GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Isolation and characterization of *CCoAOMT* in interspecific hybrid of *Acacia auriculiformis* x *Acacia mangium* - a key gene in lignin biosynthesis

S.L. Pang[1], S.S. Ong[2], H.H. Lee[3], Z. Zamri[4], K.I. Kandasamy[5], C.Y. Choong[6] and R. Wickneswari[6]

[1]Tree Breeding Unit, Applied Forest Science and Industry Development, Sarawak Forestry Corporation, Kuching, Sarawak, Malaysia
[2]Department of Medical Microbiology, Faculty of Medicine, Universiti Malaya, Kuala Lumpur, Malaysia
[3]Department of Biological Science, Faculty of Science, Universiti Tunku Abdul Rahman, Kampar, Perak, Malaysia
[4]School of Bioscience and Biotechnology, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia
[5]Malaysia Biotechnology Corporation, Kuala Lumpur, Malaysia
[6]School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

Corresponding author: R. Wickneswari
E-mail: wicki@ukmi.edu.my

Genet. Mol. Res. 13 (3): 7217-7238 (2014)
Received March 28, 2013
Accepted February 13, 2014
Published September 5, 2014
DOI http://dx.doi.org/10.4238/2014.September.5.7

ABSTRACT. This study was directed at the understanding of the function of *CCoAOMT* isolated from *Acacia auriculiformis* x *Acacia mangium*. Full length cDNA of the *Acacia* hybrid *CCoAOMT* (*AhCCoAOMT*) was 1024-bp long, containing 750-bp coding regions, with one major open reading frame of 249 amino acids. On the other

hand, full length genomic sequence of the *CCoAOMT* (*AhgflCCoAOMT*) was 2548 bp long, containing three introns and four exons with a 5' untranslated region (5'UTR) of 391 bp in length. The 5'UTR of the characterized *CCoAOMT* gene contains various regulatory elements. Southern analysis revealed that the *Acacia* hybrid has more than three copies of the *CCoAOMT* gene. Real-time PCR showed that this gene was expressed in root, inner bark, leaf, flower and seed pod of the *Acacia* hybrid. Downregulation of the homologous *CCoAOMT* gene in tobacco by antisense (AS) and intron-containing hairpin (IHP) constructs containing partial *AhCCoAOMT* led to reduction in lignin content. Expression of the *CCoAOMT* in AS line (pART-HAS78-03) and IHP line (pART-HIHP78-06) was reduced respectively by 37 and 75% compared to the control, resulting in a decrease in the estimated lignin content by 24 and 56%, respectively. *AhCCoAOMT* was found to have altered not only S and G units but also total lignin content, which is of economic value to the pulp industry. Subsequent polymorphism analysis of this gene across eight different genetic backgrounds each of *A. mangium* and *A. auriculiformis* revealed 47 single nucleotide polymorphisms (SNPs) in *A. auriculiformis CCoAOMT* and 30 SNPs in *A. mangium CCoAOMT*.

Key words: *CCoAOMT*; *Acacia* hybrid; Lignin biosynthesis; Downregulation; Single nucleotide polymorphism

INTRODUCTION

Lignin is one of the major plant cell wall components that make up 15-35% of the dry weight of wood. It confers mechanical strength and hydrophobicity to the vascular vessels (Boerjan et al., 2003). Lignin is also important in the plant defense system since it is often deposited at the wound or pathogen invasion sites to provide a physical barrier against further damage to adjacent tissues (Pagadala et al., 2009). This heteropolymer is derived from three substituted cinnamyl alcohols/monolignols, namely ρ-coumaryl, coniferyl, and sinapyl alcohols. These monolignols are synthesized through the phenylpropanoid pathway with successive hydroxylation and O-methylation of the aromatic ring. When incorporated in lignin polymer, these monolignols give rise to ρ-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) phenylpropanoid units respectively. Lignin content and composition are known to vary between the major groups of higher plants, even between species and cell types. For instance, lignins of angiosperm are composed of G and S units whereas gymnosperm lignins predominantly consist of G units. On the other hand, lignin from grasses incorporates considerable amount of H units. Besides, lignin composition is also known to change with different environmental factors (Whetten et al., 1998).

Since the first elucidation of the phenylpropanoid pathway, caffeic acid O-methyltransferase (*COMT*), which uses free acid forms of hydroxycinnamates as subtrates, was thought to be the only enzyme involved in the methylation steps in the phenylpropanoid biosynthesis pathway. However, this notion was challenged by Ye et al. (1994) when they pro-

vided the first evidence of the possibility of the involvement of caffeoyl-CoA 3-*O*-methyl transferase (*CCoAOMT*) in lignin biosynthesis in *Zinnia elegans*. *CCoAOMT* was thought to be involved in the plant defense system by synthesizing cell wall-bound ferulic acid. In parsley cell suspension cultures, the addition of fungal elicitors induced both ρ-coumaryl CoA 3-hydroxylase and *CCoAOMT*, which methylated caffeoyl CoA to form feruloyl CoA and 5-hydroxyferuloyl CoA to form sinapoyl CoA, respectively (Kneusel et al., 1989). Later, the expression of *CCoAOMT* was induced during lignifications in both *in vitro* treachery elements and lignifying tissues of *Zinnia* stems (Ye et al., 1994).

CCoAOMT exists as a multigene family (Lewis and Davin, 1999). It has been reported that this gene is encoded by one to two members in parsley (Grimmig and Matern, 1997), grapevine (Busam et al., 1997), alfalfa (Inoue et al., 1998) and aspen (Meng and Campbell, 1998), five to ten members in zinnia (Ye et al., 1994), and seven members in *Arabidopsis* (Raes et al., 2003). There are three classes of *CCoAOMT* in tobacco, and it was suggested that each class plays a role at different stages of lignification (Maury et al., 1999). *CCoAOMT* has been considered to be the key element in the biosynthesis of G units. Suppression of this gene in alfalfa (Guo et al., 2001; Chen et al., 2006) leads to a reduced lignin level and a reduction of G units without decreasing the S units. By contrast, similar studies conducted in *Nicotiana tabacum* cv. Xanthi (Zhong et al., 1998) and *Populus tremula* x *Populus alba* (Meyermans et al., 2000; Zhong et al., 2000) suggested the role of this gene in the biosynthesis of guaiacyl and syringyl lignins as the plants transformed not only demonstrated a decrease in total lignin content but also lower levels of S and G units.

A single nucleotide polymorphism (SNP) is single base variation at a unique location within a locus among different individuals of a species. A SNP occurring in exon regions may highly affect conserved amino acid sites and ultimately protein function (Polakova et al., 2005). In addition, polymorphism in the noncoding DNA sequences (regulatory domains and intron regions) can affect the level of gene expression or RNA stability, and result in quantitative variants. Both types of genetic variation may produce functional changes and affect phenotypes of an organism (Guo et al., 2004). The frequency and nature of SNPs in forest trees have received considerable attention in recent years (Zhang and Zhang, 2005). Mackay (2001) reported that non-synonymous mutations in *Acacia mangium* and *Acacia auriculiformis* may play a significant role, since changes to the protein sequence are most likely to have an effect on enzyme structure and function, therefore possibly producing a phenotypic effect (Mackay, 2001). Assaying sequence diversity in wood biosynthetic genes will lead to the identification of polymorphisms that may be associated with variation in wood quality traits. Ultimately, these SNPs can be developed into genetic markers and used in marker-assisted breeding/selection programs for improving wood quality traits in forest trees.

The first occurrence of a possible natural hybrid between *A. mangium* and *A. auriculiformis* was in 1972 in Ulu Kukut, Sabah, Malaysia. Principally, interspecific hybrids create new and different gene combinations that may not occur in nature, and these hybrids sometimes exhibit heterosis in which the hybrid is superior to both parents. This hybrid is reported to grow more vigorously than either parent (Darus and Rasip, 1989). In two years, the tree can reach 8 to 10 m and 7.5 to 9 cm DBH (diameter at breast height). The fast growth rates enable the hybrid to be harvested 2-3 years earlier than the non-hybrid. Stem volume of the *Acacia* hybrids is 2-3 times greater than that of *A. mangium* and 3-4 times greater than that of *A. auriculiformis* of the same age. F_1 hybrid trees in Vietnam produced 300-500% greater wood

volume than the parental species at 2.5-3 years. Four- to five-year-old hybrid trees produced, on average, twice the wood volume of *A. mangium* (Kha, 1996). Moreover, this species is free from heart rot disease, which made it a better choice for plantations than *A. mangium*. A very important feature of the *Acacia* hybrid is the higher pulping efficiency. The hybrid appears to have higher cellulose content and lower lignin content compared to both pure species (Yamada et al., 1990), which is beneficial to the pulp and paper industry. Besides, the paper produced from the hybrid has better mechanical strength, such as pulling and folding strength being markedly greater than the paper produced from *A. mangium* or *A. auriculiformis* wood. These features suggest that the *Acacia* hybrid has great potential to be used as a planting material in forest plantations especially for the pulp and paper industry.

