Characterization of *Cinnamyl Alcohol Dehydrogenase* gene family in lignifying tissues of *Tectona grandis* L.f.

Esteban Galeano[1]*, Tarcísio Sales Vasconcelos[1], Helaine Carrer[1]*

[1]Dept. Ciências Biológicas, ESALQ, Universidade de São Paulo, Brazil

* Corresponding author: Esteban Galeano: estebangg18@hotmail.com, Helaine Carrer: hecarrer@usp.br

Abstract

The cinnamyl alcohol dehydrogenase (CAD) enzyme catalyzes the last step of monolignols synthesis in the lignin pathway. *Tectona grandis* (teak) is a tropical tree with high valuable timber. As there is few genetic information about lignin formation in teak, the purpose of this study is to characterize members of *CAD* family in this species. As methodology, PCR amplification using cDNA samples, vector cloning, sequencing, bioinformatics analyses and gene expression studies using real time RT-qPCR were performed. As results, four members (*TgCAD1-TgCAD4*) were obtained. Comparative analyses showed that all of them have conserved residues for catalytic zinc action, structural zinc ligation, NADPH binding and substrate specificity, consistent with the mechanism of alcohol dehydrogenases. Phylogenetic analysis showed that TgCADs are present in three main classes and seven groups. Expression analyses revealed that *TgCAD1* was highly expressed in leaves and could be related with pathogen defense. *TgCAD2* was more expressed in branches and roots. Differently, *TgCAD3* and *TgCAD4* were highly expressed in juvenile and mature sapwood, suggesting a crucial role in wood development and lignin biosynthesis, with tissue-specialized expression profiles. Furthermore, *TgCAD4* could be related with teak maturation for being more expressed in sapwood of mature teak trees. As conclusion, this work is the first to characterize genes of *CAD* family in *Tectona grandis*. These genes could be interesting to develop transgenic plants for basic research and field applications.

Key words: *Tropical tree; gene characterization, phylogeny, relative expression, sapwood.*

Introduction

Lignin, a three-dimensional phenolic heteropolymer composted by *p*-coumaryl (H), coniferyl (G) and sinapyl (S) alcohols has essential roles in plant structural rigidity, pathogen defense and conduction of water and nutrients (Bonawitz and Chapple 2010; Xu et al. 2013; Tang et al. 2014a). It is the major component of secondary cell walls, for cells that will die; it constitutes the last step of cell division, expansion and elongation before cell death (Eudes et al. 2014; Laurichesse and Avérous 2014; Zeng et al. 2014).

Several biotic and abiotic stresses can activate the polymerization process which leads to lignin production (Bonawitz and Chapple 2010; Zeng et al. 2014). In the last decade, lignin has been a target for several studies due to its agricultural and economic importance, suggesting that at least ten enzymes are required for lignin biosynthesis (Xu et al. 2013). Lignin is the final product of the phenylpropanoid pathway, with cinnamyl alcohol dehydrogenase (CAD, EC 1.1.1.195) being the enzyme that catalyzes the conversion of cinnamyl aldehydes to cinnamyl alcohols (Li et al. 2006; Tang et al. 2014a).

The first step of the phenylpropanoid metabolism is the production of cinnamic acid (deamination of phenylalanine by the phenylalanine ammonia-lyase (PAL) enzyme), followed by several hydroxylations and methylations to finally obtained the lignin monomers (Barakat et al. 2009; Bonawitz and Chapple 2010). The lignin monomers are exported to the apoplast after its synthesis and being oxidized by laccases and peroxidases, producing reactive radical species, bimolecular coupling and polymer elongation (Li et al. 2006; Bonawitz and Chapple 2010).

Mansell et al. (1974) were the first to purify the CAD enzyme. Later, several CAD homologues were characterized by

genomic sequencing, even obtaining complete gene families, such as *Arabidopsis* with 9 members (Kim et al. 2004), rice with 12 CAD genes (Tobias and Chow 2005) and *Populus* with 15 members (Barakat et al. 2009; Barakat et al. 2010). *Brachypodium* (Bukh et al. 2012), camellia, (Deng et al. 2013) and melon (Jin et al. 2014) also present several *CAD* genes described. Nearly, 80 *CAD* genes in 35 plants have been studied, usually existing in a multi-gene family and with differential expression during plant development and environmental conditions (Guo et al. 2010; Tang et al. 2014a).

Among different *CAD* gene families, a strong correlation between evolutionary pattern and gene function has been described, both in angiosperms as in gymnosperms (Guo et al. 2010). The Class I bona fide *CAD* has been present in the main lineage of vascular plants since its emergence from lycophytes (Guo et al. 2010). Class II has important roles in lignin biosynthesis and plant stress resistance; the functions of Class II genes are not conserved as bona fide *CAD* (Guo et al. 2010). Class III is a monophyletic clade responsible for plant development but not related with lignin biosynthesis (Guo et al. 2010). Recently, *CAD* gene has been explored through downregulation by reverse genetics in several species, including dicot and monocot plants (Anterola and Lewis 2002; Valério et al. 2003; Vanholme et al. 2010; Bouvier d'Yvoire et al. 2013; Preisner et al. 2014). CAD antisense gene expression in poplar (Pilate et al. 2002; Baucher et al. 2003) and eucalyptus (Valério et al. 2003) trees allowed a suitable woody tissue usage for agro-industrial purposes without compromising tree growth.

Plantation forests are essential for carrying the world's demand for wood and environmental sustainability. Indeed, it appears that biotechnology application to tree improvement will aid in speeding up this process (Bonawitz and Chapple 2010; Eudes et al. 2014). Teak (*Tectona grandis* Linn. F.), from Lamiaceae family, is a deciduous tree with the most valuable commercial timber in the tropics, due to its high durability, dimensional stability and resistance to external environmental factors (Lukmandaru and Takahashi 2008). It is used for furniture, buildings, finishes, cabinets, sleepers, decorative veneers, lamination, house walls flooring, joinery, carpentry, vehicles, mining and shipbuilding (Baillères and Durand 2000; Bhat et al. 2001; Jain and Ansari 2013; Shukla and Viswanath 2014). Natural populations of teak have been found in Laos, Myanmar, Thailand, Java Islands and India, with a worldwide planted area and natural forest of 33,381 ha (2.5 million m^3 of wood) (Kollert and Cherubini 2012). Unfortunately, despite its importance, there is a lack of genetic studies regarding gene expression and characterization (Alcântara and Veasey 2013; Galeano et al. 2014; Galeano et al. 2015). Consequently, to obtain *CAD* members related to lignin biosynthesis in teak, this study amplified, cloned and characterized three partial *CAD* genes (*TgCAD2, TgCAD3 and TgCAD4*) and a complete bona fide *CAD* gene (*TgCAD1*). Their amino acid sequences were then analyzed and clustered by phylogenetical analyses. Finally, expression profiles were performed in several organs focusing on lignified tissues, including sapwood, thus making this the first study to provide several *CAD* members of *T. grandis*.

Material and Methods

Plant material

To amplify the *CAD* gene family, stem secondary xylem was used from 60-year-old trees, located in Piracicaba, São Paulo State, Brazil. The owner of the land (University of Sao Paulo) gave permission to conduct the study on this site. Bark, secondary phloem and vascular cambium (1.5 cm approx.) were removed to reach the sapwood (Supplementary Material 1), which contained functioning vascular tissue. To perform quantitative real-time PCR, in vitro leaf blades and roots of two months were collected, along with branch and stem secondary xylem from 12-year-old trees, and sapwood from 12- and 60-year old trees was collected with a Pressler borer at Diameter at breast height (DBH) (Deepak et al. 2010) (Supplementary Material 1) from fifteen trees in plantations (lat. 22°42'23"S, long. 47°37'7"W, 650 Meters Above Sea Level) located in Piracicaba, São Paulo State, Brazil. All samples were immediately frozen in liquid nitrogen and stored at -80˚C.

RNA extraction and cDNA synthesis

Using mortar and pestle, frozen samples (1.0 gr) were ground using liquid nitrogen, followed by RNA extraction (mixing five samples as one replicate) using the protocol developed by Salzman et al. (1999). Total RNA (2.0 µg) from each sample was treated with DNAse I (Promega). RNA quality was assessed performing an agarose gel and by Nanodrop ND-1000 Spectrophotometer (NanoDrop Technologies Inc., USA). Also, a PCR reaction was carried out to ensure absence of DNA contamination, using the primers and conditions described by Galeano et al. (Galeano et al. 2014) to amplify *TgEF1a* gene. PCR products were analyzed by electrophoresis on a 1 % (w/v) agarose gel with ethidium bromide staining, expecting absence of bands. Treated RNA (1.0 µg) was used to synthesize cDNA using the SuperScript™ III First-Strand Synthesis System for RT-PCR (Invitrogen) according to the manufacturer's instructions.

