5. DISCUSSION

Trees constitute the most important ligno-cellulosic woody biomass on earth that serves as an important raw material for various industrial products. The major challenge to attain sustainable bio-energy for future relies on understanding the molecular basis of growth and adaptation in woody plants to produce superior germplasm for rapid biomass production (Rathmann et al., 2010). Plantation forestry has gained importance in recent years as a renewable source of cellulosic biomass to produce raw material for pulp and bio-fuel (Hinchee et al., 2009). However, the growth of these woody perennials is delayed by long regeneration times during tree improvement programs (Grattapaglia et al., 2009; Seguin, 2011). Hence, research on reducing the generation time through breeding approaches has been a cornerstone in tree improvement programs. The use of advanced molecular technologies for accelerated domestication of tree species has been underway since two decades.

Wood property traits remain the most targeted trait in all tree improvement programs worldwide due to its industrial implications. Hence, understanding different aspects of wood properties at biochemical, physiological, genetics and genomics level has been widely documented (Carvalho et al., 2013). Further, the genomic research on herbaceous plants like Arabidopsis and rice serve as models to discover genes responsible for growth and development in trees. Nevertheless, transposing this data from annuals to woody plants becomes difficult, especially for those genes which control the wood formation (Grattapaglia et al., 2009). In recent years, several research groups are actively involved in generating insight into the molecular aspects of wood formation in tree crops like Populus (Goue et al., 2008; Wang et al., 2009), Eucalyptus (Barros et al., 2009; Solomon et al., 2010), Pinus (Yang et al., 2004; Paiva et al., 2008) Picea (Pavy et al., 2008) and Pseuodtsuga (Krutovsky and Neale, 2005; Ukrainetz et al., 2008b)

The major composition in any woody biomass is cellulose, which implies that understanding the molecular basis of cellulose biosynthesis is of vital importance. Compared to herbaceous species, the characteristics of wood in perennial plants vary with age (Plomion et al., 2001) and distinctive features like wood density, thickness of xylem vessels, cellulose biosynthetic process, cellulosic fibre size and fibre lumen differ
(Lu et al., 2008). During the process of xylogensis, several proteins have been associated with cellulose biosynthesis, out of which cellulose synthase complex (CSC) act as the major component of the rosette structure (Somerville, 2006; Joshi and Mansfield, 2007).

The cellulose synthase enzyme complex is encoded by various catalytic subunits of cellulose synthase genes, which indicate that CesA genes comprises of a multi-gene family (Richmond and Somerville, 2000). It has been reported that members of the CesA gene family play a major role in cell-type-specific expression, and their expression depends on the type of cell wall being laid down (primary or secondary cell wall) (Song et al., 2010).

_Eucalyptus_ includes some of the most widely planted hardwood trees in the world and has become a key player in high productivity plantations. They are preferred for their fast growth, short rotation, and short fibre with greater pulping yield and serves as a high quality raw material for pulp and paper industries (Rockwood et al., 2008). _Eucalyptus tereticornis_ Sm., ranks among one of the predominant species used in Indian subcontinent for production of pulp and rayon. However, the genomic studies in this species are limited to few EST submissions and identification of molecular markers (Gan et al., 2003; Poke, 2004; Balasaravan et al., 2005, 2006; Chezhian et al., 2010; Dasgupta et al., 2010; Arumugasundaram et al., 2011; Tiwari et al., 2013). Studies on understanding the gene expression patterns during wood formation/xylogensis is also not reported from this tropical species. Hence, the present study was taken up to obtain an insight into the function of CesA gene families in _E. tereticornis_.

The objective of the present study was to study the spatial expression profile of cellulose synthase gene families in _E. tereticornis_ and to isolate the developing xylem specific cellulose synthase genes. Three full length xylem specific cellulose synthase genes were found to be expressed predominantly in the developing xylem tissues, which was isolated and characterized during the course of the study. Further, their expression patterns during exogenous phyto-hormone treatment were also conducted.
5.1. Spatial expression analysis of cellulose synthase gene families in *E. tereticornis*

Gene expression studies facilitate to understand the functional roles of gene products among diverse biological processes (Taweel and Fernando, 2011). One of the simplest procedures to study gene regulation is the quantitative measurement of mRNA present in a cell at different time-points or under different conditions (Aceituno et al., 2008). The commonly used procedure for global gene expression analysis includes differential display PCR (Wei et al., 2000; Liu and Baird, 2003), serial analysis of gene expression (Matsumura et al., 2003), massively parallel signature sequencing (Busch and Lohmann, 2007), microarray (Koia et al., 2012) and RNA-Seq or transcriptome sequencing (Wong et al., 2011; Wakasa et al., 2014). The independent confirmation of differential gene expression patterns from a set of genes across various tissues and experimental conditions can be carried out through northern blot and quantitative real time PCR techniques (Dilks et al., 2003).

Cellulose is a vital polymer in the plant cell wall and identifying the function of the various CesA genes and studying its expression becomes important for understanding plant development (Cosgrove, 2005; Taylor, 2008). Molecular analysis of CesA genes from a woody species help in better understanding of the cellulose biosynthesis during xylogenesis (Lu et al., 2008). Reports reveal that various functional differences exist among different classes of CesA genes that lead to differential gene expression, regulation and catalytic functions between these genes. Hence, the functional roles of cellulose synthase genes in both the primary and secondary cell wall can be studied through understanding the differential regulation of these genes across different tissues and developmental stages (O’Brien, 2011).