In this study, we attempted to understand the involvement of the *CCoAOMT* gene in regulating the lignin content and composition during lignin biosynthesis in the *Acacia* hybrid. Full length cDNA and genomics sequence characterization as well as Southern analysis was initiated to provide a better picture of the regulatory elements and the copy number of the *CCoAOMT* gene in the *Acacia* hybrid. A partial *CCoAOMT* cDNA of the *Acacia* hybrid was transformed into tobacco using antisense and intron-containing hairpin (IHP) constructs to investigate the function of this gene. Furthermore, the polymorphism in the *CCoAOMT* gene across different *Acacia* genotypes was investigated through DNA sequencing.

MATERIAL AND METHODS

Plant materials

All tissues except roots used in this study were obtained from *A. auriculiformis* x *A. mangium* hybrid plants in Plot W, Plant Biotechnology Centre, Universiti Kebangsaan Malaysia, Bangi, Malaysia. *Acacia* hybrid 120 was chosen for full length cDNA isolation. Inner bark tissues were scraped from the cut-out bark into a clean plastic bag before dropping into liquid nitrogen. The sample was stored at -80°C until use. The inner bark tissue referred to here was a mixture of phloem, cambium, and xylem. Full length genomic *CCoAOMT* was performed using leaf tissues from *Acacia* hybrid 113. For gene expression study, the leaf, flower, inner bark and green seed pod were collected from *Acacia* hybrid 113 (7-year-old tree), while root tissues were collected from tissue culture plantlets (clone M5) obtained from Forest Research Institute Malaysia (FRIM). Florets (floral tissues) were collected and dropped directly into a plastic bag by gently running the forceps along the side of the spikes. Plantlets were gently removed from tissue culture media and washed before cutting out the root samples. All the samples were immediately dropped into liquid nitrogen and stored at -80°C until use.

RNA extraction

Total RNA was extracted using the RNeasy Midi kit (Qiagen, Germany), and RNA quality and concentration were determined by UV spectrophotometry. mRNA was isolated from total RNA using Dynabeads Oligo-(dT)15 (Dynal, Norway). To investigate the expression of the *CCoAOMT* in 18 transgenic tobacco plants, total RNA was extracted from the inner bark tissue collected from the 7th to 10th internodes of 12-week-old transformed tobacco plants using the procedure of Wang and Rhee (2001).

Partial *CCoAOMT* cDNA isolation

Primers for the isolation of partial *CCoAOMT* were designed based on the conserved region in the sequences of the gene from other woody plant species obtained from GenBank. cDNA was synthesized using Ready-To-Go™ You-Prime First-Strand Beads (Amersham Biosciences, UK). The total volume of PCR amplification was 25 μL containing 1 μL 1st strand cDNA, 1X PCR buffer (Invitrogen, USA), 1.5 mM $MgCl_2$, 0.2 mM dNTPs (Promega, USA), 0.2 μM forward primer (5'-CAACGCCAAGAACACCA-3') and reverse primer (5'-GCCGTTCCATAGGGTGT-3') and 1 U *Taq* DNA polymerase (Invitrogen). PCR amplification was carried out for 2 min at 95°C, followed by 35 cycles of 30 s at 94°C, 30 s at 54°C and 30 s at 72°C. After the final cycle, samples were incubated for a further 8 min at 72°C and then held at 4°C. The amplification was done in a GeneAmp PCR System Model 9700 (Perkin-Elmer Co. Ltd., USA). The PCR product was then cloned into pGEM-T Easy Vector (Promega, USA) and sequenced.

RACE-PCR and full length cDNA isolation

Rapid amplification of cDNA ends (RACE) method was used to obtain the 5' and 3' sequences of the *CCoAOMT* from the *Acacia* hybrid using the SMART™ RACE cDNA Amplification kit (CLONTECH Laboratories Inc., USA.). Gene specific primers (GSP) were designed on the basis of the *Acacia* hybrid specific partial *CCoAOMT* gene using the Primer Premier 5 software (PREMIER Biosoft International, Palo Alto, CA, USA). A reverse primer (5'-TGGCGAGCAAGGAGTAGCCAGTGTA-3') and a forward primer (5'-AAGGTTGGGGGTGTGATCGGGTACG-3') were designed on the basis of the partial cDNA of *CCoAOMT* obtained from the *Acacia* hybrid for 5'- and 3'-RACE, respectively. The PCR mixture was in a volume of 50 μL containing 2.5 μL 5'/3'-RACE-Ready cDNA, 0.2 μM GSP, 0.1X UPM (universal primer A mix), 1X PCR reagents, 0.2 mM dNTPs and 1X Advantage 2 polymerase mix. PCR amplification was carried out for 5 cycles of 95°C for 30 s and 72°C for 3 min, followed by 5 cycles of 94°C for 30 s, 70°C for 30 s and 72°C for 3 min, and lastly 25 cycles of 94°C for 30 s, 68°C for 30 s and 72°C for 3 min. The 5'- and 3'- RACE PCR products were then cloned into pGEM-T Easy Vector and sequenced.

To obtain the full length cDNA clone of *CCoAOMT* from the *Acacia* hybrid, primers FLCCoAF (5'-ACACGACGCAGAGTTGAAAAGGTAC-3') and FL-CCoAR (5'-TGAAGGAAATGAACTGGATATGATGAA-3') were designed on the basis of the 5'- and 3'-RACE sequences obtained above. Accordingly, 1 μL 1st strand cDNA was used in PCR amplification containing 1X High Fidelity PCR buffer (Invitrogen), 2.0 mM $MgSO_4$, 0.2 mM dNTPs (Promega), 0.2 μM each primer and 1 U Platinum® *Taq* DNA polymerase high fidelity (Invitrogen). PCR amplification was carried out for 2 min at 95°C, followed by 35 cycles of 30 s at 94°C, 30 s at 60°C, and 1 min at 72°C. The final extension was done at 72°C for 8 min, and the mix then held at 4°C. The PCR product was cloned into pGEM-T Easy Vector, and three positive clones were randomly selected for sequencing.

Tissue expression analysis

Real-time qPCR (RT-qPCR) was carried out to investigate the expression profile of the *AhCCoAOMT* in different tissues (flower, inner bark, leaf, seed pod and root) of the *Acacia*

hybrid. Primers for RT-PCR were designed on the basis of the full length cDNA sequence of *AhCCoAOMT*. The primers were RT-CCoAOMTF (5'-TTGAGAGAGATAACAGCAAAGC A-3') and RT-CCoAOMTR (5'-GGTGGCGAGCAAGGAGTAG-3'). On the other hand, the specific primers for the house-keeping gene (actin gene), ACTF (5'-GGTAACATTGTCCTCTC TGG-3') and ACTR (5'-CATCGTATTCTGCCTTCG-3'), were designed on the basis of the actin gene sequence from the *Acacia* hybrid cDNA library (Wickneswari et al., 2004). The total volume of the PCR reaction was scaled down to 10 μL. The cDNA template was diluted 200-fold before use. For standard curves, inner bark cDNA was used to prepare a 5-fold serial dilution. The PCR amplification employed 1 μL diluted cDNA as template. PCR was performed using 1X iQ™ SYBR® Green Supermix (BioRad, USA) and 0.25 μM each primer. PCR amplification was carried out at 95°C for 3 min, followed by 35 cycles of 95°C for 15 s, 60°C for 30 s and 72°C for 30 s. The same primers and PCR conditions were also employed for the investigation of expression level of the *CCoAOMT* in transformed tobacco plants. All qPCR were done in triplicate.

Characterization of full length genomic *CCoAOMT* (*AhgflCCoAOMT*)

Genomic DNA was extracted from the leaf tissues by using the CTAB method (Murray and Thompson, 1980). *CCoAOMT* full length cDNA of the *Acacia* hybrid (*AhCCoAOMT*, EU275979) was used to develop its full length genomic sequences. This full length cDNA of *CCoAOMT* was aligned with other species by BLASTN and TAIR (The Arabidopsis Information Resource). *Arabidopsis thaliana* chromosome 4 (At4g30450.1) was selected for the alignment as this sequence showed the highest similarity with the *AhCCoAOMT*. The alignment predicted that there were 4 exons in the *Acacia* hybrid. Primers were designed on the basis of the predicted exon regions for the intron amplifications from the genomic sequence of *CCoAOMT* in the *Acacia* hybrid (Table 1). Promoter regions were amplified by using the Genome Walker Universal Kit (Clontech, USA). DNA Walking SpeedUp™ Premix II (Seegene, Korea) was then used to capture the full length of the promoter region. The DNA fragments were then cloned into pGEM-T Easy Vector and sequenced. The Plant Promoter (http://www.softberry.com/berry.phtml) and Neural Network Prediction (http://www.fruitfly.org/seq_tools/promoter.html) software were used to determine the promoter region of the isolated *AhCCoAOMT* genomic sequence.

Table 1. Primers used to amplify *CCoAOMT* for SNP discovery.

Primer	Sequence (5'-3')	Annealing temperature (°C)
First segment		
CCoAOMTi1F	GAAACGAGCGTGTACCCAAG	60
CCoAOMTi3R	AGATGATGCGCCCTTGAA	60
Second segment		
CCoAOMTi3F	ACTGGAAAAGGCTGGTGTG	60
CCoAOMTFLcR	TGAAGGAAATGAACTGGATATGATGAA	60

Southern blot analysis

Five restriction enzymes (*Pvu*II, *Apa*LI, *Eco*RI, *Eco*RV, and *Hin*dIII) were selected from the list of restriction enzymes provided by the Restriction Mapper software (http://www.restric-

tionmapper.org). The full length *AhCCoAOMT* cDNA sequence was used as a probe and labeled with disodium radioisotope [^{32}P] deoxycytidine 5'-triphosphate (dCTP) using the Random Primer DNA Labeling kit (Takara, USA). Genomic DNA samples (20 µg) were digested to completion with the five restriction enzymes separately, and then ethanol-precipitated and resolved on 0.8% agarose gel. Depurinated gels were blotted on Hybond N+ membrane (Amersham, UK). Hybridization and washing were carried out following manufacturer recommendations.

