Chapter Two

Isolation, cloning and characterisation of CCR gene from *Leucaena leucocephala*
2.1 Introduction:
Lignin biosynthesis has been already discussed in detail in Chapter One of this thesis. The present chapter deals with why *Leucaena leucocephala* and CCR from *L. leucocephala* were used as candidates for this study. In its later part isolation, cloning and characterization of CCR gene from *L. leucocephala* is described in detail.

Paper industry in India mainly uses bamboo, *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 given industrial unit. In developing countries, like India, the proportionate use of bamboo and hardwood species is in the ratio of 15:85 although all these plant species are of equal importance to the paper industry.

2.1.1 *Leucaena* as a source of pulp:
*Leucaena* sp. is extensively used in India and about 25% of raw material for pulp and paper industry is contributed by this hard wood tree. To meet the increasing demand of high quality wood for paper industry, it is essential to provide designer plant species. It will thus be vital to raise plantations with elite materials and/or genetically modified plants that meet the demands of the pulp and the paper industry in economical and sustainable manner.

2.1.2 About *Leucaena leucocephala*:
*Leucaena* is a native of Central America and 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.*, 1964; Hammond *et al.*, 1989 a, b).

*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.

*Leucaena leucocephala*

**Classification:**

- Division: Magnoliophyta
- Class: Magnoliopsida
- Sub-class: Rosidae
- Order: Fabales
- Family: Fabaceae
- Sub-family: Mimosoideae
- Tribe: Mimoseae
- Genus: Leucaena
- Species: leucocephala

**Common name:** Lead tree, white popinac, subabul.

**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.
2.1.3 Lignin Biosynthesis and CCR:

Detailed Lignin biosynthesis has been discussed in Chapter One. In brief, the biosynthesis of lignin begins with the common phenylpropanoid pathway starting with the deamination of phenylalanine and leading to the cinnamoyl-CoA esters. These esters are the common precursors of a wide array of end-products such as flavonoids, coumarins and many small phenolic molecules (Hahlbrock and Scheel, 1989; Boerjan et al., 2003) that play key roles during plant development and defense. Cinnamoyl-CoA esters are then channeled into the lignin branch pathway to produce monolignols via two reductive steps using CCR and Cinnamyl Alcohol Dehydrogenase (CAD). By the action of CAD, monolignols are formed which are toxic to plant cells. They are acted upon by monolignol specific UDP-glucosyltransferase, which adds a glucose moiety to monolignols (Glycosides), and are transported to outside the cell membrane. In the cell wall the glycosides are again hydrolyzed to form Hydroxyphenyl (H), Guaiacyl (G) and Syringyl (S) units respectively. These monomers are linked together via end-wise and radical coupling reactions (Sarkanen and Ludwig, 1971; Freudenberg and Neish, 1968) to produce H, G and S lignin respectively. Here, dehydrogenative polymerization of monolignols takes place to form lignin (Boerjan et al., 2003; Dixon et al., 2001).

The first reductive step in lignin biosynthetic pathway is acted upon by CCR and it controls the over-all carbon flux towards lignin (Piquemal et al., 1998). The reduction of cinnamoyl CoA esters to cinnamaldehydes is the first metabolic step committed to monolignol formation (Lacombe et al., 1997). CCR is apparently encoded by a single gene per haploid genome in Eucalyptus (Bouget et al., 1997), poplar (Leple et al., 1998), ryegrass (Larsen, 2004; McInnes et al., 2002), Triticum (Ma, 2007) and tobacco (Piquemal et al., 1998) and by two genes in maize (Pichon M, 1998) and Arabidopsis (Lauvergeat et al., 2001). The CCR genes in various species appear as a multiple member family. In the Populus genome, there exist 8 CCR-homolog or CCR-like gene sequences (Li, 2005). Several other CCR gene sequences have been deposited in the GenBank database, but their functions have still not been demonstrated. It is proposed that all CCR enzymes have a similar catalyzing mechanism for converting the CoA ester to aldehyde in monolignol biosynthesis.
2.1.3.1 *CCR* as a candidate gene:

Out of three (H, G and S) monomer units of lignin, S and G lignins are found predominantly in angiosperms. A higher S/G ratio is desirable for paper industry as S lignin is less compact (higher degree of methoxylation) than G lignin and removal of total lignin is easier. Transgenic tobacco (Chabannes *et al.*, 2001; Abbott *et al.*, 2002) and *Arabidopsis* (Lauvergeat *et al.*, 2001; Goujon *et al.*, 2003) down regulated for *CCR* were characterized by an approximately 50% decrease in Klason lignin. The lignin S/G ratio increased (mainly because of a decrease in the G unit amount) in transgenic tobacco and was variable, depending on the growth conditions, in transgenic *Arabidopsis* (Abbott *et al.*, 2002; Goujon *et al.*, 2003 and Chabannes *et al.*, 2001). A change in 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). Transgenic plants with the lowest *CCR* activity and 50% reduced lignin showed abnormal phenotypes, such as important alterations in the fiber cell walls and loosening in the arrangement of the cellulose microfibrils that resulted in reduced cell wall cohesion (Elkind *et al.*, 1990; Pincona or b *et al.*, 2001).

Based on the above information the work was started with the objective of cloning and characterization of the *Leucaena CCR* gene and its antisense down-regulation. This study is the first instance towards isolation and characterization of lignin biosynthetic pathway gene(s) for the development of transgenic *Leucaena* plants. The study is intended to lay down criteria for the development of the transgenic plants, which would ultimately strive for: (a) reduced lignin content and (b) altered ratio of S/G lignin. 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*). The above genes of the lignin biosynthesis pathway in the target plant species could then be down regulated by 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. This thesis elaborates isolation and characterization of CCR gene, its differential expression, and down regulation.

No study was carried out on lignin biosynthetic pathway gene(s) so far in *Leucaena* and the study of Cinnamoyl Co A Reductase will help in better understanding of the pathway in *Leucaena* sp. and its manipulation. This research work deals with isolation, cloning and characterization of CCR gene from *Leucaena*. In the present study the genomic and the cDNA gene clones of CCR were isolated from *Leucaena leucocephala*. In this report, two ful-length CCR cDNAs from *Leucaena* have been isolated and characterized.

### 2.2 Materials and Methods:

#### 2.2.1 Materials:

**Chemicals:**

Ampicillin, Kanamycin, Tetracycline, Tris-buffer, 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). Megaprime labeling kit and Hybond-N⁺ membrane were obtained from Amersham (UK). α-32P-dATP and α-32P –dCTP were obtained from Bhabha Atomic Research Centre (BARC), India. X-ray films were obtained from Kodak (USA). All other chemicals and solvents of analytical grade were purchased from HIMEDIA, Qualigens Fine Chemicals and E-Merck Laboratories, India.

**Plasticware:**

Sterile disposable filter sterilization units and petri dishes were procured from Laxbro (India). Micro-centrifuge tubes and micropipette tips were procured from Axygen and Tarsons (India).

**Glassware:**

Glassware used in the experiments such as test tubes, glass bottles, petridishes, Erlenmeyer flasks and pipettes were procured from Borosil (India).
**Bacterial strains and plasmids used in the study:**

*Escherichia coli* XL-1 blue (Stratagene, USA)

*E. coli* MRF-XL-1 (Stratagene, USA)

*E. coli* XOLR (Stratagene, USA)

*E. coli* SOLR (Stratagene, USA)

λ ZAP II vector (Stratagene, USA)

pGEM-T Easy Vector Cloning vector (Promega, USA)

**Stock solutions:**

1) IPTG stock solution 200 mg/mL in sterile distilled water.

2) X-gal stock solution 20 mg/mL in dimethylformamide.

3) IPTG stock solution 1 M in sterile distilled water.

4) Kanamycin 30 mg/mL.

5) Ampicillin 100 mg/mL.

6) Tetracycline 12.5 mg/mL.

**Plant material: RACE, Isolation of full-length CCR gene:**

Xylem tissue of mature healthy plants, growing in NCL campus was harvested as and when required. Outer bark was scraped to expose the xylem tissue. The tissue was scraped and ground using liquid nitrogen and RNA was isolated for normal cDNA preparations (RACE PCR).

**2.2.2 Methods:**

**2.2.2.1 Bacterial culture conditions:**

*E. coli* strain 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.
2.2.2.2 *E. coli* transformation and selection:

LB medium (50 mL) was inoculated with 1% of the overnight grown *E. coli* culture and allowed to grow till A$_{600}$ was 0.5. The cells were harvested by centrifugation at 5,000 g for 10 min at 4 °C, suspended in 100 mM ice-cold TB Buffer and again centrifuged at 5,000 g for 10 min at 4 °C. Cells were suspended in 4.65 mL of ice-cold TB Buffer and 350 μL DMSO. Aliquots of 100 μL were frozen in liquid nitrogen and stored at –80 °C till further use. The competent *E. coli* cells, thus formed, were transformed according to Sambrook *et al.* (1989). Briefly, DNA (~50 ng in 10 μL or less) was added to 100 μL of 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 900 μL of LB broth was added and further incubated at 37 °C for 1 h. About 100 μL of the transformed competent cells was spread onto LB agar plates containing appropriate antibiotic, IPTG (4 μL per 90 mm plate) and X-gal (40 μL per 90 mm plate) (Sambrook *et al.*, 1989).

2.2.2.3 Genomic DNA extraction:

**Solutions:**

**Extraction buffer:** 100 mM Tris-HCl, pH 8.0; 20 mM Sodium EDTA; 1.4 M NaCl and 2.0% (w/v) CTAB (cetyltrimethylammoniumbromide). Store at 37 °C. Add 0.2% (w/v) β-mercaptoethanol before use.

Other solutions used were Chloroform:Isoamyl alcohol: 24:1 (v/v); autoclaved 5 M NaCl, RNase A (10 mg/mL); 95% ethanol; 75% ethanol; autoclaved TE buffer: 10 mM Tris-HCl and 1 mM EDTA pH 8.0.

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 1 g of ground tissue was extracted with 10 mL extraction buffer. The slurry was poured into a clean autoclaved 50 mL centrifuge tube and 100 mg insoluble polyvinylpolypyrrolidone (PVPP) added. The tube was gently inverted several times to thoroughly mix the slurry, incubated at 65 °C for 30 min and then allowed to cool down to room temperature. 12 mL of chloroform: isoamyl-alcohol mixture was added and the contents mixed by inverting the
tube gently till an emulsion formed. The mixture was then centrifuged at 6,000 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 5 M NaCl was added and mixed gently. Next, two volumes of cold (-20 °C) 95% ethanol was added and the sample kept at 4 °C until DNA strands appeared. The tube was centrifuged at 3,000 g for 3 min and then at 5,000 g for next 3 min. The supernatant was poured off and the DNA pellet washed with cold (4 °C) 75% ethanol and air-dried. DNA was dissolved in 300 μL of TE buffer.

