Malays. Appl. Biol. (2020) 49(1): 153–158

CHARACTERISATION OF *AmCCR1* AND *pseudoCOMTAm* PROMOTERS IN *Acacia mangium*

AZMAH ABDUL LATIF[1], SHEH MAY TAM[2] and WICKNESWARI RATNAM[1*]

[1]*School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43650 Bangi Selangor, Malaysia*
[2]*School of Biosciences, Taylor's University*
[*]*E-mail: wicki@ukm.edu.my*

Accepted 13 June 2020, Published online 30 June 2020

ABSTRACT

Acacia mangium is one of the most important species in forest plantation industry in Southeast Asia. Lignin is a complex polymer found in plant which has to be removed using harsh and toxic chemicals for the production of high quality paper. Understanding the complex pathway that underlies the regulation of lignin biosynthetic genes requires in depth knowledge of the genes involved and it's regulatory elements. Using Thermal Asymmetric Interlaced PCR, a 770bp promoter of *pseudoCOMTAm* and 619bp of *AmCCR1* promoter were isolated. Bioinformatics analysis revealed the presence of cis acting elements commonly found in lignin biosynthetic genes such as TATA box, CAAT box, W box, AC-I and AC-11 elements on both promoters.

Key words: *Acacia mangium, AmCCR1* promoter, *pseudoCOMTAm* promoter, sequence analysis

INTRODUCTION

Lignin is a complex aromatic polymer mainly found in plant secondary cell wall (Boerjan *et al*., 2003) which plays an important role in plant development (Jones *et al.,* 2001) and defence mechanism (Lauvergeat *et al.,* 2001). Being the second most abundant biopolymer after cellulose, it is a major challenge to the paper and pulp industry because of the recalcitrance to efficient pulp production (Whetten *et al.,* 1998). Plants with reduced lignin content or altered lignin composition are desirable for the industry. Because of its economic importance, tremendous effort has been put since few decades to uncover the lignin biosynthetic pathway which has been proved to be highly complex (Grima-Pettenati & Goffner, 1999; Boudet *et al.,* 2003) and the genes involved have multiple cross talk with other physiological process (Zhao & Dixon, 2011).

Lignin polymer consists of three different monomers, which are p-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units. Angiosperm's lignin consists of mainly G and S subunit. Caffeic acid O-methyltransferase (COMT) and Cinnamoyl-CoA reductase (CCR) are two enzymes involved in G and S units production (Shi *et al.,* 2010). Similar to other genes involved in lignin biosynthesis pathway, *CCR* and *COMT* exist as gene family in many plant species. However, only one *CCR* and one *COMT* gene was shown to have high expression in xylem tissues and is predicted to have major function in lignification during development (Raes *et al.,* 2003; Shi *et al.,* 2010).

Acacia mangium, an important forest tree species has an estimated worldwide plantation of 1.4 M ha (Griffin *et al.,* 2011). It is planted widely in South East Asia and used mainly for the production of pulp and paper (Griffin *et al.,* 2011). Transcriptome sequencing on *A. mangium* found important genes involved in the lignin biosynthetic pathway (Wong *et al.,* 2011). A comparative study of small interfering RNA in high lignin and low lignin plants of *A. mangium* found some conserved as well as novel microRNAs which are predicted to have some important functions in lignin and secondary cell wall biosynthesis (Ong & Wickneswari, 2011). The regulatory elements like promoter structures, transcription factors and miRNAs are important aspects to be taken into

* To whom correspondence should be addressed.

account; while studying the complex regulatory network of lignin biosynthesis.

In this study, promoter of *AmCCR1* and a new promoter of a pseudogene with high sequence similarity with *COMT* were isolated and characterised *in silico*. Both promoters are shown to have cis acting elements commonly found in lignin biosynthetic genes. Transcription factor binding site modelling shows presence of binding site for some important MYB transcription factors involved in lignin and secondary cell wall biosynthesis regulation on both promoters.

MATERIALS AND METHODS

Plant Materials, Genomic DNA Isolation

Young leave samples of *A. mangium* were obtained from Acacia plot W at Plant Biotechnology Laboratory, UKM, Malaysia. Genomic DNA was extracted using QIAGEN DNeasy Plant Mini Kit (Qiagen, Germany).

Gene Isolation

Isolation of both *AmCCR1* and *pseudoCOMTAm* promoters were done using TAIL-PCR.

Primer design for *AmCCR1*

Primers were designed based on *contigAmCCR1* sequence (Wong *et al.,* 2011). CCR1, CCR2 and CCR3 primers together with arbitrary degenerate primer (AD6) were used to amplify the promoter region using TAIL-PCR. CCRE1ME4M F and CCRE1ME4M R primers were used to amplify exon 1 until exon 4 using normal PCR. Full sequence of exon 1 was amplified using CCRE1 F and CCR E1 R. CCRE4 F1 and CCRE4 F2 together with AD6 was used to amplify exon 4 and intron 4 using TAIL-PCR. CCRI4 F and CCR14 R were used to amplify full sequence of intron 4 and exon 5 (Table 1). Melting points (Tm) for the designed gene specific primers were higher than 62°C as calculated with the formula of Mazars *et al.* (1991).

Primer design for *pseudoCOMTAm*

Two gene specific primers (GSP1 and GSP2) as designed by Sukganah *et al.* (2013) were used (Table 2). These primers were used as reverse primers, pairing with arbitrary degenerate (AD4) primer (Thanh *et al.,* 2012) as forward primer. Based on the sequence of the first fragment, new primers were designed to further amplify the 5' and 3' regions (Table 2). GSP1, GSP2 and GSP3 were used

Table 1. Primer sequence for *AmCCR1* isolation

Primer Name	Primer Sequence (5' to 3')
CCR1	CCGCGGTGGAGTGATCTCTCAGAAC
CCR2	GGTGGATTGAAGCAATGGTCCCAG
CCR3	GCCTCTCTCTAGGAGAAGCTTGACAATCC
CCRE1ME4M F	GGATTGTCAAGCTTCTCCTAGAGAGAGGC
CCRE1ME4M R	CCGCGGTGGAGTGATCTCTCAGAAC
CCRE1 F	GAGGTGGGTAATTCTGGTAGGC
CCRE1 R	CCCAATTATGAATGGCTCTCATG
CCRE4 F1	GGATTGTCAAGCTTCTCCTAGAGAGAGGC
CCRE4 F2	GCAGCATGGGATGAGGCAAAAGC
AD 6	AWGCANGNCWGANATA
CCRI4 F	GTTCTGAGAGATCACTCCACCGCGG
CCRI4 R	GTAGGGTTTTGCTCTTGG

Table 2. Primer utilized for isolation of *pseudoCOMTAm*

Primer Name	Primer Sequence (5' to 3')
GSP1	CGAGGGCTGATTTGAGAATCATGGGAAG
GSP2	TGGCTTCCTCGTCGTTGACATGGGTAGG
GSP3	GTTTTTTAATTTACCCCTATGCTACTCCAC
M1	CCACACCCTTTTGCTGGAATCATCACC
M2	GGTTCAGCCGGCGAGACTTAGATAACC
AD2	NGTCGASWGANAWGAA
AD4	AGWGNAGWANCAWAGG

in combination with AD2 to further amplify the 5' region. Primers M1, M2 and AD2 were used to further amplify the 3' region. Melting points (Tm) for the designed gene specific primers was higher than 62°C as calculated with the formula of Mazars *et al.* (1991).

TAIL-PCR Reaction

TAIL-PCR was carried out as described by Liu & Whittier (1995). The primary PCR reaction mixture consisted of 1 U of HS Taq polymerase (Takara Bio Inc., Japan), 1× Taq polymerase buffer, 200 µM dNTPs, 0.2 µM RSP1 primer, 5 µM of the AD primer, and 50 ng of genomic DNA. The secondary PCR reaction consisted of 1× Taq polymerase buffer, 200 µM dNTPs, 0.8 U of HS Taq polymerase, 0.2 µM RSP2 primer, 4 µM of the AD primer used in the primary reaction, and 50 fold dilution of the primary PCR product. The tertiary PCR mixture consisted of 1× Taq polymerase buffer supplied with the enzyme, 200 µM dNTPs, 0.5 U of HS Taq polymerase, 3 µM of the AD primer used in the previous reactions, 0.3 µM RSP3 primer, and 10 fold dilution of the secondary PCR product. The thermal cycling conditions are shown in Table 3. PCR products were visualized by electrophoresis on 1.0% (w/v) agarose gels, purified using NucleoSpin Gel and PCR Clean-up kit and sequenced.

Sequence Analysis

Sequences were assembled and characterised using BioEdit (Hall 1999) and Clustal Omega (Sievers *et al.,* 2011) before subjected to similarity search using BLASTn (Altschul *et al.,* 1990). The promoter sequences were aligned to known COMT1 promoter sequence from *Acacia auriculiformis* × *Acacia mangium* hybrid, HQ317735.1 (Sukganah *et al.,* 2013) and real *CCR* from *Leucaena leucocephala* (GU984572). Putative cis-acting elements were identified using PlantCARE (Lescot *et al.,* 2002). The ORF of the gene was detected using ExPASy Translate Tool.

RESULTS AND DISCUSSION

***AmCCR1* Sequence Analysis**

Sequence analysis shows *AmCCR1* consists of 619bp of promoter (Genbank accession number: MK353215), 5 exons and 4 introns (Figure 1). Blast N analysis shows *AmCCR1* shares 88% similarity with Real *CCR* from *Leucaena leucocephala* (GU984572). However, Tail-PCR was unable to amplify full length of intron 4 and exon 5.

Promoter of *AmCCR1* was shown to have multiple cis elements commonly present in promoters of lignin genes. TATA Box was found at -35. Xylem specific elements, ACI and ACII were found at position +60 and +61 respectively. Real *CCR* from Arabidopsis thaliana was also found to have ACI and ACII in it's promoter (Lauvergeat *et al.,* 2001). 5UTR Py-rich stretch was found at position +44. This element confers high transcriptional rate for the gene. TC rich repeats, an element involved in stress response and plant defence, was found at three positions in *AmCCR1*; -195, -386 and +15. Box S which is also involved in plant defence was found at +153. Some of these cis acting elements were also found in *CCR* from *Leucaena leucocephala* (Prashant *et al.,* 2011).

***PseudoCOMTAm* Sequence Analysis**

Using Thermal Asymmetric Interlaced PCR (TAIL-PCR), a combination of GSP1, GSP2 (Sukganah *et al.,* 2013) and AD4 (Thanh *et al.,* 2012) primers successfully amplified a 400bp fragment with 93% similarity with real *COMT* gene, *AhgCOMT1* (HQ317735) from *Acacia*

Table 3. Thermal Cycling Conditions

Reaction (primer combination)	Number of Cycles	Thermal cycling conditions
Primary PCR	1	93°C, 1 min; 95°C, 1 min
	5	94°C, 30 s; 62°C, 1 min; 72°C, 2.5 min
	1	94°C, 30 sec; 25°C, 3 min; *72°C, 2.5 min
	15	94°C, 10 s; 68°C, 1 min; 72°C, 2.5 min
		94°C, 10 s; 68°C, 1 min; 72°C, 2.5 min
		94°C, 10 s; 29°C, 1 min; 72°C, 2.5 min
	1	72°C, 5 min
Secondary PCR	12	94°C, 10 s; 64°C, 1 min; 72°C, 2.5 min
		94°C, 10 s; 64°C, 1 min; 72°C, 2.5 min
		94°C, 10 s; 29°C, 1 min; 72°C, 2.5 min
	1	72°C, 5 min
Tertiary PCR	20	94°C, 15 s; 29°C, 30 s; 72°C, 2 min
	1	72°C, 5 min

* ramping to 72°C, over 3 min.

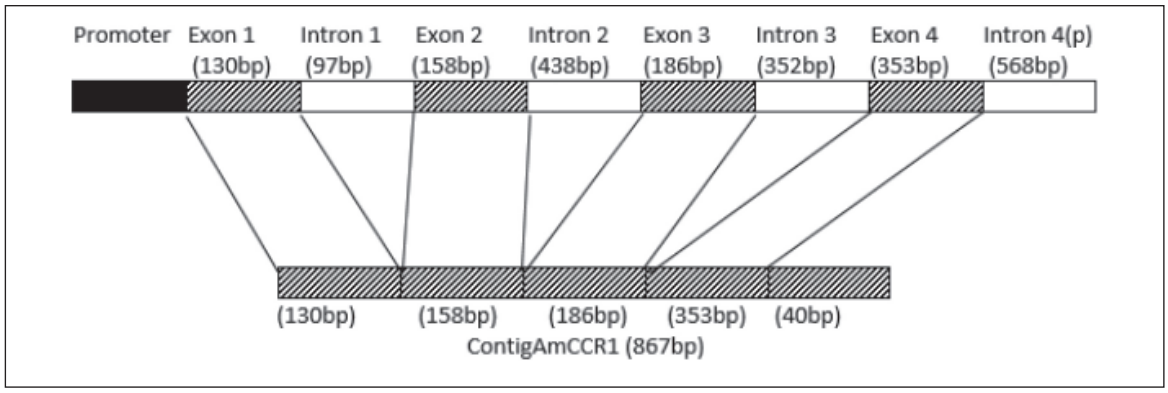

Fig. 1. The gene structure of *AmCCR1.*

Fig. 2. Cis elements of *AmCCR1.* Transcription Start Site marked in bold.

auriculiformis × *Acacia mangium* hybrid. Sequence analysis however, detected a nonsense mutation on the first exon. Thus, this gene is named *pseudoCOMTAm* (Genbank accession number: MF488717). The high sequence similarity is within 211bp in promoter region and extends to 110bp after the start codon (Figure 2).

Based on the sequence of the 400bp fragment, new primers were designed. GSP1, GSP2 and GSP3 were used in combination with AD2 to further amplify the 5' region while M1, M2 and AD2 were used to further amplify the 3' region. The total size of the promoter isolated is 770bp. PLANTCARE analysis showed presence of multiple cis acting elements (Figure 3). Core promoter element, TATA Box was found at position -22. 5UTR Py-rich stretch, a cis-acting element which confers high transcription levels were identified at -261. Cis acting elements commonly found in genes acting in lignin and phenylpropanoid pathway, ACI and ACII elements were identified at position -77 and -40 respectively. AC elements are known to interact with MYB transcription factors and function in coordinated regulation of lignin and phenyl-propanoid biosynthetic pathway genes (Zhou *et al.,* 2009).

W box which functions in defence against pathogen was found at -421. Besides that, two motifs which regulate endospermic gene expression, GCN4 and Skn-1 were found at -619 and -43. Abscisic acid responsiveness motif, ABRE was found at -110 while ARE motif which functions in anaerobic induction were found at -496. P-box, motif known to function in gibberellin response was found at +2. Multiple cis elements which function in light responsiveness like ATCT-motif, Box I, GA-motif and GT1-motif were also identified in this promoter (Figure 4). This is not surprising as lignin genes are also known to be involved in crosstalk with other plant physiological process (Rogers *et al.,* 2005).

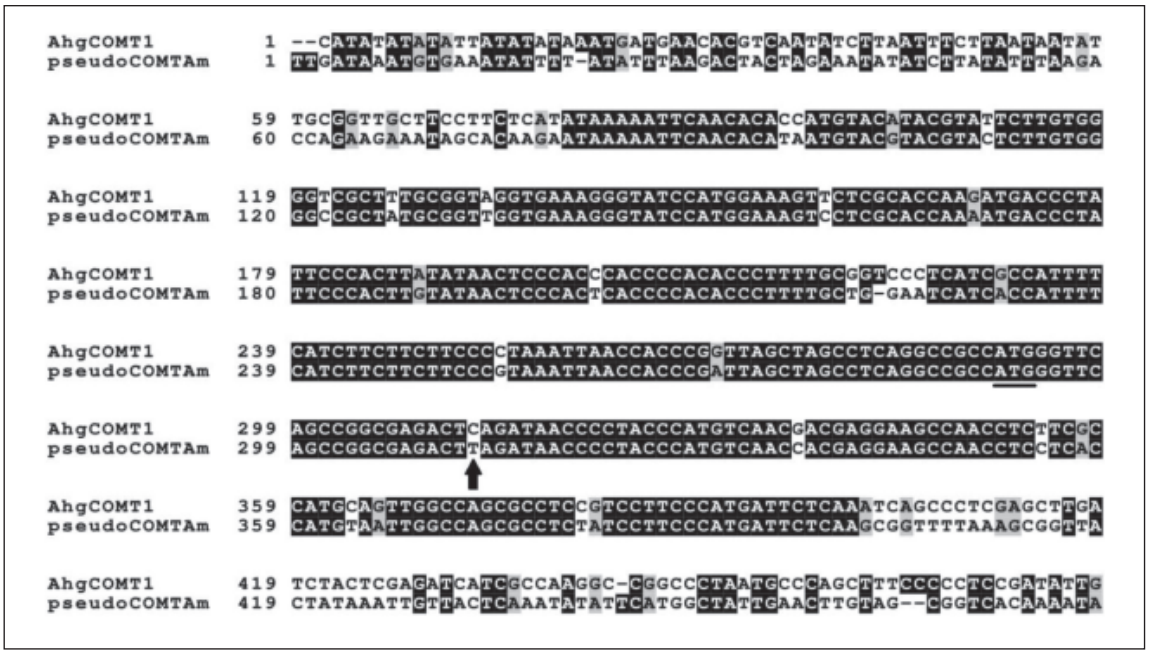

Fig. 3. Alignment of *pseudoCOMTAm* and *AhgCOMT1.* Start codon is underlined. Arrow shows mutation that creates premature stop codon.

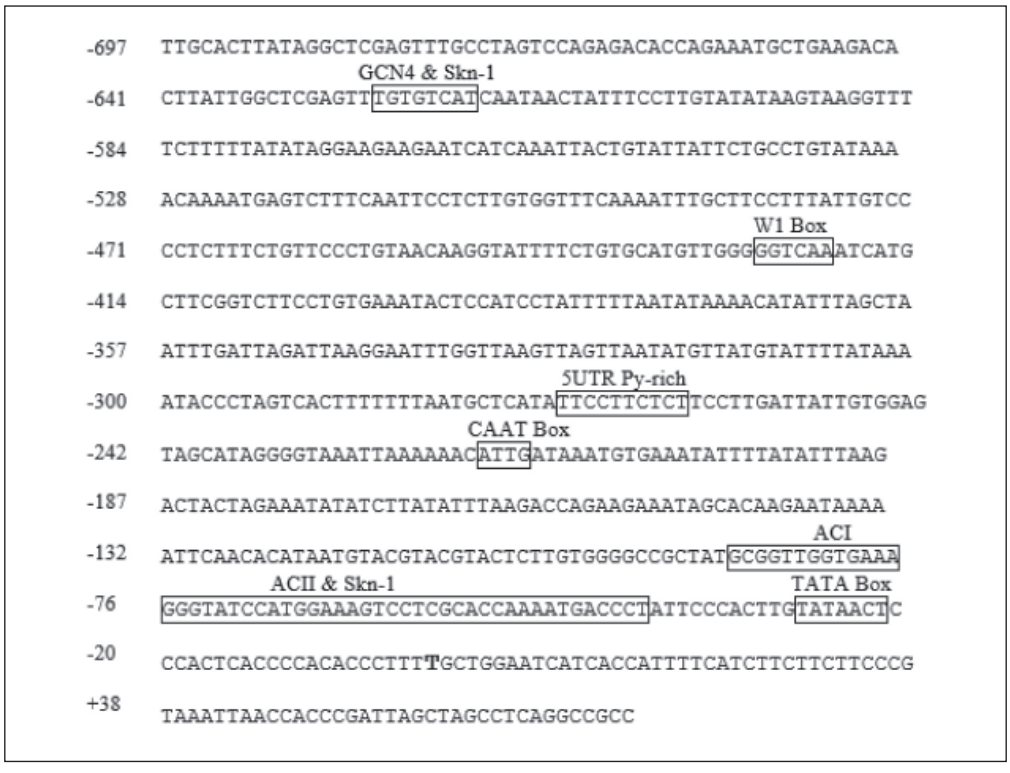

Fig. 4. cis elements of *pseudoCOMTAm.* Transcription Start Site marked in bold.

ACKNOWLEDGEMENTS

The authors would like to thank Faculty of Science and Technology, Universiti Kebangsaan Malaysia and School of Biosciences, Taylor's University for supporting this research.

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. 1990. Basic local alignment search tool. *Journal of Molecular Biology*, **215(3)**: 403-410.

Boerjan, W., Ralph, J. & Baucher, M. 2003. Lignin biosynthesis. *Annual Review of Plant Biology*, **54**: 519-546.

Boudet, A.M., Kajita, S., Grima-Pettenati, J. & Goffner, D. 2003. Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends in Plant Science*, **8(12)**: 576-81.

Griffin, A.R., Midgley, S.J., Bush, D., Cunningham, P.J. & Rinaudo, A.T. 2011. Global uses of Australian acacias – recent trends and future prospects. *Diversity and Distributions*, **17(5)**: 837-847.

Grima-Pettenati, J. & Goffner, D. 1999. Lignin genetic engineering revisited. *Plant Science*, **145(2)**: 51-65.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, **41**: 95-98.

Lauvergeat, V., Lacomme, C., Lacombe, E., Lasserre, E., Roby, D. & GrimaPettenati, J. 2001. Two *cinnamoyl-CoA reductase (CCR)* genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria. *Phytochemistry*, **57(7)**: 1187-1195.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P. & Rombauts, S. 2002. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Research*, **30(1)**: 325-327.

Liu, Y-G. & Whittier, R.F. 1995. Thermal asymmetric interlaced PCR: automatable amplification and sequencing of insert end fragments from P1 and YAC clones for chromosome walking. *Genomics*, **25(3)**: 674-681.

Mazars, G.R., Moyret, C., Jeanteur, P. & Theillet, C.G. 1991. Direct sequencing by thermal asymmetric PCR. *Nucleic Acids Research*, **19(17)**: 4783.

Ong, S.S. & Wickneswari, R. 2011. Expression Profile of small RNAs in *Acacia mangium* Secondary Xylem Tissues with Contrasting Lignin Content – potential regulatory sequences in monolignol biosynthetic pathway. *BMC Genomics*, **12**: S13.

Prashant, S., Sunita, M.S.L., Sirisha, V.L., Bhaskar, V.V., Rao, A.M., Narasu, M.L. & Kavi Kishor, P.B. 2012. Isolation of *cinnamoyl CoA reductase* and *cinnamyl alcohol dehydrogenase* gene promoters from *Leucaena leucocephala*, a leguminous tree species, and characterization of tissue-specific activity in transgenic tobacco. *Plant Cell, Tissue and Organ Culture*, **108(3)**: 421-436.

Raes, J., Rohde, A., Christensen, J., Van de Peer, Y. & Boerjan, W. 2003. Genome wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiology*, **133(3)**: 1051-1071.

Rogers, L.A., Dubos, C., Cullis, I.F., Surman, C., Poole, M., Willment, J., Mansfield, S.D. & Campbell, M.M. 2005. Light, the circadian clock, and sugar perception in the control of lignin biosynthesis. *Journal of Experimental Botany*, **56(416)**: 1651-1663.

Shi, R., Sun, Y.H., Li, Q., Heber, S., Sederoff, R. & Chiang, V.L. 2010. Towards a Systems Approach for lignin biosynthesis in *Populus trichocarpa*: Transcript Abundance and Specificity of the Monolignol Biosynthetic Genes. *Plant and Cell Physiology*, **51(1)**: 144-63.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D. & Higginsa, D.G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, **7**: 539.

Sukganah, A., Choong, C.Y., Russell, J., Neale, D. & Wickneswari, R. 2013. Nucleotide sequence analysis of two lignin genes in *Acacia auriculiformis* × *Acacia mangium* hybrid for enhancement of wood pulp quality. *Tree Genetics & Genomes*, **9(5)**: 1369-1381.

Thanh, T., Chi, V.T., Abdullah, M.P., Omar, H. & Napis, S. 2012. Efficiency of ligation-mediated PCR and TAIL-PCR methods for isolation of RbcS promoter sequences from green microalgae *Ankistrodesmus convolutus*. *Molecular Biology*, **46(1)**: 64-70.

Whetten, R.W., MacKay, J.J. & Sederoff, R.R. 1998. Recent advances in understanding lignin biosynthesis. *Annual Review of Plant Physiology and Plant Molecular Biology*, **49**: 585-609.

Wong, M.M., Cannon, C.H. & 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.

Zhao, Q. & Dixon, R.A. 2011. Transcriptional networks for lignin biosynthesis: more complex than we thought? *Trends in Plant Science* **16(4)**: 227-233.

Zhou, J., Lee, C., Zhong, R. & Ye, Z.H. 2009. MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell*, **21(1)**: 248-266.