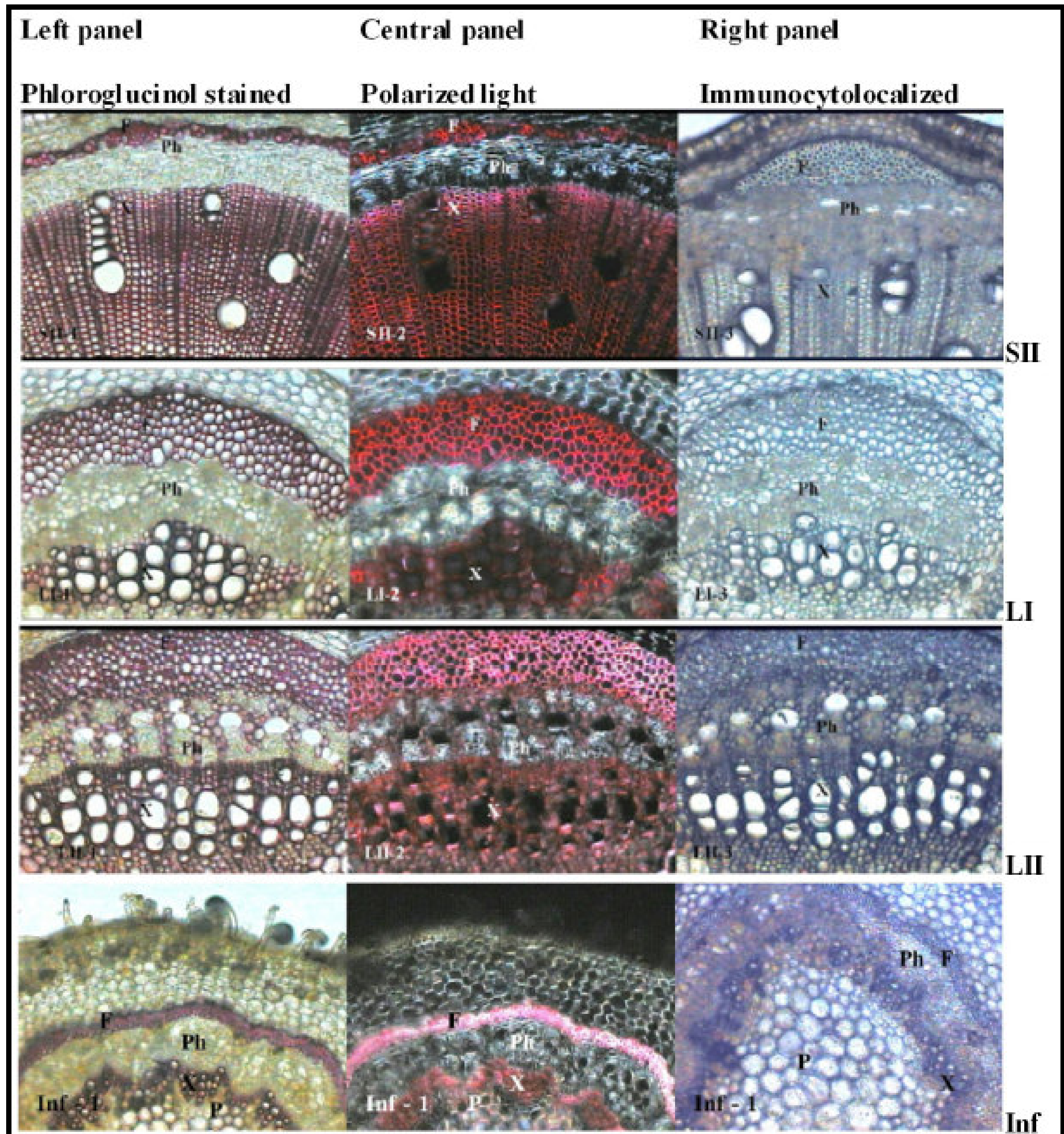


Fig. 4.6 a: Transverse sections of shoot, leaf rachis and inflorescence of *L. leucocephala*. Left panel are phloroglucinol-HCl stained sections visualized under normal light: Central panel are phloroglucinol-HCl stained sections visualized under polarized light; Right panel shows immunocytolocalization of CCoAOMT (20 X magnification).