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

### 2.2.2.4 Isolation of plasmid DNA from *E. coli* cells:

**Solutions:**

- **Soln. I (TEG Buffer):** 25 mM Tris-HCl, pH 8.0; 10 mM EDTA, pH 8.0; 50 mM Glucose
- **Soln. II:** 0.2 N NaOH, 1% SDS (freshly prepared)
- **Soln. III:** 3.0 M Potassium acetate (pH 4.8)

Chloroform, absolute ethanol, 3.0 M Sodium acetate, 70% ethanol, de-ionized 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 5-30 μg DNA per 1.5 mL culture depending on the host strain and the plasmid vector. An important feature of this protocol was the use of poly-ethyleneglycol (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 3 mL culture was centrifuged for 1 min at 4,000 g to pellet the bacterial cells. The pellet was resuspended in 200 μL of TEG buffer by vigorous pipetting, 300 μL 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 300 µL 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 g at room temperature. The supernatant was transferred to a clean tube; RNase A to a final concentration of 20 µg/mL (Sambrook et al., 1989) was added and incubated at 65 °C for 15 min. To the above solution 400 µL of chloroform was added, mixed for 30 s and centrifuged for 5 min at 12,000 g at 4 °C. The upper aqueous layer was transferred to a clean tube, 1/10th volume sodium acetate and 0.6 volume iso-propanol was added with mixing and kept at RT for 10 min. The sample was centrifuged at 12,000 g for 15 min at room temperature. The pellet was washed twice with 70% ethanol and dried under vacuum. The dried pellet was dissolved in 100 µL of de-ionized water and 100 µL of PEG-NaCl solution (20% PEG 8000 in 2.5 M NaCl) was added. The mixture was incubated on ice for 20 min and the plasmid DNA pelleted out by centrifugation at 12,000 g for 15 min at 4 °C. The supernatant was aspirated carefully, the pellet was washed with 75% ethanol and air-dried. The dried pellet was resuspended in 40 µL de-ionized water and stored at -70 °C.

2.2.2.5 Restriction digestion of DNA:
Plasmid and genomic DNA restriction digestion was set up as per manufacturer’s (Promega, USA; NEB, UK; Amersham, USA) recommendations.

2.2.2.6 Extraction and purification of DNA from agarose gel:
The restriction digested DNA or PCR amplified products were run on an agarose gel in 1X TAE buffer (see Appendix). The gel was stained with ethidium bromide (0.5 µg/mL) and viewed using a hand held long wavelength UV illuminator. The fragment of interest were excised from the gel and weighed. A 50-200 mg gel slice was transferred to a 1.5 mL micro-centrifuge tube. Purification of DNA from gel was done according to the manufacturer’s protocol (Auprep, Lifetechnologies; Axygen India Pvt. Ltd; SIGMA-ALDRICH, India.).

2.2.2.7 Total RNA Isolation:
Solutions:
TRI Reagent (SIGMA)
Chloroform and Isoamyl Alcohol (24:1); Isopropanol; 75% ethanol in DEPC treated de-ionized water; DEPC treated de-ionized water; 8 M LiCl.

RNase free environment was created and maintained as described below. The glassware and plasticware were DEPC (0.1% in water) treated overnight and autoclaved. The pestle and mortar were also DEPC treated and then baked at 300 °C for 6 h. All materials were dried in a vacuum oven.

Total RNA from different plant tissues was isolated using TRI reagent. The plant tissue was collected, washed with DEPC treated water, frozen in liquid nitrogen and crushed to a fine powder. TRI reagent (1 mL) was added to 100 mg of the fine powder and mixed thoroughly using a vortex. Chloroform: isoamyl alcohol (300 μL) was added and mixed thoroughly again using vortex. The tubes were centrifuged at 4 °C at 13,000 g for 10 min. The supernatant was transferred to 1.5 mL microcentrifuge tube and the chloroform: isoamyl alcohol step was repeated. The aqueous phase was transferred to 1.5 mL tubes and 0.6 volume isopropanol was added. It was mixed thoroughly and kept for RNA precipitation for 1 h at 4 °C. Total RNA was pelleted out by centrifugation at 13,000 g for 15 min at 4 °C. The RNA pellet was washed with 75% ethanol twice and dried in a SpeedVac (Savant, USA) centrifugal concentrator. RNA pellet was dissolved in 100 μL of DEPC treated water. 50 μL of 8 M LiCl was added, mixed and kept at -20 °C overnight. Next day total RNA was pelleted out by centrifugation at 13,000 g for 15 min at 4 °C, washed with 75% ethanol, dried in SpeedVac and dissolved in 50 μL DEPC treated de-ionized water. Purity of RNA was confirmed by measuring OD at 260/280 nm and also by visualization on 1.5% TAE agarose gel.

2.2.2.8 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 first 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.

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 manufacturer’s guidelines.

In brief, reverse transcription reactions of up to 1 μg of Total RNA performed in 20 μL reactions comprised of following components of the ImProm-II Reverse Transcription System.

The first strand reaction was set up as follows:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental RNA (1μg)</td>
<td>1-6 μL</td>
</tr>
<tr>
<td>Primer [Oligo (dT)₁₅] or Random (10 pmol)</td>
<td>1 μL</td>
</tr>
<tr>
<td>Nuclease-Free Water</td>
<td>3 μL</td>
</tr>
<tr>
<td>Final volume</td>
<td>10 μL</td>
</tr>
</tbody>
</table>

The tubes were incubated at 70 °C for 5 min and then chilled on ice for 2 min. The tubes were briefly spun in a micro-centrifuge to collect the condensate and maintain the original volume. The tubes were kept on ice until addition of the reverse transcription reaction mix was over. The reverse transcription reaction mix was prepared by combining the following components of the ImProm-II Reverse Transcription System in a sterile 1.5 mL micro-centrifuge on ice.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nuclease-free water</td>
<td>1.5 μL</td>
</tr>
<tr>
<td>ImProm-II. 5X Reaction Buffer</td>
<td>4.0 μL</td>
</tr>
<tr>
<td>MgCl₂ (15 mM)</td>
<td>2.0 μL</td>
</tr>
<tr>
<td>dNTP Mix (10 mM)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>RNasin® Ribonuclease Inhibitor (40U/μL)</td>
<td>0.5 μL</td>
</tr>
<tr>
<td>ImProm-II Reverse Transcriptase</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>Final volume</td>
<td>10.0 μL</td>
</tr>
</tbody>
</table>
An aliquot of 1.0 μg total RNA and oligo (dT)\textsubscript{15} or random hexamer primer (10 pmol) mix was added to the above reaction for a final reaction volume of 20 μL per tube. The tube was incubated at 42 °C for 1 h for cDNA first strand synthesis. Reverse transcriptase was thermally inactivated by incubation at 90 °C for 5 min prior to proceeding with PCR amplification.

2.2.2.9 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. This technique has been successfully used for various purposes like fishing out of gene(s) from genomic DNA or from cDNA population, introducing restriction sites of interest in the amplified product for directional cloning, creating sequence mismatch/ deletion/ addition resulting in mutant version of a gene or nucleotide sequence, differentiating between two alleles etc. In the present study applications of PCR were exploited for a few of the above specified applications (Mullis et al., 1992). The PCR reaction mixture and cycling conditions used were as follows:

**Reaction mixture:**

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sterile de-ionized water</td>
<td>6.7 μL</td>
</tr>
<tr>
<td>Template (100 ng/μL)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>Forward primer (7 pmol)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>Reverse primer (7 pmol)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>dNTPs (0.2 mM)</td>
<td>4.0 μL</td>
</tr>
<tr>
<td>10 X Taq Buffer A (Mg\textsuperscript{2+} 15 mM)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>Taq Polymerase (1 U/μL)</td>
<td>0.3 μL</td>
</tr>
<tr>
<td>Total volume</td>
<td>15.0 μL</td>
</tr>
</tbody>
</table>
PCR cycle conditions:

<table>
<thead>
<tr>
<th>No. of cycles</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>95 ºC</td>
<td>5 min</td>
</tr>
<tr>
<td>35</td>
<td>95 ºC 45-65 ºC (annealing temperature was dependent on primer Tm)</td>
<td>1 min 30-45 s</td>
</tr>
<tr>
<td>1</td>
<td>72 ºC</td>
<td>7 min</td>
</tr>
<tr>
<td>1</td>
<td>4 ºC</td>
<td>Hold</td>
</tr>
</tbody>
</table>

2.2.2.10 Southern Hybridization, Slot Blot Hybridisation and Random Primer Labeling:

2.2.2.10.1 Southern Hybridization:

Solutions:

20X SSC: 3 M NaCl, 0.3 M Sodium citrate (pH 7.0)
Depurination solution: 0.25 N HCl
Denaturation solution: 1.5 M NaCl, 5 M NaOH
Neutralization solution: 0.5 M Tris-HCl (pH 7.4); 3 M 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 0.5 µg/mL ethidium bromide. The gel was rinsed with de-ionized water (DW) and placed in depurination solution for 15 min. It was then rinsed with de-ionized water and immersed in denaturation solution for 30 min with gentle shaking. The gel was again rinsed with de-ionized 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 #3 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, USA) of the exact gel size was wetted with de-ionized 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 #3 paper wetted with 20X SSC was placed on the membrane followed by Whatman #3 paper pre-wetted in 2X SSC. On this paper another dry Whatman #3 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 2 h at 80 °C to immobilize DNA onto the nylon membrane. Hybridization and autoradiography were carried out as is described in the section 2.2.2.10.4.

2.2.2.10.2 Slot Blot Hybridization:

Solutions:
20X SSC; 3 M NaCl; 0.3 M Sodium citrate (pH 7.0)
3 M NaOH.

For slot blot hybridization DNA samples were diluted according to experimental requirements. The DNA samples were denatured by adding 1/10th volume 3 M NaOH and incubation at 65 °C for 10 min. An equal volume of 6X SSC was added to the denatured samples. Two layers of Whatman #3 filter paper wetted with sterile de-ionized water and 6X SSC were placed in the Slot Blot apparatus (Hoefer Scientific, USA) followed by Hybond-N+ membrane (Amersham, USA) treated in the manner as above. The Slot Blot unit was assembled and wells washed with 500 μL 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 appeared 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 2.2.2.10.4.

2.2.2.10.3 Random Primer Labeling:

Random primer labeling of the DNA probes was done using the Megaprime DNA labeling kit (Amersham, USA) Reaction (50 μL) was set up as follows:
<table>
<thead>
<tr>
<th>25 ng DNA (used as probe)</th>
<th>5.0 µL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primer solution (random hexanucleotides)</td>
<td>5.0 µL</td>
</tr>
</tbody>
</table>

Above mixture was heated in a boiling water bath for 10 min and cooled to room temperature facilitating primer annealing to the DNA. 40 µL of the following reaction mixture was added to above mixture.

<table>
<thead>
<tr>
<th>10X reaction buffer</th>
<th>5.0 µL</th>
</tr>
</thead>
<tbody>
<tr>
<td>dCTP/dATP</td>
<td>4.0 µL</td>
</tr>
<tr>
<td>dGTP</td>
<td>4.0 µL</td>
</tr>
<tr>
<td>dTTP</td>
<td>4.0 µL</td>
</tr>
<tr>
<td>α-32P-dATP/dCTP Sp. activity 3000 Ci/ mmol</td>
<td>2 to 5.0 µL</td>
</tr>
<tr>
<td>Sterile de-ionized water</td>
<td>17 to 20.0 µL</td>
</tr>
<tr>
<td>Exonuclease free Klenow fragment (2 U/µL)</td>
<td>1.0 µL</td>
</tr>
<tr>
<td>Total volume</td>
<td>50.0 µL</td>
</tr>
</tbody>
</table>

The reaction was carried out at 37 °C for 60 min and stopped by incubation in a boiling water bath for 10 min and snap chilled on ice.