Gene expression studies in different groups of CesA genes have been carried out using different procedures including standard and electronic Northern analysis, *in situ* hybridization, qRT-PCR, tissue prints, massively parallel signature sequencing, microarrays and promoter–reporter gene fusions. These studies have been conducted in both monocot and dicot species including cotton (Pear et al., 1996), *Populus* sp., (Wu et al., 2000; Tuskan et al., 2006), *Pinus* (Nairn and Haselkorn, 2005), maize (Holland et al., 2000; Dhugga, 2001), rice (Vergara and Carpita, 2001), *Arabidopsis* (Taylor et al., 1999; Fagard et al., 2000; Richmond and Somerville, 2001), *Eucalyptus* (Ranik and...
Myburg, 2006; Lu et al., 2008), Betula (Liu et al., 2012), and Acacia (Yong and Wickneswari, 2013).

In the present study, the differential expression of EtCesA genes in different tissues was analyzed using reverse northern and qRT-PCR. These techniques have been widely used to confirm gene expression across different tissues and experimental conditions (Corton and Gustaffson, 1997; Liu et al., 2003). Quantitative analysis of gene expression by reverse northern blot is a high-throughput technique for expression studies. It has been used to screen the transcript expression pattern of anthocyanin biosynthetic genes in different tissue types of wheat (Yao et al., 2005; Wang et al., 2006), fruit ripening genes of banana and mango (Godoy et al., 2009), cold responsive genes of Poncirus trifoliata (Sahin-Çevik, 2013) and disease resistance genes in Cucumis metuliferus (Lin et al., 2013b).

The alternate technique of qRT-PCR is a method of choice for high-throughput and accurate expression profiling of targeted genes. It is also useful for comparing the mRNA levels across different sample populations largely due to its high sensitivity, specificity, accuracy and reproducibility (Bustin, 2002; Bustin et al., 2005; Nolan et al., 2006). Relative quantification by qRT-PCR was used for understanding the differential expression patterns of Auxin/Indole-3-Acetic Acid (Aux/IAA) and Auxin Response Factor (ARF) genes in Populus trichocarpa (Kalluri et al., 2007), to analyze the temporal expression of stress responsive trehalose-6-phosphate in sugarcane (Almeida et al., 2013) and in spatial expression analysis of susceptible and resistant coffee genotypes for leaf miner disease (Cardoso et al., 2014). Absolute quantification of transcripts is another type of expression analysis in qRT-PCR, where cell-level variation pattern of the target gene expression can be detected during a developmental process, as reported in Ipomoea purpurea, Nicotiana benthamiana in tissue types like petals and leaves (Lu et al., 2012).

In the present study, reverse northern blot was used to study the differential expression of six classes of EtCesA transcripts in four different tissues including leaf, internodes, developing xylem and mature xylem tissues. The results of dot blot analysis revealed that the three classes of EtCesA genes, including EtCesA1, EtCesA2 and EtCesA3 expressed predominantly in secondary cell wall enriched tissues like developing xylem.
at greater intensity indicating that they were xylem specific in *E. tereticornis*. Further, the results also indicated that the *EtCesA4* and *EtCesA5* expressed at higher level in primary tissues like leaf and internodes revealing their primary cell wall specificity. Similar results of differential gene expression of *CesA* genes in primary and secondary tissues were observed in *E. grandis* (Ranik and Myburg, 2006; Lu et al., 2008) and in *P. tremuloides* (Kalluri and Joshi, 2004).

5.1.1. Selection and Validation of reference genes for qRT PCR analysis

Gene expression analysis by the qRT-PCR technique requires the use of reference gene for normalization of the quantification data. However, the major challenge in this technique of quantitative mRNA analysis is the selection of an appropriate reference gene for accurately normalizing the gene expression data (Pfaffl, 2001; Huggett et al., 2005). An ideal endogenous reference or housekeeping gene is one that is stably expressed within and among samples, including those from different tissues and cell types, developmental stages and treatment conditions (Nolan et al., 2006; Hong et al., 2008). The use of reference genes that are not stable across different conditions can have a significant impact on the results obtained, possibly leading to erroneous conclusions (Gutierrez et al., 2008; Guenin et al., 2009). Thus, the choice of an appropriate housekeeping gene for data normalization is a prerequisite in qRT–PCR experiments.

In the present study, seven reference genes including *EtAct*, *EtTub*, *EtIF4B*, *EtIDH*, *EtSAND*, *EtH2B* and *Et18srRNA* were selected based on the earlier report from different perennial species like *E. globulus* (Baova et al., 2010; Almeida et al., 2010) and *Fagus sylvatica* (Olbrich et al., 2008) and amplified in the internodes of *E. tereticornis*. The amplicon sizes of *EtAct*, *EtTub*, and *EtIF4B* was comparable to their orthologs from *E. grandis* hybrids as reported by Baova et al., (2010). The amplification of *EtSAND*, *EtIDH*, *Et18srRNA* and *EtH2B* was comparable to the size amplified in *E. globulus* (Almeida et al., 2010). Further, three algorithms including geNorm, BestKeeper and Normfinder (Vandesompele et al., 2002; Pfaffl et al., 2004, Andersen et al., 2004) were used to evaluate and select the best reference gene. These programs have been widely used in other plants for selection of reference genes as reported from wheat (Paolacci et al., 2009), flax (Huis et al., 2010); *Petunia* hybrid (Mallona et al., 2010), *Cichorium*
**intybus** (Maroufi et al., 2010), *Gossypium hirsutum* (Artico et al., 2010), *Eucommia ulmoides* (Chen et al., 2010b), *Eucalyptus globulus* (Almeida et al., 2010) and *Populus* (Xu et al., 2011a). Among the seven reference genes, these three algorithms determined *EtAct* as the best reference gene that was most stable across different tissues of *E. tereticornis* (Karpaga Raja Sundari and Ghosh Dasgupta, 2012).