S0 (1-3)- 0 day seedling shoots (when seedling fully emerges from the seed coat.), S5 (1-3)- 5 days old seedling shoots, S10 (1-3)- 10 days old seedling shoots, S15 (1-3)- 15 days old seedling shoots, SI (1-3)- shoots one season old, SII (1-3)- shoots two seasons old, LI (1-3)- Leaf rachis of one season old plant, LII (1-3)- Leaf rachis of two seasons old plant, Inf (1-3)- Inflorescence stalk. (X- Xylem, Ph- Phloem, F- Phloem fiber, C- Cortex, P- Pith)

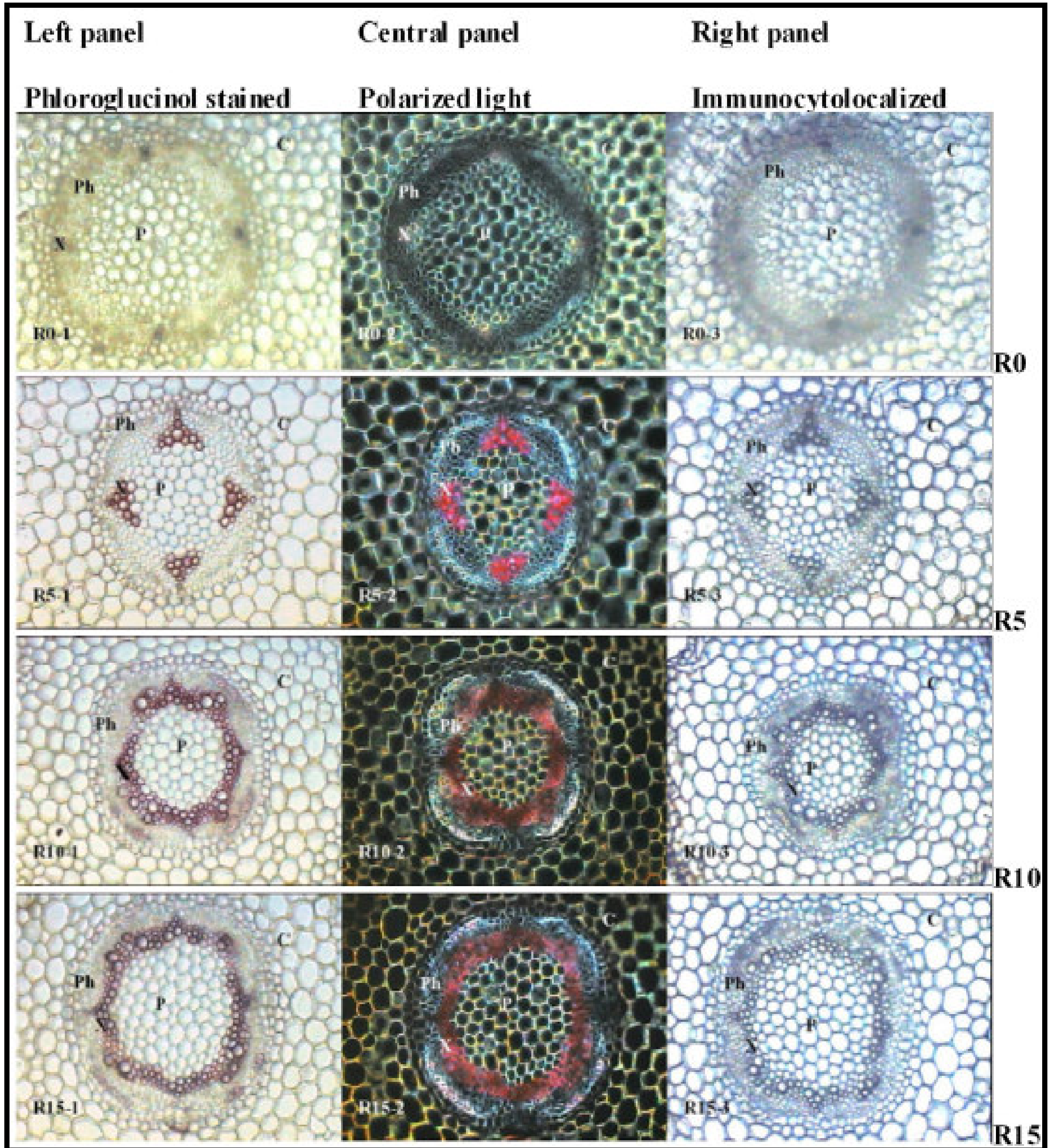


Fig 4.6 b **Contd…**

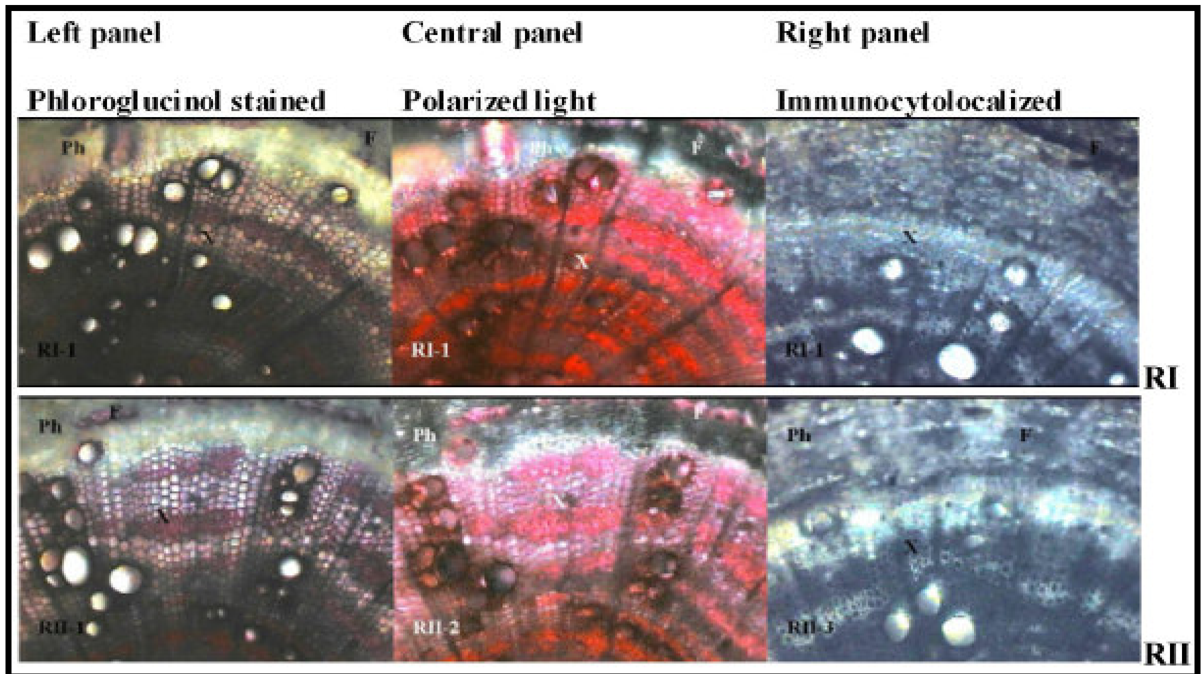


Fig. 4.6 b: Transverse sections of root of *L. leucocephala.* Left panel phloroglucinol-HCl stained sections visualized under normal light, Central panel phloroglucinol-HCl stained sections visualized under polarized light and Right panel sections immunocytolocalization of CCoAOMT (20 X magnification).