### 2.2.2.10.4 Pre hybridization and Hybridization:

**Solutions:**

- 20X SSC: 3 M NaCl; 0.3 M Sodium citrate (pH 7.0)
- Hybridization buffer: 1% BSA; 1.0 mM EDTA, pH 8.0; 0.5 M 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 2.2.2.10.1 and 2.2.2.10.2 above were pre-hybridized at 50 °C in 30 mL 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. This process was repeated two or three times, till the desired (20– 50 cps) count from membrane is achieved. The blot was checked for 20 to 50 counts per seconds by Geiger Muller counter, wrapped in Saran wrap and exposed to X-ray film at −70 °C in a cassette with intensifying screen. After four days films were developed using developer and fixer solutions described in Appendix.

2.2.2.11 Construction and screening of genomic DNA library:
Genomic DNA library was prepared using λ- ZAP II vector according to manufacturer’s guidelines (Stratagene). In Brief Genomic DNA was digested with Sau 3A restriction enzyme to generate approximately 8 to 10 kb fragments. Preliminary digestions were done to standardize the time and quantity of enzyme required to generate the required fragment of same size. Digested Genomic DNA was run on 0.8 % agarose gel and smear ranging from 8 to 10 kb was purified. Gel purified fragment were blunt ended and Eco RI adapters were ligated. These fragments were now ligated into λ-ZAP II vector (pre digested with Eco RI, and CIAP treated). The recombinant λ-ZAP II vectors were now packaged with packaging extract according to the manufacturer’s instructions. The packaged recombinant λ-ZAP II vector was used to transfect E.coli XL1-MRF strain. Twenty to thirty plates of 135 mm were plated according to transfection protocol discussed later.

Plaques appeared after 8 to 10 h. 5 ml of SM buffer was poured in each plate and kept on shaker for two hours at 4 °C. After two hours of incubation all SM buffer (refer Appendix) was pooled in one tube and kept at 4 °C for further use (Plaques stable up to 6 months at 4 °C).
2.2.2.11.1 Preparation of cells for Transfection:

Plates were poured with NZY bottom agar (refer Appendix) and tetracycline (12.5 μg/mL of media). A single colony of *E. coli* XL–Blue MRF was inoculated in 5 ml of LB media and kept on shaker overnight. 500 μl of this culture was taken and inoculated in 50 ml LB media for further sub culturing for 4 h on shaker. 25 ml of culture was taken in autoclaved tubes; these steps were performed in laminar airflow unit to avoid contamination of cells. Cells were centrifuged at 5,000 g for 10 min, supernatant was discarded and pellet was washed gently, twice, with chilled 10 mM MgSO4. Pellet was dissolved in approximately 20 mL of chilled MgSO4. These cells were kept at 4 °C for further use.

2.2.2.11.2 Transfection:

500 μl of cultured cells were taken in micro-centrifuge tube and 1 μl of λ Phage representing genomic library of *Leucaena leucocephala*, was added, mixed gently and kept in incubator at 37 °C for 15 min. In an autoclaved test tube 5 ml of melted top NZY media (containing 0.7 % agarose) and transfected bacterial cells were added and spread on plates having bottom NZY-agar media. These plates were kept in incubator at 37 °C for 8 h.

2.2.2.11.3 Plaques lifting:

The Plates with a density of 4000 – 5000 plaques per plate were used for screening. Nylon membrane (Hybond-Amersham) was placed on the plate gently with the help of forceps. It was marked asymmetrically for alignment. The membrane was peeled off the plate gently and kept in denaturing solution (refer Appendix) for 2 min followed by in neutralization solution (refer Appendix) for 5 min. It was then kept in rinsing solution (2X SSC and 0.2 M Tris HCl) for 30 s. The membrane was air-dried for about 2 h and then baked in oven at 80 °C for 2h before pre hybridization.

2.2.2.11.4 Preparation of probe, Pre-hybridization and Hybridization:

PCR was used (as described in section 2.2.2.9) to prepare radiolabelled probe. dATP having α-P³² (3000 Ci/mmol) was used as one of the dNTPs replacing normal dATP. Genomic DNA partial clone was used as a probe for screening purpose. After completion
of PCR, tubes were kept at 99 °C for 10 min. Tube was snap chilled and used for hybridization.

Pre-hybridization and hybridization was done as discussed in section 2.2.2.10.4

Necessary care was taken while working with radioactive dATP and dCTP. Lab coat, spectacles, and gloves were used while working.

2.2.2.11.5 First, Second and third round of screening:
Plaques with positive signals were cut and resuspended in 200 µl SM buffer with 10 µl chloroform. These plaques represent the λ phage having gene of interest. Plaques were kept as such for 4 h and then centrifuged at 8,000 g for 2 min. Transfection was done with phage collected after first round of screening. Subsequent screening steps i.e. secondary and tertiary screenings were followed as in case of primary screening. Single plaque was amplified and kept for further experiments.

2.2.2.11.6 Single clone excision:
For single clone excisions, following components were mixed in a Falcon 50 ml polypropylene tube:
200 µl of XL1-Blue MRF cells (A600 = 1.0)
250 µl of phage stock (containing >1 × 10⁵ phage particles)
1 µl of the ExAssist helper phage (supplied with Kit)

The mixture was incubated at 37 °C for 15 min to allow the phage to attach to the cells. 3 mL of LB broth was added with supplements in the tube and was incubated for 2.5–3 h at 37 °C with shaking. Tube was heated at 65–70 °C for 20 min to lyse the λ phage particles and the cells. It was centrifuged at 1000 g for 15 min for pelleting the cell debris. Supernatant was saved in another sterile tube and transformation was done in freshly grown E. coli SOLR cells according to manufacturer’s instructions. Plasmid was isolated from colonies, which appeared on LB agar (Kanamycin 30 µg/mL) plates and analysed.
2.2.2.12 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 and amplified by PCR using RACE. 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. 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 (Table 2.2.1). These nucleotides position the primer at the start of the A+ tail and thus eliminate the 3’ heterogeneity inherent with conventional oligo (dT) priming. 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 and a minor amount of a proofreading polymerase.

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

2.2.2.12.1 RACE cDNA preparation:

For preparation of 5'-RACE cDNA

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>RNA sample (1 μg/μL)</td>
<td>5 μL</td>
</tr>
<tr>
<td>5'-CDS primer A (12 μM)</td>
<td>1 μL</td>
</tr>
<tr>
<td>SMART II A oligo (12 μM)</td>
<td>1 μL</td>
</tr>
<tr>
<td>SMQ</td>
<td>3 μL</td>
</tr>
</tbody>
</table>

Sterile H₂O was added to a final volume of 10 μL 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:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>5X First-Strand Buffer</td>
<td>4 μL</td>
</tr>
<tr>
<td>DTT (20 mM)</td>
<td>1 μL</td>
</tr>
<tr>
<td>dNTP Mix (10 mM)</td>
<td>1 μL</td>
</tr>
<tr>
<td>MgCl₂ (25 mM)</td>
<td>3 μL</td>
</tr>
<tr>
<td>PowerScript Reverse Transcriptase (5U/ μL)</td>
<td>1 μL</td>
</tr>
<tr>
<td>Total volume</td>
<td>10 μL</td>
</tr>
</tbody>
</table>

Srivastava S, University of Pune
The contents of the tube were mixed by gentle pipetting and were briefly centrifuged. The tubes were incubated at 42 °C for 1.5 h in a hot-lid thermal cycler. The first strand reaction mixture was diluted to 100 μL with Tricine-EDTA buffer (provided with the kit) and heated at 70 °C for 7 min. The 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:

2.2.2.12.2 Primary and Nested PCR:

**Master mix for RACE PCR reaction**

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR-Grade Water</td>
<td>34.5 μL</td>
</tr>
<tr>
<td>10X Advantage 2 PCR Buffer</td>
<td>5.0 μL</td>
</tr>
<tr>
<td>dNTP Mix (10 mM)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>50X Advantage 2 Polymerase Mix</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>Total volume</td>
<td>41.5 μL</td>
</tr>
</tbody>
</table>

The above master mix was used for 5’ and 3’ RACE PCR.

The reaction was set up as follows, for 5’RACE:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ RACE cDNA</td>
<td>2.5 μL</td>
</tr>
<tr>
<td>UPM (10X)</td>
<td>5.0 μL</td>
</tr>
<tr>
<td>GSP1 (10 μM) (Sam r1)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>Master Mix</td>
<td>41.5 μL</td>
</tr>
<tr>
<td>Total volume</td>
<td>50.0 μL</td>
</tr>
</tbody>
</table>

The reaction was set up as follows, for 3’RACE:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>3’ RACE cDNA</td>
<td>2.5 μL</td>
</tr>
<tr>
<td>UPM (10X)</td>
<td>5.0 μL</td>
</tr>
<tr>
<td>GSP2 (10 μM) (SCHF2)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>Master Mix</td>
<td>41.5 μL</td>
</tr>
<tr>
<td>Total volume</td>
<td>50.0 μL</td>
</tr>
</tbody>
</table>

Srivastava S, University of Pune
Following PCR cycling conditions were used for amplifying 5’ and 3’ RACE products:

<table>
<thead>
<tr>
<th>No. of Cycles</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 cycles</td>
<td>94 °C</td>
<td>30s</td>
</tr>
<tr>
<td></td>
<td>72 °C</td>
<td>3 min</td>
</tr>
<tr>
<td>30 cycles</td>
<td>94 °C</td>
<td>30 s</td>
</tr>
<tr>
<td></td>
<td>60 °C</td>
<td>30 s</td>
</tr>
<tr>
<td></td>
<td>72 °C</td>
<td>3 min</td>
</tr>
</tbody>
</table>

After the PCR, 7 μL of the reaction mix was loaded on 1.0% agarose gel in 1X TAE buffer and checked for amplification. Nested PCR was run according to cycling conditions mentioned in section 2.2.2.9

**Nested PCR reaction mix:**

Nested 5’ RACE and 3’ RACE were performed as separate reactions.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (μL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR-Grade Water</td>
<td>41 μL</td>
</tr>
<tr>
<td>10X Advantage 2 PCR Buffer</td>
<td>5.0 μL</td>
</tr>
<tr>
<td>dNTP Mix (10 mM)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>50X Advantage 2 Polymerase Mix</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>Forward primer (SCHF2, 3’ RACE)</td>
<td>1 μL</td>
</tr>
<tr>
<td>Reverse primer (Sam r1, 5’ RACE)</td>
<td>1 μL</td>
</tr>
<tr>
<td>NUP (Table 1)</td>
<td>1.0 μL</td>
</tr>
<tr>
<td>5’ and 3’ Primary reaction</td>
<td>2.0 μL</td>
</tr>
<tr>
<td>Total volume</td>
<td>50 μL</td>
</tr>
</tbody>
</table>

Following is the list of primers (refer Table 2.2.1), which were used for the RACE:

Provided with kit
Table 2.2.1: List of Primers used for the RACE.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMART II™ A Oligonucleotide</td>
<td>AAGCAGTGGGTATCAACGCAGAGTACGCGGG</td>
</tr>
<tr>
<td>3’-RACE CDS Primer A</td>
<td>AAGCAGTGGGTATCAACGCAGAGTAC (T)_{30V} N</td>
</tr>
<tr>
<td>5’-RACE CDS Primer A (5’-CDS)</td>
<td>(T)_{25V} N</td>
</tr>
<tr>
<td>10X universal Primer A Mix (UPM)</td>
<td>Long: CTAATACGACTCCTATAGGGCAAGCA GTGGTATCAACGCAGAGT Short: CTAATACGACTCCTATAGGGC</td>
</tr>
<tr>
<td>Nested universal Primer A (NUP)</td>
<td>AAGCAGTGGGTATCAACGCAGAGT</td>
</tr>
</tbody>
</table>

2.2.2.13 Sequencing:
DNA sequencing was performed with the ABI Prism Big- Dye Terminator Cycle Sequencing Kit on the ABI Prism 3730 DNA analyzer (Applied Biosystems) at GenomeBio Biotech Pvt. Ltd., Pune, Maharashtra, India.