Similar studies of reference gene selection was reported in Peach (*Prunus persica*), wherein polyubiquitin, translation elongation factor (*TEF2*) and RNA polymerase II (*RPII*) were identified as reliable reference gene across different developmental stages and tissues (Tong et al., 2009). In *Populus*, *18srRNA* along with *Efla* gene was reported to be the most appropriate reference gene for expression studies during adventitious rooting (Xu et al., 2011a). In *Quercus suber*, *Act* and *CACs* (Clathrin adaptor complexes medium subunit family) were reported as the most stable genes using geNorm and NormFinder algorithm (Marum et al., 2012). In all the above reports, a set of five to seven reference genes were evaluated for their suitability in gene expression studies. This is in accordance with the present study, where seven reference genes were selected for analyzing the best reference gene for qRT-PCR experiments in *E. tereticornis* tissues.

In eucalypts, several studies on differential expression of genes using qRT-PCR have been reported, but most of the studies have randomly adopted known endogenous reference gene for normalization, without screening for their stability. *EgIDH* (Isocitrate Dehydrogenase) was used as reference gene for differential expression studies in *E. globulus* (Goicoechea et al., 2005; Legay et al., 2007), while *Arf* (ADP ribosylation factor), was used for normalizing the expression of primary and secondary cell wall specific cellulose synthase genes in *E. grandis* (Ranik and Myburg, 2006). Recently, there have been reports in *Eucalyptus* elucidating selection and validation of reference genes, wherein Histone (*H2B*) and Tubulin (*Tub*) genes were reported to be the most stable genes for quantifying the gene expression during adventitious rooting in *E. globulus* (Almeida et al., 2010). On the other hand, polyubiquitin C (*UBC*), α-tubulin (*α-TUB*) and elongation factor 1-α (*EF1-α*) were reported as the most stable genes for cold acclimation and de-acclimation treatments (Fernandez et al., 2010). Subsequently, *EgIDH* was identified as the most suitable housekeeping gene in *Eucalyptus* hybrid clones (*E. grandis X E. urophylla*) subjected to biotic and abiotic stress condition.
(Boava et al., 2010). However, in the present study EtAct2 was identified as the most stable gene across different tissues along with EtSAND as the best gene combination for qRT-PCR analysis. This is the first report in Eucalyptus where Actin was found to be the most suitable gene for normalization of qRT-PCR experiments. Similarly, in Fagus sylvatica and Quercus suber, Act2 was identified as the most stable housekeeping gene (Olbirch et al., 2008; Maurum et al., 2012).

5.1.2. Relative gene expression analysis of EtCesA transcripts using qRT-PCR

In general, the hexagonal rosette structure of cellulose synthesizing complex suggests that six CESA subunits exist in higher plants. These subunits form the basis for qualitative and quantitative differences in cellulose deposited in primary and secondary cell walls (Doblin et al., 2002). This also justifies the occurrences of six classes of CesA genes, with three classes associated with primary cell wall development and the other three classes involved in secondary cell wall development (Joshi, 2003). Recent reports in higher plants have substantiated the earlier hypothesis and have shown that two groups of CesA gene families exist, which are associated with either primary or secondary cell wall deposition as reported in Arabidopsis thaliana (Fagard et al., 2000), Oryza sativa (Tanaka et al., 2003), Populus tremuloides (Kalluri and Joshi, 2004; Liang and Joshi, 2004), Pinus taeda (Nairn and Haselkorn, 2005), Hordeum vulgare (Burton et al., 2004), E. grandis (Ranik and Myburg, 2006; Lu et al., 2008), Solanum tuberosum (Obembe et al., 2009), Bambusa oldhamii (Chen et al., 2010a) and Gossypium hirsutum (Kim et al., 2012).

In the present study, the relative expression of the three EtCesA (EtCesA1, EtCesA2 and EtCesA3) transcripts studied using qRT-PCR indicated their predominant expression in the developing xylem tissues. The expression of the EtCesA3 transcripts was 87 fold higher in developing secondary xylem tissues, in comparison to the primary tissues like leaf and internodes (Karpaga Raja Sundari and Ghosh Dasgupta, 2014), which was analogous to the 50 fold abundant expression of EgrCesA3 in the secondary xylem tissues of E. grandis (Ranik and Myburg, 2006). Correspondingly, the expression of EtCesA1, EtCesA2 and EtCesA3 in developing xylem tissues was similar to the xylem specific predominant expression of PtCesA1, PtCesA2 and PtCesA3 in Pinus taeda (Nairn and Haselkorn, 2005), PtrCesA3 in Populus tremuloides (Kalluri and Joshi,
2004), BplCesA4, BplCesA7 and BplCesA8 in Betula platyphylla (Liu et al., 2012), Ll-8CesA in Leucaena leucocephala (Vishwakarma et al., 2012) and SpCesA1 in Shorea parvifolia (Lau et al., 2009) which revealed the expression of specific groups of the CesA genes during the secondary cell wall formation in woody perennials.