R0 (1-3)- 0 day old seedling root (when seedling fully emerges from the seed coat), R5 (1-3)- 5 days old seedling root, R10 (1-3)-10 days old seedling root, R15 (1-3)-15 days old seedling root, RI (1-3)- one season old plant root, RII (1-3)- two seasons old plant root. (X- Xylem, Ph- Phloem, F- Phloem fiber, C-Cortex, P- Pith)

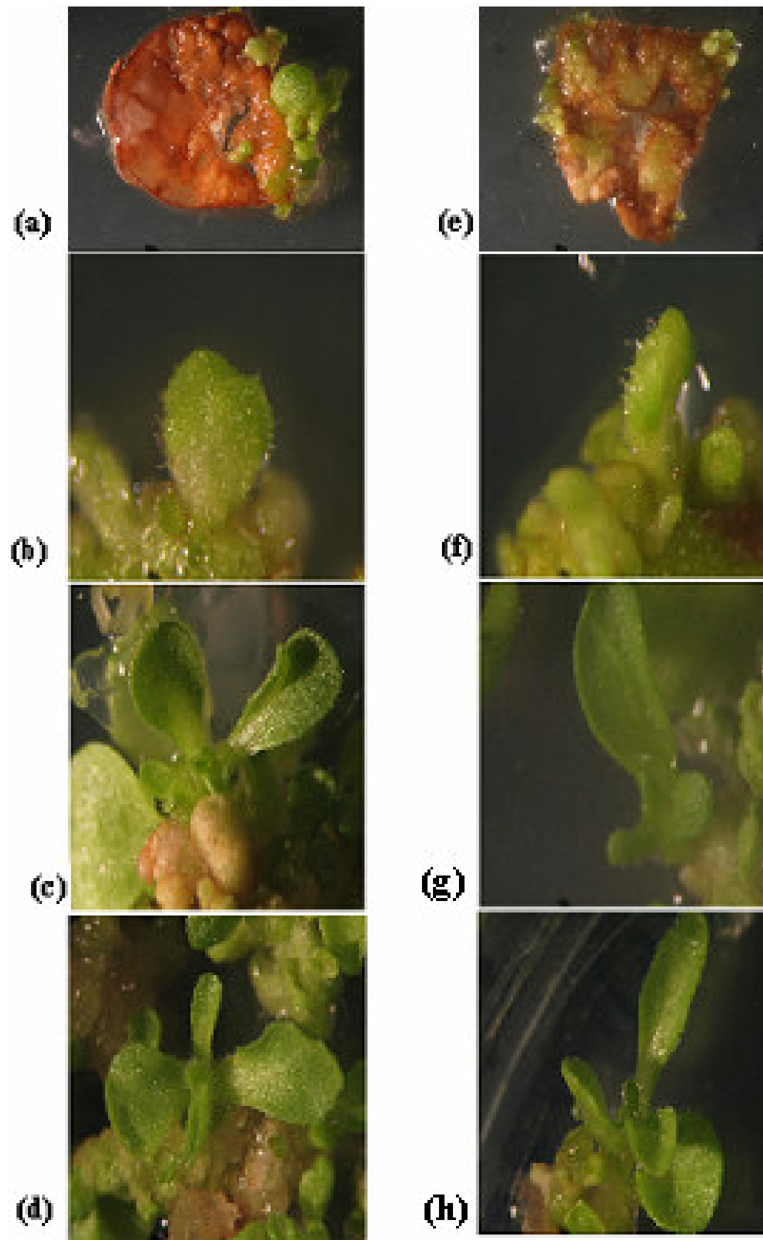


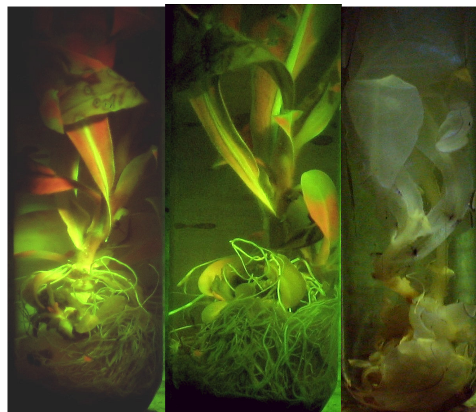
Fig. 5.3: Putative transformed tobacco shoots regenerated *in vitro* on selection medium. CProG1 (a - d) and CProG2 (e - h). Shoot bud induction after 2 weeks (a and e), proliferation of shoot bud (b, c, f and g) and shoots after 8 weeks (d and h).

(Wc) (P1) (P2)

Fig. 5.4: Transformed tobacco shoots on root induction medium – untransformed control (Wc), CProG1 (P1) and CProG1 (P2) transformation events.

