



Identification of the *CAD* gene from *Eucalyptus urophylla* GLU4 and its functional analysis in transgenic tobacco

B.W. Chen^{1,2}, Y.F. Xiao², J.J. Li², H.L. Liu², Z.H. Qin², Y. Gai¹ and X.N. Jiang¹

¹School of Biology Science and Biotechnology, Beijing Forestry University, Beijing, China

²Guangxi Key Laboratory of Superior Timber Trees Resource Cultivation, Guangxi Forestry Research Institute, Nanning, China

Corresponding authors: B.W. Chen / X.N. Jiang
E-mail: gfri_bwchen@163.com / jiangxn@bjfu.edu.cn

Genet. Mol. Res. 15 (4): gmr15049062

Received August 9, 2016

Accepted October 4, 2016

Published December 2, 2016

DOI <http://dx.doi.org/10.4238/gmr15049062>

Copyright © 2016 The Authors. This is an open-access article distributed under the terms of the Creative Commons Attribution ShareAlike (CC BY-SA) 4.0 License.

ABSTRACT. Cinnamyl alcohol dehydrogenase (*CAD*) catalyzes the final step in lignin biosynthesis. The genus *Eucalyptus* belongs to the family Myrtaceae, which is the main cultivated species in China. *Eucalyptus urophylla* GLU4 (GLU4) is widely grown in Guangxi. It is preferred for pulping because of its excellent cellulose content and fiber length. Based on GLU4 and *CAD* gene expression, a *Eucalyptus* variety low in lignin content should be obtained using transgenic technology, which could reduce the cost of pulp and improve the pulping rate, and have favorable prospects for application. However, the role and function of *CAD* in GLU4 is still unclear. In the present study, *EuCAD* was cloned from GLU4 and identified using bioinformatic tools. Subsequently, in order to evaluate its impact on lignin synthesis, a full-length *EuCAD*

RNAi vector was constructed, and transgenic tobacco was obtained via *Agrobacterium*-mediated transformation. A significant decrease in *CAD* expression and lignin content in transgenic tobacco demonstrated a key role for *EuCAD* in lignin biosynthesis and established a regulatory role for RNAi. In our study, the direct molecular basis of *EuCAD* expression was determined, and the potential regulatory effects of this RNAi vector on lignin biosynthesis in *E. urophylla* GLU4 were demonstrated. Our results provide a theoretical basis for the study of lignin biosynthesis in *Eucalyptus*.

Key word: *CAD* gene; Lignin; Cellulose; *Eucalyptus urophylla*; RNAi

INTRODUCTION

Lignin is the main component of plant secondary cell walls. It is a class of complex polymers mainly generated via the polymerization of the *p*-hydroxyphenyl lignin monomer (H), guaiacyl lignin monomer (G) and syringyl lignin monomer (S) (Vanholme et al., 2010), which are major components of plant secondary cell walls, second only to cellulose in their natural abundance (Battle et al., 2000). Lignin is covalently attached to semi-cellulose, which enhances the mechanical strength of the cell wall. It is important for the maintenance of normal plant structure, translocation of water and nutrients, and defense against invasive pathogens (Tronchet et al., 2010).

However, the presence of lignin increases the cost of paper manufacturing, reduces the rate of cellulose utilization, and leads to serious environmental pollution. It represents the major source of pollution from the paper manufacturing industry. Therefore, genetic engineering to improve the quality of plant material in paper manufacturing is an area of great interest internationally (Chapple et al., 2007).

The various monomers involved in lignin biosynthesis are derived from cinnamic acid (Figure 1), and the key synthetic enzymes include 4-coumarate: CoA ligase (4CL), cinnamoyl-CoA reductase (CCR), cinnamyl alcohol dehydrogenase (CAD), *p*-coumarate-3-hydroxylase (C3H), caffeic-acid-*O*-methyltransferase (COMT), and ferulate-5 hydroxylase (F5H). In lignin synthesis, coumaric acid is catalyzed into a series of CoA forms by 4CL, reduced to aldehyde substrates by CCR, and finally reduced to alcohol substrates by CAD. These alcohol substrates generate different lignin monomers. Hydroxylases and transferases such as C3H (Ralph et al., 2006), COMT (Jouanin et al., 2000), and F5H (Nair et al., 2000) catalyze the conversion of CoA in series to generate aldehydes and alcohols. The regulation of these genes diverts lignin synthesis to form specific monomers. However, 4CL (Kajita et al., 1996; Hu et al., 1999), CCR (Ralph et al., 1998; Chabannes et al., 2001), and CAD (MacKay et al., 1997; Halpin et al., 1998; Sibout et al., 2005) are involved in the biosynthesis of each type of lignin monomer. Therefore, the inhibition of any of these three genes triggers changes in the composition of lignin monomers, and ultimately reduces the total lignin content. This represents the main experimental strategy used to regulate lignin content. CAD is particularly important, because it catalyzes the final step in lignin biosynthesis, which is considered to be a sign of lignification (Li et al., 2001).

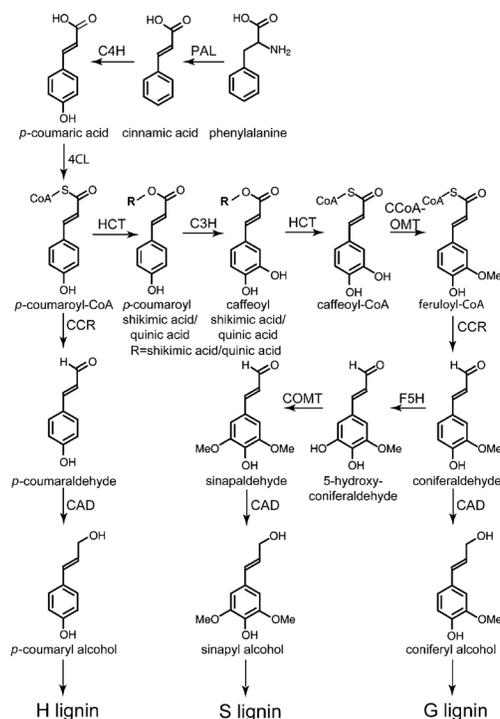


Figure 1. Biosynthetic pathway of lignin (Vanholme et al., 2010).

Since the discovery of *CAD* by Gross et al. (1973), growing evidence has demonstrated that *CAD* genes usually occur as part of a multi-gene family. At present, the number of *CAD* genes reported in various species is as follows: 17 in *Arabidopsis* (Kim et al., 2004), 12 in *Oryza* (Tobias and Chow, 2005), 15 in *Populus* (Barakat et al., 2009), 8 in *Gossypium* (Fan et al., 2009), 3 in *Camellia* (Deng et al., 2013), 14 in *Sorghum* (Saballos et al., 2009), and 11 in *Triticum* (Ma, 2010). However, in some plants, such as *Picea abies* L., *CAD* occurs as a single gene (Galliano et al., 1993).

Most *CAD* genes are located in duplicate blocks (Barakat et al., 2009). In some species, multiple *CAD* genes have been reported with different tissue specificities or substrate preferences. Barakat et al. (2009) found that two *CAD* genes in poplar were specifically expressed in the xylem. Kim et al. (2004) identified *AtCAD4* and *AtCAD5*, which exhibited the highest activity: *AtCAD5* catalyzed five substrates while *AtCAD4* weakly catalyzed mustard aldehyde. Fan et al. (2009) demonstrated that, although *GhCAD1* and *GhCAD6* shared the highest degree of similarity and catalyzed different substrates, *GhCAD6* played a major role in lignin biosynthesis, whereas *GhCAD1* acted as its redundant counterpart. Deng et al. (2013) showed that *CsCAD2* exhibited higher activity than that by the other two genes. In alfalfa, *MsCAD2* only used cinnamaldehyde as a substrate. However, *MsCAD1* showed a higher affinity for various aldehydes (Brill et al., 1999). In wheat, *TaCAD1* and *TaCAD2* preferred coumaraldehyde as the substrate, and *TaCAD3* preferred sinapaldehyde (Mitchell et al., 1999).

In other species, one or more *CAD* genes have been found to be highly homologous. However, the sequence lacks a functional domain and represents a false gene (as in *Triticum*).