Amplification of *TgCAD1*

As a first strategy for amplifying *CAD* genes in teak, *CAD* sequences of different plants were grouped. Consequently, clustal alignment (http://www.ebi.ac.uk/Tools/msa/clustalw2) of *Citrus sinensis* (HQ841075.1), *Eucalyptus urophylla* (FN393570.1), *Populus trichocarpa* (XM_002313839.1), *Acacia auriculiformis x Acacia mangium* (EU275981.1), *Arabidopsis thaliana* (Z31715.1), *Bambusa multiplex* (FJ787493.1) and *Picea sitchensis* (BT071217.1) was performed to detect conserved domains of *Cinnamyl alcohol dehydrogenase* (CAD) gene (Supplementary Material 2) in order to manually design the degenerated primers (Supplementary Material 3), followed by a PCR using 100 ng of cDNA of stem secondary xylem from 60-year-old trees. Single fragments corresponding to the expected size were excised, purified using Fragment CleanUp® (Invisorb, USA) and cloned using the CloneJet™ PCR Cloning Kit (Thermo Scientific, USA) and *DH5a*™ competent cells (Life Technologies, USA). Three colonies were sequenced with the 3100 Genetic Analyzer (Applied Biosystems, USA) using the specific primers of the pJET1.2/Blunt vector. Sequences were blasted (blastx),

translated to amino acid sequences (http://web.expasy.org/translate/) and submitted to PFAM (http://pfam.sanger.ac.uk/search) to confirm *CAD* domains followed by the design of specific primers to amplify the internal region (Table 1). Previous analysis of the partial *TgCAD1* sequence showed this gene as *bona fide CAD* (data not showed). To obtain the ends of the *TgCAD1* gene, the 3´ and 5´ RACE System for Rapid Amplification of cDNA Ends Kit (Life Technologies, USA) was performed, using the partial *CAD* sequence. For the 5´- end, the first strand of cDNA was prepared with 5µg total RNA, GSP1 primer (100 nM; 5´-GCA CAC GAG ATC ACT AT TTC A-3´), $MgCl_2$ (2.5 mM), dNTPs (400 µM), DTT (10 mM), SuperScript II RT (200 units), incubated 42 °C for one hour, 70 °C to stop reaction and finally added 1.0 µl RNase mix. SNAP column was used to purify cDNA, followed by an addition of Terminal deoxynucleotidyl transferase (TdT) and dCTP (200 µM) to create binding tails for the abridged anchor primer (AAP). PCR of tailed cDNA was performed using nested GSP2 primer (400 nM; 5´-CCC TGG AAC CAA AGG GTA TT-3´), Abridged Anchor Primer (400 nM), MgCl2 (1.5 mM), dNTPs (200 µM), Taq DNA polymerase (2.5 units) and tailed cDNA (5 µl). For the 3´- end, the first strand of cDNA was obtained using 5µg total RNA, $MgCl_2$ (2.5 mM), dNTPs (500 µM), DTT (10 mM), SuperScript II RT (200 units), incubated 42 °C for one hour, 70 °C to stop reaction and finally added 1 µl RNase H. PCR reaction was performed adding $MgCl_2$ (1.5 mM), dNTPs (200 µM), Taq DNA polymerase (2.5 units), Gene-Specific Primer GSP (200 nM; 5´-TTC ATC AGG TCA GGG GTG AG-3´), Universal Amplification Primer UAP (200 nM; 5´-CUA CUA CUA CUA GGC CAC GCG TCG ACT AGT AC-3´) and cDNA (2 µl). Both 5' and 3' RACE fragments were cloned and sequenced to complete the *TgCAD1* gene.

Amplification of *TgCAD2, TgCAD3, TgCAD4*

As a second strategy to amplify other *CAD* genes in teak, *Populus trichocarpa CAD* gene family was used for being an arboreal plant and having the largest number of members reported so far. Therefore, the cDNA sequences from thirteen *P. trichocarpa CAD* genes were used (*PtrCAD1, PtrCAD2, PtrCAD3, PtrCAD5, PtrCAD6, PtrCAD7, PtrCAD8, PtrCAD9, PtrCAD10, PtrCAD12, PtrCAD13, PtrCAD14, PtrCAD15*) (Supplementary Material 4) (Barakat et al. 2010; Shi et al. 2010) available at JGI (http://genome.jgi.doe.gov/) to design specific primers above the conserved domains (Supplementary Material 5). A pool of cDNA was obtained with stem secondary xylem from 60-year-old trees, branch secondary xylem from 12-year-old trees, leaves and roots from two-month-old in vitro plantlets and used for the amplifications by PCR reactions using 100 ng of the pool. Gel fragments were purified (Fragment CleanUp®, Invisorb, USA), cloned (CloneJet™ PCR Cloning Kit, Thermo Scientific, USA), sequenced (3100 Genetic Analyzer, Applied Biosystems, USA) and analyzed by blastx. Translation (http://web.expasy.org/translate/) and domain evaluation (http://pfam.sanger.ac.uk/search) were also performed.

Characterization of TgCAD1 protein

TgCAD1 was aligned with the initial sequences used to obtain this gene (*Citrus sinensis, Eucalyptus urophylla, Populus trichocarpa, Acacia auriculiformis x Acacia mangium, Arabidopsis thaliana, Bambusa multiplex* and *Picea sitchensis*). Elements of the secondary structure, domains and relevant residues, catalytic and structural sites were found following the annotations described by Tang et al. (Tang et al. 2014a), Jin et al. (Jin et al. 2014), Larroy et al. (Larroy et al. 2002) and Youn et al. (Youn et al. 2006). ProtParam Tool (http://web.expasy.org/protparam/) was used to explore protein characteristics.

Phylogeny and characterization of CAD family in *Tectona grandis*

Domains of TgCAD1, TgCAD2, TgCAD3, TgCAD4, and several members of the *CAD* family from *Arabidopsis thaliana, Oryza sativa, Lolium perenne, Zea mays, Nicotiana tabacum, Medicago sativa, Saccharum officinarum, Picea abies, Pinus taeda, Eucalyptus globulus, Theobroma cacao, Populus trichocarpa* and *Vitis vinifera* were prepared and aligned with Clustal W. Neighbor joining method was used with 10,000 bootstrap replicates, Poisson Model and Pairwise Deletion, all in Mega 6 program, in order to build the phylogenetic tree. For the dendrogram, classes were characterized following the methods of Guo et al. (Guo et al. 2010) and Barakat et al. (Barakat et al. 2010), and groups (colored) were defined following the methods of Barakat et al. (Barakat et al. 2010). Secondary structure elements, important relevant residues, and catalytic and structural sites were found according to the methods described by Tang et al. (Tang et al. 2014a), Jin et al. (Jin et al. 2014), Larroy et al. (Larroy et al. 2002) and Youn et al. (Youn et al. 2006) annotations. Peptide signals and protein subcellular localization was performed following the method described by Jin et al. (Jin et al. 2014), using Signal IP4.1 program (http://www.cbs.dtu.dk/services/SignalP/) and CELLO program v.2.5 (http://cello.life.nctu.edu.tw/), respectively.

Gene expression of the *CAD* family in Tectona grandis by qRT-PCR

Primers for qRT-PCR (Supplementary Material 6) were designed above *TgCAD1, TgCAD2, TgCAD3* and *TgCAD4* sequences, using OligoPerfectTM Designer (Life Technologies, USA). Primer specificity was evaluated with the melting curve and the amplification efficiencies with the standard curve using three cDNA dilutions from the leaf sample (Supplementary Material 7). The mix for qRT-PCR contained cDNA from each sample (125 ng), primers (50 µM), SYBR Green PCR Master Mix (Applied Biosystems, USA) (12.5 µl), finally adding water for a total volume of 25 µl. The StepOnePlus™ System (Applied Biosystems, USA) was used for the PCR reactions as follows: 2 min at 50˚C, 2 min at 95˚C, 45 cycles of 15 s at 95˚C, 1 min at 65˚C, using 96-well optical reaction plates (Applied Biosystems, USA). Each reaction was done with technical replicates, with negative control (absence of template); leaf was used as calibrator and *EF1α* as control gene (Galeano et al. 2014). Statistical analyses were performed in SAS at 95 % confidence level using F-test for ANOVA and LSD for pair comparison.

