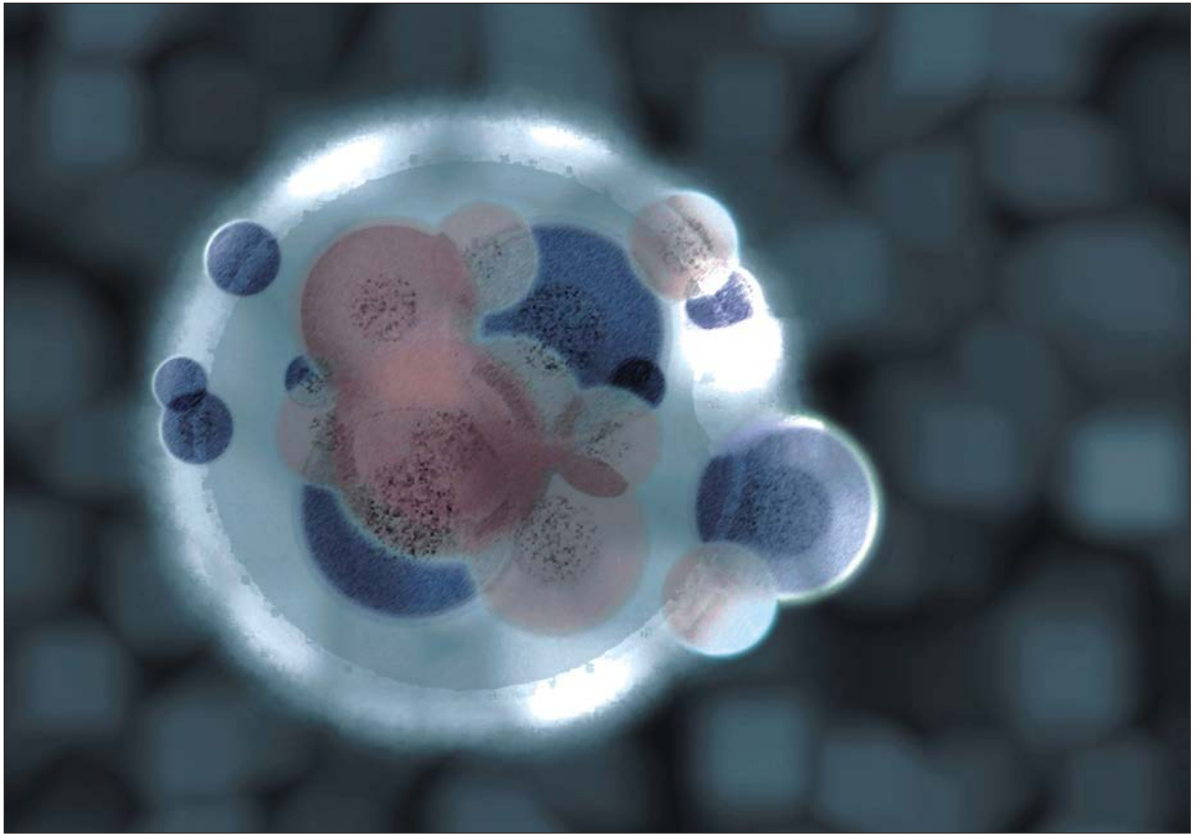

Journal of Biological Sciences

ISSN 1727-3048

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Biological Sciences 8 (3): 687-690, 2008
ISSN 1727-3048
© 2008 Asian Network for Scientific Information

Extensive DNA Sequence Variations in Two Lignin Genes, *Cinnamate 4-Hydroxylase* and *Cinnamyl Alcohol Dehydrogenase* from *Acacia mangium* and *Acacia auriculiformis*

M.S. Nur Fariza, S.L. Pang, C.Y. Choong and R. Wickneswari
School of Environmental and Natural Resource Sciences, Faculty of Science and Technology,
Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor Darul Ehsan, Malaysia