Several genes resemble *CAD*-like genes on the basis of sequence analysis, molecular markers, and bioinformatics (as in *Triticum* and *Arabidopsis*). The function of these genes is yet to be systematically analyzed.

The role of *CAD* in *Eucalyptus* is still unclear. We retrieved 14 complete *CAD* genes from eight species of *Eucalyptus* deposited in GenBank (Table 1). The sequences of these 16 genes share more than 97% homology, which is very high. The *CAD* genes reported for each *Eucalyptus* species were mostly single or different forms of a single gene. For instance, the two *CAD* sequences from *Eucalyptus urophylla* were mRNA and genomic DNA sequences. No clear evidence regarding other members of the *CAD* gene family in *Eucalyptus* is available.

Table 1. Gene and cDNA sequence of cinnamyl alcohol dehydrogenase (*CAD*) in *Eucalyptus*.

Latin name	GenBank accession No.
<i>Eucalyptus globulus</i>	AF038561.1
	AF109157.1
<i>Eucalyptus camaldulensis</i>	GQ916948.1
	GU109374.1
<i>Eucalyptus urophylla</i>	FN393570.1
	GQ387647.1
<i>Eucalyptus gunnii</i>	X65631.1
	X75480.1
	X88797.1
<i>Eucalyptus botryoides</i>	D16624.1
<i>Eucalyptus saligna</i>	AF294793.1
<i>Eucalyptus pyrocarpa</i>	AB591256.1
<i>Eucalyptus pilularis</i>	AB591254.1
	AB591253.1

Studies investigating the regulation of *CAD* have been conducted in many plants. For example, in the maize *bm1* mutant, the total lignin content decreased by 20%, with a 65% reduction of G and S monomers (Halpin et al., 1998). A natural *CAD* mutant in pine exhibited a 9% decrease in total lignin content; however, the conifer aldehyde content increased (MacKay et al., 1997). Similarly, *CAD4* and *CAD5* double mutants in *Arabidopsis* exhibited a 94% reduction in lignin, resulting in serious logging in the whole plant (Sibout et al., 2005).

Antisense inhibition of *CAD* expression in transgenic tobacco reduced *CAD* activity by 90%, resulting in increased cinnamaldehyde content, although the overall lignin content was not altered (Halpin et al., 1994). *CAD* enzyme activity was suppressed by 70% following antisense inhibition of *CAD* in transgenic poplar, and enhanced aldehyde content did not substantially alter the total lignin content (Baucher et al., 1996). Antisense inhibition of *CAD* in transgenic alfalfa compromised *CAD* activity, and although the total lignin content remained unaltered, levels of the S monomer decreased (Baucher et al., 1999).

RNA interference (RNAi) is a valuable tool used to elucidate gene function (Abdurakhmonov et al., 2016). Fornalé et al. (2012) used RNAi to downregulate the expression of *CAD* in maize and found that the S/G ratio in the stem cell walls slightly decreased, while the total lignin content was not affected, and the lignin content in the midrib cell wall decreased by 6.4%. Downregulation of *CAD* by RNAi decreased the lignin content by 40% in transgenic flax (Wróbel-Kwiatkowska et al., 2007). RNAi inhibited *CAD* activity in *Pinus radiata* to approximately 30% (Möller et al., 2005).

Most studies investigating *CAD* gene regulation have shown a decrease in enzyme activity under the conditions studied. However, the degree of lignin decrease was not consistent,

and the S/G ratio was altered in a few cases. These varying results presumably reflected the multi-gene status of the *CAD* family.

Eucalyptus originated in Australia, and is the general term used to refer to plants belonging to the genus *Eucalyptus* of the family Myrtaceae (Xiang et al., 2006). Currently, it is cultivated in more than 100 countries worldwide and was introduced into the tropical and subtropical regions of China more than 100 years ago (Tang and Li, 2002). Because of the rapid growth, high yield, and desired properties of *Eucalyptus* wood, it has become the main tree species used as a short-term industrial raw material in pulp and paper manufacturing, and in the design of artificial and fiber boards (Xie, 2003).

E. urophylla GLU4 (GLU4) is widely grown in Guangxi. It is the preferred pulping species because of its excellent cellulose content and fiber length. Based on GLU4 and *CAD* gene expression, a *Eucalyptus* variety low in lignin content should be obtained using transgenic technology, which could reduce the cost of pulp and improve the pulping rate, which has favorable prospects for its application.

Currently, the molecular mechanisms associated with *CAD* gene expression in GLU4 are still unclear. However, the high degree of gene homology present in *Eucalyptus* was conducive to *CAD* cloning in GLU4 in the present study.

After isolating the *CAD* gene from GLU4, we investigated its role in lignin synthesis, and verified the genetic transformation. The regeneration and genetic transformation period of GLU4 is prolonged using currently available technology. Therefore, it is desirable to select an appropriate model plant for genetic transformation. Tobacco is one such model plant that has been extensively investigated. The *CAD* genes of *Eucalyptus* share 76% homology with two complete *CAD* gene sequences from tobacco, suggesting that tobacco is suitable for genetic transformation.

Thus, the *CAD* gene was cloned from GLU4, and the full-length RNAi vector was constructed and transformed into tobacco. Subsequently, *CAD* gene expression, lignin content, and anatomy of the stem in the transgenic plants were determined. In this preliminary study, we investigated the molecular basis and biological function of the *CAD* gene in GLU4, to serve as a reference for further research into transgenic *Eucalyptus*.

MATERIAL AND METHODS

Gene cloning of *EuCAD*

The *CAD* DNA sequences of several species were obtained from GenBank. One-hundred-base pair sequences before the start and termination codons were analyzed using BLAST to identify highly homologous fragments. The primers CAD-F (5'-ctttgagcaaaaatggcagctcttg-3') and CAD-R (5'-gggaaaggacaaaactaatcaagc-3') were designed based on the homology of the fragments using the Vector NTI software.

The stem tissue, representing 3 to 4-week-old tissue-cultured GLU4 seedlings, was collected, and the genomic DNA was extracted using the Hi-DNAsecure Plant Kit DP350 (Tiangen, China). Total RNA was extracted using the RNAPrep Pure Plant Kit DP441 (Tiangen). The cDNA was obtained from total RNA using the RNA LA PCR Kit (AMV) Ver.1.1 (TaKaRa, Japan). All samples were stored at -80°C.

The full-length *CAD* sequences of genomic DNA (gDNA) and cDNA were amplified using gDNA and cDNA as a template, respectively. The PCR amplification conditions were as follows: 95°C for 5 min, followed by 28 cycles of 95°C for 30 s each, 60°C for 30 s, and 72°C for

2 min 30 s for gDNA (1 min 30 s for cDNA), followed by 72°C for 10 min, and 4°C for 30 min.

The PCR products derived from gDNA and cDNA amplification were detected by electrophoresis on 0.8% agarose gel and then purified using the TIANquick Midi Purification Kit DP204 (Tiangen). The band was ligated into a pMD-18T vector using the pMD18-T Vector Cloning Kit (TaKaRa, Japan), and the ligation product was used to transform *Escherichia coli* strain DH5 α . Colony PCR was performed with randomly selected positive colonies using CAD-F1/CAD-R1 primers, and the confirmed colonies were sequenced. The correct plasmid was designated pCADg (carrying genomic DNA sequence) and pCADr (carrying cDNA sequence).

Bioinformatic analysis

Sequence homology was analyzed using BLASTn in NCBI. The open reading frame was obtained using ORF Finder in NCBI, which was then translated into a protein sequence. Homologous conserved regions and active sites of protein sequences were analyzed by BLASTp. The exons and introns were identified from the gDNA and cDNA sequences using the AlignX software in VectorNTI. ProtParam in ExPaSy was used to determine amino acid composition, isoelectric points, and hydrophilic/hydrophobic properties.

TMHMM and SignalP provided by Denmark Technology University were used for the identification of transmembrane protein domains and signal peptides. The secondary structure was analyzed using PSIPRED of ExPaSy. The homologous tertiary structure of genes and amino acid sequences was modeled and analyzed using the SWISS-MODEL software. Multiple protein sequence alignments of *EuCAD* and CAD derived from other species were performed using ClustalW, and the phylogenetic tree was constructed using MEGA 5.0 with the neighbor-joining algorithm.