Results

Amplification of *TgCAD1*

Although lignified tissues are difficult to macerate, protocol from Salzman et al. (Salzman et al. 1999) was efficient to perform RNA extraction from stem secondary xylem. An EST fragment of 985 bp was amplified using the primers "TgCAD1f" and "TgCAD1r" (Supplementary Material 3), performing the standard RT-PCR process, followed by sequencing the internal portion of the gene using internal primers "inTgCAD1f" and "inTg-CAD1r"(Supplementary Material 3). After verifying the domains of the partial sequence, the principle of the Rapid Amplification of cDNA ends (RACE) was used. It is based in the amplification of the 5′- and 3′- ends of the gene from the partial cDNA using special adapters in the 5′CAP and poli(A) regions and designing specific primers. Therefore, from the partial *TgCAD1* EST, we obtained the complete sequence of *TgCAD1* (NCBI Accession number KX058438), composed by 1071 bp, which could be translated into a protein sequence encoding 357 amino acid residues using ExPASy tool.

Amplification of *TgCAD2, TgCAD3, TgCAD4*

From the primers designed above using thirteen *CAD* genes from *Populus trichocarpa* (Supplementary Material 4 and 5), three *CAD* members in teak were amplified, called as *TgCAD2, TgCAD3* and *TgCAD4* (Supplementary Material 8 and 9) from *PtCAD2, PtCAD5* and *PtCAD15*, respectively. We obtained 889, 697 and 813 nucleotides for *TgCAD2* (NCBI Accession number KX058439), *TgCAD3* (NCBI Accession number KX058440) and *TgCAD4* (NCBI Accession number KX058441), respectively, which are partial sequences of the gene. The deduced amino acid sequences of TgCAD2, TgCAD3 and TgCAD4 (296, 218 and 250 amino acids, respectively) (Supplementary Material 9) showed the highest homology with *Mimulus guttatus* (86, 79 and 73 % of identity, respectively), followed by *Eucalyptus grandis* (79 % identity with TgCAD2), *Perilla frutescens* (77 % identity with TgCAD3) and *Olea europaea* (72 % identity with TgCAD4) using blastp. The CAD protein sequences of teak were aligned against each other by clustalW, but only TgCAD3 and TgCAD4 showed high identity (68 %). The other comparisons were between 42-48 % of identity.
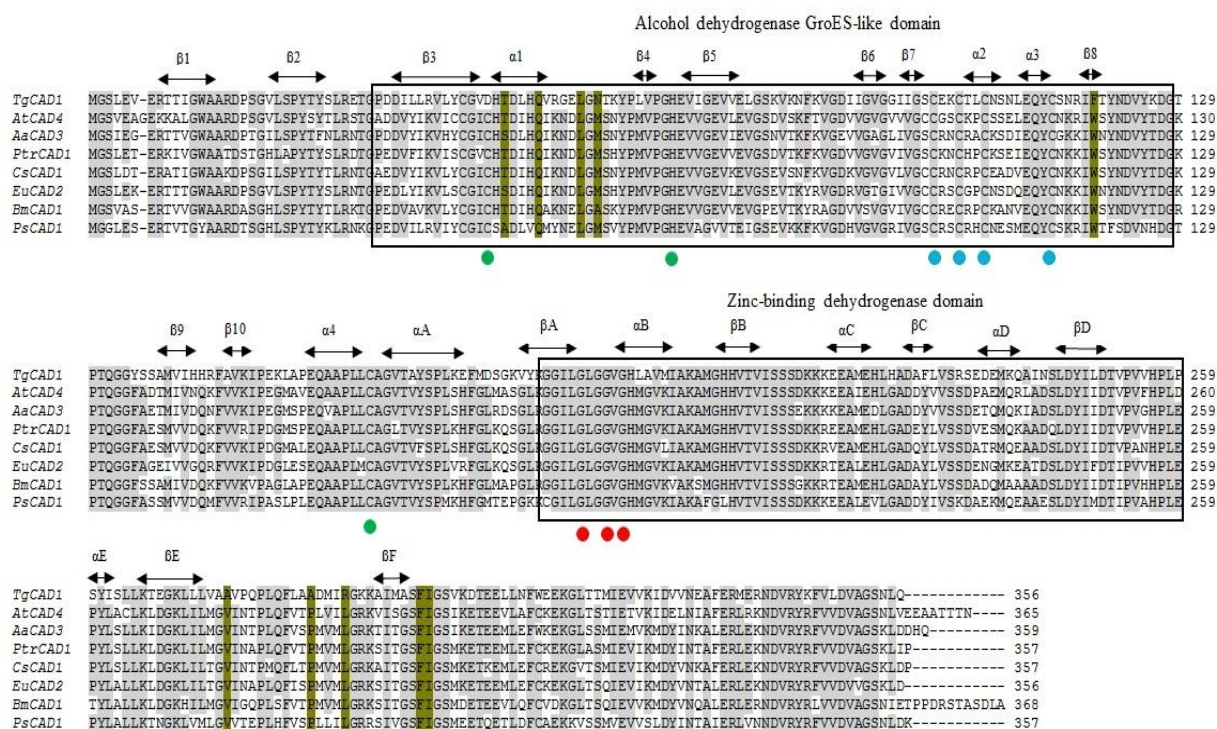

Figure 1

Comparison of amino acid sequences of the Cinnamyl alcohol dehydrogenase gene in the species *Citrus sinensis* (CsCAD1, Genebank accession number ABM67695.1), *Eucalyptus urophylla* (EuCAD2, Genebank accession number ACU77870.1), *Populus trichocarpa* (PtrCAD1, Genebank accession number XP_002313875.1), *Acacia auriculiformis x Acacia mangium* (AaCAD3, Genebank accession number ABX75855.1), *Arabidopsis thaliana* (AtCAD4, Genebank accession number CAP09029.1), *Bambusa multiplex* (BmCAD1, Genebank accession number ADG02378.1), *Picea sitchensis* (PsCAD1, Genebank accession number ABK27071.1) and *Tectona grandis* (TgCAD1, Genebank accession number ABK27071.1). Elements of the secondary structures (alpha and beta) are defined above sequences. Domains of nucleotide ligation are squared. Conserved residues of glycine (GXGGXG) are determined by red dots. Identical amino acids are shaded with grey. Conserved residues which determine substrate ligation are shaded in brown. Green dots define places of catalytic zinc action while blue dots define cysteine residues involved in the structural zinc ligation.
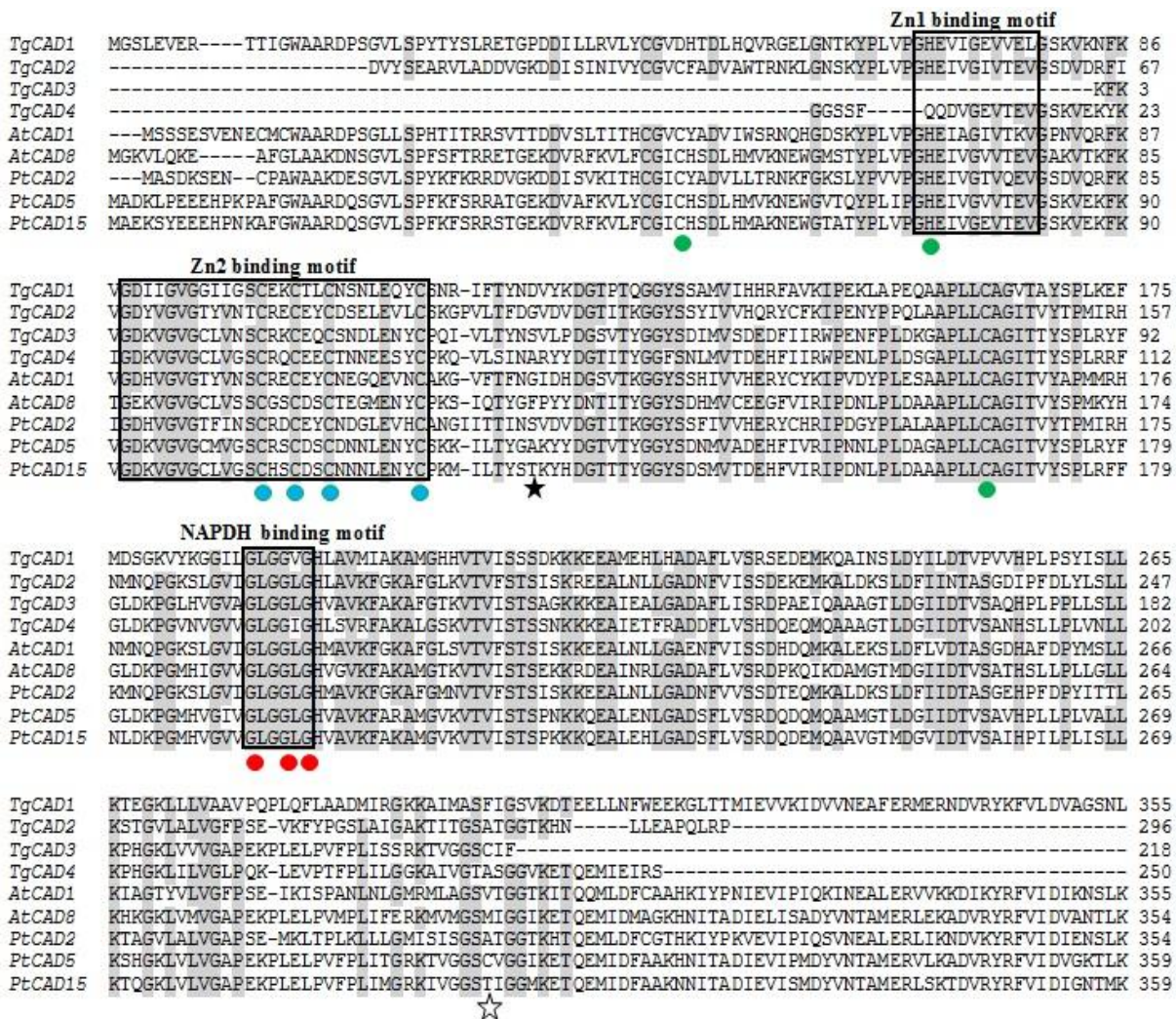
Figure 2