On the other hand, the relatively higher expression of EtCesA4 and EtcesA5 with 8 to 19 fold in primary cell wall tissues of E. tereticornis is comparable to the abundant expression of EgrCesA4 and EgrCesA5, with up to 15 fold in primary wall specific CesA genes of E. grandis (Ranik and Myburg, 2006). This result was also in agreement with the expression pattern of primary cell wall specific PtrCesA4 genes in Populus trichocarpa (Kalluri and Joshi, 2004) and PtrCesA6 and PtrCesA7 genes in P. tremuloides (Samuga and Joshi, 2004). There are also reports in Arabidopsis where AtCesA1 (Arioli et al., 1998), AtCesA3 (Scheible et al., 2001) and AtCesA6 (Fagard et al., 2000; Scheible et al., 2001) were expressed largely in primary walls of root and hypocotyls and thus involved in primary cell wall cellulose biosynthesis. Similarly, GhCes3 and GhCesA5 genes of Gossypium hirsutum (Kim and Triplett, 2007) and BplCesA3 genes of Betula Platypylla showed preferential expression in young primary tissues (Liu et al., 2012). The findings reveal that functionally two groups of CesAs exist with distinct role during primary and secondary wall biogenesis in trees and this functional distinctness of CesAs are conserved between annual herbaceous plants and perennial trees (Kalluri and Joshi, 2004).

5.2. Isolation of full length xylem specific EtCesA genes

The identification, cloning and characterization of differentially expressed genes could provide crucial insight into the molecular mechanisms involved in processes such as growth, development and differentiation. The discovery of the three CesA genes expressing in secondary cell wall of many higher plants support the fact that functional rosette complex of CESA requires the presence of at least three different functional co-enzymes (Gardiner et al., 2003; Taylor et al., 2003). It has also been reported that the developing xylem tissues are the appropriate tissue system to understand the assembly of cellulose synthase complex that involves enriched secondary cell wall deposition (Gardiner et al., 2003). Reports also reveal that secondary cell wall contain higher amounts of cellulose in comparison to primary walls, with a greater degree of
polymerization and crystallinity (Mellerowicz et al., 2001). Hence, characterization of CesA genes, primarily responsible for development of secondary cell wall is of vital importance. In the present study, three full length EtCesA genes predominantly expressed in developing xylem tissues of E. tereticornis were isolated and characterized.

5.2.1. Characterisation of EtCesA Genomic sequence

In general, the size of the CesA genes vary from 3.5-6.5 kb as reported from different plant species like Arabidopsis, Populus and Eucalyptus. These genes have 10-14 introns in the genomic sequence and encode proteins with 975-1088 amino acid residues (Richmond, 2000). This was evidenced in the present study where the genomic sequence of the three EtCesA genes varied from 5532 bp to 6523bp, which was comparable to its orthologs from E. globulus [5579 to 6414 bp] (Negishi et al., NCBI 2011) and E. urophylla [5464 to 6718 bp] (Favreau et al., NCBI 2013). The presence of 13 introns in EtCesA genes were concordant to the similar number of introns in CesA genes from Arabidopsis and Populus (Kalluri and Joshi, 2003; Richmond, 2000).

It was predicted that the intron-exon boundaries are highly conserved across CesA genes from different plants (Richmond, 2000). However, there was a difference in gene size between Arabidopsis and Populus, mostly due to the loss of introns and smaller size of introns in Arabidopsis (Liang and Joshi, 2004). The size of the introns in CesA genes from Populus is larger (91–483 bp) as reported by Liang and Joshi, (2004), which corresponds with the size of introns (32-608 bp in length) predicted in EtCesA genes. Reports suggest that the size of introns varies among different organisms due to their inherent mutational processes that cause insertions and deletions in gene sequences (Petrov et al., 2000, Petrov, 2002).

Recent reports also reveal that the larger size of intron may cause alternative splicing (AS) events, which can lead to alterations in protein function through modulating protein structure (Seo et al., 2011). This was illustrated by analyzing functional implications of AS events in wood-forming xylem tissues of Populus and Eucalyptus. It revealed that large proportion (25.0% in Populus and 26.8% in Eucalyptus) of AS events resulted in protein domain modifications involved in wood formation process.
which included specifically Cellulose_synt, Glyco_transf and Glyco_hydro, that could have significant impact in cell wall biosynthesis (Xu et al., 2014).

5.2.2. Comparative analysis of EtCesA CDS and their deduced amino acid sequence

In the present study, three full length CDS of EtCesAs expressed in developing xylem tissues were isolated with the sizes of 2940 bp in EtCesA1, 3114 bp in EtCesA2 and 2406 bp in EtCesA3 comprising of 979-1040 amino acids. The sizes of the full length CDS and deduced amino acid sequences were comparable to their orthologs from other tree species including Betula Platypylla (BplCesAs) with 2958-3255 bp size (Liu et al., 2012), Leuceana (Ll-CesA) with 3228 bp size (Vishwakarma et al., 2012), Populus tremuloides (PtCesA) with 2934 bp size (Wu et al., 2000), Eucalyptus grandis (EgrCesAs) with 2934 to 3135 bp size (Ranik and Myburg, 2006), Shorea parvifolia (SpCesA) of 3120 bp size (Lau et al., 2009) and Acacia mangium (AaxmCesA) with 3186 bp size (Yong and Wickneswari, 2013). Comparative analysis of EtCesA CDS with their orthologs from other eucalypt species revealed the presence of insertions and deletions (InDels) in the nucleotide sequence of EtCesA. However, it was predicted that the InDels in E. urophylla CesA sequence was larger which causes the variation in nucleotide sequence at the molecular level. This could also be substantiated to the phylogenetic distinctness of E. urophylla with their orthologs from other eucalypt and corresponded to the phylogenetic tree where E. urophylla formed a separate cluster in the cladogram (Figure 4.12).