Construction of full-length RNAi vector

Sequencing analysis revealed that pCADr carried the cDNA sequence of the *CAD* gene between the *Sall* and *XbaI* restriction sites, with a *KpnI* restriction site downstream of the *XbaI* site.

The *EuCAD* full-length RNAi fragment was constructed via enzyme digestion and ligation based on the deduced structure (Figure 2), with the following primers: CAD-F2 (5'-gggggtaccactagctctttgagcaaaaatgggcagctcttg-3') and CAD-R2 (5'-gggtctagagggaaaggacaaaac taatcaagc- 3'), using the constructed cDNA sequence of the *CAD* gene, and *KpnI*, *XbaI*, and *SpeI* restriction sites (underlined in sequence).

The *CAD* gene was amplified under the specified conditions using CAD-F2/CAD-R2 primers and pCADr as the template. Amplification products were purified using Kit DP204 (Tiangen) and digested with *KpnI* and *XbaI* endonucleases. The pCADr plasmid was extracted using the TIANpure Mini Plasmid Kit DP104 (Tiangen, China) and digested with *KpnI* and *XbaI*, followed by purification with Kit DP204 (Tiangen), respectively. The products were ligated and used to transform the *E. coli* strain DH5 α .

Colony PCR was performed using the CAD-R2 primer and the universal primer RV-M of the Pmd18-T Vector. Positive colonies were sequenced using the universal primers RV-M/M13-47. The correct plasmid was designated as pCADi.

The plant expression vector pBDRG (Zhang et al., 2008) was extracted using the TIANpure Mini Plasmid Kit DP104 (Tiangen, China) and digested with *SpeI* and *Sall*.

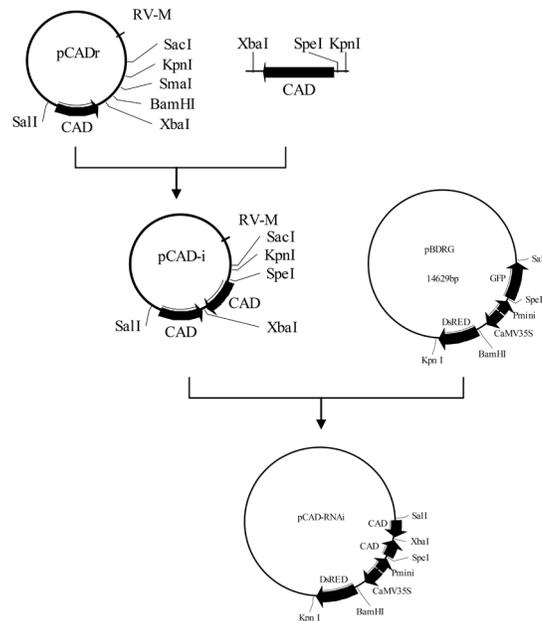


Figure 2. Schematic of pCAD-RNAi showing the design of pCADi and pCAD-RNAi.

The digestion product was detected via electrophoresis on 0.8% agarose gel, and the separated bands were purified by the Kit DP214 (Tiangen).

The pCADi plasmid was extracted using the Kit DP104 (Tiangen) and digested with *SpeI* and *SalI*. The digestion product was detected via electrophoresis on 0.8% agarose gel, and the RNAi band was cut and purified with the Kit DP214 (Tiangen).

The digested pBDRG vector was ligated to the RNAi fragment, and the ligation product was used to transform *E. coli* strain DH5 α . Colony PCR was performed using the CAD-F2/CAD-R2 primers. The positive colonies were sequenced using the primers conGFP-F and conGFP-R of pBDRG. The correct plasmid was designated as pCAD-RNAi.

Generation of transgenic tobacco

pCAD-RNAi was introduced into *Agrobacterium* LBA4404 according to the method described by Han (2013). The *Agrobacterium* was inoculated into LB liquid medium (rifampicin 50 mg/L, kanamycin 50 mg/L), and cultured at 28°C to an OD₆₀₀ of approximately 0.5. Cultures were centrifuged at 4000 rpm for 10 min and resuspended in MS liquid medium. The leaves of aseptic tobacco seedlings were cut into two fragments of 0.5 cm each, soaked in the suspension for 10 min, transferred into co-culture medium, and incubated in the dark to induce callus formation for 3 days at 28°C. The callus was transferred into differentiation medium, cultured at 28°C, and illuminated for 12 h day. One-week later, the callus sprouted the shoot, which was transferred into the strengthening shoot medium, and cultured for 2-3 weeks. Robust seedlings were selected and transferred into rooting medium, and cultured. Strong and healthy roots were selected and transplanted to red soil. The transgenic plants were numbered, and the total DNA extracted from the leaves was used for PCR.

The co-culture medium consisted of MS + 1.5 mg/L 6-BA+ 0.1 mg/L NAA; the differentiation medium comprised MS + 1.5 mg/L 6-BA+ 0.1 mg/L NAA+ 100 mg/L kanamycin + 200 mg/L cefotaxime; the strengthening shoot medium was MS + 0.1 mg/L 6-BA+ 0.01 mg/L NAA + 100 mg/L kanamycin + 200 mg/L cefotaxime; and the rooting medium had MS + 50 mg/L kanamycin+ 200 mg/L cefotaxime.

Quantitative RT-PCR (qRT-PCR)

The root, leaf, young stem, and semi woody stem tissues of GLU4 were collected, and expression of the *CAD* gene was detected. The leaves of transgenic tobacco plants were collected, and the *CAD* expression was detected with the wild type tobacco used as a control.

P-EuCAD-F (5'-taccctgtgggtcacctct-3')/P-EuCAD-R (5'-tcaagcatctcctctgtttct-3') primers were designed based on the *EuCAD* sequence. The internal control primers P-Euactin-F (5'-gcagtctcccagattgttg-3')/P-Euactin-R (5'-tgccatatttctccatgtcgt-3') were designed based on the *actin* gene (AB505624.1), and the only complete *actin* gene sequence in *Eucalyptus* was retrieved from GenBank. The P-EuCAD-F/P-EuCAD-R and P-Euactin-F/P-Euactin-R primers were used to determine the expression of the *EuCAD* gene in GLU4.

P-CAD-F (5'-gtatggcaccagaacaagcag-3')/P-CAD-R (5'-ccaatgctcttctctcttat-3') primers were designed according to the homologous sequences of the two complete *CAD* genes (X62343.1, X62344.1), which were retrieved from GenBank. The tobacco *actin* gene (EU938079) was used as a reference sequence, and P-actin-F (5'-ctggaatccatgagactactacaa-3')/P-actin-R (5'-aaccgcactgagcacaata-3') primers were designed to determine the expression of the *CAD* gene in tobacco in addition to the P-CAD-F/P-CAD-R primers.

Total RNA from sampled tissues was extracted using the RNeasy Pure Plant Kit DP441 (Qiagen) and the cDNA was synthesized by the RNA LA PCR reverse transcription kit (TaKaRa). qPCR was performed using Power SYBR Green PCR Master Mix (ABI, USA). The conditions for qPCR were as follows: 95°C for 2 min, followed by 40 cycles of 95°C for 10 s each, and 60°C for 40 s. Gene expression levels were determined using three biological replicates and three technical replicates for each treatment.

Evaluation of cellulose and lignin content

Stems from the third to seventh node of transgenic tobacco and control (wild type) plants were collected. The leaves were shed from the stems and cut into small pieces, dried at 60°C for 24 h, and triturated to determine lignin and cellulose contents.

The lignin content was determined by the Klason method. Each 1-g sample was extracted using a benzene/alcohol mixture (2:1) and then air dried for 6 h. The sample was transferred into a flask containing 15 mL 75% sulfuric acid cooled to 15°C, and the mixture was incubated at 20°C for 2 h with an occasional vortex and mix. The contents were rinsed with distilled water. The total volume was adjusted to 560 mL using distilled water, which was boiled for 4 h with constant addition of water to maintain the total volume, and incubated until acid-insoluble lignin deposits were obtained. The pellet was filtered using constant weight quantitative filter paper of known weight, and washed with distilled water until the elute was clear when in contact with 10% barium chloride and of a neutral pH. The filter paper was incubated in a 105°C oven to constant weight. The lignin content (%) was calculated using the formula: (lignin residue weight after drying / total dry sample weight) x 100%.

The cellulose content was determined using the nitric acid method. Each 1-g sample was extracted with 25 mL nitric acid ethanol mixture (4:1) in a flask with a reflux condenser, and the flask was placed in a boiling water bath for 1 h with occasional vortexing. The mixture was incubated until the residues were obtained. The resulting supernatant was transferred to an 1G2 glass filter under constant weight, and dried under a vacuum in order to collect the residue in the filter, which was transferred back into the flask. Another 25-mL nitric acid ethanol mixture (4:1) was used to transfer the residue from the filter and mouth of the flask. After repeating the above steps three times, the total flask content was transferred to the filter, washed with 10 mL of the nitric acid/ethanol mixture and rinsed until the pH of the elute was neutral. Finally, after two ethanol rinses, the filtrate was dried and the filter was incubated in an oven at 105°C to constant weight. The cellulose content (%) was determined using the formula: (cellulose residue weight after drying / total dry sample weight) x 100%.

Anatomy of stem tissue

The stem tissue from the fifth to seventh node of 9-month old tobacco seedlings was collected and fragmented, soaked in FAA solution, and held at 4°C. The samples were used to generate paraffin sections, which were subsequently stained with safranin and fast green as previously described (Li, 1987). The slides were visualized using a Nikon research optical microscope, and photographed and analyzed using the NIS-Elements software.

Statistical analysis

Data were statistically analyzed using the SPSS 13.0 software (SPSS, Chicago, IL, USA). Data between the control and RNAi transgenic lines were compared. Graphs were generated using the Origin 8.0 software.

RESULTS

Gene cloning and sequence analysis of *EuCAD*

The primers used in this study were designed based on homology analysis of *Eucalyptus CAD* genes retrieved from GenBank (Table 1), with which the *EuCAD* gene was amplified from *E. urophylla* GLU4 using RT-PCR (Figure 3). The sequencing results showed that the genomic DNA and cDNA of *EuCAD* were 2200- and 1102-bp long, respectively (KF467162 and KF467163), and the sequences are listed in [Figure S1](#).

Homologous alignment demonstrated that the *EuCAD* sequence shared 97% homology with *Eucalyptus CADs* ([Figure S2](#)), and contains an intact conserved CAD1 domain, NADP-binding site of CAD, characteristic NADP-binding site, and two Zn-binding sites, which are consistent with the known enzymatic function of CAD. Using the online tools of TMHMM and SignalP, no transmembrane structure or signal peptide sequence was observed, indicating that the sequence is not part of a membrane protein or secretory protein, and therefore, should be localized to the cytoplasm, which is also consistent with its known function. The alignment of gDNA and cDNA revealed the presence of five exons and four introns in *EuCAD*, and the fourth exon and third intron were the longest.

Analysis of secondary structure with PSIPRED of ExPaSy revealed 8 alpha helices and

16 beta sheet structures within the protein. In addition, the homologous tertiary structure of EuCAD was modeled from the 7th to the 356th amino acid using SWISS-MODEL with the Arabidopsis AtCAD4 used as the modeling template. AtCAD4 and AtCAD5 are two isoforms of ethanol dehydrogenase isolated from *Arabidopsis*, and AtCAD4 is mainly involved in the biosynthesis of lignin monomers (Youn et al., 2006). The sequence identity between the target and reference proteins was 76.86%. The model e-value of $9.65e^{-140}$ suggests high reliability of the model, and therefore, it is reasonable to speculate that EuCAD and AtCAD4 possess catalytic functions.

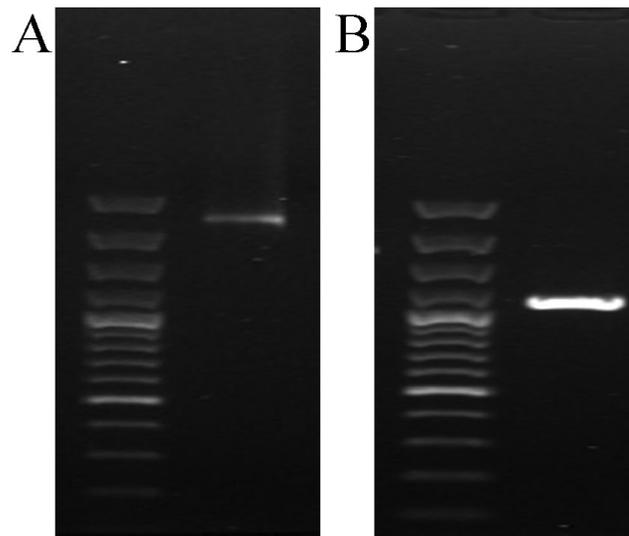


Figure 3. PCR amplification of *EuCAD*. The PCR amplification product of genomic DNA (gDNA) (A) and cDNA (B) of *EuCAD* using CAD-F/CAD-R primers. gDNA and cDNA of *EuCAD* was 2200- and 1102-bp, respectively. Thermo Scientific GeneRuler 100 bp Plus DNA Ladder (SM0321) was used.

Phylogenetic analysis

We retrieved 2428 terrestrial plant CAD sequences from the NCBI protein database, and 29 sequences containing complete coding sequences were selected for phylogenetic analysis.

The results showed that *EuCAD* and CADs from *Eucalyptus* clustered together with CADs from *Citrus sinensis*, *Hevea brasiliensis*, and *Populus*. Phylogenetic analysis of model plants indicated that *EuCAD* clustered with tobacco, and was distinct to *Arabidopsis*, but still shared 76% homology with the latter (Figure 4). Phylogenetic analysis showed that CAD was more evolutionarily conserved, which may be related to its constitutive growth function.

Construction of *EuCAD* full-length RNAi vector

The universal primer RV-M was located 97-bp away from the *KpnI* restriction site. A 1.2-kb fragment was amplified from the positive clone of pCADi using the RV-M and CAD-R2 primers. As shown in Figure 5, the target band was obtained from No.1 clones, which were considered as the positive clone of pCADi. The sequencing results confirmed that a forward CAD sequence was inserted between the *KpnI* and *XbaI* restriction sites, which formed a full-length RNAi fragment.

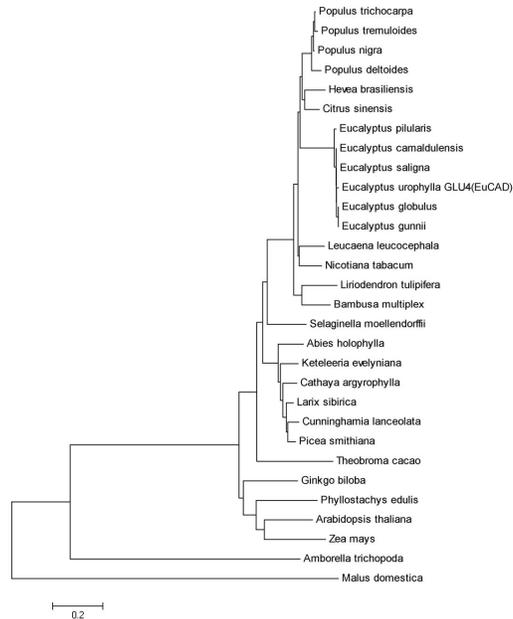


Figure 4. Phylogenetic analysis of EuCAD. Twenty-nine sequences containing complete coding sequences were selected along with EuCAD for phylogenetic analysis by MEGA 5.0 with the neighbor-joining algorithm.

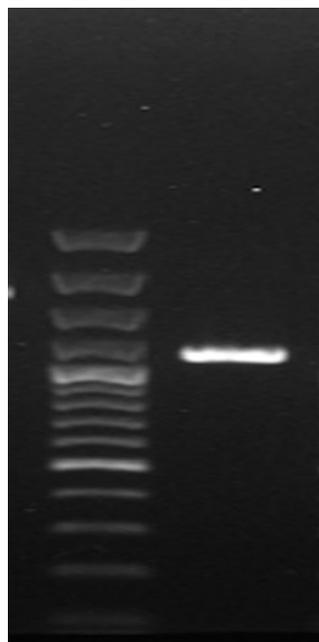


Figure 5. Colony PCR amplification of pCADi. The PCR amplification product of the No. 1 clone using RV-M and CAD-R2 primers, measured approximately 1.2 kb in length. Thermo Scientific GeneRuler 100 bp Plus DNA Ladder (SM0321) was used.

To construct pCAD-RNAi, the 2.2-kb full-length RNAi fragment was obtained from PCADi products following digestion with *SpeI* and *SalI* endonucleases (Figure 6) and introduced into pBDRG. PCR validation showed that target bands of around 1.1 kb were obtained from No. 1, No. 2, and No. 3 clones using the primers CAD-F2/CAD-R2 (Figure 7), suggesting that they were positive clones.

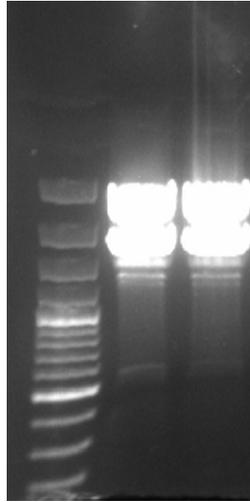


Figure 6. pCADi endonuclease digestion products using *SpeI* and *SalI*. PCADi was digested with *SpeI* and *SalI*. The products include a full-length RNAi fragment measuring approximately 2.2 kb in length, and the vector backbone measuring approximately 2.6 kb in length. Thermo Scientific GeneRuler 100 bp Plus DNA Ladder (SM0321) was used.

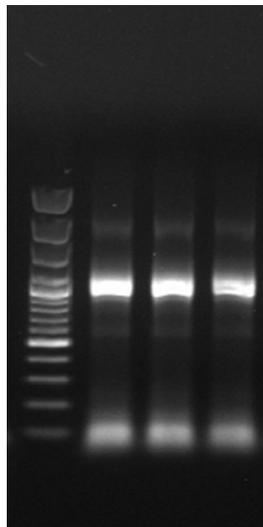


Figure 7. Colony PCR amplification of pCAD-RNAi. The PCR amplification products of No. 1, No. 2, and No. 3 clones using CAD-F2/CAD-R2 primers, measuring approximately 1.1 kb in length. Thermo Scientific GeneRuler 100 bp Plus DNA Ladder (SM0321) was used.

The positive clones were further sequenced using the primers conGFP-F/conGFP-R. The results showed that a full-length RNAi fragment of EuCAD was indeed inserted between the *SpeI* and *SalI* restriction sites, which indicated successful construction of the full-length RNAi vector pCAD-RNAi.

Generation of transgenic tobacco

Two complete CAD genes (X62343.1 and X62344.1) were retrieved from GenBank. Results of homologous alignment showed that the sequences share 76% homology with CADs in tobacco ([Figure S3](#)); therefore, transgenic tobacco was generated by *Agrobacterium*-mediated leaf disk transformation to investigate the function of EuCAD in lignin synthesis (Figure 8).

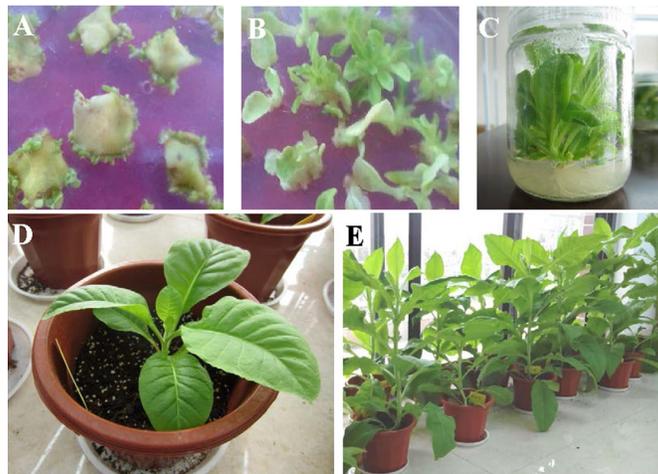


Figure 8. Various stages of transgenic tobacco generation. **A.** After co-cultivation, the leaf callus was transferred into differentiation medium to generate buds. **B.** Buds were transferred to strengthening shoot medium to develop buds. **C.** Transgenic plants were grown on rooting medium containing kanamycin. **D.** Transgenic seedlings grown in the soil after 1 month. **E.** Transgenic plantlets grown in the soil after 4 months.

Total DNA was extracted from seedling leaves and verified by PCR using the primer pair CAD-F1/CAD-R1. The *EuCAD* gene was amplified from the positive transgenic seedling via insertion of the *EuCAD* full-length RNAi fragment, with a length of approximately 1.1 kb. The PCR products derived from lines A3 and A6 were consistent with these expectations (Figure 9), suggesting that A3 and A6 represented transgenic plants. Transgenic and wild-type seedlings were grown in the soil and observed. No significant difference was observed in the growth status (Figure 9) and stem radius between two transgenic lines and the wild-type control (Figure 10).

Expression of *EuCAD* in various tissues of eucalyptus

To determine the expression of *EuCAD* in various tissues of eucalyptus, the root, leaf, young stem, and semi-woody stem tissue were collected from half-year-old seedlings of GLU4 to determine the *EuCAD* mRNA levels.

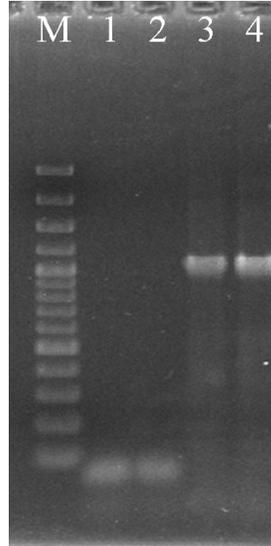


Figure 9. PCR amplification of transgenic tobacco. gDNA was extracted from A3 and A6 leaves, and verified by PCR with the primer pair CAD-F1/CAD-R1. The PCR amplification product of A3 (*lane 3*) and A6 (*lane 4*) was approximately 1.1-kb long. *Lane 1* was the negative control (wild type) and *lane 2* was the negative control with H₂O as template. Thermo Scientific GeneRuler 100 bp Plus DNA Ladder (SM0321) was used.

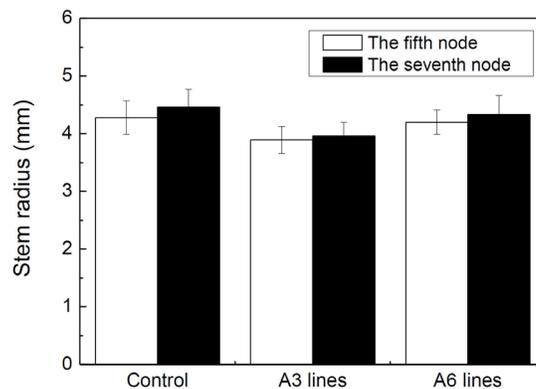


Figure 10. Stem radius in transgenic tobacco. The stem tissue from the fifth and seventh node of 9-month-old tobacco seedlings was used to generate paraffin slides. The stem radius was measured by NIS-Elements. The values are reported as means \pm SE, calculated from three sets of experiments.

The highest expression of *EuCAD* was observed in semi-woody stems (Figure 11), followed by young stems, and roots. The low expression of *EuCAD* in the root, corresponded to the physiological structure of root. The level of lignification was relatively low in the leaf and germinating young stem, resulting from the low expression of *EuCAD*. The plant growth and increased lignification in the stem suggested that the expression of *EuCAD* was significantly enhanced ($P < 0.01$, *t*-test). Overall, the tissue specific expression of *EuCAD* was consistent with its known function.

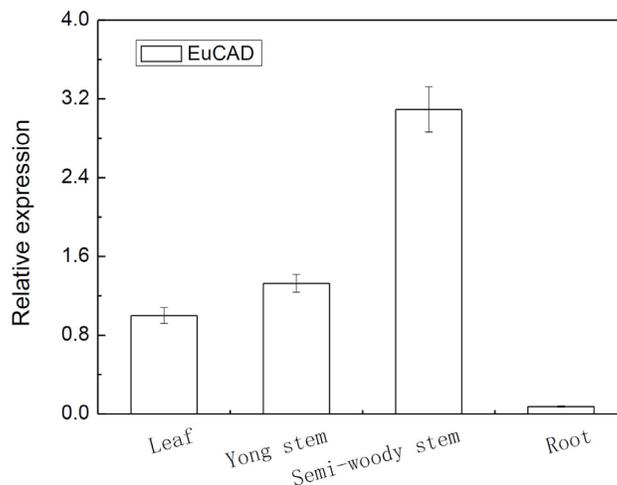


Figure 11. Expression of *EuCAD* in various tissues. The root, leaf, young stem, and semi-woody stem tissue were collected from half-year old seedlings of eucalyptus GLU4 for qRT-PCR. The values are reported as means \pm SE, calculated from three sets of experiments.

Effect of transformation on gene expression

The leaf tissue of both 9-month-old transgenic and wild-type control seedlings was used to determine the level of *EuCAD* transcription. The level of *EuCAD* expression in both transgenic lines was decreased compared with that in the wild type: a 31.17% reduction in line A3 and a 28.84% reduction in line A6 were observed (Figure 12, $P < 0.01$, *t*-test) indicating that tobacco *CAD* gene expression was significantly inhibited by *EuCAD* RNAi.

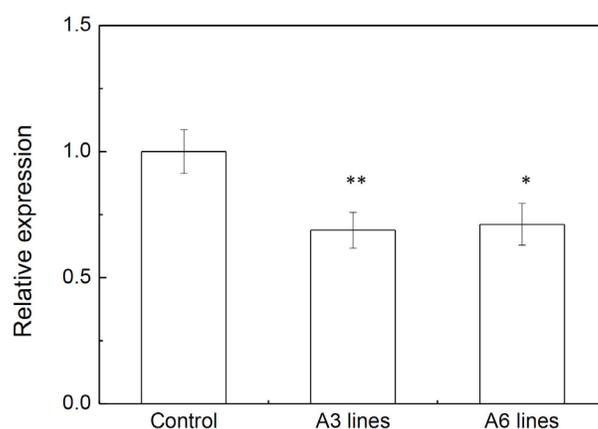


Figure 12. Effect of transformation on gene expression. Transgenic and wild-type seedlings were grown for 9 months, and the leaf tissue was collected to determine the transcript level of *EuCAD*. The values are reported as means \pm SE, calculated from three sets of experiments. Statistical differences between control and RNAi transgenic lines (* $P < 0.05$ and ** $P < 0.01$, independent samples *t*-test).

Effect of transformation on lignin and cellulose levels

The Klason method is a classical method used for the determination of lignin content. It is based on the removal of cellulose with 72% sulfuric acid. The lignin content is calculated following filtration, washing, drying, and weighing. The nitric acid method was used to measure the cellulose content. The samples were treated with nitric acid and ethanol solution. The residue containing lignin, and hemicellulose was dissolved, and the cellulose residue remained.

The content of cellulose and lignin in tobacco stems was determined using the Klason and nitric acid methods. Compared with the wild type, the lignin content in the two transgenic lines decreased sharply while the cellulose content increased significantly (Figure 13, $P < 0.01$, t -test). The lignin reduction and cellulose enhancement were 58.89 and 24.77%, respectively, for line A3, and 59.72 and 28.50%, respectively for line A6 ($P < 0.01$, t -test). The lignin content decreased and the cellulose content increased more significantly in the A3 line than in the A6 line.

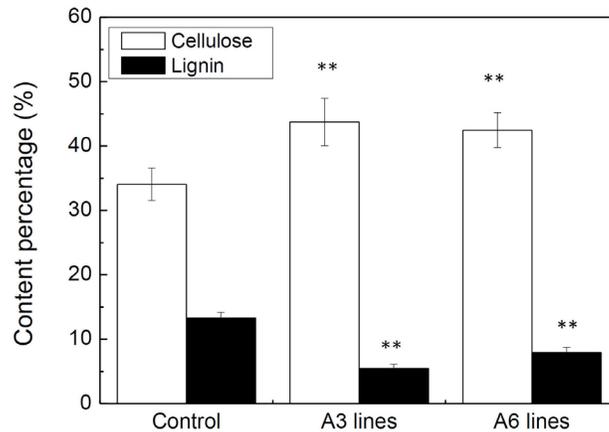


Figure 13. Lignin and cellulose content in transgenic tobacco. Transgenic and wild-type seedlings were grown for 9 months, and the stem tissue was used to determine the cellulose and lignin content. The values are reported as means \pm SE, calculated from three sets of experiments. Statistical differences between control and RNAi transgenic lines (** $P < 0.01$, independent samples t -test).

Anatomy of stem tissue

The microscopic structure suggested a decrease in xylem thickness of A3 and A6 lines, with the variation in xylem thickness in A3 plants being more significant, especially in the seventh quarter (Figure 14). The diameter and thickness of the xylem were measured by the NIS-Elements software, and statistical analyses were performed.

The degree of reduction in xylem thickness in line A3 was higher than that in line A6. The difference in the seventh node was more marked, with a reduction of 44.34% compared with 32.28% in the fifth node, while the decrease of the fifth and seventh nodes was similar in line A6, at 21.14 and 23.15%, respectively (Figure 15, $P < 0.01$, t -test).

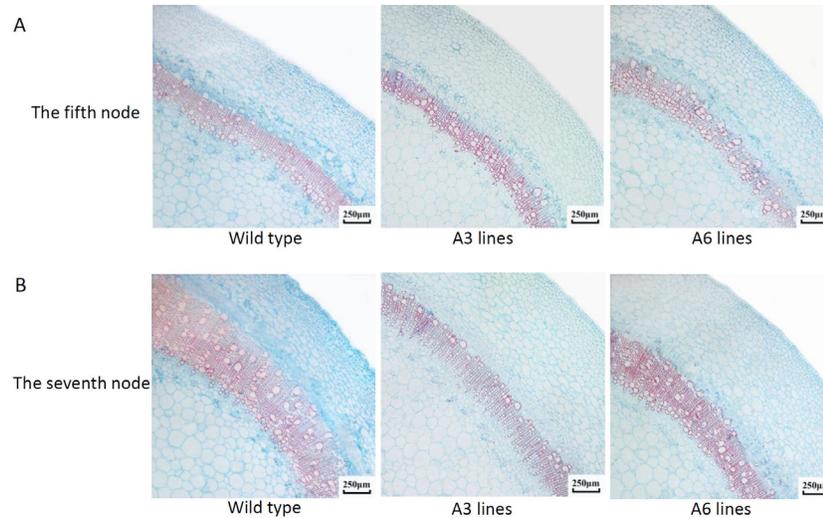


Figure 14. Microstructure of the stem in transgenic tobacco. The slides of the fifth node (**A**) and seventh node (**B**) of tobacco seedlings were stained with safranin and fast green for microscopic observation.

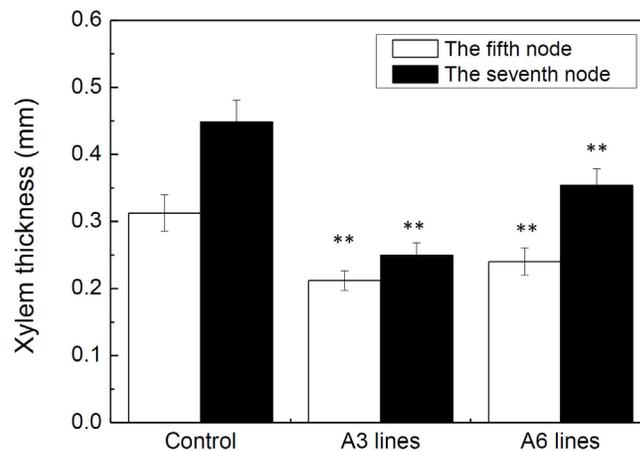


Figure 15. Xylem thickness in transgenic tobacco. The stem tissue from the fifth and seventh nodes of 9-month-old tobacco seedlings was used to generate paraffin slides. Xylem thickness was measured by NIS-Elements. The values are reported as means \pm SE, calculated from three sets of experiments. Statistical differences between control and RNAi transgenic lines (** $P < 0.01$, independent samples t -test).

In the stem, except for the relatively small decrease in the seventh node of the A6 line (18.06%), the two nodes in line A3 and the fifth node of line A6 exhibited a significant reduction of one-third compared with the wild type (Figure 16, $P < 0.01$, t -test). These results show that lignin synthesis was significantly inhibited in transgenic tobacco seedlings transformed with EuCAD.

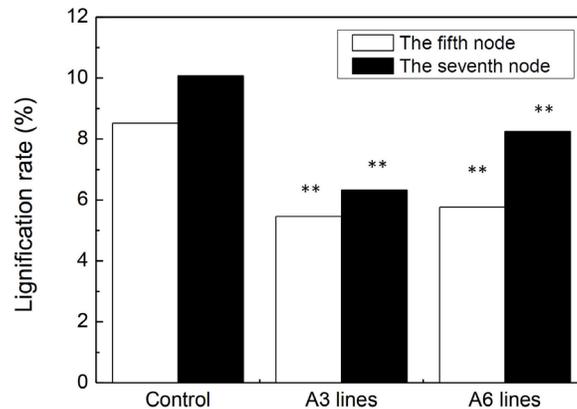


Figure 16. Lignification ratio of the stem in transgenic tobacco. The stem tissue from the fifth and seventh nodes of the 9-month-old tobacco seedlings was analyzed on slides. The xylem thickness and stem radius were measured by NIS-Elements. The lignification ratio represents the ratio of xylem thickness to the area of stem cross-section. The values are reported as means \pm SE, calculated from three sets of experiments. Statistical differences between control and RNAi transgenic lines (** $P < 0.01$, independent samples *t*-test).

DISCUSSION

In this study, the *EuCAD* gene was isolated and bioinformatic analysis of *Eucalyptus* GLU4 was conducted. A full-length RNAi vector was constructed, and the transgenic tobacco plants were obtained by *Agrobacterium*-mediated transformation. A variety of phenotypes were tested in transgenic tobacco plants. Gene expression was determined using primers based on the homologous region of two *CAD* sequences in tobacco. The results showed that tobacco *CAD* expression was significantly decreased. Investigation of the cell-wall components showed that the lignin content decreased significantly, and this was accompanied by a significant increase in cellulose content. Anatomical investigation also revealed that the xylem was significantly thinner in the transgenic plants than in the wild type. Additionally, the growth status and radius of the stems in the transgenic plants were not significantly different from those in the wild type. These results showed that transformation of tobacco with the full-length *EuCAD* RNAi fragment strongly inhibited the expression of *CAD* and affected lignin synthesis. However, it did not affect the growth status of tobacco. These findings are of significance as a standard of reference, although several issues need to be resolved.

RNAi induced by double-stranded RNA specifically inhibits the expression of target genes at the post-transcriptional level. It is a common phenomenon in animals, plants, and microorganisms. Because of the specificity and high efficiency of the regulatory target sequence, RNAi has become the preferred method of post-transcriptional regulation (Abdurakhmonov et al., 2016). Transformation with RNAi fragments designed to target specific genes could trigger RNAi and downregulate target gene expression; however, RNAi cannot remove the gene from genome. Therefore, the target gene is still expressed to an extent, which suggests that the phenotypic change is not as obvious as that observed in deletion mutants.

The length of RNAi fragments used in previous studies varies, suggesting that there are no strict requirements for fragment length. However, the specificity of shorter RNAi fragments is higher, and the nonspecific effect on receptor plants is smaller, and is also easier

to manipulate. In the present study, in order to verify the function of *EuCAD*, the full-length sequence of the gene was used to construct the RNAi vector. In future studies, segmentation of the *EuCAD* sequence and creation of small-fragment RNAi vectors is needed to further identify the core functional regions of the gene.

Similar to other *Eucalyptus* species, a highly homologous *EuCAD* gene was successfully obtained by RT-PCR in the present study. It is not known if other members of the CAD family exist in GLU4. A whole genome search is needed, using a probe designed according to the core functional domain of the *EuCAD* gene. The existence of the CAD gene family in GLU4 cannot be confirmed until further data are available.

Tobacco is a model plant, and is used to study the genetic transformation of plants. It facilitates the understanding of target gene function. Although *EuCAD* and tobacco *CAD* are highly homologous, and RNAi of *EuCAD* significantly inhibited lignin synthesis in tobacco, it is difficult to determine whether a similar approach was effective in GLU4. Tobacco and GLU4 differ physiologically and metabolically as the result of differences between the species. The role of *EuCAD* needs to be verified in the GLU4 genetic transformation system in future studies.

As described previously, despite their high homology, the gene expression pattern, function, catalytic activity, and substrate preference of CAD family members differ, warranting further studies investigating *EuCAD*. The expression pattern of *EuCAD* was initially clear in this study. However, the catalytic activity and substrate preference of the encoding protein are not clear. The catalytic activity and substrate preference of *EuCAD* need to be investigated using prokaryotic expression systems. Determination of the content of coumaryl alcohol, sinapinic alcohol, and coniferyl alcohol in transgenic GLU4 plants facilitates study on the role of *EuCAD* in lignin synthesis.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by funding from the National Natural Science Foundation of China (#31100473) and the Guangxi Forestry Science and Technology Project [(#(2014)033 and #(2016)11)].

REFERENCES

- Abdurakhmonov IY, Ayubov MS, Ubaydullaeva KA, Buriev ZT, et al. (2016). RNA Interference for Functional Genomics and Improvement of Cotton (*Gossypium* sp.). *Front. Plant Sci.* 7: 202. <http://dx.doi.org/10.3389/fpls.2016.00202>
- Barakat A, Bagniewska-Zadworna A, Choi A, Plakkat U, et al. (2009). The cinnamyl alcohol dehydrogenase gene family in *Populus*: phylogeny, organization, and expression. *BMC Plant Biol.* 9: 26. <http://dx.doi.org/10.1186/1471-2229-9-26>
- Battle M, Bender ML, Tans PP, White JW, et al. (2000). Global carbon sinks and their variability inferred from atmospheric O₂ and delta¹³C. *Science* 287: 2467-2470. <http://dx.doi.org/10.1126/science.287.5462.2467>
- Baucher M, Chabbert B, Pilate G, Van Doorselaere J, et al. (1996). Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. *Plant Physiol.* 112: 1479-1490. <http://dx.doi.org/10.1104/pp.112.4.1479>
- Baucher M, Bernard-Vailhé MA, Chabbert B, Besle JM, et al. (1999). Down-regulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (*Medicago sativa* L.) and the effect on lignin composition and digestibility. *Plant Mol. Biol.* 39: 437-447. <http://dx.doi.org/10.1023/A:1006182925584>

- Brill EM, Abrahams S, Hayes CM, Jenkins CL, et al. (1999). Molecular characterisation and expression of a wound-inducible cDNA encoding a novel cinnamyl-alcohol dehydrogenase enzyme in lucerne (*Medicago sativa* L.). *Plant Mol. Biol.* 41: 279-291. <http://dx.doi.org/10.1023/A:1006381630494>
- Chabannes M, Barakate A, Lapierre C, Marita JM, et al. (2001). Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) in tobacco plants. *Plant J.* 28: 257-270. <http://dx.doi.org/10.1046/j.1365-313X.2001.01140.x>
- Chapple C, Ladisch M and Meilan R (2007). Loosening lignin's grip on biofuel production. *Nat. Biotechnol.* 25: 746-748. <http://dx.doi.org/10.1038/nbt0707-746>
- Deng WW, Zhang M, Wu JQ, Jiang ZZ, et al. (2013). Molecular cloning, functional analysis of three cinnamyl alcohol dehydrogenase (CAD) genes in the leaves of tea plant, *Camellia sinensis*. *J. Plant Physiol.* 170: 272-282. <http://dx.doi.org/10.1016/j.jplph.2012.10.010>
- Fan L, Shi WJ, Hu WR, Hao XY, et al. (2009). Molecular and biochemical evidence for phenylpropanoid synthesis and presence of wall-linked phenolics in cotton fibers. *J. Integr. Plant Biol.* 51: 626-637. <http://dx.doi.org/10.1111/j.1744-7909.2009.00840.x>
- Fornalé S, Capellades M, Encina A, Wang K, et al. (2012). Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. *Mol. Plant* 5: 817-830. <http://dx.doi.org/10.1093/mp/ssr097>
- Galliano H, Cabané M, Eckerskorn C, Lottspeich F, et al. (1993). Molecular cloning, sequence analysis and elicitor/ozone-induced accumulation of cinnamyl alcohol dehydrogenase from Norway spruce (*Picea abies* L.). *Plant Mol. Biol.* 23: 145-156. <http://dx.doi.org/10.1007/BF00021427>
- Gross GG, Stöckigt J, Mansell RL and Zenk MH (1973). Three novel enzymes involved in the reduction of ferulic acid to coniferyl alcohol in higher plants: ferulate: Co A ligase, feruloyl-Co A reductase and coniferyl alcohol oxidoreductase. *FEBS Lett.* 31: 283-286. [http://dx.doi.org/10.1016/0014-5793\(73\)80123-1](http://dx.doi.org/10.1016/0014-5793(73)80123-1)
- Halpin C, Knight ME, Foxon GA, Campbell MM, et al. (1994). Manipulation of lignin quality by downregulation of cinnamyl alcohol dehydrogenase. *Plant J.* 6: 339-350. <http://dx.doi.org/10.1046/j.1365-313X.1994.06030339.x>
- Halpin C, Holt K, Chojecki J, Oliver D, et al. (1998). Brown-midrib maize (bm1) - a mutation affecting the cinnamyl alcohol dehydrogenase gene. *Plant J.* 14: 545-553. <http://dx.doi.org/10.1046/j.1365-313X.1998.00153.x>
- Han BT (2013). Cinnamoyl-coenzyme A reductase (CCR) function analysis and regulation of the gene in tobacco: effects on lignin. DOCTOR, Beijing Forestry University.
- Hu WJ, Harding SA, Lung J, Popko JL, et al. (1999). Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. *Nat. Biotechnol.* 17: 808-812. <http://dx.doi.org/10.1038/11758>
- Jouanin L, Goujon T, de Nadaï V, Martin MT, et al. (2000). Lignification in transgenic poplars with extremely reduced caffeic acid O-methyltransferase activity. *Plant Physiol.* 123: 1363-1374. <http://dx.doi.org/10.1104/pp.123.4.1363>
- Kajita S, Katayama Y and Omori S (1996). Alterations in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate: coenzyme A ligase. *Plant Cell Physiol.* 37: 957-965. <http://dx.doi.org/10.1093/oxfordjournals.pcp.a029045>
- Kim SJ, Kim MR, Bedgar DL, Moinuddin SG, et al. (2004). Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA* 101: 1455-1460. <http://dx.doi.org/10.1073/pnas.0307987100>
- Li L, Cheng XF, Leshkevich J, Umezawa T, et al. (2001). The last step of syringyl monolignol biosynthesis in angiosperms is regulated by a novel gene encoding sinapyl alcohol dehydrogenase. *Plant Cell* 13: 1567-1586. <http://dx.doi.org/10.1105/tpc.13.7.1567>
- Li ZL (1987). Plant production technology. The Science Publishing Company, Beijing.
- Ma QH (2010). Functional analysis of a cinnamyl alcohol dehydrogenase involved in lignin biosynthesis in wheat. *J. Exp. Bot.* 61: 2735-2744.
- MacKay JJ, O'Malley DM, Presnell T, Booker FL, et al. (1997). Inheritance, gene expression, and lignin characterization in a mutant pine deficient in cinnamyl alcohol dehydrogenase. *Proc. Natl. Acad. Sci. USA* 94: 8255-8260. <http://dx.doi.org/10.1073/pnas.94.15.8255>
- Mitchell HJ, Hall SA, Stratford R, Hall JL, et al. (1999). Differential induction of cinnamyl alcohol dehydrogenase during defensive lignification in wheat (*Triticum aestivum* L.): characterisation of the major inducible form. *Planta* 208: 31-37. <http://dx.doi.org/10.1007/s004250050531>
- Möller R, Steward D, Phillips L, Flint H, et al. (2005). Gene silencing of cinnamyl alcohol dehydrogenase in *Pinus radiata* callus cultures. *Plant Physiol. Biochem.* 43: 1061-1066. <http://dx.doi.org/10.1016/j.plaphy.2005.11.001>
- Nair RB, Joy RW, 4th, Kurylo E, Shi X, et al. (2000). Identification of a CYP84 family of cytochrome P450-dependent mono-oxygenase genes in *Brassica napus* and perturbation of their expression for engineering sinapine reduction in the seeds. *Plant Physiol.* 123: 1623-1634. <http://dx.doi.org/10.1104/pp.123.4.1623>

- Ralph J, Hatfield RD, Piquemal J, Yahiaoui N, et al. (1998). NMR characterization of altered lignins extracted from tobacco plants down-regulated for lignification enzymes cinnamylalcohol dehydrogenase and cinnamoyl-CoA reductase. *Proc. Natl. Acad. Sci. USA* 95: 12803-12808. <http://dx.doi.org/10.1073/pnas.95.22.12803>
- Ralph J, Akiyama T, Kim H, Lu F, et al. (2006). Effects of coumarate 3-hydroxylase down-regulation on lignin structure. *J. Biol. Chem.* 281: 8843-8853. <http://dx.doi.org/10.1074/jbc.M511598200>
- Saballos A, Ejeta G, Sanchez E, Kang C, et al. (2009). A genomewide analysis of the cinnamyl alcohol dehydrogenase family in sorghum [*Sorghum bicolor* (L.) Moench] identifies SbCAD2 as the brown midrib6 gene. *Genetics* 181: 783-795. <http://dx.doi.org/10.1534/genetics.108.098996>
- Sibout R, Eudes A, Mouille G, Pollet B, et al. (2005). CINNAMYL ALCOHOL DEHYDROGENASE-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of *Arabidopsis*. *Plant Cell* 17: 2059-2076. <http://dx.doi.org/10.1105/tpc.105.030767>
- Tang YH and Li ZH (2002). Early selection in cold-resistant eucalyptus I. Researching summaries on early selection in cold-resistant eucalyptus. *J. Cent. South For. Univ.* 20: 5.
- Tobias CM and Chow EK (2005). Structure of the cinnamyl-alcohol dehydrogenase gene family in rice and promoter activity of a member associated with lignification. *Planta* 220: 678-688. <http://dx.doi.org/10.1007/s00425-004-1385-4>
- Tronchet M, Balagué C, Kroj T, Jouanin L, et al. (2010). Cinnamyl alcohol dehydrogenases-C and D, key enzymes in lignin biosynthesis, play an essential role in disease resistance in *Arabidopsis*. *Mol. Plant Pathol.* 11: 83-92. <http://dx.doi.org/10.1111/j.1364-3703.2009.00578.x>
- Vanholme R, Demedts B, Morreel K, Ralph J, et al. (2010). Lignin biosynthesis and structure. *Plant Physiol.* 153: 895-905. <http://dx.doi.org/10.1104/pp.110.155119>
- Wróbel-Kwiatkowska M, Starzycki M, Zebrowski J, Oszmiański J, et al. (2007). Lignin deficiency in transgenic flax resulted in plants with improved mechanical properties. *J. Biotechnol.* 128: 919-934. <http://dx.doi.org/10.1016/j.jbiotec.2006.12.030>
- Xiang DY, Chen JB, Ye L and Shen WH (2006). Development Present Situation, Problems and Countermeasures of *Eucalyptus* plantation in Guangxi (in Chinese). *Guangxi Forest. Sci.* 35: 7.
- Xie YJ (2003). Primary Studies on Sustainable Management Strategy of *Eucalyptus* Plantation in China. *World Forest. Res.* 16: 6.
- Youn B, Camacho R, Moinuddin SG, Lee C, et al. (2006). Crystal structures and catalytic mechanism of the *Arabidopsis* cinnamyl alcohol dehydrogenases AtCAD5 and AtCAD4. *Org. Biomol. Chem.* 4: 1687-1697. <http://dx.doi.org/10.1039/b601672c>
- Zhang CX, Gai Y and Jiang XN (2008). Construction of a bidirectional promoter and its transient expression in *Populus tomentosa*. *Frontiers Forest. China* 29: 119-122.

Supplementary material

Figure S1. *Eucalyptus urophylla* clone GLU4 cinnamyl alcohol dehydrogenase mRNA, complete CDS (The start and the stop codon are underlined).

Figure S2. The alignment of EuCAD cDNA and EuCAD gDNA with CAD genes of *Eucalyptus*. The start codon and the stop codon are marked in red boxes, and the primers sequence used in cloning are underlined.

Figure S3. The alignment of EuCAD cDNA with CAD genes of tobacco. The primers sequence used in RT-PCR are underlined.