Amino acid sequence alignment of the *Tectona grandis* CAD proteins and other experimentally proved CADs from *Arabidopsis thaliana* (Genebank database) and *Populus trichocarpa* (JGI database). AtCAD1 (AY288079), AtCAD8 (AY302080), PtCAD2 (LG_XVI0159), PtCAD5 (LG_XVI2049), PtCAD15 (LG_IX000475). Multiple alignments were performed with the ClustalW2 software program. Zn1, Zn2, and NADPH binding motifs are squared. Conserved glycine residues for substrate specificity (GXGGXG) are determined by red dots. Identical amino acids are shaded with grey. Green dots define coordinating residues of catalytic zinc ion action while blue dots define coordinating cysteine residues involved in the structural zinc ion ligation. White star indicates key Phe299 residue for substrate specificity and black star indicates key Asp123 residue for substrate binding. Most of the alignment information was identified according to the results described by Jin et al. (Jin et al. 2014) and Tang et al. (Tang et al. 2014a).

Characterization of TgCAD1 protein

Using Protparam tool from ExPASy, TgCAD1 presented 38.96 kDa of molecular mass, a theoretical pI of 5.7 and a protein formula of $C_{1739}H_{2771}N_{459}O_{519}S_{17}$. Also, the CAD1 protein from teak was classified as stable (instability index of 27.07) and positive value for hydrophobicity (hydropathicity value of -0.020). Hydrophobic and hydrophilic residues represented 44.4 % and 30.9 %, respectively. The remaining residues were neutral. *TgCAD1* exhibited the alcohol dehydrogenase GroES-like domain" and the "Zinc-binding dehydrogenase" domains (Figure 1). With blastp, TgCAD1 protein showed highest identity with *Ricinus communis* and *Prunus mume* (71 %) followed by *Vitis vinifera* and *Jatropha curcas* (70 %). When the *TgCAD1* gene was aligned with *CADs* from *Arabidopsis thaliana*, *Citrus sinensis*, *Eucalyptus urophylla*, *Populus trichocarpa*, *Acacia auriculiformis x Acacia mangium*, *Bambusa multiplex* and *Picea sitchensis*, several motifs were conserved among all species, although some features appear to be specific to teak, which are described later. TgCAD1 protein presented a dimer with two zinc ions per subunit (Youn et al. 2006) when compared with AtCAD4 and AtCAD5, with several amino acids involved in catalytic and structural zinc ligation (Larroy et al. 2002).
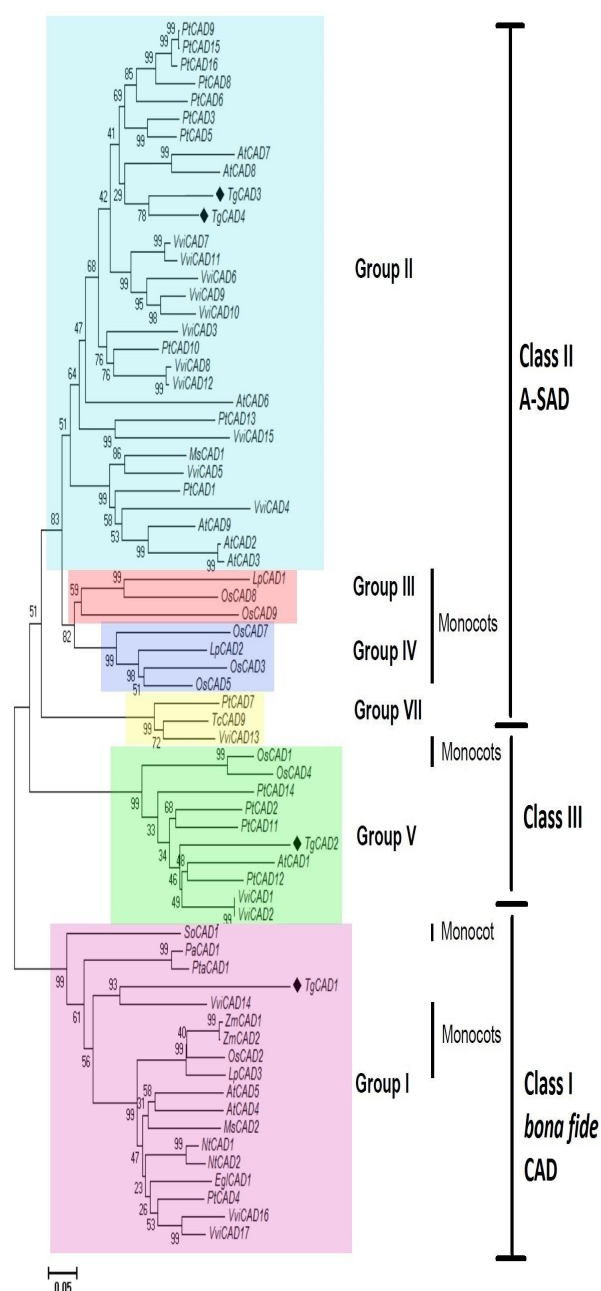
Figure 3

Phylogenetic tree of monocots, dicots and *Tectona grandis* CAD proteins. The neighbor-joining method was used with 10000 bootstraps. Teak CAD proteins are represented with a diamond: TgCAD1 (KX058438), TgCAD2 (KX058439), TgCAD3 (KX058440), TgCAD4 (KX058441). The bar indicates the evolutionary distance of 0.05. NCBI Accession numbers of sequences used to build the tree are *Arabidopsis thaliana:* AtCAD1 (AY288079), AtCAD2 (AY302077), AtCAD3 (AY302078), AtCAD4 (AY302081), AtCAD5 (AY302082), AtCAD6 (AY302075), AtCAD7 (AY302079), AtCAD8 (AY302080) and AtCAD9 (AY302076); *Oryza sativa*: OsCAD1 (AAN09864), OsCAD2 (BK003969), Os-CAD3 (AAP53892), OsCAD4 (BK003970), OsCAD5 (BK003971), OsCAD7 (CAE05207), OsCAD8 (BK003972) and OsCAD9 (AAN05338); *Lolium perenne*: LpCAD1(AAL99535), LpCAD2 (AAL99536) and LpCAD3 (AAB70908); *Zea mays*: ZmCAD1