2.2.2.14 Bioinformatics analysis:
Nucleotide and amino acid sequence analysis was done using software pDRAW 32 and ClustalX 1.8. Primer designing was done by aligning CCR sequences in NCBI database using ClustalX. LI-CCR gene sequences were characterized with the Genscan software and homology was verified by database searching at the National Center for Biotechnology Information server using BLAST algorithm (http://www.ncbi.nlm.nih.gov). The deduction of the amino acid sequences, calculation of the theoretical molecular mass and pI, as well as prediction of sub-cellular localization was performed with ExPASy Proteomic tools provided at http://www.expasy.ch/tools/. Global alignment of two nucleotide or amino acid sequences and percentages of identity were calculated using the EMBOSS Pairwise Alignment algorithms (http://www.ebi.ac.uk/emboss/). Multiple alignments of the amino
acid sequences were carried out with the Clustal W1.8 program (http://www.ebi.ac.uk/clustalw/). The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the taxa analyzed (Felsenstein, 1985). Branches corresponding to partitions reproduced in less than 50% bootstrap replicates are collapsed. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling 1965) and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated from the dataset (Complete deletion option). Phylogenetic analyses were conducted in MEGA4 (Tamura, 2007). Online bioinformatics analysis facility available at www.justbio.com, Mat-Inspector (for promoter analysis.) and www.expasy.org was used for detailed analysis.

2.2.2.15 Estimation of CCR gene copy number:
Genomic DNA was isolated using Lodhi et al. method. Gene copy number was estimated by southern hybridization and Real time PCR analysis. DNA was digested with four restriction enzymes, Hinc II, Nde I, Bgl I and Sac I. One of the enzymes, Sac I does not cut inside the gene while the other three have one site in the gene. Southern hybridization was done using an ≈800 bp fragment (from 5’ of CCR coding region and UTR) as a probe at 62 °C hybridization temperature (Sambrook et al., 1989). Gene copy number was also estimated using Taqman based PCR chemistry. Standard dilutions of CCR gene cloned in pGEM-T vector was used to make a standard graph The reaction was run in triplicates. Primers and probe were designed from Eurogenetec, Belgium and care was taken to design it from one single exon (Table 2.2.2). Genomic DNA was quantified and dilutions were made according to the dilutions made for standard graph (1.81 pg of DNA per haploid genome of L. leucocephala). Number of copies initially present in amount of gDNA used in Taqman PCR was compared with standard graph and copy number was estimated (Freeman et al., 1999).
Table 2.2.2: Primers used to determine Gene copy Number using Taqman probe.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCR2 Taqprobe Modifications 3’- Black Hole Quencher-3™ 597.63 5’-Cy5Indocarbocyanin ® 533.63</td>
<td>ACT CCT TCT GCC TCC GGT CGT TAT</td>
</tr>
<tr>
<td>CCR2F2</td>
<td>TGT TGC ATT AGC CCA TGT TCT TG</td>
</tr>
<tr>
<td>CCR2R2</td>
<td>GAG GAT CTC GAC CAG TTC TCC</td>
</tr>
</tbody>
</table>

2.3 Results:

2.3.1 The presence of the CCR gene in *L. leucocephala*:

The presence of Cinnamoyl Co A Reductase (*CCR*) gene in *Leucaena leucocephala* was established by slot blot analysis. Genomic DNA of good integration was isolated from the leaves of *L. leucocephala*.

**Slot Blot Hybridization:**

The genome size of *L. leucocephala* has ~1.81 pg of DNA per haploid genome. 1.81 pg corresponds to a single representation of *L. leucocephala* genomic DNA, 0.18 µg and 0.36 µg representing the genome 1x10⁵ and 2x10⁵ times respectively was spotted on Hybond N⁺ membrane (Amersham, USA). The CCR cDNA gene clone from tobacco (1061 bp; a kind gift from Dr. Claire Halpin, Dundee, UK) was spotted as standard dilutions representing 1x10⁵ and 2x10⁵ copies in duplicates. The blot was probed with the radiolabelled tobacco *CCR* gene. Positive signal obtained under high stringency hybridization conditions was indicative of the presence of the *CCR* gene in *L. leucocephala*. Based on signal intensity it was also inferred that in *L. leucocephala CCR* belonged to a gene family represented possibly by two members (Fig. 2.3.1).
2.3.1 Slot Blot Hybridization for \( CCR \) gene in \( L. \) leucocephala:

Slot Blot of \( L. \) leucocephala genomic DNA hybridized with radiolabelled tobacco \( CCR \) gene (U20736):

Lane A1, A2 and A3, A4 signals from \( 2 \times 10^5 \) and \( 1 \times 10^5 \) copies of genomic DNA representations of \( L. \) leucocephala in duplicates, respectively. Lane B1, B2 and B3, B4 signals from \( 2 \times 10^5 \) and \( 1 \times 10^5 \) copies of tobacco \( CCR \) gene clone (each blotted in duplicates, respectively).

2.3.2 Isolation of partial fragment of \( CCR \) gene:

After establishment of presence of \( CCR \) gene in \( L. \) leucocephala, the next step was to isolate and characterize the gene in the same plant. This could be done in three ways:

1. PCR based approach
2. cDNA/gDNA library screening
3. Reverse genetics

In the present study, PCR based approach and cDNA/gDNA library screening were used as a tool for isolation of \( CCR \) gene from \( L. \) leucocephala. In preliminary attempts, PCR based method was followed to achieve above objective.

2.3.2.1 Total RNA isolation and cDNA preparation:

RNA was isolated from xylem tissue of \( L. \) leucocephala (Fig. 2.3.2) and cDNA first strand was prepared using above stated protocol section (2.2.2.8).
2.3.2.2 Primer designing:
All known CCR amino acid and cDNA sequences from NCBI GenBank database were aligned using CLUSTAL X software. Few conserved domains were recognized in the aligned sequences. Alignments were also performed using different combinations of all known sequences e.g.
1. Sequences reported from one type of family.
2. All tree species.
3. All full-length sequences reported.

On the basis of above alignments 10-12 pairs of primers were designed. These primers were used to perform PCR using gDNA and cDNA as a template.

2.3.2.3 PCR, Cloning and sequence analysis:
Above sets of primers, in all possible combinations were used in PCR reactions using cDNA and gDNA as a template. Out of all the primer sets only one set gave amplification with estimated size (Table 2.3.1).

<table>
<thead>
<tr>
<th>S. No</th>
<th>Name</th>
<th>Sequence: 5’- 3’</th>
<th>Tm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SCHF2</td>
<td>GATGAGGTGG TTGACGAGTC TT</td>
<td>64 °C</td>
</tr>
<tr>
<td>2</td>
<td>Samr1</td>
<td>TTAGCTGAGCCAGTGAGGTACTT</td>
<td>62 °C</td>
</tr>
</tbody>
</table>

Table 2.3.1: Primers used for isolation of partial CCR gene fragment.
Using above primers, ~250bp cDNA fragment and ~600 bp gDNA fragment were amplified (Fig. 2.3.3).

![PCR gel image](image)

**Fig. 2.3.3: PCR amplified fragment of CCR gene run on 1% agarose gel:** Lane M: 100 bp ladder Bangalore Genei, Lane 1: 250 bp cDNA fragment, Lane 2: 600 bp gDNA fragment.

The above two fragments were cloned in pGEM-T easy vector (Fig. 2.3.4) and sequenced. Sequencing revealed that cDNA fragment is 230 bp and gDNA fragment is 585 bp.
2.3.2.4 Sequence analysis of cDNA and gDNA Fragments:

The above cloned fragments of cDNA and gDNA were assigned as **pCCR230** and **pCCR585**. BLAST result of both the sequences showed significant similarity with other known *CCR* sequences in database.

**pCCR585**: Genomic DNA nucleotide sequence

```
1  GATGAGGTGG TTGACGAGTC TTGCTGGAGC AATTTGGATT ATTGCAAGAA
51  CACAAAGGTA TATACTTCCA ACCTACCCCC TCTCTTCTCA ATCATAAATC
101  CCCCATAATT TTGTTGACAC GTGATCTTGG AGGTCAATCA AGAGACAATA
151  GTTCTAAATT CTATTTTTAG AAGTAATATT ATATGCATAT GTCCCTCCCC
201  AGACTCCATC TGATAGAGAG TGGCTCGCCC TAACGCTCAT TTTGAAACTC
251  GAGTAAAGGAGCTGCTCAGT GCATTAGTCT CTCTAGGAAG GATAGATGCA
301  CATGACCTTA TCTTGGAGATT TTGAATCATG TCACATCATG CACTATATTA
351  GCAATCTTAC TTCTGGGGCTA AGCTCATGCT CATGGATACA ATATGAGGGA
401  TGGATTTTGC AGAATTGGCTA TTGCTATGGG AAGGATGGGG CAAGAAAATC
451  AGCATGGGGAT GAGGAAAAG CAAGAGGGGT GGATTTGGTT GTGGTGATAC
501  CATTTTTGCT TGTTGGGACCA TTGGTTCACA CCACCATGAA TGCAAGCACA
551  ATTCACATCC TCAAGTACCT CACTGGCTCA GCTAA
```
The sequence consists of an incomplete exon from nucleotide 1 to 57, a complete intron from nucleotide 58 to 412 (in red), one partial exon from nucleotide 413 to 585. Spidey hosted on www.ncbi.nlm.nih.gov/tools has been used for gene/exon prediction and to predict the splice sites. The intron splice junctions followed the GT and AG rule and were confirmed through cDNA analysis.

**pCCR230: cDNA nucleotide sequence**

<table>
<thead>
<tr>
<th></th>
<th>GATGAGGTGG TTGACGAGTC TTGCTGGAGC AATTGGGAAT ATTGCAAGAA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CACAAAGAAC TGGTATTGCT ATGGGAAGGC AGTGGCAGAG CAAGCAGCAT</td>
</tr>
<tr>
<td>51</td>
<td>GGGATGAGGC AAAAGCAAGA GGGGTGGATT TGGTTGTGGT GAATCCAGTT</td>
</tr>
<tr>
<td>101</td>
<td>TTGGTGTTGG GACCATTGCT TCAACCACC ATGAATGCAA GCACAATTCA</td>
</tr>
<tr>
<td>151</td>
<td>CATCCTCAAG TACCTCACTG GCTCAGCTAA</td>
</tr>
</tbody>
</table>

Alignment of the spliced gDNA sequence and cDNA sequence showed marginal sequence differences, which may be due to sequencing error. Alignment of the spliced partial gDNA and the cDNA sequence with known CCR sequences in NCBI GenBank Databases is shown below, which suggests that the sequence is a partial fragment of CCR gene from *L. leucocephala*.