Abstract: This study examined DNA sequence variations of coding regions of *Cinnamate 4-hydroxylase* (*C4H*) and *Cinnamyl Alcohol Dehydrogenase* (*CAD*) in *Acacia mangium* and *A. auriculiformis*. cDNA fragments of *C4H* and *CAD* with size 1.5 and 1.3 kb, respectively were cloned into pGEM-T Easy Vector and were sequenced. Twenty eight Single Nucleotide Polymorphisms (SNPs) were identified in the coding region of *C4H* of which 8 caused changes in the amino acids or nonsynonymous mutations and 20 were synonymous mutations. Thirty two SNPs were detected in coding region of *CAD*. Of these, 12 were nonsynonymous mutations and 20 were synonymous mutations. Two *A. mangium* individuals (M20 and M22) and two *A. auriculiformis* individuals (A6 and A3) were used as parents for generating F_1 mapping populations. Nucleotide sequence alignment of coding region of *CAD* detected 28 and 22 SNPs from A3×M22 parental combination and A6×M20 parental combination respectively. Nucleotide sequence alignment of coding region of *C4H* identified 6 and 23 SNPs for A3×M22 and A6×M20 parental combinations, respectively. For parental combination A3×M22, 11 nonsynonymous mutations were detected while for A6×M20 parental combination 23 nonsynonymous mutations were detected from *CAD* gene. Amino acid sequence alignment of *C4H* detected 4 amino acid variations from each parental combination. The putative SNPs can be developed as SNPs markers for Quantitative Trait Loci (QTL) detection. Selecting favourable alleles from progenies which produce desirable lignin profiles would be advantageous in tree breeding programmes for plantation establishment.

Key words: *Acacia mangium*, *Acacia auriculiformis*, *cinnamate 4-hydroxylase* (*C4H*), *cinnamyl alcohol dehydrogenase* (*CAD*), single nucleotide polymorphism

INTRODUCTION

The wood of *Acacia mangium* and *A. auriculiformis* is suitable in pulp and paper making as well as for manufacturing of furniture, Medium Density Fiberwood (MDF), veneer and plywood. These species were firstly introduced to Malaysia in 1960s and then have been planted extensively in South East Asia because of their fast growth, good form and economic potential of the wood (Sahri *et al*., 1993). Spontaneous hybrids of *A. mangium* and *A. auriculiformis* were first reported to occur in *A. mangium* plantations in Sabah (Sim, 1987). *Acacia* hybrids have shown superior characteristics in growth, adaptation to different types of soils, resistance to disease and higher pulp yield over the parental species (Pinso and Nasi, 1991), showing their importance to the pulp and paper making industry.

Lignin is formed by intracellular synthesis of the monolignol precursors (ρ-coumaryl, coniferyl and sinapyl) each giving rise to the hydroxyphenyl (H), guaiacyl (G) and syringyl (S) lignin units respectively. For the production of high quality paper, lignin has to be removed from the cellulose by chemical pulping and the rate of delignification is proportional to the S/G ratio (Piquemal *et al*., 1998). The benefits of removing as much lignin as possible to avoid residual lignin causes discoloration and reduces paper brightness (Chiang *et al*., 1988). Cinnamate 4-hydroxylase (C4H) and cinnamyl alcohol dehydrogenase (CAD) are two enzymes that play important role in the beginning and the end of the lignin biosynthesis pathway. C4H catalyzes the hydroxylation of cinnamate to 4-coumarate while the CAD catalyzes the final reduction of cinnamaldehydes to ρ-coumaryl, coniferyl and sinapyl alcohols (Lewis, 1999). Down regulation of *C4H* in *Nicotiana tabacum* was reported to reduce the lignin content by over 90% and decrease S/G ratio (Blee *et al*., 2001). The lignin of plants with low *CAD* activity was more extractable in alkaline

Corresponding Author: Wickneswari Ratnam, School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor Darul Ehsan, Malaysia Tel: 603-89213840 Fax: 603-89253357

chemical pulping process (Baucher *et al*., 1996). *CAD* was demonstrated to be a useful target for paper pulping quality improvement in dicotyledons (O'Connell *et al*., 2002; Pilate *et al*., 2002).