Southern hybridization was also carried out to find out the copy number of the AS and IHP transgene fragment expression cassette in the genome of transgenic tobacco plants. Three plants for each expression cassette that showed positive results by PCR assay were selected for Southern analysis. The [^{32}P]-labeled 720-bp long *npt*II gene fragment probe was used to hybridize with the genomic DNA of the three transgenic tobacco plants. The genomic DNA was digested with *Eco*RI restriction enzyme, which cuts once in the plasmids.

Construction of transformation vectors

Partial *AhCCoAOMT* isolated from the *Acacia* hybrid in sense and antisense orientations were assembled in the suppression vector pHANNIBAL (Wesley et al., 2001). For the construction of AS of coding sequence, a 332-bp sequence was amplified using a forward primer (5'-GACTCTAGACACTGGCTACTCCTTGCT-3') and a reverse primer (5'-TATATCGATTCCATAGGGTGTTGTCGT-3'), containing *Xba*I and *Cla*I sites (underlined), respectively. The resultant PCR fragment was cloned into pHANNIBAL to produce the AS expression cassette. The same 332-bp sequence was again amplified with another pair of primers (forward primer 5'-AATGAATTCCACTGGCTACTCCTTGCT-3' and reverse primer 5'-GAAGGTACCTCCATAGGGTGTTGTCGT-3') containing *Eco*RI and *Kpn*I recognition sites (underlined), respectively. This PCR fragment was then cloned into the AS expression cassette to give rise to the IHP expression cassette. The AS and IHP expression cassettes were then subcloned as *Not*I fragments into the pART27 vector (provided by Dr. G. Allen, CSIRO, Australia).

Plant transformation

Nicotiana tabacum cultivar White Burley was used as a host for plant transformation. Leaves were excised from 7-week-old sterile tobacco plants. Leaf discs were obtained by punching leaves with a paper punch (6 mm in diameter). The leaf discs with the lower epidermis down were put on callus induction medium containing 1X Gamborg's B5 salts and vitamins, 30 g/L sucrose, 8 g/L agar, 0.1 mg/mL IBA and 0.2 mg/mL BAP. The leaf discs were then precultured in a culture room for two days before transformation. Transformation with *Agrobacterium tumefaciens* strain LBA 4404 was done as reported by Horsch et al. (1985). The presence of the expression cassettes in the transgenic lines was confirmed by PCR amplification with a forward primer nptF (5'-GAAGGGACTGGCTGCTATTGGG-3') and a reverse primer nptR (5'-GTAAGTGTTGGATCGGGTGGG-3') annealing to the *npt*II gene.

Histochemical lignin analysis

Stem samples were collected from 3th to 4th internodes and fixed with acetic acid: 95% ethanol (1:3). Stem samples were cross-sectioned 60-100 µm thick using a SM2010 R sliding

microtome (Leica, Germany). For Wiesner reaction, the cross-sectioned stem samples were stained with phloroglucinol-HCl (1% phloroglucinol in 95% ethanol) for 2 min. A few drops of hydrochloric acid were added to the stem samples before observation under a dissection microscope (Leitz Diaplan, USA). For the Maüle test, the method described by Chapple et al. (1992) was applied.

Analytical pyrolysis

The 5th-6th internodes of stems from the selected controls and transformed tobacco lines were used in the pyrolysis analysis following the procedures described in TAPPI T 264 om-88. Pyrolysis of the stem tissue was performed using a Shidmadzu PYR-4A pyrolyzer (Shimadzu, Japan) and the pyrolyzed products were analyzed with a gas chromatograph mass spectrometer GCMSQP5050a (Shimadzu). The sample was pyrolyzed at 550°C for 10 s. Chromatographic separation was performed using capillary column BP1701 (50 m x 0.22 mm, 0.25-µm film, J&W Scientific, USA). For the gas chromatograph, the injector was set at 250°C and detector was set at 280°C. The temperature program was started at 45°C with 4 min isothermal, followed by a heating rate of 4°C/min until 280°C, then hold for 15 min. The compounds were identified by their mass spectra and retention time by comparison with those reported in the literatures (del Río et al., 2005; Ibarra et al., 2005) and in the Wiley and NIST computer libraries. Relative peak molar areas (obtained by dividing the peak area by the molecular weight) were calculated for each lignin pyrolyzed product. The summed molar areas of the relevant peaks were normalized to 100%. Lignin content was calculated as (H+G+S)/(H+G+S+C) (Rodrigues et al., 2001). S/G ratio was calculated by dividing the total percentage of syringyl lignin by the total percentage of guaiacyl lignin.

SNP screening

SNPs in the *AhgflCCoAOMT* sequence were screened across eight different individuals of *A. mangium* and *A. auriculiformis* from different provenances in Australia (**Table S1**) using two sets of primers. Primer set one was CCoAOMTi1F (5'-GAAACGAGCGTGTACC CAAG-3') and CCoAOMTi3R (5'-AGATGATGCGCCCTTGAA-3') whereas primer set two was CCoAOMTi3F (5'-ACTGGAAAAGGCTGGTGTG-3') and CCoAOMTFLcR (5'-TGAAGGAAATGAACTGGATATGATGAA-3'). Only DNA regions from exon 1 to exon 3 and from exon 3 to polyA tail of the *CCoAOMT* were used in the analysis.

RESULTS

Characterization of the full length cDNA of *AhCCoAOMT*

To isolate the partial fragment of *CCoAOMT* from the *Acacia* hybrid, a pair of primers was designed at the conserved regions of this gene obtained from other woody plant species. After cloning, sequencing and blast analysis, we managed to obtain a 317-bp fragment of *CCoAOMT*. On the basis of this short fragment, primers for 5'- and 3'-RACE were designed. The 5'- and 3'-RACE fragments were then manually joined to the partial *CCoAOMT* fragment to obtain the hypothetical full length sequence of this gene. Again, this hypothetical full length sequence was used to design primers that flanked the start and stop codons for the amplification of full length cDNA *CCoAOMT* of the *Acacia* hybrid.

The full-length cDNA of *CCoAOMT* sequence of the *Acacia* hybrid (*Ah*CCoAOMT, EU275979) was 1024-bp long, containing a 750-bp open reading frame which encodes a polypeptide of 249 amino acids. The encoded polypeptide had a predicted molecular weight of 28.18 kDa and an isoelectric point of 5.46. BLAST analysis at the nucleotide level revealed that the *AhCCoAOMT* had 70% similarity to *Leucaena leucocephala CCoAOMT* (gb|DQ431233.2). BLAST output also revealed that AhCCoAOMT and *Gossypium hirsutum* CCoAOMT (gb|ACQ59095.1) had 96% similarity at the protein level. The comparative sequence analyses suggested that the *AhCCoAOMT* is a novel SAM-dependent OMT.

Multiple amino acid sequence alignment of the AhCCoAOMT with other plant CCoAOMT amino acid sequences revealed several regions of high similarity (Figure 1). In all CCoAOMTs, an amino acid sequence motif GXXXGYS was found to be conserved as a SAM-binding domain (Ye et al., 1994). Another three domains that are proposed to be involved in the SAM-binding of OMTs as well as other CCoAOMT-specific sequence motifs (Joshi and Chiang, 1998) were also found in the AhCCoAOMT amino acid sequence, i.e., motif A [L-X-X-L-V-(DK)-V-G-G-X-X-G], motif B [V-(AP)-X-(APG)-D-A-X-X-X-X-(WV)-(VI)] and motif C [A-(LI)-(AP)-X-X-(GP)-(KR)-(VI)-(EI)-(LI)] at positions 178-189, 199-211 and 224-233, respectively (Figure 1). Plant OMTs contain eight conserved motifs (A-H). Motifs A, B and C are common in plant OMTs, while motifs D, E, F, G and H are considered *CCoAOMT* signature sequences (Joshi and Chiang, 1998).

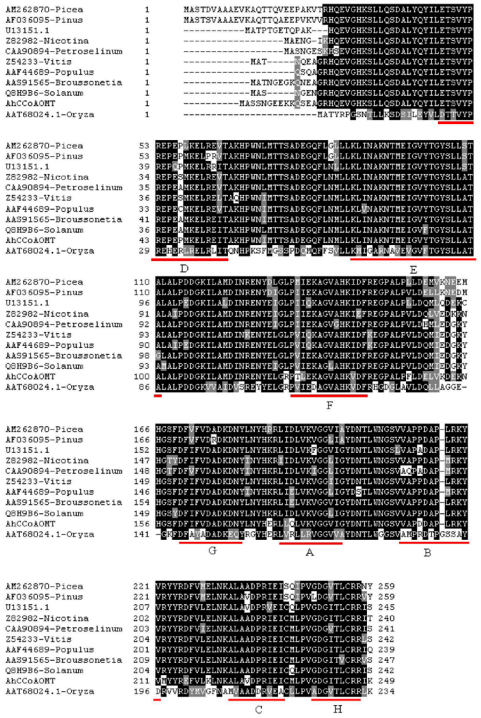

Figure 1. Multiple alignment of the deduced amino acid sequence of AhCCoAOMT with other higher plant CCoAOMT proteins; conserved motifs A-H found in the CCoAOMTs are underlined; the GenBank accession No. AM262870 = *Picea abies*, Z82982 = *Nicotiana tabacum*, Z54233 = *Vitis vinifera*, AF036095 = *Pinus taeda*, U13151 = *Zinnia elegans*, CAA90894 = *Petroselinum crispum*, Q8H9B = *Solanum tuberosum*, AAF44689 = *Populus tomentosa*, AAS91565 = *Broussonetia papyrifera*, AAT68024.1 = *Oryza sativa*, and EU275979 = *AhCCoAOMT.*

Characterization of the full length genomic sequences of *CCoAOMT*

Full length genomic *CCoAOMT* sequence of the *Acacia* hybrid (*AhgflCCoAOMT*) (accession No. JN227537) was 2548-bp long and consisted of three introns: intron 1 (85 bp), intron 2 (956 bp) and intron 3 (92 bp), 4 exons and a promoter of 391 bp (Table 2).

Table 2. Full length genomic *CCoAOMT* sequence of the *Acacia* hybrid (*AhgflCCoAOMT*; GenBank accession No. JN227537).

Region	Predicted region
Promoter	1-391
Exon 1	392-620 (5' UTR 293-438)
Intron 1	621-705
Exon 2	706-844
Intron 2	845-1800
Exon 3	1801-1932
Intron 3	1933-2024
Exon 4	2025-2548 (3' UTR 2316-2548)

*Apa*LI, *Eco*RI, *Eco*RV, and *Pvu*II restriction enzymes (enzymes that do not cut *AhgflCCoAOMT*) and *Hin*dIII restriction enzyme (enzyme that cuts *AhgflCCoAOMT*) were employed to demonstrate the copy number of *AhgflCCoAOMT* present in the *Acacia* hybrid (Figure 2). *Apa*LI generated five bands of 2, 3.5, 4.8, 6 and 10 kb in length; *Eco*RI produced three bands of 4, 5.5 and 6 kb in length; *Eco*RV generated five bands of 2, 3, 4, 5.5 and 10 kb in length; and *Pvu*II produced five bands of 6, 8, 12, 15 and 16 kb in length. On the other hand, *Hin*dIII produced six bands of 3, 3.5, 4, 6, 8, and 9 kb in length. Several weak signals were also detected in this study. This suggests that more than three copies of *AhgflCCoAOMT* might have been present in the *Acacia* hybrid.

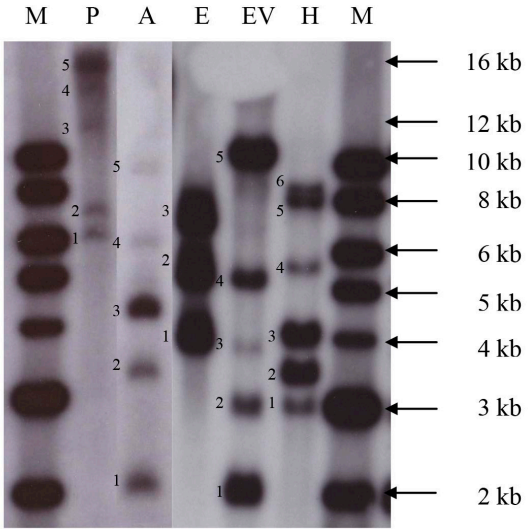

Figure 2. Southern analysis on AhgflCCoAOMT-1 using restriction enzymes *Pvu*II (P), *Apa*LI (A), *Eco*RI (E), *Eco*RV (EV), and *Hin*dIII (H); *lane M* = 2-log marker.

Regulatory element in *AhgflCCoAOMT* gene

Plant Promoter Prediction and Neural Network were used to find regulatory elements in the *AhgflCCoAOMT* (Figure 3). Transcription start site (TSS) was found at 69 bp upstream of the ATG start codon. A TATA box was found in the upstream region of the transcription start site (-25 bp) and A CAAT box was found at -186 bp of the transcription start site. Cis regulatory elements were also detected in the *AhgflCCoAOMT* (Figure 3). The presence of ACCTA elements in AC-I, AC-III, box 1, box 2 and L box promoters suggests their roles in phenylpropanoid pathway genes (Hatton et al., 1995; Wanner et al., 1995). The L box is recognized as an AC-rich box element (Hatton et al., 1995), and it was found at position -137 of the *AhgflCCoAOMT*. The L box is believed to play an important role in increasing the expression of monolignol biosynthesis pathway genes (Hatton et al., 1995; Mohan et al., 2006). An AC-rich element is present in most promoters of lignin biosynthetic genes in *Arabidopsis* and *Populus* (Raes et al., 2003). An MYB binding motif was present at -131 (CACGAACC) in the *AhgflCCoAOMT*. Sablowski et al. (1994) reported that a P box is conserved in the promoter phenylalanine ammonia-lyase (*PAL*), 4-coumarate:CoA ligase (4CL) and several flavonoid biosynthetic genes. We detected the presence of a G box (CCAGTG) at position -229 in the *AhgflCCoAOMT* promoter.

GAGTTTAGGTCCAGCGTCCGTGGGGGGGGTACGCATGTTGGGACCCACCCTTAACAATA

CAGTATTGGCTTCAGGAAAAATATAACAGTATTTTTAAAAAATTAAAAATTACAAGCCA

GCATACTTTGCTGGACCGCCAGTGGAAGGAACAATAATTGCGCCTTATCACTGAGCCCA

CCATCTGATTGCTGGCAAAATTTATTTTTTAAAAAAAATTATTATTTTATAAGAAAAAT

TAGACTGGTCCTACCCAACTGGAGTGGAGCGTGGGACCAGGTTCACGCATCCGAAGGAG

TTCACCTACCACGAACCGGTTCAACCCGGACTTGGCCTGTCTTATTTATAAGGTCATCT

TCATTGATCTTC**A**TCTCCAAGCTCAGCTTCAACCATCACACGACGCAGAGTTGAAAAGG

TACAAGTTTGAGGGGAGTATTA**ATG**

Figure 3. Regulatory element in the promoter region of CCoAOMT. Start codon ATG is in bold. TATA box is indicated in box; cis regulatory elements AC (AC/L), P box (P), G box (G), and E box (E) are underlined.

Tissue expression pattern analysis

We conducted a gene expression study of *AhCCoAOMT* in five different tissues (flower, leaf, inner bark, root and seed pod) using real-time qPCR to investigate the tissue-specific expression pattern in the *Acacia* hybrid. *AhCCoAOMT* was expressed in all the tissues tested (Figure 4). The expression of *AhCCoAOMT* in all tissues was compared with expression in the calibrator (inner bark tissue). The highest level of the *AhCCoAOMT* transcript was found in root tissue, which was 4.25-fold higher compared with inner bark, and the transcript level in leaf tissue was slightly higher compared with inner bark tissue. The lowest expression level of this gene was detected in the seed pod, which was 1.42-fold lower than in the calibrator

(Figure 4). The real-time qPCR results revealed that this gene was also highly expressed in the flower (two-fold higher than in inner bark), suggesting that the *AhCCoAOMT* may be involved in the synthesis of uncharacterized flower-specific compounds (Martz et al., 1998) apart from lignin biosynthesis in the *Acacia* hybrid.

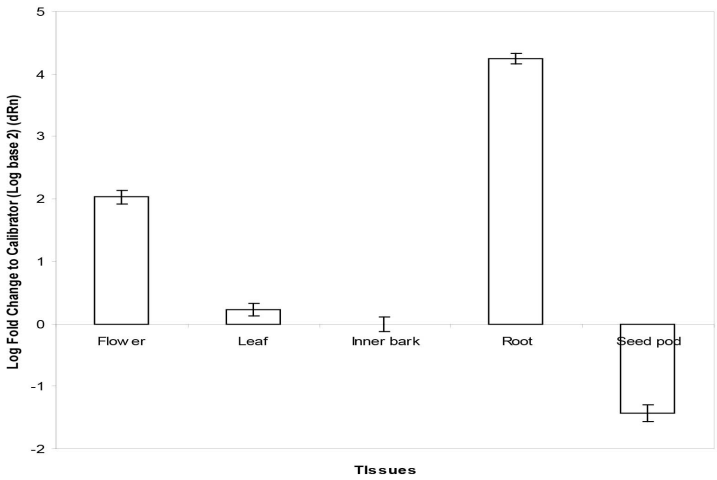

Figure 4. Relative quantity chart for AhCCoAOMT transcript levels in different tissues of the *Acacia* hybrid.

Donwregulation of tobacco homologous *CCoAOMT* gene

To gain insight into the function of *AhCCoAOMT* in lignin biosynthesis, the tobacco homologous *CCoAOMT* was downregulated using the partial fragment of *AhCCoAOMT* by AS and IHP silencing strategies (Figure 5). Shoot regeneration occurred within two to five weeks. After four weeks, shoots with stem were carefully excised from calli and transferred to rooting medium. PCR was carried out using nptF and nptR primers to detect the presence of transgene in the tobacco plant genomes. Plants shown to be positive by PCR assay were transferred to soil. Three plants for each transgenic line and control were selected for Southern hybridization and real-time qPCR to determine the expression of *CCoAOMT*.

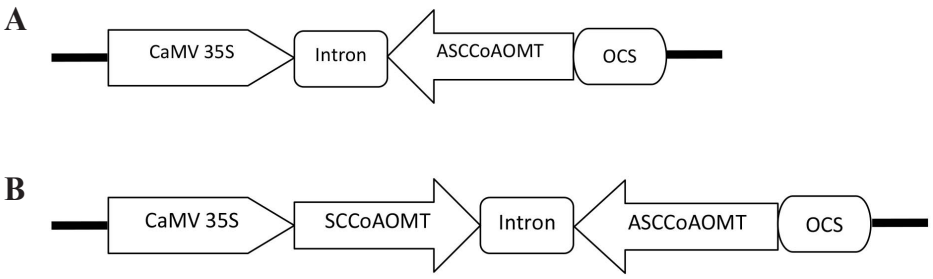

Figure 5. Expression cassettes of transformation vectors constructed using partial fragment of AhCCoAOMT. A = Antisense (AS) construct and B = Intron-containing hairpin (IHP) construct; CaMV 35S = cauliflower mosaic virus 35S promoter; ASCCoAOMT = antisense *CCoAOMT*; SCCoAOMT = sense *CCoAOMT*; OCS = octopine synthase terminator; diagram is not drawn to scale.

Southern hybridization was done on the negative control lines 1-3, pART-PO lines T9, T10 T11, pART-HAS78 lines T3, T4, T5 and pART-HIHP78 lines T4, T5, T6. Figure 6 shows the autoradiograms of the Southern hybridization. The autoradiograms depicted the different patterns of signals for the transgenic plants. No hybridization signal was observed in genomic DNA isolated from the negative controls. In pART-HAS78 line T4 and pART-HIHP78 line T4, two signals were prominent which reflected the integration of two copies of the transgene in the genomes of the transgenic plants. For the rest of the plants, only one single prominent signal was observed, which showed the presence of one transgene in the genomes of the transgenic plants. These results confirmed the integration of the constructed expression cassettes in the genomes of the transgenic tobacco plants.

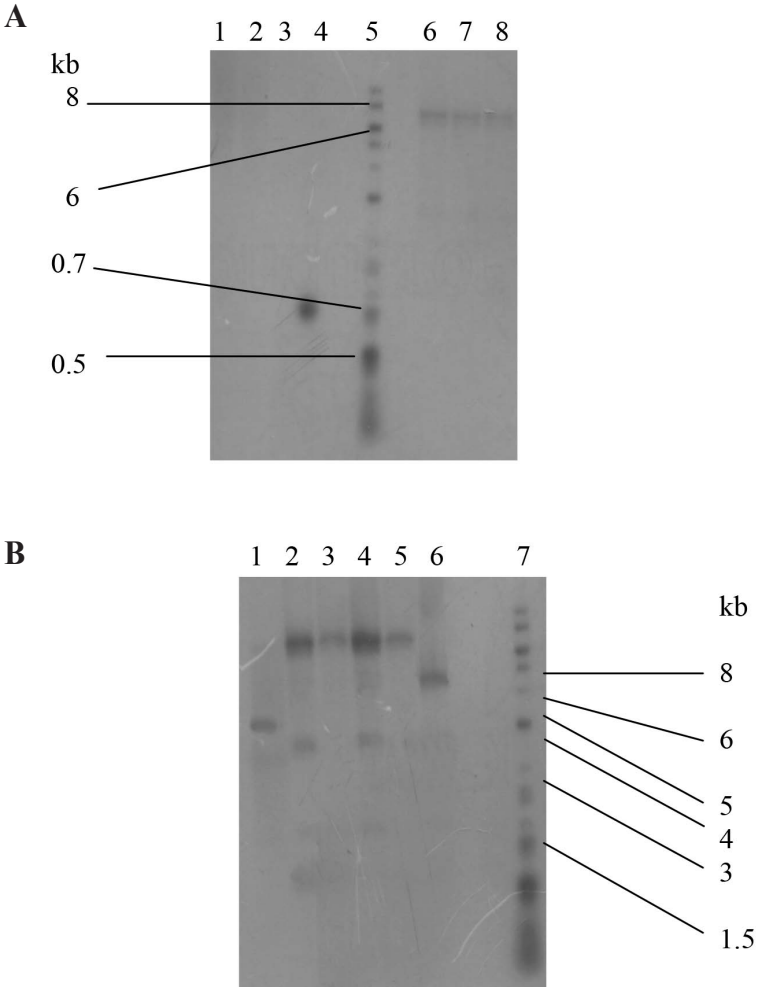

Figure 6. Autoradiograms of Southern hybridization; the different patterns of signal suggest the multiple integrations of the expression cassettes in the plant genome; A) *Lanes 1-3* = negative control lines 1-3; *lane 4* = nptII gene fragment; *lane 5* = 2-log DNA ladder (New England, USA); *lanes 6-8* = pART-PO lines T9-T11. B) *lanes 1-3* = pART-HAS78 lines T3-T5; *lanes 4-6* = pART-HIHP78 lines T4-T6; *lane 7* = 2-log DNA ladder (New England, USA).

Expression of *CCoAOMT* in the negative and positive controls as well as the transgenic plants carrying pART-HAS78 and pART-HIHP78 is shown in Figure 7. The expression levels of *CCoAOMT* in the negative control plants ranged 1.41-1.95, while in the positive control plants, the expression levels ranged at 1.35-1.70. The calculated average expression of *CCoAOMT* in the controls was 1.55. Generally, the level of *CCoAOMT* transcript was lower than in the controls in all the transgenic plants carrying the expression cassette of pART-HAS78 or pART-HIHP78. The transformed plants with AS construct had expression levels of 1.23, 0.83 and 1.26 for transgenic lines T3, T4 and T5, respectively, which represented 79.4, 53.5 and 81.3% of the average expression in the control plants. For the plants transformed with IHP construct, the expression levels of *CCoAOMT* were 0.78, 1.23 and 0.49 for transgenic lines T4, T5 and T6, respectively. These expression levels represented 50.3, 79.4 and 31.6% of the average expression in the control plants.

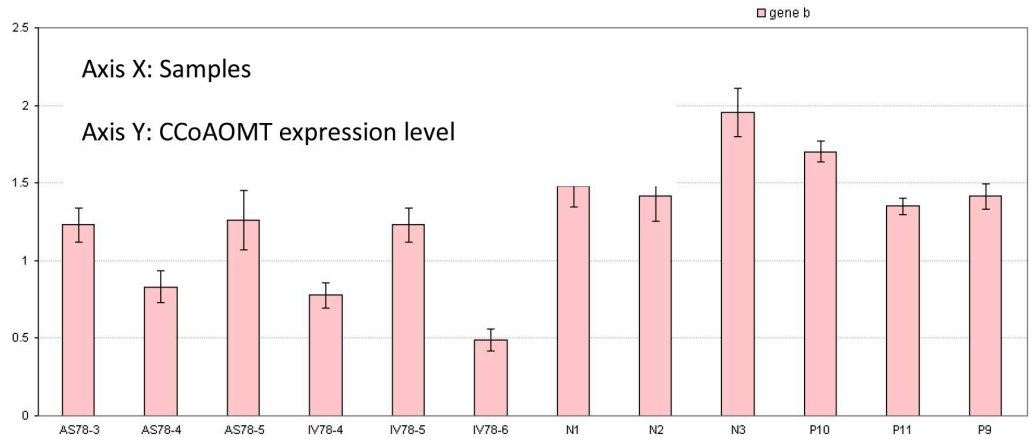

Figure 7. Expression levels of *CCoAOMT* in selected transgenic tobacco lines and controls. AS78-3-5 = pART-HAS78 lines T3-T5; IV78-4-6 = pART-HIHP78 lines T4-6; N1-3 = negative controls lines 1-3; P9-11 = pART-PO lines T9-11.

Histochemical lignin analysis

To investigate potential changes in lignin chemistry accompanying *CCoAOMT* down-regulation, stem samples were collected from the 3rd and 4th internodes from the same tobacco transgenic lines and control plants as those selected for Southern hybridization and RT-qPCR. The intensity of phloroglucinol staining is often taken as a reflection of lignin content, although the reagent appears most specific for conideraldehyde end groups in lignin (Lewis and Yamamoto, 1990).

Using phloroglucinol, the xylem in both the negative and positive controls was stained strongly (Figure 8). In contrast, staining of the pART-HAS78 and pART-HIHP78 stem sections gave a weak positive reaction restricted to the tracheids in the primary xylem and isolated vessel elements in the secondary xylem. The reduction in the intensity of the phloroglucinol stain suggests that the transgenic plants had less lignin content or fewer lignins with aldehyde groups. pART-HIHP78 line T6 showed strong reduction in total lignin content. Syringyl lignin

can be distinguished from guaiacyl lignin *in situ* by the Maüle reagent (Nakano and Meshitsuka, 1992). The chlorination of syringyl unit leads to a wine red color, whereas the guaiacyl units produce a light to dark-brown color (Nakano and Meshitsuka, 1992). The different staining intensity between the controls and the transgenic plants suggests some changes in lignin composition. The controls were more intensely stained wine-red in color than were all transgenic tobacco plants (Figure 9). This appears to be mostly due to the lower lignin content in the transgenic plants. In most of the transgenic plants, larger vessels appeared mostly brown in color, indicating that vessels consisted of mainly G units and the smaller vessels contained more S units.

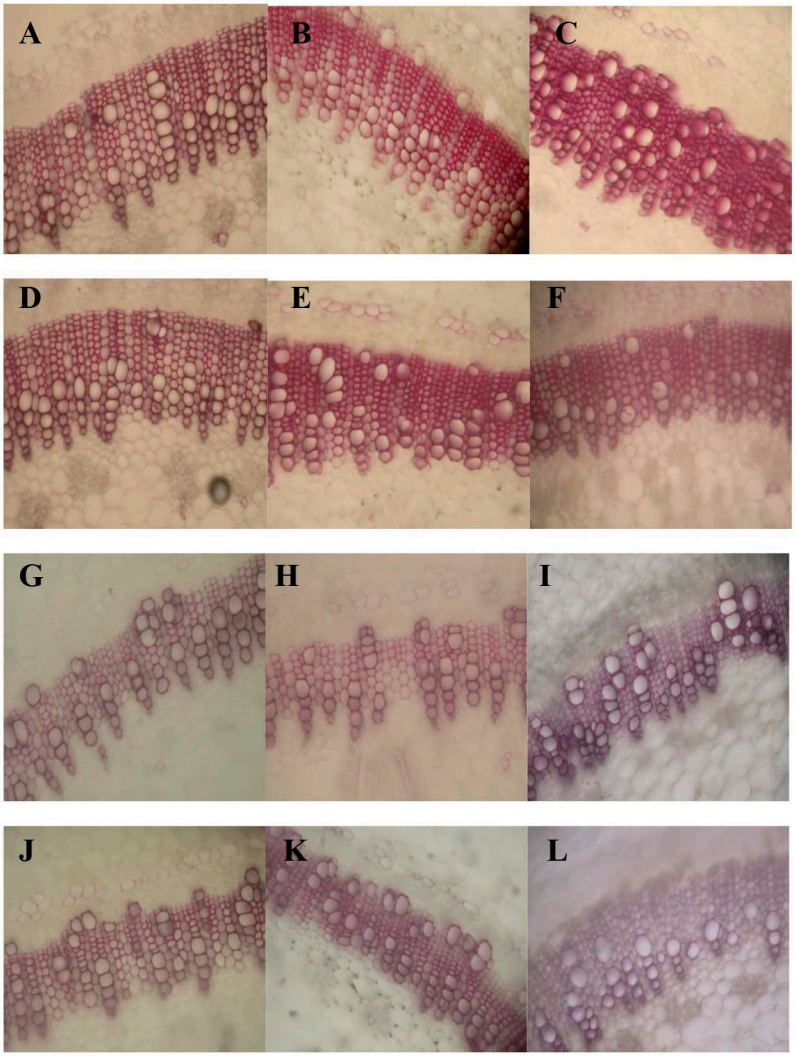

Figure 8. Histochemical analysis of lignin in stem cross-sections using Wiesner reagent. A-C = negative control lines 1-3; D-F = pART-PO lines T9-11; G-I = pART-HAS78 lines T3-5; J-L = pART-HIHP78 lines T4-6.

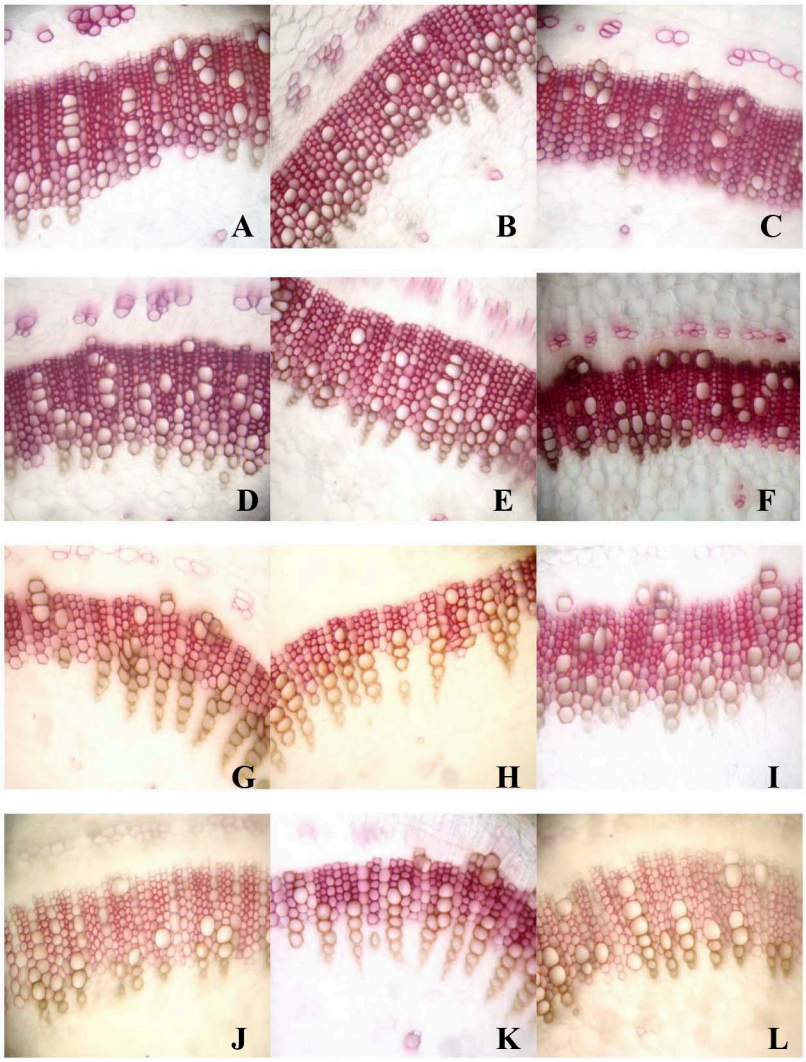

Figure 9. Histochemical analysis of lignin in stem cross-sections using Maüle reagent. A-C = negative control lines 1-3; D-F = pART-PO lines T9-11; G-I = pART-HAS78 lines T3-5; J-L = pART-HIHP78 lines T4-6.

Analytical pyrolysis

On the basis of the expression levels of *CCoAOMT* and Southern analysis of the T0 tobacco transformed lines, one plant for each construct was selected for further analysis using pyrolysis-GC/MS. The plants selected were the negative control line 3, positive control pARTPO line T10, pART-HAS78 line T3 and pART-HIHP78 line T6. These selected transgenic tobacco plants showed repression in *CCoAOMT* expression by using real-time qPCR analysis and carried a single copy of the respective gene construct as shown by Southern hy-

bridization. The important degradation products were classified as carbohydrate-(C), guaiacyl lignin-(G), syringyl lignin-(S) and ρ-hydroxyphenyl lignin (H)-derived products (**Table S2**).

Pyrolysis GC/MS was performed in triplicate for all the plants selected. According to the pyrolysis-GC/MS results (Table 3), lignin contents and S/G ratios in the transgenic tobacco plants were generally lower than in both control plants. By combining pyrolysis data with the real-time qPCR data, the expression level of *CCOAOMT* in pART-HAS78 line T3 was estimated to be reduced by 37% compared to the average expression of the controls, and this lower expression reduced the estimated lignin content by 24% relative to the controls. This lignin reduction was also detected by phloroglucinol staining, as less intense red color was observed in the transgenic plants.

The estimated S unit as well as the S/G ratio was also reduced. On the other hand, the downregulation of the *CCoAOMT* expression level by 75% in pART-HIHP78 line T6 caused a strong decrease in the estimated lignin content by 56% compared to the negative control. The lignin reduction was also manifested in the phloroglucinol staining, where only a few vessels gave a weak positive reaction. This repression resulted in a total loss of S units in the lignin. The estimated area for G unit in this plant was only 0.80% (representing 13.8% of the lignin content) compared to 8.55% (representing 65% in the lignin content) in the negative control plants.

Single nucleotide polymorphisms

The *AhgflCCoAOMT* gene sequence of 2.4 kb in size was used in the development of the SNPs. Polymorphisms were detected in *AhgflCCoAOMT* gene sequence across eight individuals of *A. auriculiformis* and *A. mangium*. A total of 47 SNPs and seven insertions/deletions (INDEL) were detected in *A. auriculiformis* (**Table S3**). On the other hand, 30 SNPs and one INDEL were detected in *A. mangium* (**Table S4**). Six non-synonymous mutations were detected in the *A. auriculiformis CCoAOMT* sequences (**Table S5**), while the *A. mangium CCoAOMT* sequences had seven non-synonymous mutations (**Table S6**). A total of five synonymous mutations and four synonymous mutations were found in each *A. auriculiformis* and *A. mangium CCoAOMT*. High level of synonymous and non-synonymous mutations can be developed as SNP markers for quantitative trait loci (QTL) detection of the variation in lignin content in different genotypes (Nur Fariza et al., 2008).

DISCUSSION

The full length cDNA *AhCCoAOMT* isolated from the *Acacia* hybrid was 1024-bp long encoding a polypeptide of 249 amino acids. Amino acid sequence analysis of the AhCCoAOMT revealed that this gene is closely related to the CCoAOMT of *L. leucocephala, Populus tomentosa*, *Populus tremuloides*, and *Eucalyptus gunnii*. The amino acid sequence of *AhCCoAOMT* also exhibits the important motifs as that of the *CCoAOMT* protein of other plant species. These analyses suggest the successful isolation of a putative *CCoAOMT* from the *Acacia* hybrid. Since the plant material used in this study was a hybrid, two allelic forms of *CCoAOMT* could be present. However, we did not detect any sequence variation between three randomly selected clones. Our *de novo* transcriptome sequencing also revealed a high level of homology between *A. mangium* and *A. auriculiformis* (Wong et al., 2011).

CCoAOMT is shown to be temporally and spatially regulated during lignification in different cell types (Ye et al., 1994). This gene was found to be highly expressed in lignifying tissues, especially in xylem ray parenchyma (Zhong et al., 2000). In *Zinnia*, *CCoAOMT* was barely detected in leaves, but reported to accumulate to high level in stems, roots and flower buds, and markedly increased in mesophyll cells undergoing lignifications (Ye et al., 1994). *CCoAOMT* has also been reported to be involved in plant defense reactions by synthesizing wall-bound forms of ferulic acid.

To study the function of *AhCCoAOMT*, a partial fragment of this gene was employed in the AS and IHP constructs to downregulate the tobacco homologous *CCoAOMT* gene. Downregulation of *CCoAOMT* by 37% in tobacco using the AS construct resulted in a 24% reduction in lignin content in stems. In pART-HAS78 line T3 carrying the AS constructs, the reduction in lignin was accompanied by a decrease in S units, resulting in a decrease in S/G ratio. However, the lignin of this transgenic line consisted of 64% G units, which was not much different compared with the negative and positive controls, of which the lignin contained 65 and 67% G units, respectively. A slight increase in H units was detected in this transgenic line.

A severe reduction in lignin content (56%) was also observed in the tobacco plants transformed with the IHP construct. For pART-HIHP78 line T6 carrying the IHP construct, the expression of *CCoAOMT* was downregulated by 75% compared to the controls. This reduction in lignin (56%) was accompanied by a marked decrease in both S and G units, and resulted in a substantial enrichment of H units. The lignin of this transgenic line contained 86% H units and 14% G units, and no S units were detected. These observations suggest that the pathway leading to the synthesis of G and S units was severely affected in pART-HIHP78 line T6, and apparently the cells affected were unable to compensate for the reduction in G and S units by increasing the amount of H units. In *Arabidopsis*, *ccomt* 1 knockout mutants show a reduction in G unit accompanied by enrichment in H and S units (Do et al., 2007). However, in pART-HIHP78 line T6, Maüle staining indicated the presence of at least low levels of S units. This might have been due to the difference in the plant internodes/segments used for analytical pyrolysis and histochemical analysis. Maüle staining and acetyl bromide-soluble lignin content of the downregulated OMT and CCR in transgenic perennial ryegrass plants showed different lignin content and S/G ratio in different internodes of the same transgenic plant (Tu et al., 2010).

The pyrolysis GC/MS analysis of the transgenic tobacco plants demonstrated that S units were preferentially decreased over G units, which resulted in a decrease in the S/G ratio. *CCoAOMT* is reported to be involved in the methylation of caffeoyl-CoA and 5-hydroxyferuloyl-CoA to form feruloyl-CoA and sinapoyl-CoA, which are the precursors for G and S units, respectively. Until lately, however, downregulation of this gene in most plant species reported stronger depletion of G units than S units. Apart from that, the *in vivo* ability of *CCoAOMT* to methylate 5-hydroxyferuloyl-CoA is still uncertain (Atanassova et al., 1995). On the other hand, *COMT* is able to methylate hydroxycinnamic acids. These lead to the commonly accepted fact that *COMT* is responsible for 5' methylation reactions, whereas *CCoAOMT* is responsible for 3' methylation reactions. However, the recent research on *Linum usitatissimum* showed that *CCoAOMT* is able to 5'-methylate 5-hydroxyconiferaldehyde to sinapaldehyde, a reaction leading to S unit formation (Day et al., 2009). This 5'-methylation ability of *CCoAOMT* could explain the finding obtained in this study.

In transgenic alfalfa, downregulation of *CCoAOMT* leads to reduced lignin content, a reduction in G units without reduction in S units (Pincon et al., 2001; Guo et al., 2001; Chen et al., 2006). These findings demonstrate the importance of *CCoAOMT* in the methylation at

the C3 position of the aromatic ring, which leads to the formation of G units. Guo et al. (2001) proposed that the formation of S unit was independent of G units. Besides, downregulation of this gene in poplar (Meyermans et al., 2000; Zhong et al., 2000) and tobacco (Zhong et al., 1998) also suggested the involvement of this gene in the biosynthesis of S units, since both G and S units were reduced. Similar research was also conducted in *L. usitatissimum* where the S/G ratio was shown to increase in one transformed line whereas another line showed a slight decrease (Day et al., 2009). These results again suggested the importance of *CCoAOMT* in S unit formation.

Substrate affinity analysis has revealed that *CCoAOMT* shows higher methylation rate with caffeoyl-CoA than 5-hydroxyferuloyl-CoA (Ye et al., 1994; Busam et al., 1997; Martz et al., 1998; Zhong et al., 2000). *CCoAOMT* is reported to exist in a multigene family that can be grouped into class I and II (Raes et al., 2003). In *Arabidopsis*, there is apparently only one member found in class I *CCoAOMT,* and other *CCoAOMT*-like genes are grouped into class II (Do et al., 2007). However, it has been reported in tobacco that *CCoAOMTs* are encoded by at least six isoforms which can be grouped into three classes, i.e. class I, II and III (Martz et al., 1998; Maury et al., 1999). Substrate affinity analyses of the tobacco *CCoAOMTs* revealed that class III enzymes appear to accept almost exclusively caffeoyl-CoA as a substrate. Most of the class I enzymes are able to methylate caffeoyl-CoA more efficiently than 5-hydroxyferuloyl-CoA, whereas class II enzymes show higher affinity toward 5-hydroxyferuloyl-CoA compared to caffeoyl-CoA (Maury et al., 1999). The results of our study suggest that the *AhCCoAOMT* most likely belongs to class II *CCoAOMT* of tobacco because its application in downregulation resulted in increased depletion of S units rather than G units.

In tobacco, transcript *CCoAOMT2* and another three *CCoAOMT* transcripts were detected in healthy leaf, stem, root and some flower parts, e.g., petals, pistils and stamens, with transcript gene 1 being detected in leaf with TMV infection or elicitor treatment, whereas transcript gene 3 being found highly accumulated in stem, root and flower tissues of healthy tobacco (Martz et al., 1998). This might suggest that the different transcripts of *CCoAOMT* in tobacco display differential expression. Southern analysis of the *AhCCoAOMT* suggested that this gene was organized in a small multigene family in the genome of the *Acacia* hybrid (Figure 2).

SNP discovery revealed that *CCoAOMT* in *A. auriculiformis* has higher variation than *CCoAOMT* in *A. mangium*. The average SNP in the *CCoAOMT* in *A. mangium* was one in every 62 bp whereas one in every 42 bp for *CCoAOMT* in *A. auriculiformis*. This result was comparable to the frequency of SNP in *CCoAOMT* from *Eucalyptus nitensis*, which was one in every 59 bp (Zhang and Zhang, 2005). Synonymous and non-synonymous mutations may result in the differential expression of certain transcripts, which may play different roles during lignin biosynthesis and indirectly affecting the S/G units. In addition, high level of polymorphisms detected in *CCoAOMT* sequence from *A. mangium* and *A. auriculiformis* individuals may play important regulatory roles during lignin biosynthesis. Although polymorphism in the pathway genes involved in lignifications is very important, limited data on the effects of polymorphism of *CCoAOMT* gene on lignin content and composition has limited our understanding of their regulatory roles. In *Populus trichocarpa*, 13 differences in amino acid sequences have been reported, which may be responsible for the difference in the enzyme-substrate interactions of *CCoAOMT1* and *CCoAOMT2*, where *CCoAOMT2* shows stronger interactions with substrates and products (Phogat et al., 2010). Amino acid differences were also found in *CCoAOMT1* and *CCoAOMT2* of *L. leucocephala*, and docking analysis indicated that caffeoyl CoA binds with higher affinity to *CCoAOMT2* than *CCoAOMT1* (Pagadala et al., 2009).

Since the *CCoAOMT* gene plays an important role in regulating lignin content and composition, the presence of various copies of this gene as revealed in this study and the high level of polymorphisms detected across various different genotypes suggest that the *CCoAOMT* gene may be a potential candidate for improvement of wood pulp quality in the *Acacia* hybrid. The presence of multiple regulatory element sites as revealed in this study strengthens its regulatory roles.

ACKNOWLEDGMENTS

Research supported by the IRPA program (Grant #01-02-02-0010-PR0003/03-01) from the Ministry of Science, Technology and Innovation, Malaysia. We thank the Commonwealth Scientific and Industrial Research Organization (CSIRO), Plant Industry, Australia for providing the primary and binary cloning vectors (pHANNIBAL and pART27).

Supplementary material

REFERENCES

Atanassova R, Favet N, Martz F, Chabbert B, et al. (1995). Altered lignin composition in transgenic tobacco expressing O-methyltransferase sequences in sense and antisense orientation. *Plant J.* 8: 465-477.

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

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

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

Chen F, Srinivasa Reddy MS, Temple S, Jackson L, et al. (2006). Multi-site genetic modulation of monolignol biosynthesis suggests new routes for formation of syringyl lignin and wall-bound ferulic acid in alfalfa (*Medicago sativa* L.). *Plant J.* 48: 113-124.

Darus A and Rasip AG (1989). A note on *Acacia* hybrids in a forest plantation in Peninsular Malaysia. *J. Trop. For. Sci.* 2: 170-171.

Day A, Neutelings G, Nolin F, Grec S, et al. (2009). Caffeoyl coenzyme A O-methyltransferase downregulation is associated with modifications in lignin and cell-wall architecture in flax secondary xylem. *Plant Physiol. Biochem.* 47: 9-19.

del Río JC, Gutiérrez A, Hernando M, Landín, et al. (2005). Determining the influence of eucalypt lignin composition in paper pulp yield using Py- GC/MS. *J. Anal. Appl. Pyrol.* 74: 110-115.

Do CT, Pollet B, Thevenin J, Sibout R, et al. (2007). Both caffeoyl Coenzyme A 3-O-methyltransferase 1 and caffeic acid O-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in *Arabidopsis*. *Planta* 226: 1117-1129.

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

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

Guo M, Rupe MA, Zinselmeier C, Habben J, et al. (2004). Allelic variation of gene expression in maize hybrids. *Plant Cell* 16: 1707-1716.

Hatton D, Sablowski R, Yung MH, Smith C, et al. (1995). Two classes of cis sequences contribute to tissue-specific expression of a PAL2 promoter in transgenic tobacco. *Plant J.* 7: 859-876.

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, et al. (1985). A simple and general method for transferring genes into plants. *Science* 227: 1229-1231.

Ibarra D, del Río JC, Gutiérrez A, Rodríguez IM, et al. (2005). Chemical characterization of residual lignins from eucalypt paper pulps. *J. Anal. Appl. Pyrol.* 74: 116-122.

Inoue K, Sewalt VJ, Murray GB, Ni W, et al. (1998). Developmental expression and substrate specificities of alfalfa caffeic acid 3-O-methyltransferase and caffeoyl coenzyme A 3-O-methyltransferase in relation to lignification. *Plant Physiol.* 117: 761-770.

Joshi CP and Chiang VL (1998). Conserved sequence motifs in plant S-adenosyl-L-methionine-dependent methyltransferases. *Plant Mol. Biol.* 37: 663-674.

Kha LD (1996). Studies on Natural Hybrids of *Acacia mangium* and *A. auriculiformis* in Vietnam. In: Tree Improvement for Sustainable Tropical Forestry. Proceedings QFRI-IUFRO Conference vol. 2. (Dieters MJ, Matheson AC, Nikles DG, Harwood CE, et al., ed.). Queenland Forestry Research Institute, Gympie, 328-332.

Kneusel RE, Matern U and Nicolay K (1989). Formation of trans-caffeoyl-CoA from trans-4-coumaroyl-CoA by Zn2+-dependent enzymes in cultured plant cells and its activation by an elicitor-induced pH shift. *Arch. Biochem. Biophys.* 269: 455-462.

Lewis NG and Yamamoto E (1990). Lignin: occurrence, biogenesis and biodegradation. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 411: 455-496.

Lewis NG and Davin LB (1999). Lignans: Biosynthesis and Function. In: Comprehensive Natural Products Chemistry. (Barton DHR NKM-COe, ed.). Elsevier, London, 639-712.

Mackay TF (2001). Quantitative trait loci in *Drosophila*. *Nat. Rev. Genet.* 2: 11-20.

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

Maury S, Geoffroy P and Legrand M (1999). Tobacco O-methyltransferases involved in phenylpropanoid metabolism. The different caffeoyl-coenzyme A/5-hydroxyferuloyl-coenzyme A 3/5-O-methyltransferase and caffeic acid/5-hydroxyferulic acid 3/5-O-methyltransferase classes have distinct substrate specificities and expression patterns. *Plant Physiol.* 121: 215-224.

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

Meyermans H, Morreel K, Lapierre C, Pollet B, et al. (2000). Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon downregulation of caffeoyl-coenzyme A O-methyltransferase, an enzyme involved in lignin biosynthesis. *J. Biol. Chem.* 275: 36899-36909.

Mohan D, Pittman Jr CU and Steele PH (2006). Pyrolysis of wood/biomass for bio-oil: A critical review. Energy Fuels. *Energy Fuels* 20: 848-889.

Murray MG and Thompson WF (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8: 4321-4325.

Nakano J and Meshitsuka G (1992). The Detection of Lignin. In: Methods in Lignin Chemistry. (Dence CW and Lin SY, eds.). Springer-Verlag, New York, 23-31.

Nur Fariza MS, Pang SL, Choong CY and Wickneswari R (2008). Extensive DNA sequence variations in two Lignin Genes, Cinnamate 4-Hydroxylase and Cinnamyl alcohol Dehydrogenase from *Acacia mangium* and *Acacia auriculiformis*. *J. Biol. Sci.* 8: 687-690.

Phogat N, Vindal V, Kumar V, Inampudi KK, et al. (2010). Sequence analysis, in silico modeling and docking studies of caffeoyl CoA-O-methyltransferase of *Populus trichopora*. *J. Mol. Model.* 16: 1461-1471.

Pincon G, Maury S, Hoffmann L, Geoffroy P, et al. (2001). Repression of O-methyltransferase genes in transgenic tobacco affects lignin synthesis and plant growth. *Phytochemistry* 57: 1167-1176.

Polakova KM, Kucera L, Laurie DA, Vaculova K, et al. (2005). Coding region single nucleotide polymorphism in the barley low-pI, alpha-amylase gene Amy32b. *Theor. Appl. Genet.* 110: 1499-1504.

Raes J, Rohde A, Christensen JH, Van de Peer Y, et al. (2003). Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiol.* 133: 1051-1071.

Rodrigues J, Graca J and Pereira H (2001). Influence of tree eccentric growth on syringyl/guaiacyl ratio in *Eucalyptus globulus* wood lignin assessed by analytical pyrolyisis. *J. Anal. Appl. Pyrol.* 58: 481-489.

Sablowski RW, Moyano E, Culianez-Macia FA, Schuch W, et al. (1994). A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. *EMBO J.* 13: 128-137.

Pagadala NS, Arha M, Reddy PS, Kumar R, et al. (2009). Phylogenetic analysis, homology modelling, molecular dynamics and docking studies of caffeoyl-CoA-O- methyl transferase (CCoAOMT 1 and 2) isoforms isolated from subabul (*Leucaena leucocephala*). *J. Mol. Model.* 15: 203-221.

Tu Y, Rochfort S, Liu Z, Ran Y, et al. (2010). Functional analyses of caffeic acid O-Methyltransferase and Cinnamoyl-CoA-reductase genes from perennial ryegrass (*Lolium perenne*). *Plant Cell* 22: 3357-3373.

Wang MH and Rhee HI (2001). Improved technique for isolating RNA from tobacco tissues. *Plant Mol. Biol. Rep.* 19: 187-188.

Wanner LA, Li G, Ware D, Somssich IE, et al. (1995). The phenylalanine ammonia-lyase gene family in *Arabidopsis thaliana*. *Plant Mol. Biol.* 27: 327-338.

Wesley SV, Helliwell CA, Smith NA, Wang MB, et al. (2001). Construct design for efficient, effective and high-throughput gene silencing in plants. *Plant J.* 27: 581-590.

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

Wickneswari R, Cheong PL, Pang SL, Choong CY, et al. (2004). Analysis of a Batch of Expressed Sequence Tags From Inner Bark cDNA Library of A. *Mangium* x A. *Auriculiformis* Interspecific Hybrid (in Malay). In Proceedings of the 7th National Biology Symposium, Highlands, 477-480.

Wong MM, Cannon CH and Wickneswari R (2011). Identification of lignin genes and regulatory sequences involved in secondary cell wall formation in *Acacia auriculiformis* and *Acacia mangium* via de novo transcriptome sequencing. *BMC Genomics* 12: 342.

Yamada N, Khoo KC and Mohd MY (1990). Sulphate pulping characteristics of *Acacia* hybrid, *Acacia mangium* and *Acacia auriculiformis* from Sabah. *J. Trop. Forest Sci.* 4: 206-214.

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

Zhang DQ and Zhang ZY (2005). Single nucleotide polymorphisms (SNPs) discovery and linkage disequilibrium (LD) in forest trees. *Forest Stud. China* 7: 1-14.

Zhong R, Morisson WH, Negrel J and Ye ZH (1998). Dual methylation pathways in lignin biosynthesis. *Plant Cell* 10: 2033-2046.

Zhong R, Morrison WH, Himmelsbach DS, Poole FL, et al. (2000). Essential role of caffeoyl coenzyme A O-methyltransferase in lignin biosynthesis in woody poplar plants. *Plant Physiol.* 124: 563-578.