(a) P1 P2 (b) P1 P2

Fig. 5.5: (a) Hardening of transformed plant CProG1 (P1) and CProG2 (P2), (b) transformed plant CProG1 (P1) and CProG2 (P2) in pots.



(a) (b) (c)

Fig. 5.6: Transgenic tobacco plants CProG1 (a), CProG2 (b) and untransformed control tobacco plant (c) illuminated at 395nm using NightSea GFP flashlight.

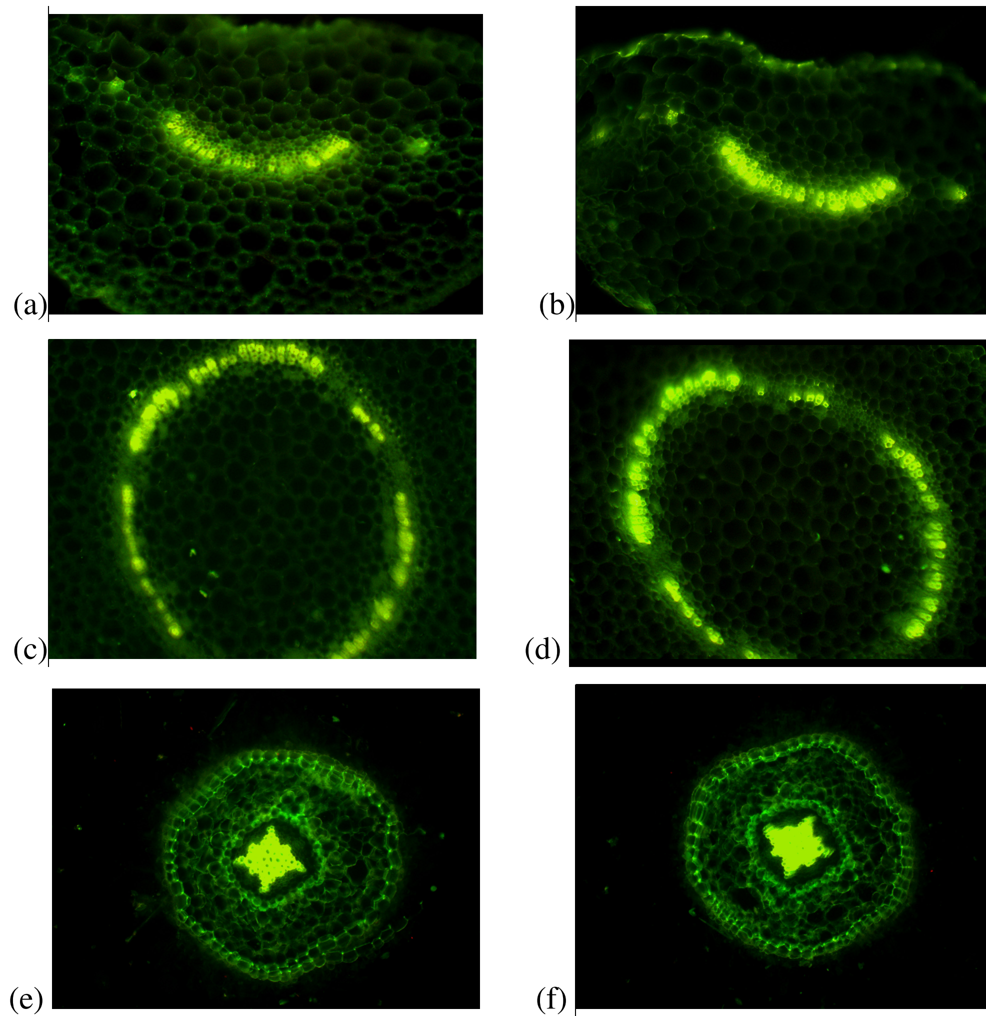


Fig. 5.8: Transverse sections of transgenic tobacco plant visualized in 20X magnification: through leaf midrib of CProG1 (a) and CProG2 (b); stem of CProG1 (c) and CProG2 (d); and root of CProG1 (e) and CProG2 (f).