(AJ005702) and ZmCAD2 (Y13733); *Nicotiana tabacum*: NtCAD1 (X62343) and NtCAD2 (X62344); *Medicago sativa*: Ms-CAD1 (AAC35846) and MsCAD2 (AAC35845); *Saccharum officinarum*: SoCAD1(CAA13177); *Picea abies*: PaCAD1(CAA51226); *Pinus taeda* PtaCAD1(CAA86072); *Eucalyptus globulus*: Egl-CAD1 (AF038561); *Theobroma cacao*: TcCAD9 (EOY15101.1). *Populus trichocarpa* (JGI database): PtCAD1 (estExt_fgenesh4_pg.C_LG_I2533), PtCAD2 (estExt_fgenesh4_pg.C_ LG_XVI0159), PtCAD3 (estExt_fgenesh4_pm.C_LG_VI0462), PtCAD4 (estExt_Genewise1_v1.C_LG_IX2359), PtCAD5 (estExt_Genewise1_v1.C_LG_XVI2049), PtCAD6 (eugene3.00011775), PtCAD7 (eugene3.00020162), PtCAD8, (eugene3.00091019) PtCAD9 (eugene3.20690001), PtCAD10 (grail3.0004034803), PtCAD11 (gw1.VI.1869.1), PtCAD12 (gw1. XI.816.1), PtCAD13 (LG_I002927), PtCAD14 (LG_III001697), PtCAD15 (LG_IX000475) and PtCAD16 (LG_IX000970). *Vitis vinifera* (Plant Genome Database): VviCAD1 (TA39092_29760_ D1b), VviCAD2 (GSVIVP00000463001), VviCAD3 (GS-VIVP00002954001), VviCAD4 (GSVIVP00008718001), VviCAD5 (GSVIVP00008719001), VviCAD6 (GSVIVP00011478001), VviCAD7 (GSVIVP00011479001), VviCAD8 (GS-VIVP00011484001), VviCAD9 (GSVIVP00011638001), VviCAD10 (GSVIVP00011639001), VviCAD11 (GSVIVP00011640001), VviCAD12 (GSVIVP00013365001), VviCAD13 (GS-VIVP00014356001), VviCAD14 (GSVIVP00019568001), VviCAD15 (GSVIVP00029747001), VviCAD16 (GS-VIVP00036661001) and VviCAD17 (GSVIVP00036664001). Classes were identified according to the results described by Guo et al. (Guo et al. 2010) and Barakat et al. (Barakat et al. 2010). Groups (colored) were defined according to the results described by Barakat et al. (Barakat et al. 2010).

Therefore, the TgCAD1 nucleotide binding domain was composted by six parallel β-sheet chains (βA to βF) flanked by five α-helixes (αA to αE), while the catalytic domain consisted in a core of antiparallel β-sheet chains (β1 to β9) with six helical segments (α1 to α6) (Figure 1).

Phylogeny and characterization of CAD family in *Tectona grandis*

We performed ClustalW with all of the TgCAD proteins with the amino acid sequences of *Populus trichocarpa* used to amplify *TgCAD* genes (PtCAD2, PtCAD5, PtCAD15) and two *Arabidopsis thaliana* protein sequences (AtCAD1, AtCAD8). Later, we found several conserved characteristics in teak proteins previously identified by Jin et al. (Jin et al. 2014) and Tang et al. (Tang et al. 2014a). TgCAD1 and TgCAD2 exhibited the Zn1 binding motif GHEx$_2$Gx$_5$Gx$_2$V and the catalytic sites for zinc ion action (C47, H69 and C163 with TgCAD1 as reference) (Figure 2). All the four CAD proteins of *T. grandis* had the Zn2 structural motif with the coordinating cysteine residues for zinc ion ligation (C100, C103, C106 and C114 with TgCAD1 as reference) and the NADPH coenzyme binding motif GLGG[VL]G (usually called Rossmann Fold) (Figure 2). Those motifs defined TgCADs as alcohol dehydrogenases (zinc-dependent), belonging to the medium-chain dehydrogenase/redutase family. Only TgCAD1 presented the Phe299 and Asp123 residues as determinants of substrate specificity and binding, respectively (Jin et al. 2014).

The phylogenetic tree revealed seven groups and three classes of CAD with all the sequences used, including the localization of Monocots (Figure 3). Classes and groups were identified according to the classification described by Guo et al. (Guo et al. 2010) and Barakat et al. (Barakat et al. 2010), respectively. TgCADs were classified in three groups and three classes, supported by high bootstrap values (Figure 3). TgCAD1 was in the group I, Class I (bona fide CAD). TgCAD2 was in the group V class III. TgCAD3 and TgCAD4 belonged to group II class II (A-SAD).

TgCADs expression in different tissues

To explore the expression levels of the four *TgCADs*, teak leaves, roots, stem secondary xylem from 12- and 60-year-old trees and branch secondary xylem from 12-year-old trees were collected. By quantitative real-time PCR, different expression levels of all *TgCAD* genes were detected for all tissues sampled (Figure 4). *Ef1α* was the endogenous control for this experiment (Galeano et al. 2014). *TgCAD1* and *TgCAD3* were expressed at very low level in tissues, and *TgCAD1* showed lower expression level in leaf and stem secondary xylem at both ages. Also, *TgCAD1* is three-fold more expressed in leaves than root and branch secondary xylem and ten-fold more expressed than stem secondary xylem at both ages (Figure 4). Distinctively, *TgCAD2* is almost six-fold more expressed in branch secondary xylem and root than leaf and stem secondary xylem. *TgCAD3* and *TgCAD4* genes presented similar expression, with more expression in roots compared to the other tissues. In addition, Figure 4 shows that *TgCAD4* had higher expression in all tissues. *TgCAD4*, particularly, exposed almost 160-fold more expression compared to the rest of the tissues, with weak transcript levels. Regarding expression of this gene family in secondary xylem, *TgCAD2* and *TgCAD4* showed the highest expression level, followed by *TgCAD3* and *TgCAD1*, although the last three genes did not show statistical significance with the F-test. *TgCAD4* had 20-fold more expression level in branch secondary xylem than leaf.

Furthermore, for a better understanding of the *TgCADs* genes function, sapwood tissue was collected from mature and young *T. grandis* trees and contrasted with leaves in the quantitative real-time PCR experiments (Figure 5). *TgCAD1* and *TgCAD2* genes had low expression levels in sapwood coming from both plant ages (12- and 60-year-old teak trees) with a statistical grouping difference evaluated by F-test. In contrast, *TgCAD3* and *TgCAD4* presented high expression levels in sapwood in both ages, but only *TgCAD4* showed statistical differences, with 60-year-old trees presenting almost 300-fold and two-fold more expression than leaf and 12-year-old trees, respectively. *TgCAD3* presented approximately 9-fold more expression in sapwood from mature than young teak trees.