It is reported that the deduced proteins from members of the CesA gene family are characterized by the presence of different domains that share significant sequence homology with other glycosyl transferase family. A conserved, extended N-terminal region containing zinc fingers resembling LIM/Ring domains followed by the CSR-I region signifies the plant CESA proteins that are involved in the oligomerization of CESAs (Richmond and Somerville, 2000; Saxena et al., 2001). The zinc-binding domain possesses conserved motifs like ‘CXXC’ motif that distinguishes CESA from Cellulose synthase like genes, where these motifs are absent (Kurek et al., 2002). All these structural characteristic of CESA protein was documented in the present investigation, where the translated amino acid sequences of all the three EtCESAs
harboured the signature domains, including a single Zinc binding domain, CxxC motifs, glycosyl transferase catalytic domain, QxxRW motif and trans-membrane helices as reported from CESA proteins of *Arabidopsis thaliana* (Taylor et al., 1999), *Zea mays* (Holland et al., 2000), *E. grandis* (Ranik and Myburg, 2006) and *Populus tremuloides* (Wu et al., 2000). Additionally, the ‘X’ residues in ‘CXXC’ motif was distinctly identified in each classes of EtCESA protein as ‘CNTC’ residues in ETCESA1, ‘CRVC’ residues in ETCESA2 and ‘CEIC’ motif in ETCESA3. These ‘CXXC’ residues identified in EtCESA matches exactly with the similar ‘X’ residues in the corresponding classes of CESA from other eucalypts species like *E. grandis*, *E. camaldulensis* and *E. globulus* as shown in multiple sequence alignment of CESA protein (Figure 4.11). This indicates that any un-known class of CESA protein can be identified through these ‘CXXC’ motifs.

The predicted CESA membrane protein contains a cytosolic N terminus followed by two membrane-spanning domains (TM), while the central portion of the protein is cytosolic comprising of six predicted trans-membrane domains at the C terminus (Joshi, 2003; Joshi et al., 2004). Thus, all members of the cellulose synthase super family possess conserved motifs surrounding three conserved Asp (D) residues with eight predicted trans-membrane helices (Doblin et al., 2002; Somerville, 2006). This was predicted in the present study, where eight TM domains were identified (with two TM helices in the N-terminus region and six TM helices in the carboxyl terminal region) in the EtCESA1, EtCESA2 and EtCESA3 proteins similar to the TM helices in the CESA proteins reported from *Arabidopsis thaliana* (Taylor et al., 1999), *Zea mays* (Holland et al., 2000), *E. grandis* (Ranik and Myburg, 2006), *Populus tremuloides* (Wu et al., 2000), *Leucaena leucocephala* (Vishwakarma et al., 2012) and *Betula platyphylla* (Liu et al., 2012).

5.2.3. Analysis of UTRs in *EtCesA* genes

Un-translated regions (UTR) of eukaryotic mRNAs contain motifs that are vital for regulation of gene expression at the post-transcriptional level. Specific interactions between RNA-binding proteins and *cis*-acting elements in 5’- and 3’-untranslated regions (UTRs) are responsible for regulating essential biological activities such as mRNA localization, mRNA turnover, stability and translation efficiency (Mignone et
Many mRNAs from different eukaryotic systems can be modulated in their translational efficiency by signals encoded either in the 5′ or 3′ untranslated region (Eveland et al., 2008). Earlier reports have demonstrated that cis-regulatory elements located in both the 5′ and 3′ UTRs or the coding regions of transcripts can enhance or regulate gene expression in plant and animals by interacting with trans-acting factors (Kertesz et al., 2006). The significant motifs present in the 5′ and the 3′ UTR are also involved in regulating the translation of genes in response to environmental stresses such as heat shock response, drought and salinity (Floris et al., 2009).

In the present study, the 5′ and 3′ UTR of all the three EtCesA genes were identified and analyzed by PLACE and PlantCARE UTR analysis tools. The sizes of both 5′ and 3′ UTRs of EtCesA1 and EtCesA3 gene were comparable to the length of UTRs in E. grandis, E. globulus, E. nitens and E. urophylla (Ranik and Myburg, 2006; Creux et al., 2013). On the contrary, the 5′ UTR of EtCesA2 was larger (469 bp) due to the presence of two introns in the 5′ UTR. This is analogous to the larger size of 5′ UTR of CesA2 gene in E. grandis retrieved from Phytozome database (Eucgr.A01324 www.Phytozome.org), which harboured two introns.

The presence of introns in UTR is reported to regulate tissue-specific expression of genes in dicots and cereal monocots (Kamo et al., 2012). Additionally, it is reported that the introns at the 5′ UTR stimulate translation efficiency by increasing the specific gene mRNA level at the post-transcriptional stage, as reported in both Gladiolus and Arabidopsis (Kamo et al., 2012). Although, there were introns in 5′ UTR of EtCesA2 gene, no intron was observed in 3′ UTR of EtCesA2 gene which can be substantiated by the report that the introns abundance in 5′ UTRs are greater in comparison to its less prevalence in 3′ UTR (Pesole et al., 2001; Hong et al., 2006).