![Alignment of sequences](alignment.png)

Srivastava S, University of Pune
2.3.3 Screening of genomic DNA Library:

To fish out the full-length gene the second option, screening of genomic DNA library was done. The partial 585 bp genomic DNA fragment was used to screen genomic DNA library in \( \lambda \) phage (ZAP II, Stratagene.). Screening was performed as described in 2.2.2.11. Ten plates with average number of plaques around 5000 were used for primary screening. Few positive plaques were picked and used for secondary and subsequently for tertiary screening (Fig. 2.3.5).

Secondary and tertiary screening was performed at a higher hybridization temperature (60 °C). Good positive signals were obtained. In the tertiary screening, transfection was performed with an aim to obtain isolated plaques. The positive plaques from tertiary screening were cut individually and excision was performed according to the manufacturer’s protocol. Excised plaque was further characterized to be a 5 kb fragment. This 5 kb insert was later characterized to be a part of a calcium dependent protein kinase.

Due to major failure of genomic library screening, next step to fish out full-length gene was performed. Information of partial cDNA sequence showing significant similarity with known orthologous \( CCR \) sequence was exploited to perform RACE. RACE reactions were performed as described in section 2.2.2.12 of this chapter.
Fig. 2.3.5: Screening of genomic library of *L. leucocephala*: Arrow shows positive signals. A, B and C: Signals from primary screening. D: Signals from secondary screening. E: Signals from tertiary screening.
2.3.4 Rapid Amplification of cDNA Ends (RACE):

2.3.4.1 5’ and 3’ RACE:

SMART-RACE cDNA amplification kit (Clontech Laboratories Inc., Mountain View, USA) and Total RNA isolated from xylem tissue were used to amplify the 5’ and 3’ ends of *L. leucocephala CCR* (Li-CCR) cDNA. The 5’ and 3’ ends of Li-CCR were amplified in two rounds of PCR with the Li-CCR gene-specific primers (Table 2.3.1) and with primers provided with the kit (Table 2.2.1). The first PCR for 5’ end amplification of *Leucaena CCR* gene was performed with the Samr1 and UPM primers under following conditions: 3 min at 95 ºC, 35 cycles of 1 min at 95 ºC, 30 s at 45 ºC, 3 min at 72 ºC, and a final extension step for 10 min at 72 ºC. The PCR product of the first round PCR was 40 fold diluted and used as template in the nested PCR amplification with Samr1 and NUP primers at the annealing temperature of 55 ºC and under the same PCR conditions as in the case of first PCR. The first PCR for 3’ end amplification of *Leucaena CCR* gene was performed with SCHF2 and UPM primers under following conditions: 3 min at 95 ºC, 35 cycles of 1 min at 95 ºC, 1 min at 58 ºC, 3 min at 72 ºC, and a final extension step for 7 min at 72 ºC. The PCR product of the first round PCR was 40 fold diluted and used as template in the second round of nested PCR with SCHF2 and NUP primers for Li-CCR at the annealing temperatures of 60 ºC (Lacombe *et al.*, 1997; Ma, 2007).

Secondary reaction of 5’ and 3’ RACE yielded an 850 bp and 825 bp fragments respectively (Fig. 2.3.6). These two fragments were cloned in pGEM-T easy vector and six clones from each 5’ and 3’ RACE product were sequenced. Two 5’ Clones and three different 3’ clones were characterized. On sequence validation it was found that two type of 5’UTR and one type of 3’ UTR were present. These clones were assigned as pGEM5’A, pGEM5’C, pGEM3’ (Genbank Accession Nos. 5’A-EU195226, 5’C-EU195225, 3’-DQ986908).
Fig. 2.3.6: 5’ and 3’ RACE PCR reaction: Lane M: Marker (100 bp ladder, Bangalore Genei, India), Lane 1 and 2: Secondary PCR product of 3’ RACE reaction, Lane 3 and 4: Secondary PCR product of 5’ RACE reaction.

Analysis of two 5’ clones showed that they are 850 (Sam5’ A) and 844 bp (Sam 5’ C) fragments with coding region starting at 177th and 157th base, respectively. Coding region showed significant similarity to known CCR sequence in NCBI GenBank. UTR of 176 and 156 bases was identified in these fragments (Fig. 2.3.7 a, b and c).

2.3.4.2 Sequence analysis of 5’ RACE clones:

5’ UTR region and Start codon have been highlighted in blue and red color respectively.

Nucleotide sequence of pGEM5’A, 5’ RACE product (GenBank Acc. No.EU195226):

1 GGTCAAGCCA GAGATCATCT TCTCCACTCT CCAACTTCTT CCTCTTTTCT CTCTCCTTCT
61 CACCAACCCC CGCACAGAGA ATTTTCTCAT TTCCCTCCTT TGCTCTATCC TAGGCTCTGT
121 TCTGATTCCA AATTCGTCTC TGTTTACAGA TAATCTTATT AACACTACCC ACCACCATGC
181 CTGCTGCCGC CCCAGCCCC ACCGCCGCTA ACACCACTTC ATCCAGTTCC GGGCAACCG
241 TCTGCGTCAC AGGGGCGGCT GGCTCATCAT CCTCTTTGAT TGTCAAGCTC TTGCTAGAGA
301 GAGGCTACAC TGTTAGAGGC ACCGTCAGAA ATCCAGATGA TTCTAAGAAC TCTCACTTGA
361 AAGAGTTGGA AGGAGCGAGAG GAGAGGCCTAA CTCTTCATAA GGTTGATCTT CCTGATCTTG
421 AATCTGTGAA AGCTGTTATC AATGGTTGTG ATGGCATCAT TCACACGGCT TCTCCAGTC
481 CATCACACCC CGAAGAGATG GATCAGCAGG CGGATGAATGG AGCAAAGAAT GTGATCATCG
541 CAGCTGCAGA AGCGAAAGTG AGAAGAGTAG TGTTCACGTC ATCCATTGGA GCCGTCTACA

Srivastava S, University of Pune
Nucleotide sequence of pGEM5’C, 5’ RACE product (GenBank Acc. No. EU195225):

1 GGAGAAGTCA TCTCTCCACT CTCACTCTTC CTCTCTCTCG TTTCAACCAA CCCCGCACAG
61 AGAATTTTCTC ACTCCCTCTTT TTGCTCTATC CTACGCTTCG TTCTGATTCC AAATCTTCTC
121 TGTATACAGA TAGTCTTTATT AACACTACCC ACCACCATGC CTGCTGCCGC CCCCGCCGCC
181 GCTAACACCA CCTCATCAGG TTCCGGCCAA ACCGTCTGCG TCACAGGCGC CGGTGGCTTC
241 ATCGCCTCTT GGATTGTCAA GCTCTTGCTA GAGAGAGACT ACACTGTCAG AGGCACCGCC
301 AGAAATCCCA ATGATTCTAA GAACGCACAC TTAAAAGAGT TGGAAGGAGC AGAGGAGAGG
361 CTAACTCTCTC ATAAGGGTGA TCTTCTTGTG ATGGGAATCTG TGAAAGCTGC TATCAATGGC
421 TGTGATGGCG TCATTCACAC GGCTTCTCCA GTCACAGACA ACCCCGAAGA GATGGTGGAG
481 CGGCGGTGA ATGGAGCAAA GAATGTGATC ATCGCAGCTG CAGAAGCGAA AGTGAGAAGA
541 GTAGTGTTCA CGTCATCCAT TGGAGCCGTC TACATGGACC CCAGCAGGAA CATTGATGAG
601 GTGGTTGACG AGTCTTGCTG GAGCAATTTG GAATATTGCA AGACCACAAA GAACTGGTAT
661 AGCTATCGGC AGGCGGTGAA ATCCCTGCTG ATGCTGAAAC AGTGAGAAGG GATGGTGGAG
721 TGGTTGACG AGTCTTGCTG GAGCAATTTG GAATATTGCA AGACCACAAA GAACTGGTAT
781 GCAAGCAGCC TCTCCATCCCTT CAGTGACACCT ATGAGCTGAG CTGCCCTGTAG TATCAATGGC

Deduced amino Acid sequence of coding region of pGEM5’A RACE product:

Met P A A A P A P T A A N T T S S G S G Q T V C V T G A G G F I A S W I V K L L
V I I A A A E A K V R R V V F T S I G A V Y Met D P S R N I D E V V D E S C W S
P V L V L G P L L Q S T Met N A S T I H I L K Y L T G S A
Deduced amino Acid sequence of coding region of pGEM5’C RACE product:

Met P A A P A A A N T S S G S G Q T V C V T G A G G F I A S W I V K L L L E
A A A E A K V R R V V F T S S I G A V Y Met D P S R N I D E V V D E S C W S N L
L V L G P L L Q T T Met N A S T V H I L K Y L T G S A

Alignment of deduced amino acid sequence of pGEM5’A and pGEM5’C showed few differences at amino acid level too. The sequence alignment showed major difference near the start codon, which strengthens the possibility of two different isoforms of CCR gene present in L. leucocephala (Fig. 2.3.7 c).

2.3.4.3 Analysis of two 5’ UTRs:
Inspection of both the UTRs revealed that there is a possibility of two different UTR, driving two different CCR genes. Some motifs were characterized which were already reported in promoters of plant origin. Motif like P$GAGA/GAGABP.01 and P$IBOX/IBOX.01 (IUPAC name for motifs) were characterized only in 5’C UTR and they were missing in 5’A UTR. This analysis supported the information of differential expression of two CCR genes, which is already reported for maize and Arabidopsis (Table 2.3.2).
<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
<th>Position</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>P$GAGA/GAGABP.01</td>
<td>(GA)n/(CT)n binding proteins (GBP, soybean; BBR, barley)</td>
<td>20 - 44(5°C)</td>
<td>gaaacgAGAGagaggaagagtgaga</td>
</tr>
<tr>
<td>P$MIIG/PALBOXL.01</td>
<td>Cis-acting element conserved in various PAL and 4CL promoters</td>
<td>61 – 75 (5’A)</td>
<td>gggggttGGTGaga</td>
</tr>
<tr>
<td></td>
<td></td>
<td>58 - 72(5°C)</td>
<td>gggggttGGTGaga</td>
</tr>
<tr>
<td>P$MYBS/MYBST1.01</td>
<td>MybSt1 (Myb Solanum tuberosum 1) with a single myb repeat</td>
<td>104 – 120(5’A)</td>
<td>tgtctATCCtagcct</td>
</tr>
<tr>
<td></td>
<td></td>
<td>82 - 98(5°C)</td>
<td>tgtctATCCtagcctt</td>
</tr>
<tr>
<td>P$IBOX/IBOX.01</td>
<td>I-Box in rbcS genes and other light regulated genes</td>
<td>109 - 125(5°C)</td>
<td>aaacaGAGAagattttgg</td>
</tr>
<tr>
<td>P$MIIG/ATMYB84.01</td>
<td>R2R3-type myb-like transcription factor (IIG-type binding site)</td>
<td>166 – 180(5’A)</td>
<td>tgtgtggGGTGAgg</td>
</tr>
<tr>
<td></td>
<td></td>
<td>143 - 157(5°C)</td>
<td>tgtgtggGGTGAp</td>
</tr>
</tbody>
</table>