Discussion

Characterization of *CAD* gene family

In our study, a bona fide *CAD* cDNA of 1071 bp and a protein of 357 amino acids was obtained from *Tectona grandis*. The Rossmann foldable structure of βαβ from TgCAD1 was similar to previous predictions in Arabidopsis thaliana AtCAD5, *Sorghum bicolor* SbCAD4 and *Ginkgo biloba* GbCAD1 (Youn et al. 2006; Sattler et al. 2009; Cheng et al. 2013; Tang et al. 2014a). Also, SKL motif is usually present in the carboxy terminus of several CAD proteins (Tobias and Chow 2005), it is a target sequence for peroxisomes and it plays a role in subcellular localization. Our amino acid sequence analysis found no homology between the SKL residues of TgCAD1 with other species (Figure 1), advising absence of this protein in the peroxisome, as it was found by Cheng et al. (2013) for *Lolium perenne*. There are ten residues to be involved in stabilizing the aromatic ring of cinnamaldehydes with pi-bonding, related by Youn et al. (Youn et al. 2006) and Tang et al. (Tang et al. 2014a). TgCAD1 showed the presence of T49, Q53, L58, F299 and I300 residues when the sequences were compared in alignment (Figure 1), but differences were found in N60, F119, A276, A286 and R290 residues (from TgCAD1) instead of M60, W119, V276, P286 and L290 residues (from AtCAD4). After using Signal IP 4.1 software, no typical signal peptides were found in any of the *TgCADs*. Subcellular localization prediction of TgCADs (Supplementary Material 10) showed that they may exist in the cytoplasm. Moreover, TgCADs contained the conserved GLGGVG motif, all the Zn-catalytic center, Zn-binding site and key residues for correct activity of monolignols. This observation is consistent with CAD genes of *Pennisetum purpureum* (Tang et al. 2014a), *Brachypodium distachyon* (Bukh et al. 2012; Trabucco et al. 2013), *Brassica chinensis* (Zhang et al. 2010) and *Oryza sativa* (Tobias and Chow 2005). Furthermore, the four teak CAD proteins were part of three groups and three classes using amino acid sequences (Figure 3). Class I is comprised by *bona fide CAD* genes, Class II is comprised by *A-SAD* type genes and Class III is poorly understood (Tang et al. 2014b). *TgCAD1* is in group I and Class I, *TgCAD2* in group V class III. *TgCAD3* and *TgCAD4* are in group II class II. Class II and III are known to contain monocots (groups III and IV in class II and group V in class III) and eudicots angiosperm sequences (Figure 3). It means that their evolution happened before monocot and eudicot differentiation (Barakat et al. 2009; Tang et al. 2014a). Gene duplication in the land plants ancestor of *CAD-like* genes in Class II and III (such as PtCAD10 and AtCAD6) has been described previously (Barakat et al. 2010). *TgCAD1*, a bona fide CAD, was grouped close to AtCAD4 and AtCAD5, and these genes from Arabidopsis participate in pathogen-infected tissues (Tronchet et al. 2010). Several genes involved in lignin biosynthesis belong to Class II, such as OsCAD7 (Li et al. 2009), AtCAD7 and AtCAD8 (Kim et al. 2004) and PtCAD10 (Barakat et al. 2009). *TgCAD3* and *TgCAD4* are presumably involved in lignin biosynthesis. *TgCAD2* belongs to Class III, where important genes such as PtCAD2, AtCAD1, OsCAD1 and OsCAD4 are located (Figure 3). A single *CAD* gene cannot be responsible for the lignin biosynthesis in plants, but rather several members of the family work together and may respond differently according to the external signals. There is a functional redundancy in which some members compensate functions from other genes, due to metabolic networks complexities (Kim et al. 2004; Tang et al. 2014a). It means that some *CAD-like* genes such as *TgCAD2, TgCAD3* and *TgCAD4* could perform *bona fide CAD* gene functions.
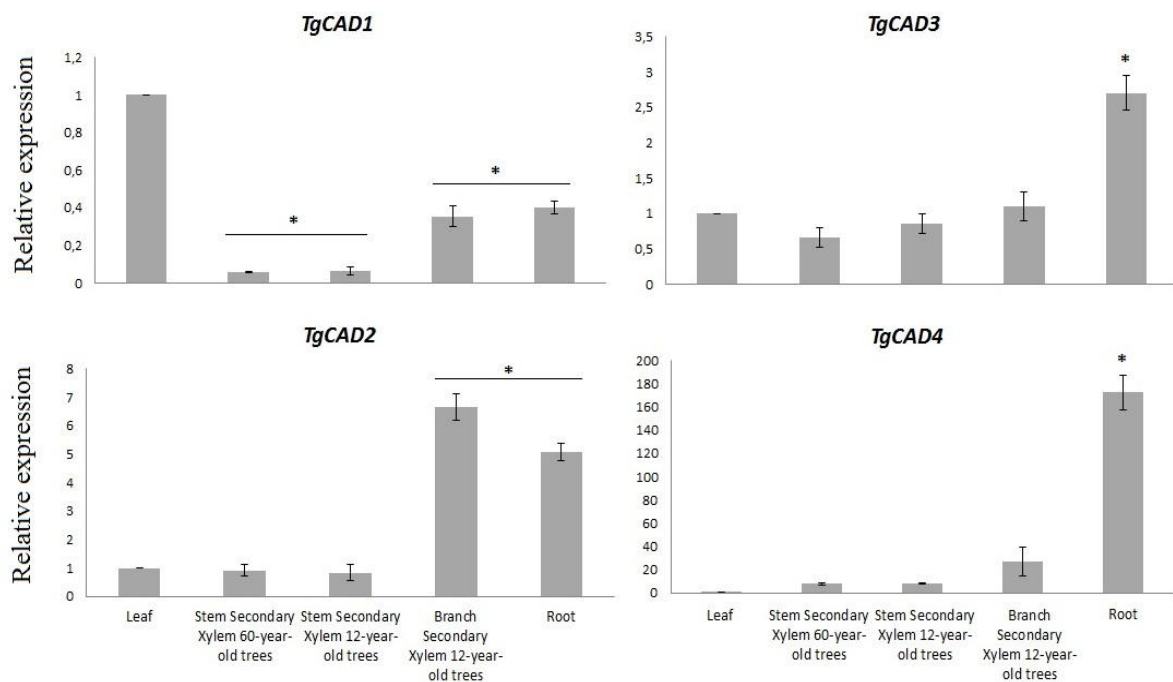
Figure 4
Expression of teak *CAD* gene family. Relative quantification of gene expression was studied in different tissues (leaf, root, stem and branch secondary xylem from different ages), shown at the bottom of the diagrams. The name of each gene is specified at the top of the histograms. ± means SE of the biological triplicates and technical replicates. * shows $p<0.05$ according to F-test. Y axis indicates the relative expression level of each gene compared to the control tissue (leaves). EF1α was the endogenous control used according to the results described by Galeano et al. (Galeano et al. 2014).
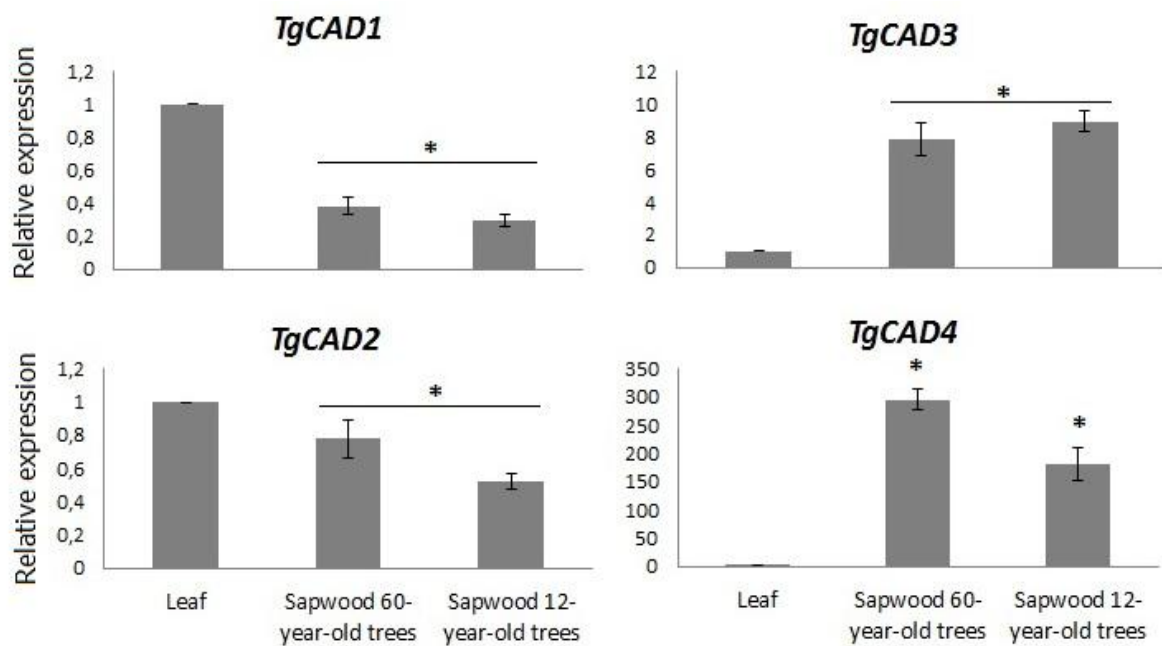

Figure 5
Expression of teak *CAD* gene family in sapwood at different developmental stages. The name of each gene is specified at the top of the histograms. ± means SE of the biological triplicates and technical replicates. * shows $p<0.05$ according to F-test. Y axis indicates the relative expression level of each gene compared to the control tissue (leaves). EF1α was the endogenous control used according to the results described by Galeano et al. (Galeano et al. 2014).

Differential expression of *TgCAD* gene family

Presumably, *TgCAD* genes, which all are expressed in several tissues, especially in root, are involved in distinct functions, such as lignin biosynthesis and secondary cell wall formation. Although *TgCAD1* is part of the *bona fide CAD* group and it is close to *AtCAD4* and *AtCAD5* (*Arabidopsis* genes related to lignin biosynthesis, biotic stress and high expression in root) (Tronchet et al. 2010), it seems to have low relation with lignin biosynthesis due to its low expression in lignified tissues (Figure 4 and 5). Since leaves are susceptible to pathogens and herbivorous insects, some genes such as *AtCAD4* and *AtCAD5* are highly induced by biotic attack (Tronchet et al. 2010). Presumably, *TgCAD1* could be responsible for pathogen defense in leaves and highly expressed with biotic stress, since this gene was clustered with *AtCAD4* and *AtCAD5* and it had high expression in leaves. Consequently, further gene expression studies of *TgCAD1* under pathogen treatment must be done. *TgCAD2* and *TgCAD3* were more expressed in root and branch secondary xylem, while *TgCAD4* showed a significant expression level in root, followed by branch. Curiously, when evaluating expression in sapwood (which is composted primary of secondary xylem), *TgCAD1* and *TgCAD2* had no significant expression levels while *TgCAD3* and *TgCAD4* presented substantial transcripts level at both ages, with *TgCAD4* being 300-fold more expressed in 60-year-old teak trees when compared with leaves. Presumably, *TgCAD4* could be related to teak maturation and secondary wall deposition at latter stages of the tree. In melon, *CmCAD1-3* presented high expression in roots, young stems and during sapwood development (Jin et al. 2014), as *TgCAD3* (Figure 4 and 5). *CmCAD4* was considered a pseudogene, being not expressed or expressed at low levels in lignified tissues (Jin et al. 2014), as observed in *TgCAD1*. Additionally, *Liriodendron tulipifera LtuCAD1* (Xu et al. 2013), B*rachypodium distachyon BdCAD1* (Trabucco et al. 2013; Bouvier d'Yvoire et al. 2013), *Ginkgo biloba GbCAD1* (Cheng et al. 2013) and *Lolium perenne LpCAD1-3* (Lynch et al. 2002) were expressed in lignifying tissues, as observed in *TgCAD4*. In *Populus*, the *PtCAD10* gene, which is close to *TgCAD3* and *TgCAD4* phylogenetically (Figure 3), was 100-times more expressed in xylem compared to other *CAD* genes. However, *PtCAD9*, from the group II, was preferentially expressed in leaves and xylem (Barakat et al. 2009). Above all, *AtCAD8* and *AtCAD7*, the most closely related amino acid sequences to *TgCAD3* and *TgCAD4* (Figure 3) have been suggested to be involved in lignin biosynthesis (Kim et al. 2004). Particularly, several plants exhibit synergistic control over lignin production, such as *AtCAD4* and *AtCAD5* (Trabucco et al. 2013). This could be the case of *TgCAD3* and *TgCAD4*. *PtCAD12*, closely related with *TgCAD2*, presented expression in all tissues but it was greatly expressed in leaves (Barakat et al. 2009). It is noteworthy that the same authors found that *PtCAD2, PtCAD3, PtCAD5, PtCAD6, PtCAD11, PtCAD14* and *PtCAD15* have no expression differences between tissues without treatments. *CAD* gene duplications have been reported due to the nearby clustering between sequences (Barakat et al. 2009; Barakat et al. 2010), as *TgCAD3* and *TgCAD4* seem to be. Notably, *Populus CAD* genes can change their functions and expression profiling under different stress conditions due to innumerable specialized motifs (MeJA, wound, defense responsiveness, ethylene) present in their sequences (Barakat et al. 2009). Therefore, the function of some bona fide *CAD/CAD-like* genes, such as *TgCAD1*, might not be associated with secondary xylem formation but could be induced with biotic or abiotic stress, as it has been observed in other plants, e.g. herbivore damage in Populus leaves (Barakat et al. 2009). In other cases, high expression levels of a *CAD* gene from *Picea sitchensis* have been observed in sapwood compared to bark after pathogen inoculation, meaning a fast response by this tissue (Deflorio et al. 2011). In addition, CAD family genes can present highly divergent functions, as the case of the *Camellia sinensis CAD* members (Deng et al. 2013). Finally, a functional specialization in *CAD* teak family can be explained due to differences in clustering along with the expression profiles in the tissues tested. A duplication in *TgCAD3* and *TgCAD4* genes could have happened, but maintaining similar functions.

Conclusions

The methodology was successful in the identification and characterization of conserved domains of *CAD* genes in teak. *TgCAD1* could be related with pathogen defense and biotic stress, with a very low expression level in lignified tissues when compared with leaves. TgCAD3 and TgCAD4 proteins exhibited a Zn-catalytic center and binding sites along with cofactor binding motif. Also, they are clustered with lignin-related genes and have high relative expression levels in secondary xylem and in sapwood during teak maturation. As a result, they seem to be duplicated and be responsible for the synergistical control of the monolignol biosynthesis. Lignin biosynthesis studies are essential for plant transformation aiming to improve growth rates and adaptability of trees, such as teak. Indeed, functional studies of *TgCAD2, TgCAD3* and *TgCAD4* in *planta* are essential for understanding their relation to secondary xylem formation and cell wall lignin deposition. *CAD* gene family characterization in teak supports the idea that monolignol formation process is conserved among distinct species of vascular plants.

Acknowledgements

References

Alcântara BK, Veasey EA (2013) Genetic diversity of teak (Tectona grandis L. f.) from different provenances using microsatellite markers. Rev Árvore 37:747–758. https://doi.org/10.1590/s0100-67622013000400018

Anterola AM, Lewis NG (2002) Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity. Phytochemistry 61:221–294. https://doi.org/10.1016/s0031-9422(02)00211-x

Baillères H, Durand PY (2000) Non-destructive techniques for wood quality assessment of plantation grown teak. Bois Forêts dês Trop 263:17–29.

Barakat A, Bagniewska-Zadworna A, Choi A, et al (2009) The cinnamyl alcohol dehydrogenase gene family in Populus: phylogeny, organization, and expression. BMC Plant Biol 9:26. https://doi.org/10.1186/1471-2229-9-26

Barakat A, Bagniewska-zadworna A, Frost CJ, Carlson JE (2010) Phylogeny and expression profiling of CAD and CAD-like genes in hybrid Populus (P. deltoides × P. nigra): evidence from herbivore damage for subfunctionalization and functional divergence. BMC Plant Biol 10:100. https://doi.org/10.1186/1471-2229-10-100

Baucher M, Halpin C, Petit-Conil M, Boerjan W (2003) Lignin: Genetic Engineering and Impact on Pulping. Crit Rev Biochem Mol Biol 38:305–350. https://doi.org/10.1080/10409230391036757

Bhat KM, Priya PB, Rugmini P (2001) Characterisation of juvenile wood in teak. Wood Sci Technol 34:517–532. https://doi.org/10.1007/s002260000067

Bonawitz ND, Chapple C (2010) The genetics of lignin biosynthesis: connecting genotype to phenotype. Annu Rev Genet 44:337–63. https://doi.org/10.1146/annurev-genet-102209-163508

Bouvier d'Yvoire M, Bouchabke-Coussa O, Voorend W, et al (2013) Disrupting the cinnamyl alcohol dehydrogenase 1 gene (BdCAD1) leads to altered lignification and improved saccharification in Brachypodium distachyon. Plant J 73:496–508. https://doi.org/10.1111/tpj.12053

Bukh C, Nord-Larsen PH, Rasmussen SK (2012) Phylogeny and structure of the cinnamyl alcohol dehydrogenase gene family in Brachypodium distachyon. J Exp Bot 63:6223–6236. https://doi.org/10.1093/jxb/ers275

Cheng H, Li L, Xu F, et al (2013) Expression patterns of a cinnamyl alcohol dehydrogenase gene involved in lignin biosynthesis and environmental stress in Ginkgo biloba. Mol Biol Rep 40:707–21. https://doi.org/10.1007/s11033-012-2111-0

Deepak MS, Sinha SK, Rao R V (2010) Tree-ring analysis of teak (Tectona grandis L . f .) from Western Ghats of India as a tool to determine drought years. Emirates J Food Agric 22:388–397. https://doi.org/10.9755/ejfa.v22i5.4826

Deflorio G, Horgan G, Woodward S, Fossdal CG (2011) Gene expression profiles, phenolics and lignin of Sitka spruce bark and sapwood before and after wounding and inoculation with Heterobasidion annosum. Physiol Mol Plant Pathol 75:180–187. https://doi.org/10.1016/j.pmpp.2011.02.002

Deng W-W, Zhang M, Wu J-Q, 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 170:272–82. https://doi.org/10.1016/j.jplph.2012.10.010

Eudes A, Liang Y, Mitra P, Loque D (2014) Lignin bioengineering. Curr Opin Biotechnol 26:189–198. https://doi.org/10.1016/j.copbio.2014.01.002

Galeano E, Vasconcelos TS, Ramiro DA, et al (2014) Identification and validation of quantitative real-time reverse transcription PCR reference genes for gene expression analysis in teak (Tectona grandis L.f.). BMC Res Notes 7:464. https://doi.org/10.1186/1756-0500-7-464

Galeano E, Vasconcelos TS, Vidal M, et al (2015) Large-scale transcriptional profiling of lignified tissues in Tectona grandis. BMC Plant Biol 15:221. https://doi.org/10.1186/s12870-015-0599-x

Guo D-M, Ran J-H, Wang X-Q (2010) Evolution of the Cinnamyl/Sinapyl Alcohol Dehydrogenase (CAD/SAD) gene family: the emergence of real lignin is associated with the origin of Bona Fide CAD. J Mol Evol 71:202–18. https://doi.org/10.1007/s00239-010-9378-3

Jain A, Ansari S a. (2013) Quantification by allometric equations of carbon sequestered by Tectona grandis in different agroforestry systems. J For Res 24:699–702. https://doi.org/10.1007/s11676-013-0406-1

Jin Y, Zhang C, Liu W, et al (2014) The cinnamyl alcohol dehydrogenase gene family in melon (Cucumis melo L.): bioinformatic analysis and expression patterns. PLoS One 9:e101730. https://doi.org/10.1371/journal.pone.0101730

Kim S-J, Kim M-R, Bedgar DL, et al (2004) Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in Arabidopsis. Proc Natl Acad Sci U S A 101:1455–60. https://doi.org/10.1073/pnas.0307987100

Kollert W, Cherubini L (2012) Teak resources and market assessment 2010 (Tectona grandis Linn. F.). Rome

Larroy C, Fernandez MR, Gonzalez E, et al (2002) Characterization of the Saccharomyces cerevisiae YMR318C (ADH6) gene product as a broad specificity NADPH-dependent alcohol dehydrogenase : relevance in aldehyde reduction. Biochem J 361:163–172. https://doi.org/10.1042/bj3610163

Laurichesse S, Avérous L (2014) Progress in Polymer Science Chemical modification of lignins : Towards biobased polymers. Prog Polym Sci 39:1266–1290. https://doi.org/10.1016/j.progpolymsci.2013.11.004

Li L, Lu S, Chiang V (2006) A Genomic and Molecular View of Wood Formation. CRC Crit Rev Plant Sci 25:215–233. https://doi.org/10.1080/07352680600611519

Li X, Yang Y, Yao J, et al (2009) FLEXIBLE CULM 1 encoding a cinnamyl-alcohol dehydrogenase controls culm mechanical strength in rice. Plant Mol Biol 69:685–97. https://doi.org/10.1007/s11103-008-9448-8

Lukmandaru G, Takahashi K (2008) Variation in the natural termite resistance of teak (Tectona grandis Linn. fil.) wood as a function of tree age. Ann For Sci 65:708. https://doi.org/10.1051/forest:2008047

Lynch D, Lidgett A, Mcinnes R, et al (2002) Isolation and characterisation of three cinnamyl alcohol dehydrogenase homologue cDNAs from perennial ryegrass (Lolium perenne L .). J Plant Physiol 159:653–660. https://doi.org/10.1078/0176-1617-0776

Mansell RL, Gross GG, Stöckigt J, et al (1974) Purification and properties of cinnamyl alcohol dehydrogenase from higher plants involved in lignin biosynthesis. Phytochemistry 13:2427–2435. https://doi.org/10.1016/s0031-9422(00)86917-4

Pilate G, Guiney E, Holt K, et al (2002) Field and pulping performances of transgenic trees with altered lignification. Nat Biotechnol 20:607–612. https://doi.org/10.1038/nbt0602-607

Preisner M, Kulma A, Zebrowski J, et al (2014) Manipulating cinnamyl alcohol dehydrogenase (CAD) expression in flax affects fibre composition and properties. BMC Plant Biol 14:50. https://doi.org/10.1186/1471-2229-14-50

Salzman RA, Fujita T, Hasegawa PM (1999) An Improved RNA Isolation Method for Plant Tissues Containing High Levels of Phenolic Compounds or Carbohydrates. Plant Mol Biol Report 17:11–17.

Sattler SE, Saathoff AJ, Haas EJ, et al (2009) A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the Sorghum brown midrib6 phenotype. Plant Physiol 150:584–95. https://doi.org/10.1104/pp.109.136408

Shi R, Sun Y-H, Li Q, et al (2010) Towards a systems approach for lignin biosynthesis in Populus trichocarpa: transcript abundance and specificity of the monolignol biosynthetic genes. Plant Cell Physiol 51:144–63. https://doi.org/10.1093/pcp/pcp175

Shukla SR, Viswanath S (2014) Comparative study on growth, wood quality and financial returns of teak (Tectona grandis L.f.) managed under three different agroforestry practices. Agrofor Syst 88:331–341. https://doi.org/10.1007/s10457-014-9686-5

Tang R, Zhang X-Q, Li Y-H, Xie X-M (2014a) Cloning and in silico analysis of a cinnamyl alcohol dehydrogenase gene in Pennisetum purpureum. J Genet 93:145–158. https://doi.org/10.1007/s12041-014-0355-2

Tang X, Xiao Y, Lv T, et al (2014b) High-Throughput Sequencing and De Novo Assembly of the Isatis indigotica Transcriptome. PLoS One 9:e102963. https://doi.org/10.1371/journal.pone.0102963

Tobias CM, Chow EK (2005) Structure of the cinnamyl-alcohol dehydrogenase gene family in rice and promoter activity of a member associated with lignification. Planta 220:678–88. https://doi.org/10.1007/s00425-004-1385-4

Trabucco GM, Matos D a, Lee SJ, et al (2013) Functional characterization of cinnamyl alcohol dehydrogenase and caffeic acid O-methyltransferase in Brachypodium distachyon. BMC Biotechnol 13:61. https://doi.org/10.1186/1472-6750-13-61

Tronchet M, Balagué C, Kroj T, 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. https://doi.org/10.1111/j.1364-3703.2009.00578.x

Valério L, Carter D, Rodrigues JC, et al (2003) Down regulation of Cinnamyl Alcohol Dehydrogenase , a lignification enzyme , in Eucalyptus camaldulensis. Mol Breed 12:157–167.

Vanholme R, Demedts B, Morreel K, et al (2010) Lignin biosynthesis and structure. Plant Physiol 153:895–905. https://doi.org/10.1104/pp.110.155119

Xu Y, Thammannagowda S, Thomas TP, et al (2013) LtuCAD1 Is a Cinnamyl Alcohol Dehydrogenase Ortholog Involved in Lignin Biosynthesis in Liriodendron tulipifera L., a Basal Angiosperm Timber Species. Plant Mol Biol Report 31:1089–1099. https://doi.org/10.1007/s11105-013-0578-z

Youn B, Camacho R, Moinuddin SG a, et al (2006) Crystal structures and catalytic mechanism of the Arabidopsis cinnamyl alcohol dehydrogenases AtCAD5 and AtCAD4. Org Biomol Chem 4:1687–97. https://doi.org/10.1039/b601672c

Zeng Y, Zhao S, Yang S, Ding S (2014) Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. Curr Opin Biotechnol 27:38–45. https://doi.org/10.1016/j.copbio.2013.09.008

Zhang L, Wang G, Chang J, et al (2010) Effects of 1-MCP and ethylene on expression of three CAD genes and lignification in stems of harvested Tsai Tai (Brassica chinensis). Food Chem 123:32–40. https://doi.org/10.1016/j.foodchem.2010.03.122