5.2.3.1. Role of EtCesA UTR motifs in xylogenesis

Many motifs were predicted by PLACE and PLANT CARE analysis tools in 5′ and 3′ UTR of EtCesA1, EtCesA2 and EtCesA3 sequences, out of which some significant motifs suggest the role of CesA genes in various regulatory processes like xylogenesis, stress response and hormone regulation. The important motifs in 5′ UTR of EtCesA2 involved in xylogenesis included ‘LTRECOREATCOR15’ motif found to regulate the promoter of wood formation genes in Pinus taeda (Palle, 2010) and ‘BS1EGCCR’
motif that regulates vascular expression of Cinnamoyl-CoA reductase (CCR) gene in *E. gunnii* (Lacombe *et al.*, 2000) and wood formation genes of *Pinus taeda* (Palle, 2010). In addition, the occurrence of ‘MYBST1’ motif in the 3’UTR of *EtCesA1* was reported to act as a master switch for secondary wall formation in *Arabidopsis* and regulate all the three secondary wall-associated CesA genes in *Arabidopsis* (Oh *et al.*, 2003; Kim *et al.*, 2012).

The motifs in the 5’UTR of *EtCesA1* involved in hormone regulation process was ‘SURECOREATSULTR11’, found to regulate auxin responsive element in secondary xylem tissues of *Populus* (Bao *et al.*, 2009). The other important motifs in the 3’UTR of *EtCesA1, EtCesA2* and *EtCesA3* with similar role in auxin, gibberellin and brassi-steroid signalling included ‘ARFAT’ motif, P-box and TATC-box, which are reported in the promoter of *GhCesA4* gene in cotton and involved in fibre elongation (Triplett *et al.*, 2007; Kim *et al.*, 2011). Additionally, the presence of ‘ARR1AT’ motif in the 3’UTR of *EtCesA2*, reported as stress-responsive element in *Arabidopsis* (Ishida *et al.*, 2008) was found in the promoter of *EgrCesA3* during tension stress in *E. grandis*, indicating that this motif in *CesA* gene might play a role in plant response to mechanical stress (Lu *et al.*, 2008).

### 5.3. Phylogenetic analysis of *EtCesA* genes

Phylogenetic analysis of the deduced amino acid sequences of all the three *EtCesA* were compared across different eucalypts species. All the three classes of *EtCESA* grouped closely as distinct cluster with their corresponding classes of CesA genes expressed in secondary xylem tissues of *E. camaldulensis* [*EcCesA1, 2 and 3*] (Lin *et al.*, 2013a). On the other hand, the translated products from xylem specific CesA genes of *E. grandis* [*EgrCesA1, 2 and 3*] (Ranik and Myburg, 2006) and *E. globulus* [*EglCesA1, 2 and 3*] formed a close sub-group. This clustering pattern indicate that these sequences of EtCESA1, 2 and 3 are closely related to their orthologs from *E. camaldulensis*, a tropical species when compared to their orthologs from temperate taxas like *E. grandis* and *E. globulus*. On the contrary, CESA from *E. urophylla* grouped as a separate clade.
This close grouping of both the tropical eucalypts species is in accordance with the taxonomical hierarchy of the genus *Eucalyptus*, where *E. camaldulensis* and *E. tereticornis* are grouped together under the subgenus Symphomyrtus belonging to section Exsertaria, whereas *E. grandis* and *E. urophylla* comes under the section Transversaria (Pryor and Johnson, 1971; Balasaravanan *et al*., 2005). There are also reports which reveal that *E. grandis* show greater genetic relatedness to *E. tereticornis* and *E. camaldulensis* than *E. urophylla*, although it comes under the section Transversaria (Balasaravanan *et al*., 2005). This is also consistent with the present study where *E. grandis* grouped closely with classes of EtCESA than with *E. urophylla*. This is corroborated by the difference in natural distribution of *E. urophylla*, which has Indonesian origin as reported by the microsatellite study of Payn (2008). On the other hand, *E. globulus* belongs to section Maidenaria under the subgenus symphomyrtus. The close grouping of EgICESA with EgrCESA is in accordance with the previous report by De Araujo *et al*., 2002. The out-group of bacterial cellulose synthase from *Acetobacter xylinum* clearly separated from all the secondary xylem specific eucalypts CESA proteins, indicating its phylogenetic distance. Additionally, the bacterial CESA formed a separate branch with primary wall specific genes of CESA, indicating the homology of plant cellulose synthase with bacterial CESA due to its evolutionary origin (Arioli *et al*., 1998; Taylor *et al*., 1999).

The other dendrogram generated by the phylogenetic analysis of all the three EtCESAs with the deduced amino acid sequences from other plant species revealed significant similarity with the known CesA genes from dicots and monocots. This un-rooted radial tree revealed that all the three classes of EtCESA protein sequences grouped closely with their orthologs from both short-lived plant species and woody perennials, indicating that the underlying mechanisms of cellulose biosynthesis are conserved in plants (Joshi and Mansfield, 2007). This tree also revealed that many of the clades contain *Eucalyptus* CESA grouped with their orthologs from both monocots and dicots including PtCESA1, 2 and 3 of *P. trichocarpa* (Kalluri and Joshi, 2004), AtCESA4, 7 and 8 of *A. thaliana* (Richmond and Somerville, 2000) and ZmCESA10, 11 and 12 of *Zea mays* (Holland *et al*., 2000), indicating that the CesA orthologs are more similar than the paralogs, a characteristic feature of plant CESAs protein (Vergara and Carpita, 2001).
Similar phylogenetic pattern with clustering of xylem specific CesA orthologs was observed in other tree species like *Betula platyphylla* (Liu *et al.*, 2012), *P. tremuloides* (Liang and Joshi, 2004), *E. grandis* (Ranik and Myburg, 2006, Lu *et al.*, 2008), *Leuceana* (Vishwakarma *et al.*, 2012) and *Acacia* (Yong and Wickneswari, 2013). This also supports the hypothesis that various subclasses of CESA protein may have ascended relatively early in the evolution of these genes, before the divergence of both monocot and dicot lineage (Holland *et al.*, 2000; Appenzeller *et al.*, 2004). Additionally, all the monocot CESAs from different plants like rice, maize and barley formed a separate subgroup within the phylogenetic tree, suggesting that structural evolution of CesA gene continued even after divergence of monocot from dicots (Dhugga, 2001, Vergara and Carpita, 2001). The outgroup of bacterial Cellulose synthase from *Acetobacter xylinum* formed a separate cluster, indicating the evolutionary endo-symbiotic origin of plant cellulose synthase genes from bacteria (Arioli *et al.*, 1998; Taylor *et al.*, 1999; Delmer *et al.*, 1999).

**5.4. Molecular characterization of EtCesA genes during hormone signaling**

Wood formation or xylogenesis is under strict developmental-stage-specific-transcriptional control. Hormones including auxin, ethylene, gibberellins and cytokinins have a major role in transcription signaling during xylogenesis. A complex interplay between hormones is also essential for optimal growth and development in plants (Bjorklund *et al.*, 2007; Nieminen *et al.*, 2008; Sorce *et al.*, 2013). Auxin is an important regulator in the process of xylogenesis for the growth and development of primary and secondary vascular tissues which induces vascular differentiation in callus and explants (Mattsson *et al.*, 1999). This role of auxin (IAA) in the cambial zone differentiation was studied in different tree species like *Populus tremula* (Hellgren *et al.*, 2004; Israelsson *et al.*, 2005), *Pinus sylvestris* (Ugglä *et al.*, 1996, 1998, 2001) and hybrid aspen (Tuominen *et al.*, 1997). Gibberellins (GAs) act synergistically with auxin in stimulating cambial growth and also play a role in the processes of xylem elements expansion (Israelsson *et al.*, 2005). The significant role of GA in inducing xylem fibre elongation and relative abundance of fibers was reported in *Eucalyptus globulus* (Ridoutt *et al.*, 1996) and transgenic *Populus* (Eriksson *et al.*, 2000). Brassino-steroids have a role in vascular differentiation process during wood formation (Carlsbecker and
This hormone was also found to stimulate tracheary element differentiation in the cambial region of pine trees (Kim et al., 1990) and regulate cellulose biosynthesis as reported in Arabidopsis (Xie et al., 2011).

5.4.1. Differential Expression of EtCesA genes in response to Auxins

Auxins are historically known to regulate diverse developmental processes which affect cambial activity and xylem development during wood formation (Sundberg et al., 2000) as reported in hybrid aspen [Populus tremula L. X Populus tremuloides Michx] (Tuominen et al., 1997) and Pinus sylvestris L. (Uggla et al., 1998). Recently, functional studies using transgenic Populus have provided further evidence on the role of auxin in the regulation of secondary xylem development (Nilsson et al., 2008). In the present study, techniques like in situ hybridization and quantitative real time PCR were used to analyze the differential expression pattern of EtCesA transcripts. The technique of mRNA in situ hybridization facilitates the direct visualization of gene expression within cells and tissues (Javelle et al., 2011). This expression analysis system has been used to characterize the spatial expression pattern of xylem specific cellulose synthase genes in Populus tremuloides (Wu et al., 2000), lignin biosynthetic genes in E. gunnii (Hawkins et al., 2003) and CesA genes in different tissues of E. grandis (Lu et al., 2008). The qRT-PCR analysis was used to study the relative expression pattern in hormone signalling genes during brassinosteroid and abscisic acid signalling in Arabidopsis thaliana (Zhao et al., 2013; Yang et al., 2014).

In the present study, the expression pattern of developing xylem specific EtCesAs were evaluated during the exogenous application of the natural (IAA) and synthetic auxin (NAA) in rooted cuttings. A significant up-regulation of all the three xylem specific EtCesAs was observed with intense dark black spots during IAA treatment in the early formed xylem tissue region, when compared to less denser spots in NAA treated sections during in situ expression analysis. Similarly, the expression analysis by qRT-PCR also revealed that the relative expression of EtCesA genes in IAA treated cDNA samples was up to 9 fold abundant in comparison to NAA treatment (1 to 2 fold).
In *A. thaliana*, the up-regulation of specific CesAs during endogenous auxin signalling was reported by Goda *et al.*, 2004. However, no direct evidence of CesA expression during exogenous application of auxins is reported in woody perennials. It is reported that synthetic auxin like NAA had a lesser effect on the secondary wall cellulose biosynthesis in cotton ovule/fibre culture (*Gossypium hirsutum*), when compared to the natural auxin, IAA (Singh *et al.*, 2009). The expression patterns of genes involved in the secondary wall cellulose synthesis like *GhCesA1* and *GhCesA2* showed significant up-regulation on exogenous application of IAA, while limited expression was reported on NAA application in cotton (Singh *et al.*, 2009). Additionally, it was reported that auxins not only regulate the expression levels of cellulose synthase, but also control the transition between fibre elongations to fibre thickening in cotton (Kim *et al.*, 2011). This report is similar to the present study, which revealed the significant up-regulation of xylem specific CesAs during IAA treatment when compared to NAA treatment.

Another report on the expression pattern of *GhCesA* promoter in transgenic *Arabidopsis* revealed that NAA down-regulated the secondary xylem specific *GhCesA4* expression and showed differential regulation at different developmental stages of cotton fibre (Kim *et al.*, 2011). In bamboo (*Bambusa oldhamii*), the mRNA levels of the four primary cell wall specific *BoCesA2, 5, 6* and 7 genes decreased significantly upon treatment with NAA, suggesting that synthetic auxin like NAA had a significant effect only on rooting of cuttings but had negative impact on shoot proliferation and development (Chen *et al.*, 2010a). The physiological and cellular difference in response to both the natural and synthetic auxin can be attributed to the difference in their mechanisms of uptake (Delbarre *et al.*, 1996), rate of metabolism (Beyer and Morgan, 1970) and interaction with the auxin receptor TIR1 (Tan *et al.*, 2007). Further, there are reports which revealed that the response of CesAs to the naturally abundant IAA could be attributed to the presence of Auxin response elements (*AuxRE*) in the promoter region as documented in *GhCesA4*, the gene involved in secondary wall synthesis in cotton (Wu *et al.*, 2009).

5.4.2. Differential Expression of *EtCesA* genes in response to Gibberellins

Earlier reports revealed that gibberellins could be used to regulate the direction of cellulose microfibrils by stabilizing cortical microtubule array in the shoot and root (Shibaoka, 1993; Inada et al., 2000). GAs are reported to play two tissue-specific roles in the wood forming zone as demonstrated in hybrid poplar by Mauriat and Moritz (2009), where they stimulate xylogenesis in the cambial zone and fiber elongation in the developing xylem. This is well established in cotton, where exogenous application of GA₃ caused an increase in cellulose synthesis, resulting in modifications of cell wall chemistry with increased cellulose production (Seagull et al., 2000). Another report in Arabidopsis reveal that the enhanced xylem deposition accelerated secondary growth and shift to xylogenesis occurs during GA signalling (Ragni et al., 2011).

Funada et al. (2008) reported that exogenous application of GA induced tension wood formation in four tree species, Fraxinus mandshurica var. japonica, Quercus mongolica var. grosseserrata, Kalopanax pictus and Populus sieboldii. It was identified that GA signaling caused increase in cellulose content and changes in the orientation of cellulose microfibrils. In the present study, significant induction of xylem specific EtCesA 1, 2 and 3 genes were observed in the GA treated sections at the region of early formed xylem tissues with intense dark spots during in situ hybridization and up to 6 fold higher expression compared to control in qRT PCR analysis, supporting the observation made by Funada and coworkers (2008). The increased cellulose deposition could be explained through the up-regulation of the xylem specific CesA genes, as observed in the present study, which would result in increased biosynthesis of cellulose.

5.4.3. Differential Expression of EtCesA genes in response to Brassino-steroids

Brassino-steroids (BRs) include plant steroid hormones that regulate different aspects of plant development including vascular differentiation (Turner et al., 2007). The typical role of BR dependent plant vascular development was illustrated in Zinnia culture where elevated levels of expression of transcription factors during secondary wall formation was documented (Fukuda, 2004; Yamamoto et al., 2007). It was reported that this hormone induced transcription factors, which in turn regulated the cambial cell differentiation and xylem cell specific gene expression (Turner et al., 2007; Lucas et al., 2013). Similar function was also identified in Lepidium sativum, where the exogenous application of Brazinazole (BR inhibitor) inhibited the development of secondary
xylem, indicating that BR participates in vascular differentiation (Nagata et al., 2001) and also control tissue-type-specificity of vascular cell proliferation (Carlsbecker and Helariutta, 2005).

The role of BRs in regulation of cellulose biosynthetic genes were studied in Arabidopsis, wherein it was found that the content of cellulose and the expression level of primary xylem specific cellulose synthases was induced by BR (Xie et al., 2011). This was also supported by another report in Arabidopsis, where the BR-deficient or BR-perceptional mutants contained less cellulose than the wild type, which confirmed that BRs regulate CesA gene expression (Oh et al., 2011). Additionally, BR was also reported to up-regulate the expression of secondary xylem specific cellulose synthase genes including CesA4, CesA7 and CesA8 in Arabidopsis (Oh et al., 2011). However, in the present study the expression of the xylem specific CesA in E. tereticornis was not induced significantly by 2, 4-epiBL, which may be due to the difference in response of secondary xylem tissues of perennials like eucalypts from Arabidopsis during hormone treatment. However, further investigation on expression patterns of primary cell wall specific CesA genes is necessary to clearly elucidate the effect of brassinosteroids on regulating cell wall biosynthesis in eucalypts.