Table 2.3.2: Nucleotide sequence analysis of 5’A and 5’C UTRs.
Fig. 2.3.7 a: Alignment of 5’A and 5’C UTR region.
Fig. 2.3.7 b: Alignment of nucleotide sequence of pGEM5’-A and pGEM5’-C clone
Fig. 2.3.7 c: Alignment of amino acid sequence of coding region of pGEM5’A and pGEM5’C
2.3.4.4 Sequence analysis of 3’ clone:

Analysis of 3’ clone showed that it is 825 bp fragments with coding region ending at 564th base. Coding region showed significant similarity to known CCR sequence in NCBI GenBank. A UTR of 260 bases was identified in this fragment. 3’ UTR region is highlighted in red. This clone was assigned as pGEM3’A

Nucleotide sequence of pGEM3’A clone (GenBank Acc. No. DQ986908):

1 GATGAGGTGG TTGACGAGTC TTGCTGGAGC AATTTGGAAT ATTGCAAGAA
51 CACAAAGAAC TGGTATTGCT ATGGGAAGGC AGTGGCAGAG CAAGCAGCAT
101 GGGATGAGGC AAAAGCAAGA GGGGTGGATT TGGTTGTGGT GAATCCAGTT
151 TTGTTGTTGG GACCATTGCT TCAAACCACC ATGAATGCAA GCACACTCTC
201 CATCCTCAAG TAGCTCAGTG CTTCTGCAAA GACCTATGGA ATGCGCCACTC
251 AGGCCTATTG TCATGTTAAG GATGTTGCAAT TAGCCCATGT TCTTGTATTC
301 GAGACCTCCTT CTGCCCTCGG TCGTTATCTA TGTTCGGAGA GGTCTCTCCA
351 CGGTGGAAGAA CTGGTGCAAGA TCCTCGCCAA ATATTTCCCA GAATACCCCA
401 TTCTCTACCAAA ATGTTTCCGA GAGAAGAATC CAAGAGCAAA ACCCTACACA
451 TTCTCTACA AAGGGCTGAA AATTTTAGGA TAGATTACTT ATTACAGCCTCA
501 TCAGTGGCTAA TACGAGACCAG TAAAGACCGT GCAGGACAAA GCGCATCTTC
551 TCCTCTCCAC CAAGTAAACAG AGACTCTGAA TATTAAATCC TAGACACTTTT
601 ATTTAGTTCTA TACTTCCCAA TATATTATTT ATAAAGAAAG AAGAAGAAGA
651 AGTCTCTTCTT AAGAATTTGG TGGCTCCATA TGAAGCTACTA ATTTATAGAT
701 CGATCATATGAA GAACATGTTG AATATCTTCA TCTGATCATG ACAATCTCTT
751 GAAACGGTTTT TCCTGTGAAAC CAGTGGGAC AATCAATATG TTTGTTTATA
801 AAAAAAAAAAAAAAAAAAAAAAAAAAA

Deduced amino acid sequence of coding region of 3’ RACE Product:

DEVDSECSWSNLYESKKNWYCYGKAVAEQA
AAWDDEAKARGVDDLVVVPNVLVLGPLLLQT
NASIHLKYLTSKAKTYANATQAAYVHVVDVAL
AHVLVYTEPSASGRYLCSSESSLHRGELEVILAK
KYFPEYPISPTKCSDEKNPRAKPYTFSNSNKRLKD
LGLETFPTPVHQCLYDTVKSLQDKGHLLLPKT*
3’ UTR of CCR gene:
Putative poly-adenylation site is highlighted (grey).

1 CAGAGATCTG TCAGATTTAA TCCTACACAC TTTATTTAGT TACTACTTCC
51 AAGTATTATT ATTATTAAAG AAGAAGAAGA AGAAGCAGTT TTTAGCAATT
101 TGGTGCGCTCC ATATGAAGCT ACTAATTTAA GATAGAATTA TGTGAACATG
151 TTGAATATCT TCATCTGTAT CAGACAATTC CTTGTAACGG TTTTTCTGTG
201 TAACAGTGGGT GACAATCAAT ATGTTTTGTTT TAAAAAATAAAAAAA
251 AAAAAAAAA

2.3.5.1 Amplification of full-length CCR gene:
Coding region of RACE products showed significant similarity to known CCR sequences in NCBI GenBank. Difference in 5’ UTR regions allowed designing two different forward and one reverse primer. Using combination of two different sets of primers, PCR was performed to amplify two different full-length CCR genes.

First strand cDNA was prepared from RNA and full-length CCR was amplified using cDNA as a template. Amplified fragment was of ~1 kb and was cloned in pGEM-T easy vector and sequenced (Fig. 2.3.8).

![Fig. 2.3.8: Amplification of full-length CCR gene](image)

Lane1 to 5: PCR amplification of CCR1, Lane M: 1 kb ladder, Lane 5 to 10: PCR amplification of CCR 2.

2.3.5.2 Sequence analysis of CCR 1 CCR 2:
Two CCR fragments were characterized and designated as pGEMCCR1 of 1011 bases (GenBank Accession No.DQ986907) and pGEMCCR2 of 1005 bases (GenBank Accession No. DQ986908).
Deduced Protein sequence showed that CCR1 is 336 AA and CCR2 is 334 AA long (the amino acid composition is given in Table 2.3.3). They were designated as Ll-CCRH1 and Ll-CCRH2 respectively (Fig. 2.3.10).

Nucleotide Sequence of Full-length Ll-CCRH1 gene:

```
1 - ATGCCTGCTGCCCGCCGCAGCCCCACCGCCGCTAAACACCACCTCATCAGTGGTGCTCCGCGC - 60
 1 - M P A A A A A N T T S S G S Q - 20
61 - ACCGTCTGCAGCAACAGGGTCGCTATCTCTCTCTGTATGTGATCTACGCTCTTGCTAC - 120
21 - T V C V T G A G G F I A S W I V K L L - 40
121 - GAGAGGCTCTACACTTTAGAGCGACACCTCAAGATTCTAGAACCTCTC - 180
41 - E R G Y T V R G T V R N P D S K N S H - 60
181 - TTGAAAGGCTTTGGAAGAGCAAGGAGGCTCTCATCTTAAGGTTATCTTCTTGAT - 240
241 - CTGAATCTGTGAAAGCTGTATGCTGATGCTACTTGACCCATCTCAACCCGTCTCTCCA - 300
81 - L E S V A V I N G C D G I I T A S P - 100
301 - GTACAGACAAACCCCAAGAGATGGTGAAGGCCGCGGTGAAATGGAATGATATCTC - 360
101 - V T D N P E E M V E P A V N G A K N V I - 120
361 - ATGCAGCTGAGACAGGCGAAATGAGATGTGCTCTGATCCATCTGAGCCGTC - 420
121 - I A A A A E A K V R R V V F T S S I G A V - 140
421 - TACATGAGCCCCAGGACAAATGATGAGTGCTTGACCCATCTGTACCTGAGCAATTG - 480
141 - Y M D P S R N I D E V V D S C W S N L - 160
481 - GAATATGCAAAACCAAAAGACTGTATTGCTATGGGAAGCCGAGCTGAGCAAGCA - 540
161 - E Y C K N T K N W Y C Y G K A V A E Q A - 180
541 - GCATGGGATGAGCCAAAGCAAGAGGTGATGGGATGCTATCAGCAGTCTCTTGTACG - 600
181 - A W D E A K A R G V D L V V V N P V L V - 200
581 - ACTGCCGCTGTCCATGAACCTATGCTAAACGGCAGCTGAGCAGCAAGCA - 660
191 - T G S A K T Y A N A T Q A Y V H V K D V - 220
661 - GCATGGGATGAGCCAAAGCAAGAGGTGATGGGATGCTATCAGCAGTCTCTTGTACG - 720
221 - T G S A K T Y A N A T Q A Y V H V K D V - 240
721 - GCATGGGATGAGCCAAAGCAAGAGGTGATGGGATGCTATCAGCAGTCTCTTGTACG - 780
251 - A L A H V L V Y E I P S A S G R Y L C S - 260
781 - GAGGGATCCTCTCCACCGCTTGGAGAATCTGTATGAGATCTCCTGCCGCGTCCTATATGTATTC - 840
261 - E S S L E V L V E I L A K Y F P E Y - 280
841 - CCAAATCTCACAAATGTTCCCGAGAACGCCAGCACAAAGCAAGACAAAAACGTACACATTTC - 900
281 - P I P T K C S D E K N P R A K A Y T F S - 300
901 - AACGAGGTAGCTGACAGGTATTAGAGATTTACACAGGATCCCTGCTCTCTCTCTCTC - 960
301 - N K R L K D L G L E F T P V H Q C L Y D - 320
961 - ACCGTTAAGGCCTGACAGGAACAAGGCGCTTCTCTCTCTCTCTCTCTCACCAGTGAA - 1011
321 - T V K S L Q D K G H L P L P T K * - 340
```
Nucleotide Sequence of Full-length Ll-CCRH2 gene:

1 - ATGCCTGCTGCCGCCGCCGCCGCCGCTAACACCACCTCATCAGGTTCCGGCCAAACCGTC - 60
1 - M  P  A  A  A  P  A  A  A  A  N  T  S  S  S  G  S  G  Q  T  V - 20
61 - TGCCTCACAGGCGGCCGCTGCTCACCTTCCTTTGGTGTAGCTCTTGCTAGAGAG - 120
21 - C  V  T  G  A  M  G  F  S  A  W  I  V  K  L  L  E  R - 40
121 - GACTACACTGTAGAGCGAGCACCAGCAAAATCAATGAGAGACTTACTCTCTTCTTCTGATCTGGAA - 240
41 - D  Y  T  V  R  G  T  A  R  N  P  D  D  S  K  N  A  H  L  K - 180
181 - ELEGENAIELMTLHKLVDLLE - 80
241 - TCTTGAAGAGCTCATATCAATGAGCCTGTATGAGCCCTACACCGGCTTCTCCAGTCACA - 300
81 - S  V  K  A  A  I  N  G  C  D  G  V  I  H  T  A  S  P  V  T - 100
301 - GACACCGAGCAGGCAAGATGGGCTTGCAGCGCTACACCTTCACACCTGGAGGCTGCTCTACATG - 360
101 - D N P E E M V E P A V N G A K N V I A - 120
361 - GCTGCAAGACGCCAAGGAGATGGAGAGGCTTCTAGCTCTACCAGTCATCCATTCAGTCACA - 420
121 - A A A E A K V R R V V F T S S I G A V Y M - 140
421 - TCTGTGAAAGCTGCTATCAATGGCTGTGATGGCGTCATTCACACGGCTTCTCCAGTCACA - 480
141 - D P S R N E V V D E S C W S N L E Y - 160
481 - TGCAAGACACCAAGAATGTGTTAGCTTATAGGGAAGCGAGCTGAGAGCAAGGCTGACG - 540
161 - C K T T K N W Y C Y G K A V A E Q A A W - 180
541 - GATGAGGCAAAAGCAAGGAGGGTAGTTGTTGGTAGGATACCCAGTTTGTGTTGAGGA - 600
181 - D E A K A R G V D L V V V N P V L V L G - 200
601 - CCATGCTTCAAGACCCATTGAAGGAGCGAAGACACATCAATCCTACATCTCAGTCCG - 660
201 - P L L Q T T M N A S T I H I L K Y L T G - 220
661 - TCTGCCAGACCTATGCAAATGGCCACTAGGCTATGTTAGAGAGATGTCATAGAA - 720
221 - S A K T Y A N A T Q A Y V H V K D V A L - 240
721 - GCCCAGTGCCTCTGTTGTTAGAGACATCCCTCTCCCGCGGTGTCTATGTATGCCGAGGA - 780
241 - A H V L V E E P S A S G R Y L C S E S - 260
781 - TCTCTCCACGGTGAGAACATGTCGAGATCTGCCGCCGCTGTTTCTATGTATGCCAGAGT - 840
261 - S L H R G E L V E I L A K Y F P E Y P I - 280
841 - CCTACCAAGTGTCGGACAGAGAATCAAGAGGAAAGACCCCTACACAGATCTCTCAAGAG - 900
281 - P T K C S D K N P R A K P Y T F S N K - 300
901 - AAGCTGAAGATTTAGGAGATTAGTATCAGGTTTCAATCAATACAGACAGGCTT - 960
301 - R L K D L G L E F T P V H Q C L Y D T V - 320
961 - AAGAGGCTGAGAAGAAGGAGCTTCTGCTCTCCACGCAAGTTA - 1005
321 - K S L Q D K G H L P L P T K * - 340

Srivastava S, University of Pune
<table>
<thead>
<tr>
<th>S. No</th>
<th>Amino acid (Aa)</th>
<th>Number (CCRH1)</th>
<th>Number (CCRH2)</th>
<th>Percent (CCRH1)</th>
<th>Percent (CCRH2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>35</td>
<td>37</td>
<td>10.42</td>
<td>11.08</td>
</tr>
<tr>
<td>2</td>
<td>C</td>
<td>8</td>
<td>8</td>
<td>2.38%</td>
<td>2.40</td>
</tr>
<tr>
<td>3</td>
<td>D</td>
<td>16</td>
<td>17</td>
<td>4.76</td>
<td>5.09</td>
</tr>
<tr>
<td>4</td>
<td>E</td>
<td>22</td>
<td>22</td>
<td>6.55</td>
<td>6.59</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>5</td>
<td>5</td>
<td>1.49</td>
<td>1.50</td>
</tr>
<tr>
<td>6</td>
<td>G</td>
<td>20</td>
<td>19</td>
<td>5.95</td>
<td>5.69</td>
</tr>
<tr>
<td>7</td>
<td>H</td>
<td>9</td>
<td>9</td>
<td>2.68</td>
<td>2.69</td>
</tr>
<tr>
<td>8</td>
<td>I</td>
<td>14</td>
<td>12</td>
<td>4.17</td>
<td>3.59</td>
</tr>
<tr>
<td>9</td>
<td>K</td>
<td>23</td>
<td>23</td>
<td>6.85</td>
<td>6.89</td>
</tr>
<tr>
<td>10</td>
<td>L</td>
<td>31</td>
<td>31</td>
<td>9.23</td>
<td>9.28</td>
</tr>
<tr>
<td>11</td>
<td>M</td>
<td>4</td>
<td>4</td>
<td>1.19</td>
<td>1.20</td>
</tr>
<tr>
<td>12</td>
<td>N</td>
<td>16</td>
<td>15</td>
<td>4.76</td>
<td>4.49</td>
</tr>
<tr>
<td>13</td>
<td>P</td>
<td>18</td>
<td>18</td>
<td>5.36</td>
<td>5.39</td>
</tr>
<tr>
<td>14</td>
<td>Q</td>
<td>6</td>
<td>6</td>
<td>1.79</td>
<td>1.80</td>
</tr>
<tr>
<td>15</td>
<td>R</td>
<td>12</td>
<td>12</td>
<td>3.57</td>
<td>3.59</td>
</tr>
<tr>
<td>16</td>
<td>S</td>
<td>24</td>
<td>22</td>
<td>7.14</td>
<td>6.59</td>
</tr>
<tr>
<td>17</td>
<td>T</td>
<td>22</td>
<td>24</td>
<td>6.55</td>
<td>7.19</td>
</tr>
<tr>
<td>18</td>
<td>V</td>
<td>33</td>
<td>32</td>
<td>9.82</td>
<td>9.58</td>
</tr>
<tr>
<td>19</td>
<td>W</td>
<td>4</td>
<td>4</td>
<td>1.19</td>
<td>1.20</td>
</tr>
<tr>
<td>20</td>
<td>Y</td>
<td>14</td>
<td>14</td>
<td>4.17</td>
<td>4.19</td>
</tr>
</tbody>
</table>

**Table 2.3.3: Amino acid composition of deduced CCR sequence.**

2.3.5.3: *In-silico* Restriction digestion analysis of Ll-CCRHs:

Restriction digestion was performed using online software NebCutter. The restriction pattern suggested that out of frequently used hexa-cutters, *Hinc* II, *Pvu* II, *Taq* I, *Stu* I, *Acc* I and *Bgl* I cut both the sequence but *Hind* III exclusively cuts only Ll-CCRH1 and this restriction site was absent in Ll-CCRH2 (Fig. 2.3.9). pGEMCCRH1 and pGEMCCRH2 were double digested with *Eco* RI and *Hind* III and run on 1% agarose (Fig. 2.3.9 C).
Fig. 2.3.9: Restriction digestion analysis: A and B: *In-silico* restriction analysis of L1-CCRH1 and 2; C: Restriction digestion of L1-CCRH1 and 2 by Hind III, Lane M: Marker, Lane 1: L1-CCRH1 and lane 2: L1-CCRH2.
Fig. 2.3.10: Alignment of L1-CCRH1 and L1-CCRH2 nucleotide sequence:

Continued to next page
**Fig. 2.3.10: Alignment of L1-CCR1 and L1-CCR2 nucleotide sequence.**

Srivastava S, University of Pune
2.3.5.4 Bioinformatic analysis of deduced \textit{CCR} proteins:

The \textit{L. leucocephala} \textit{CCR} cDNAs were conceptually translated and protein sequence was analyzed. The predicted molecular weight and pI for \textit{L1-CCRH1} and \textit{L1-CCRH2} were estimated to be 36.5 kD / 6.32 and 36.3 kD / 6.16, respectively and identified as cytosolic proteins (www.expasy.org/tools; Fig. 2.3.11). When comparative primary structure analysis of the deduced \textit{L1-CCR} s with orthologous \textit{CCR} sequences was performed using Pfam and rps-blast search, putative domains such as 3-beta-hydroxysteroid dehydrogenase/isomerase (3Beta_HSD), NADH-flavin reductase, and NAD dependent epimerase/dehydratase for binding cofactors were observed in both the \textit{L1-CCR} amino acid sequences (Lacombe \textit{et al.}, 1997; McInnes \textit{et al.}, 2002; Larsen, 2004). Apart from these domains, it contains conserved domains for various types of reductase and dehydrogenase activities (McInnes \textit{et al.}, 2002; Lacombe \textit{et al.}, 1997). This data classifies the protein as a member of oxido-reductase family (Fig. 2.3.12a). \textit{L1-CCR}s were aligned with other \textit{CCR} sequences and found to be similar to dihydro-kaempeferol reductase from \textit{Arabidopsis}, which also possesses the above mentioned conserved domains indicating that these domains are conserved evolutionarily in plant systems. Site for NADP binding (marked with solid blue arrow) and exon-intron junction was identified (McInnes \textit{et al.}, 2002) and a signature sequence of NWYCYGK (marked by thin red arrow) was also observed (Fig. 2.3.12b) (Larsen, 2003; Larsen, 2004). Hydropathic plot was plotted with a window size of 6 and hydrophobic regions were characterized by positive values. The hydrophilic regions were characterized as negative values and they correspond to exposed part of globular protein, which may be putative antigenic epitope (Fig. 2.3.13). As both the genes showed very little to no difference, they were designated as two isoforms of \textit{CCR} gene in \textit{Leucaena} genome. Considering the identical molecular weight, isoelectric point, conserved domains and homology between two genes only \textit{L1-CCRH2} gene was over-expressed and characterized further.
Deduced amino acid sequence LI-CCRH 1 Gene:

MPAAPAAPTAAANTSSGSQTVCTGAGGFIAWSIVKLLL DERGYTVRGTVRNPDDSKNSHLOLEGEAEERLTLHKVDLLD LSVKAVINGCDGIHTASPVTVDNPEEMVEPAVNGAKNVII AAEAKVRVVFSTSSIGAVYMPSRNIDEVVEDSCEWSNLE YCKNTKNWYCAGKAVAEQAAWDTEAKARGVDLVVVNPVL VLGPLLQSTMNASTIHILKYLTSAKTYANATQAYVHVKD VALAHVLYEIPSASGRYLCSESSLHRGEILAKYFPYEY PIPTKCSDEKNPRAKYTFSNKRLKDGLFEFTPVHQCLYDT VKSLQDKGHLPLPTK

Deduced amino acid sequence LI-CCRH 2 Gene:

MPAAPAAPTAAANTSSGSQTVCTGAGGFIAWSIVKLLL DYTVRGTARNPDDSKNAHLLEGEAEERLTLHKVDLLDLE SVKAANICDGIHTASPVTVDNPEEMVEPAVNGAKNVIIA AAEAKVRVVFSTSSIGAVYMPSRNIDEVVEDSCEWSNLEY CKYTKNWYCAGKAVAEQAAWDTEAKARGVDLVVVNPVLV LGPLLQSTMNASTIHILKYLTSAKTYANATQAYVHVKDVALAHVLYEIPSASGRYLCSESSLHRGEILAKYFPYEYP IPTKCSDEKNPRAKYTFSNKRLKDGLFEFTPVHQCLYDTV KSLQDKGHLPLPTK
Fig. 2.3.11: Alignment of Ll-CCR1 and Ll-CCR2 amino acid sequence.
Fig. 2.3.12 a: Putative conserved domains in Ll-CCR1.