This study was undertaken to determine the feasibility of using two selected lignin genes, *C4H* and *CAD* to find Single Nucleotide Polymorphisms (SNPs) in parental combinations of *A. mangium* and *A. auriculiformis* used for generation of segregating mapping populations for fibre length and wood density. Presence of single base differences among the individuals of same species can change the coding of amino acid and protein function resulting in changes in lignin content and composition. SNPs can serve as genetic markers and can be used to characterize genetic resources, map plant genome and specific genes (Dreher *et al*., 2003). SNP markers have been developed in *Eucalyptus nitens* and demonstrated that linkage disequilibrium mapping of lignin genes can be used to identify alleles associated with wood quality traits in natural populations (Thumma *et al*., 2005). SNP discovery and estimation of linkage disequilibrium have been completed in loblolly pine (*Pinus taeda*) for nineteen candidate genes for wood formation including *C4H* and *CAD* (Brown *et al*., 2004).

MATERIALS AND METHODS

Inner bark tissues of two individuals from *A. mangium*, M20 and M22 and *A. auriculiformis*, A3 and A6 were used in this study. Parental combination of A3× M22 and A6×M20 were used in crosses to produce *Acacia* hybrid mapping population.

Total RNA was extracted from the inner bark tissues using RNeasy Midi Kit (Qiagen, Germany). Total RNA was used as template to synthesize the first strand cDNA and further amplified using a specific primer by reverse transcriptase PCR (RT-PCR) approach. Two primer pairs were designed from full length cDNA of *C4H* and *CAD* from *Acacia* hybrid (Pang *et al*., 2005). The primer pairs were forward primer (FL-C4H-F: 5' GTC CAC CTT TCT CAG CAG TAT CAA 3') and reverse primer (FL-C4H-R: 5' CAA TCT CCA AAT CGC CAA CA 3') to generate the sequence of *C4H*. *CAD* was amplified using forward primer (FL-CAD-F: 5' GAC ATT CTT TCT TCT TCT TCT T 3') and reverse primer (FL-CAD-R: 5' GAC AAA CAT CTG TGA GGC AT 3'). Each reaction consisted of 2 µL of 2.5 ρ-mol forward and reverse primers, 1.0 µL of dNTP mix, 2.5 µL of 10×PCR buffer (200 mM Tris-HCl pH 8.8, 100 mM KCl, 1% Triton X-100, 100 mM $(NH_4)_2SO_4$, 1 µg mL^{-1} BSA), 0.75 µL of 100 mM $MgCl_2$, 0.04 U YEA Taq DNA Polymerase (Yeastern Biotech Co.) and sterile distilled water to make up 25 µL. *C4H* amplification was performed in Eppendorf Thermal Cycler (Germany) for 2 min at 95°C, 35 cycles of 30 sec at 94°C, 30 sec at 55°C and 45 sec at 70°C, followed by final step of 10 min at 70°C. *CAD* amplification was carried out by initial denaturing at 95°C for 2 min, followed by 30 cycles of 94°C for 30 sec, 46°C for 30 sec and 70°C for 45 sec with the final extension time for 10 min.

PCR products were ligated into pGEM-T Easy Vector (Promega) and transformed into competent cell of DH5α from *Escherichia coli*. Positive clones were cultured in LB Broth supplemented with ampicillin (50 µg μL^{-1}) for overnight. The recombinant plasmids were purified using Wizard Plus SV Minipreps DNA Purification System (Promega). Six to eight positive clones per gene per individual were sent for sequencing using ABI 3730 DNA Analyzer (Applied Biosystem, USA)

The *C4H* and *CAD* sequences for the four parental trees (M20, M22, A3, A6) were aligned using the CLUSTALW (www.ebi.ac.uk/clustalw/) and visually inspected for base changes. The nucleotide sequence alignments were translated into protein in all reading frames using ORF finder (http://www.ncbi.nlm.nih.gov). The open reading frames of both genes were aligned among four individuals using CLUSTALW to detect the nonsynonymous and synonymous mutations.

RESULTS AND DISCUSSION

cDNA fragments of *C4H* and *CAD* of 1.5 and 1.3 kb, respectively (Fig. 1) were cloned into pGEM-T Easy Vector and were sequenced. Six to eight complete cDNA clones for each gene and each individual were sequenced and aligned using CLUSTALW. A3 did not show any nucleotide variation while the others (A6, M20 and M22) showed variations among the clones for *C4H*. *CAD* nucleotide sequences showed 20 heterozygous loci which were from A3, M20 and M22. A6 did not show any heterozygous loci.

The consensus sequences of A3, A6, M20 and M22 for both genes were aligned together to detect the SNPs. In total 28 SNPs were found in the coding region of *C4H*. Of these, 6 SNPs were detected in A3×M22 parental combination while 23 SNPs were detected in A6 x M20 parental combination. Within the coding region of *CAD*, 32 SNPs were identified. Of these, 28 SNPs were detected in A3×M22 parental combination and 22 SNPs were detected in A6×M20 parental combination. The cDNA sequence of *CAD* was more polymorphic than that of *C4H*. On average, one SNP occurred for every 53 bp in *C4H* compared to one SNP for every 39 bp in *CAD* (Table 1). The occurrence of SNPs detected in this study

Table 1: The number of SNPs in *C4H* and *CAD*

Feature	*C4H*	*CAD*
Total No. of SNPs	28	32
A3×M22	6	28
A6×M20	23	22
Size of sequenced fragment (bp)	1481	1256
Occurrence of SNPs	One SNP per 53 bp	One SNP per 39 bp

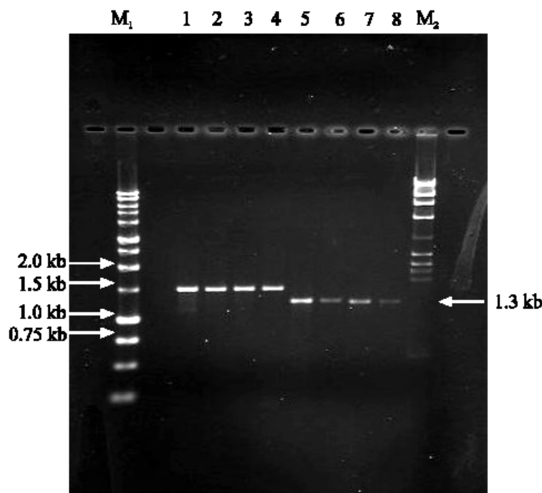

Fig. 1: Agarose gel electrophoresis of RT-PCR products of *C4H* (lane 1-4) and *CAD* (lane 5-8). M_1 is a 1 kb marker and M_2 is a Lambda HindIII marker

for *A. mangium* and *A. auriculiformis* are much higher than that detected for coding regions in humans (Cargill *et al*., 1999), *Zea mays* (Zhu *et al*., 2003) and barley (Kanazin *et al*., 2002). The occurrence of SNPs in *Eucalyptus globulus* was also reported to be much lower in the coding region of *CAD*; one SNP occurred for every 147 bp (Poke *et al*., 2003) compared to that in *A. mangium* and *A. auriculiformis*. *A. mangium* and *A. auriculiformis* are considered as pioneer species. The high occurrences of SNPs in these species are possible because they have to adapt physiologically to a variety of unpredictable environmental conditions such as infertile acid soils outside their native range. High genetic variation at many physiologically important gene loci may be a strategy for tree species to maintain a higher level of adaptability (Ziehe *et al*., 1999).

Eight and 12 nonsynonymous mutations were detected in *C4H* and *CAD*, respectively (Table 2). Out of these, 4 mutations were detected in A3×M22 parental combination for *C4H* whereas the remaining 4 mutations were detected in A6×M20 parental combination. Within the open reading frame of *CAD*, 11 and three nonsynonymous mutations were detected in A3×M22 and A6×M20 parental combination, respectively. The synonymous mutations detected were 20 in both genes.

Table 2: The number of synonymous and nonsynonymous mutations detected in *C4H* and *CAD*

Feature	*C4H*	*CAD*
Total No. of non-synonymous mutations	8.000	12.000
A3×M22	4.000	11.000
A6×M20	4.000	3.000
Total No. of synonymous mutations	20.000	20.000
Total No. of amino acids	480.000	359.000
Proportion of non-synonymous mutations (dN)	0.017	0.033
Proportion of synonymous mutations (dS)	0.042	0.056

The proportion of nonsynonymous mutations, dN were 0.017 and 0.033 for *C4H* and *CAD*, respectively. The proportions of synonymous mutation, dS were 0.042 and 0.056 for *C4H* and *CAD*, respectively. The proportion indicates that synonymous mutations were more common than nonsynonymous mutations for both genes. The results were similar to Cargill *et al.* (1999) and Ramos-Onsins *et al.* (2008) where the proportions of synonymous mutations were higher than non-synonymous mutations in human coding regions and nine loci involved in the phenylpropanoid pathway of *Arabidopsis thaliana*, respectively. SNPs of candidate genes for wood properties in *P. pinaster*. showed one nonsynonymous mutation out of 24 polymorphic sites detected (Pot *et al*., 2006). Gonzalez-Martinez *et al*. (2007) also reported that a nonsynonymous substitution for *CAD* in loblolly pine was in strong association with earlywood specific gravity and causing lignin modification in wild trees.

The putative SNPs can be developed as SNP markers for Quantitative Trait Loci (QTL) detection in *Acacia* hybrid mapping populations after validation using segregation analysis. A highly polymorphic SNP marker allows the selection of alleles in breeding programmes to produce trees with more favourable lignin profiles for pulping.

ACKNOWLEDGMENTS

This research was supported by funds for postgraduate studies from the Faculty of Science and Technology, Universiti Kebangsaan Malaysia.

REFERENCES

Baucher, M., B. Chabbert, G. Pilate, J. VanDoorsselaere and M.T. Tollier, 1996. Red xylem and higher lignin extractiblity by down regulating a cinnamyl alcohol dehydrogenase in poplar (*Populus tremula*× *Populus alba*). Plant Physiol., 12 (4): 1479-1490.

Blee, K., J.W. Choi, A.P. O'Connell, S.C. Jupe, W. Schuch, N.G. Lewis and G.P. Bolwell, 2001. Antisense and sense expression of cDNA coding for CYP73A15, a class II cinnamate 4-hydroxylase, leads to a delayed and reduced production of lignin in tobacco. Phytochemistry, 57 (7): 1159-1166.

Brown, G.R., G.P. Gill, R.J. Kuntz, C.H. Langley and D.B. Neale, 2004. Nucleotide diversity and linkage disequilibrium in loblolly pine. Proc. Natl. Acad. Sci. USA., 101 (42): 15255-15260.

Cargill, M., D. Altshuler, J. Ireland, P. Sklar, K. Ardlie and N. Patil, 1999. Characterization of single nucleotide polymorphism in coding regions of human genes. Nat. Genet., 22: 231-238.

Chiang, V.L., R.J. Puumala, H. Takeuchi and R.E. Eckert, 1988. Comparison of softwood and hardwood kraft pulping. Tappi J., 71 (9): 173-176.

Dreher, K., M. Khairallah, J.M. Ribaut and M. Morris, 2003. Money matters (I): Costs of field and laboratory procedure associated with conventional and marker assisted maize breeding at CIMMYT. Mol. Breed., 11 (3): 221-234.

Gonzalez-Martinez, S.C., N.C. Wheeler, E. Ersoz, C.D. Nelson and D.B. Neale, 2007. Associations genetics in *Pinus taeda* L. I. wood property traits. Genetics, 175: 399-409.

Kanazin, V., H. Talbert, D. See, F. DeCamp, E. Nevo and T. Blake, 2002. Discovery and assay of single nucleotide polymorphisms in barley (*Hordeum vulgare*). Plant Mol. Biol., 48 (5-6): 529-537.

Lewis, N.G., 1999. A 20th century roller coaster ride: A short account of lignification. Curr. Opin. Plant Biol., 2 (2): 153-162.

O'Connell, A., K. Holt, J. Piquemal, J. Grima-Pettenati, A. Boudet, B. Pollet, C. Lapierre, M. Petit-Conil, W. Schuch and C. Halpin, 2002. Improved paper pulp from plants with suppressed cinnamoyl-CoA reductase or cinnamyl alcohol dehydrogenase. Transgenic Res., 11 (5): 495-503.

Pang, S.H., C.Y. Choong and R. Wickneswari, 2005. Isolation and characterization of *C4H*, CCoAOMT and *CAD* from interspecific hybrid *Acacia mangium*× *Acacia auriculiformis*. Proceeding of Acacia Research in Malaysia. Seminar on Current Updates on Acacia Genomics and Breedings, www.ukm.my/acacia, pp: 7-10.

Pilate, G., E. Guiney, K. Holt, M. Petit-Conil, C. Lapierre, J.C. Leple, B. Pollet, I. Mila, E.A. Webster, H.G. Marstorp, D.W. Hopkons, L. Jouanin, W. Boerjan, W. Schuch, D. Cornu and C. Halpin, 2002. Field and pulping performances of transgenic trees with altered lignification. Nat. Biotechnol., 20 (6): 607-612.

Pinso, C. and R. Nasi, 1991. The potential use of *Acacia mangium*×*Acacia auriculiformis* hybrid in Sabah. In: Breeding Technologies for Tropical Acacias, Carron, L.T. and K.M. Aken (Eds.). Canberra: Australian Centre for International Agricultural Research No. 37, pp: 17-21.

Piquemal, J., C. Lapierre, K. Myton, A. O'connell, W. Schuch and J. Grima Pettenati, 1998. Down-regulation of cinnamoyl CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. Plant J., 13 (1): 71-83.

Poke, F.S., R.E. Vaillancourt, R.C. Elliot and J.B. Reid, 2003. Sequence variation in two lignin biosynthesis genes, *Cinnamoyl CoA reductase* (*CCR*) and *Cinnamyl alcohol dehydrogenase* 2 (*CAD*2). Mol. Breed., 12 (2): 107-118.

Pot, D., J.C. Rodrigues, P. Rozenberg, G. Chantre, J. Tibbits, C. Cahalan, F. Pichavant and C. Plomion, 2006. QTLs and candidate genes for wood properties in maritime pine (*Pinus pinaster* Ait.). Tree Genet. Genomes, 2: 10-24.

Ramos-Onsins, S.E., E. Puerma, D. Balana-Alcade, D. Salguero and M. Aguade, 2008. Multilocus analysis of variation using a large empirical data set: Phenylpropanoid pathway genes in *Arabidopsis thaliana*. Mol. Ecol., 17 (5): 1211-1223.

Sahri, M.H., F.H. Ibrahim and N.A.A. Shukor, 1993. Anatomy of *Acacia mangium* grown in Malaysia. IAWA. J., 14: 245-251.

Sim, B. L., 1987. Research on *Acacia mangium* in Sabah: A review. ACIAR Proc., 16: 164-166.

Thumma, B., R., M.F. Nolan, R. Evans and G.F. Moran, 2005. Polymorphisms in *cinnamoyl CoA reductase* (*CCR*) are associated with variation in microfibril angle in *Eucalyptus* sp. Genetics, 171: 1257-1265.

Zhu, Y.L., Q.J. Song, D.L. Hyten, C.P. Van Tassell, L.K. Matukumalli, D.R. Grimm, S.M. Hyatt, E.W. Fickus, N.D. Young and P.B. Cregan, 2003. Single nucleotide polymorphism in soybean. Genetics, 163: 1123-1134.

Ziehe, M., H.H. Hattemer, R. Muller-Starck and G. Muller-Starck, 1999. Genetic Structures as Indicators for Adaptation and Adaptational Potentials. In: Forest Genetics and Sustainability, Matyas, C. (Ed.). Kluwer Academic Publication, Dordrecht, The Netherlands, pp: 75-89.