3Beta_HSD: 3- Beta-hydroxysteroid; Polysacc_synt_2: Polysaccharide synthatase; NAD_binding_4: NAD binding domain; adh_short: Aldehyde dehydrogenase; AdoHcyase superfamily; Epimerase; NmrA: NmrA like protein; KR: KR like protein

Fig. 2.3.12 a: Putative conserved domains Ll-CCR2.
Fig. 2.3.12 b: Alignment of L1-CCR with orthologous CCR sequences from different plant system:

---

Srivastava S, University of Pune
Fig. 2.3.12 b: Alignment of LI-CCRs with orthologous CCR sequences from different plant system: Putative NADP binding site, marked in solid blue arrow, Thin red arrow marks CCR signature sequence and solid black arrow denotes Exon-Intron junction.
Fig. 2.3.13: Hydropathic plot. Window size of 6 suggested a good value for finding putative surface-exposed regions (Hydrophilic region characterized by positive values). Both the CCRs showed exactly similar putative antigenic epitopes (www.justbio.com).
2.3.5.5 Phylogenetic analysis of CCRs:

The closest in silico match for Ll-CCR1 and Ll-CCR2 proteins was Eucalyptus globulus and Populus tomentosa, which showed 76-77% identity and 84-86% similarity at the amino acid level respectively. Ll-CCR protein sequences share overall 54-66% identities with CCR of monocotyledonous plants and 77-65% identities with CCR of dicotyledonous plants, respectively. CCR sequence also showed 38 and 39% identities with dihydro-flavanoid reductases from Arabidopsis thaliana and Zea mays respectively and only 40% identities with Acacia mangium x Acacia auriculiformi. Ll-CCRH1 and Ll-CCRH2 bears 70% identity to AT-CCR1 and around 68% identity to AT-CCR2. Interestingly, they showed 70% identity with Pinus taeda. To investigate the evolutionary relationship between Ll-CCRs and CCRs from other plants, a phylogenetic tree (Neighbor-Joining) was constructed based on the deduced amino acids sequence. Z. mays-Dihydroflavanol-reductase (Zm-DFR) and A. thaliana Dihydrokempferol-reductase (At-DKFR) were taken as an out-group for analyzing evolutionary significance and CCR protein sequences from other plant species were used to deduce the similarity index. The tree can be divided into three clusters. The first cluster comprises of monocotyledonous plant species, which can be divided in two groups. The first group comprises of CCRs from Lolium perene, Zea mays-CCR1, Hordeum vulgare and Saccharum officinarum and other group consist of Zea mays-CCR2 and Oryza sativa-CCR. The second cluster comprises of dicotyledonous plant species, which again were further, divided in two groups. First group comprises of P. tomentosa, E. Globulus, Solanum tuberosum, Codonopsis lanceolata, Lycopersicon esulentum and Fragaria x ananassa; whereas the second group comprises of A. thaliana CCR1, A. thaliana CCR2 and L. leucocephala CCRH1 and CCRH2. The third cluster comprises of Oryza sativa and Z. mays-CCR2. Acacia mangium x Acacia auriculiformi (A. mangium). A. mangium, O. sativa and Z. mays-CCR2 are found to be least evolved CCR genes in their respective families. At-DKFR and Zm-DFR are treated as out-group. Phylogenetic analysis clearly demonstrated that Ll-CCRHs are closely related to CCRs of dicots than monocots and is distantly related to gymnosperm (P. taeda) and DFRs (Fig. 2.3.14).
Fig. 2.3.14: Phylogenetic tree made using the Neighbor-Joining method: Numbers at nodes represent the bootstrap values. The evolutionary history was inferred using the Neighbor-Joining method. *Z. mays* Dihydroflavanol-reductase and *A. thaliana* Dihydrokaempferol-reductase were taken as an out-group for analyzing evolutionary significance. CCR protein sequence from other plant species were used to deduce the similarity index. Monocot group consist of *Lolium perene* (AAG09817), *Zea mays*-CCR1 (CAA66707), *Hordeum vulgare* (AAN71760) and *Sacharrum officinarum* (CAA13176) and other group consist of *Zea mays*-CCR2 (AAO42621) and *Oryza sativa*-CCR (BAF08006). Dicot group comprises of *P. tomentosa* (AAR83344), *E. Globulus* (AAT74878), *Solanum tuberosum* (AAN71671), *Codonopsis lanceolata* (BAE48787), *Lycopersicon esculentum* (AAY41879), *Fragaria x ananassa* (AAP46143), *A. thaliana CCR1* (AAG48822), *A. thaliana CCR2* (AAG53687) and *L. leucocephala CCRH1* (ABL01801) and CCRH2 (ACB45309). *Acacia mangium* x *Acacia auriculiformi* (*A. mangium* AAY86360), *O. sativa* and *Z. mays*-CCR2 are found to be least evolved CCR genes in their respective families. At-DKFR (ABF74722) and Zm-DFR (AAK52955) are treated as out-group. Phylogenetic analysis clearly demonstrated that L1-CCRHs are closely related to CCRs of dicots than monocots and is distantly related to gymnosperm (*P. taeda*-AAL47684) and DFRs. GenBank accession numbers given in brackets.
2.3.6.1 Gene copy number-Southern Hybridization:

To further validate the results from slot blot experiment and to understand the distribution of the CCR gene in the *L. leucocephala* genome, Southern hybridization was performed. A 25 µg aliquot of *L. leucocephala* genomic DNA was restriction digested individually with the restriction enzymes *Nde* I, *Hinc* II, *Sac* I and *Bgl* I, which are predicted not to cut (*Sac* I) or cut once (*Nde* I, *Bgl* I and *Hinc* II) within the sequence. As shown in Fig. 2.3.15, four bands were detected in the *Hinc* II and *Nde* I digest, while two were recorded in the *Sac* I digest. *Bgl* I showed only three distinct bands, which may be due to the reason that fourth band, may be a very small and of low intensity, which is not seen in blot. These results suggest that Ll-CCRH2 is present as a two-copy number gene. Southern hybridization was done using a part of coding region and part of 5' UTR. An approximately 800 bp fragment was used as a probe for hybridization at 62 °C. Banding pattern in Southern hybridization suggested that at least two copy of CCR gene is present in *L. leucocephala* genome.

Fig. 2.3.15: Southern hybridization analysis: Four distinct band are seen in *Hinc* II and *Nde* I lane, two bands in *Sac* I lane, and three bands in *Bgl* I lane.
2.3.6.2 Gene copy number -Real time PCR

Copy number was also estimated using Real-Time PCR analysis. Brilliant II PCR master mix (Stratagene, USA) was used for the reactions. Dilution of CCR gene in pGEMT vector ranging from $1 \times 10^5$ to $7 \times 10^5$ copies was treated as standard. Known amount of *Leucaena* genomic DNA was used in each reaction, which was extrapolated to estimate copies of CCR gene present in *Leucaena* genome. The gene copy number in *Leucaena leucocephala* was estimated to be two in numbers (Fig. 2.3.16).

![Standard Curve](image)

**Fig. 2.3.16: Estimation of gene copy number using Real-Time PCR:** Solid triangle shows (1) $1 \times 10^5$, (2) $2 \times 10^5$ and (3) $3 \times 10^5$ copies of *Leucaena* gDNA in duplicates. Solid square represents $1 \times 10^5$ to $7 \times 10^5$ copies of CCR gene in pGEMT easy vector (Standard graph).
2.4 Discussion:

CCR is one of the key genes involved in lignin biosynthesis. Hydroxycinnamoyl-CoA esters of general phenyl propanoid pathway, when acted upon by CCR, become destined to form respective monolignols. CCR activity is found to be generally low in plants so it is hypothesized that it may play a crucial role as a rate limiting step in regulation of lignin biosynthesis (Ma, 2007). CCR gene was first reported from E. gunnii (Piquemal et al., 1997). CCR gene has also been characterized from maize, ryegrass, Arabidopsis and several other plants (Pichon et al., 1998; McInnes et al., 2002; Lauvergeat et al., 2001). Two cDNAs have been identified to encode CCR gene in maize (Pichon et al., 1998), ryegrass (McInnes et al., 2002) and Arabidopsis (Lauvergeat et al., 2001) and their chemical properties have been analysed. Down regulated tobacco plants with anti-sense CCR gene showed a strong decrease in lignin content with an alteration in development. These plants when crossed with homozygous CAD gene down regulated lines, showed hybrid with reduced lignin content without affecting the plant development (Chabannes et al., 2001). The above information suggests that there is a need for a detailed study of CCR gene. Apart from Euclyptus, poplar and Norway spruce, only few CCR genes have been cloned and characterized from tree species. Leucaena is a tree species which is extensively used in paper and pulp industry in India. To achieve transgenic L. leucocephala down regulated with CCR gene with reduced/altered lignin content which can be utilized properly in paper and pulp industry, it was neccessary to isolate and characterise CCR gene from it.

In this study, two cDNAs encoding CCR gene have been isolated which were designated as Li-CCRH1 and Li-CCRH2 (Fig. 2.3.10). Li-CCRHs have almost identical NADP binding motif found in other CCR sequences in NCBI GenBank database. Signature amino acid sequence from reported orthologous CCRs i.e. NWYCYGK was also present (Fig. 2.3.12; Fig. 2.3.13). Conserved domains of various other epimerase, reductase and dehydrogenase were also characterized on analysis of multiple sequence alignment of Li-CCRHs with other orthologous CCR protein sequences (Lacombe et al., 1997; McInnes et al., 2002). Phylogenetic analysis of Li- CCR grouped it among other CCRs from dicots.
As both CCR homologs exhibited similar property; for further study one of the CCR homologs i.e Ll-CCRH2 was characterized.

Studies have shown that there are eight CCR like sequences present in Arabidopsis (Anterola and Lewis, 2002) and poplar (Li et al., 2005) genome. In the same way; another key gene in lignin biosynthesis i.e Cinnamyl alcohol dehydrogenase (CAD) has also been found to be present as different isoforms. Three different CAD2 isoforms were isolated from Eucalyptus gunnii. Different CAD2 isoforms exhibited different affinities for different substrates which suggested that a modulation of monolignol pattern could be regulated through differential CAD activities (Hawkins and Boudet, 1994; Boudet, 2000). This could be also true in case of L. leucocephala as two different cDNAs encoding for CCR were characterized with 97% identities and showed very little difference between them (Fig. 2.3.10 and Fig. 2.3.11). According to earlier studies of ZM-CCR2 and AT-CCR2, it has been found that these isoforms mainly express in case of pathogen infection and/or wound or stress conditions but in present study of L. leucocephala all experiments were performed under normal conditions (no known stress given) hence both isoforms, Ll-CCRH1 and Ll-CCRH2 are hypothesized to be involved exclusively during constitutive lignification.

Gene copy number was estimated to be two in L. leucocephala. Considering the analysis of Ll-CCRH2 and Ll-CCRH1 gene it can be suggested that it belongs to CCR I gene of other plant species. The possibility of one type of CCR gene (as two isoforms) in Leucaena genome also can not be ruled out as in case of Eg-CCR (Piquemal et al., 1997) which is present as one copy in number and differentially expresses in normal and stressed conditions.
2.5 Conclusion:
Deduced amino acid sequence of cDNA from *Leucaena leucocephala* contains consensus sequences for NADP-binding site, *CCR* signature sequence (NWYCYGK). Other conserved domains responsible for dehydrogenase and reductase activity of known *CCR* genes were also present in above two cDNAs. All the above data classifies *CCR* to be a member of family oxido-reductase. The phylogenetic analysis of deduced amino acid sequence shows that it is having a close resemblance with other known *CCR* genes from dicots. The gene is present as two copies in the *Leucaena* genome. Considering the molecular mass of the deduced Ll-CCR protein and predicted involvement in constitutive lignification, it can be concluded that two cDNAs isolated from *Leucaena* encodes for cinnamoyl-CoA reductase (Ll-CCR).

2.6 References:


