Lawrence Berkeley National Laboratory

LBL Publications

Title

Expression Pattern of Two Paralogs Encoding Cinnamyl Alcohol Dehydrogenases in Arabidopsis. Isolation and Characterization of the Corresponding Mutants

Permalink

https://escholarship.org/uc/item/6zk8s7g6

Journal

Plant Physiology, 132(2)

ISSN

0032-0889

Authors

Sibout, Richard
Eudes, Aymerick
Pollet, Brigitte
et al.

Publication Date

2003-06-01

DOI

10.1104/pp.103.021048

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial License, available at https://creativecommons.org/licenses/by-nc/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Expression Pattern of Two Paralogs Encoding Cinnamyl Alcohol Dehydrogenases in Arabidopsis. Isolation and Characterization of the Corresponding Mutants[1]

Richard Sibout, Aymerick Eudes, Brigitte Pollet, Thomas Goujon, Isabelle Mila, Fabienne Granier, Armand Séguin, Catherine Lapierre, and Lise Jouanin*

Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, 1055 du PEPS, P.O. Box 3800, Quebec, Canada G1V 4C7 (R.S., A.S.); Biologie Cellulaire, Institut National de la Recherche Agronomique (INRA), 78026 Versailles cedex, France (R.S., A.E., T.G., L.J.); Chimie Biologique, INRA-Institut National d'Agronomie de Paris-Grignon, 78850 Thiverval-Grignon, France (B.P., I.M., C.L.); and Génétique, INRA, 78026 Versailles cedex, France (F.G.)

Studying Arabidopsis mutants of the phenylpropanoid pathway has unraveled several biosynthetic steps of monolignol synthesis. Most of the genes leading to monolignol synthesis have been characterized recently in this herbaceous plant, except those encoding cinnamyl alcohol dehydrogenase (CAD). We have used the complete sequencing of the Arabidopsis genome to highlight a new view of the complete CAD gene family. Among nine *AtCAD* genes, we have identified the two distinct paralogs *AtCAD-C* and *AtCAD-D*, which share 75% identity and are likely to be involved in lignin biosynthesis in other plants. Northern, semiquantitative restriction fragment-length polymorphism-reverse transcriptase-polymerase chain reaction and western analysis revealed that *AtCAD*-C and *AtCAD-D* mRNA and protein ratios were organ dependent. Promoter activities of both genes are high in fibers and in xylem bundles. However, *AtCAD-C* displayed a larger range of sites of expression than *AtCAD-D*. Arabidopsis null mutants (*Atcad-D* and *Atcad-C*) corresponding to both genes were isolated. CAD activities were drastically reduced in both mutants, with a higher impact on sinapyl alcohol dehydrogenase activity (6% and 38% of residual sinapyl alcohol dehydrogenase activities for *Atcad-D* and *Atcad-C*, respectively). Only *Atcad-D* showed a slight reduction in Klason lignin content and displayed modifications of lignin structure with a significant reduced proportion of conventional S lignin units in both stems and roots, together with the incorporation of sinapaldehyde structures ether linked at Cβ. These results argue for a substantial role of *AtCAD-D* in lignification, and more specifically in the biosynthesis of sinapyl alcohol, the precursor of S lignin units.

Lignin is a complex phenolic polymer whose structure is vital to functions such as imparting rigidity to plant organs and as a physical barrier to invading pests. Its presence in cell wall confers to vessels hydrophobic properties that facilitate conduction of water, photo-assimilates, and minerals to different parts of the plant. Lignin structure and composition differ widely at the interspecies level as well as cell types and at the subcellular cell wall level (Donaldson, 2001). Striking differences are mostly observable between gymnosperms and angiosperms. These taxa contain different qualitative and quantitative proportions of monolignols or cinnamyl alcohols representing the main lignin monomers. The formation of cinnamyl alcohols from the corresponding cinnamoyl-CoA esters requires two enzymatic modifications of the carbonate chain of the phenolic precursors. The first step is catalyzed by cinnamoyl CoA reductase, and the second step is catalyzed by cinnamyl alcohol dehydrogenase (CAD). CAD leads to the conversion of hydroxy-cinnamaldehydes to the corresponding alcohols. The relative proportions of these cinnamyl alcohols is an important factor for lignin structural traits and mechanical properties (Baucher et al., 1998; Mellerowicz et al., 2001).

CAD was one of the first enzymes studied in the lignin synthesis pathway (Mansell et al., 1974; Wyrambik and Grisebach, 1975). Since then, many *CAD* cDNAs have been isolated in different plant species (for review, see Dixon et al., 2001). Initially, CAD was believed to be multispecific, catalyzing the reduction of the different cinnamyl-aldehydes. The discovery of isozymes in *Eucalyptus gunii* (Grima-Pettenati et al., 1993), alfalfa (*Medicago sativa*; Brill et al., 1999), and aspen (*Populus tremuloides*; Li et al., 2001) with different affinities for various substrates has lead to the hypothesis that multiple substrate specificities were related to various physiological roles. In parallel, expression of *CAD* cDNAs in *Escherichia coli* or yeast (*Saccharomyces cerevisiae*) was carried out to determine substrate specificity of some cloned *CAD* cDNA, which shared high sequence similarity to known CAD proteins. However, in most

[1] This work was supported in part by GENOPLANTE (grant no. Af1999011) and by the National Biotechnology Strategy of Canada (to A.S.).

* Corresponding author; e-mail jouanin@versailles.inra.fr; fax 33–1–30–83–3099.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.103.021048.

 © 2003 American Society of Plant Biologists

cases, these experiments led to conflicting results. Meanwhile, subsequent in depth analysis suggested unsuspected functions for these proteins (Somssich et al., 1996; Goffner et al., 1998). Together, these studies indicate that, if heterologous protein expression is useful in determining biochemical profiles, other approaches are needed to confirm a biological function. Genetic approaches could be essential in investigating biological roles of a specific enzyme in planta.

This point of view has led to the design of experiments aimed at down-regulating or overexpressing *CAD* genes in transgenic plants to analyze repercussions on lignin content and/or structure. Halpin et al. (1994) obtained tobacco (*Nicotiana tabacum*) lines with low residual CAD activity as a consequence of down-regulation of *CAD*. The xylem of these transgenic plants exhibits a red coloration, and their lignins incorporate cinnamyl-aldehydes (Ralph et al., 1998). Two- and 4-year-old *CAD* antisense transgenic poplars contain less lignins than control plants (Lapierre et al., 1999; Pilate et al., 2002) and show important modifications of their lignin composition (increase of free phenolic compounds and accumulation of sinapaldehyde). Surprisingly, despite a reduction of the sinapyl to coniferyl alcohol ratio, no cinnamaldehydes were detected by thioacidolysis in *CAD* antisense alfalfa (Baucher et al., 1999).

However, the specificity and extent of gene disruption through such gene silencing by antisense or sense strategies sometimes may be difficult to evaluate. Furthermore, expression of gene target paralogs could be damaged. Knockout mutants present an alternative way to determine the role of a gene. Natural mutants of *CAD* have been characterized. First, maize (*Zea mays*) *bm1-2* showed a mutation in the *CAD* gene, resulting in a 20% reduction of lignin content with no alteration of the S to G ratio (Halpin et al., 1998). Second, a loblolly pine (*Pinus taeda* L.) line harboring a mutated allele of the *CAD* gene was identified (MacKay et al., 1997) and characterized (Ralph et al., 1997; Lapierre et al., 2000). This mutant presented a slight reduction of lignin content associated with a modified lignin structure including incorporation of coniferaldehyde and a high level of dihydroconiferyl alcohol, an unusual lignin intermediate.

A straightforward approach to study a complete gene family is now possible with Arabidopsis because its genome is completely sequenced (Arabidopsis Genome Initiative, 2000). In addition, Arabidopsis, despite its small size, is now well considered as a relevant model to study cell wall formation (Reiter, 1998), including that of lignified secondary cell wall (Boudet, 2000; Turner et al., 2001; Goujon et al., 2003a). Thus far, however, few Arabidopsis mutants involved in this pathway have been identified and characterized. Three mutant lines (*fah1*, *ref8*, and *ref3*) for genes encoding cytochrome P450-type enzymes (Chapple et al., 1992; Ruegger and Chapple, 2001; Franke et al., 2002a, 2002b) were identified. Jones et al. (2001) have demonstrated that the *irx4* mutant corresponds to a mutation in the *CCR* gene. The *Atomt1* line mutated in *AtOMT1* gene encoding the enzyme responsible for the methylation of S-unit precursors has been characterized recently (Jouanin et al., 2001; Goujon et al., 2003b).

The search for Arabidopsis *CAD* mutants could constitute a unique opportunity to investigate the *CAD* gene family. Tavares et al. (2000) have listed eight *AtCAD* genes in Arabidopsis and proposed to group them in a multigene family based on nucleotide similarities. We have used the availability of the complete Arabidopsis genome sequence to reexamine the *AtCAD* family. In the present study, we report on the expression patterns of two *AtCAD* gene paralogs and on the isolation of the corresponding mutants (*Atcad-C* and *Atcad-D*). Although consequences of mutations have a strong but different impact on total cinnamyl alcohol activities in several organs, lignin quality is significantly modified in only one mutant.

RESULTS

Predicted Amino Acid Sequences of the *AtCAD* Gene Family and Phylogenetic Analysis

Screening of GenBank entries identified 17 putative *CAD* genes in the Arabidopsis genome (data not shown). Among these putative genes, only nine of the corresponding translated proteins share conserved cofactor and zinc-binding sequences specific for the CAD enzyme. Tavares et al. (2000) previously have listed eight *CAD* genes and named them *CAD1*, *Eli3-2*, and *CADL-A* to *CADL-F*. For clarity, we named the members of this family *AtCAD*. The ninth putative *CAD* gene identified as a result of the complete genome sequencing of Arabidopsis (Arabidopsis Genome Initiative, 2000) was named *AtCAD-G*. *Eli-3-2* (Somssich et al., 1996) was renamed *AtCAD-B2* because of its high identity to *AtCAD-B1*.

Analysis of the nine Arabidopsis CADs at the amino acid level revealed a diversified small family with highly conserved clusters. Only 26% of the amino acids are conserved on an overall total length of 383. However, some *At*CADs are rather closely related, such as *At*CAD-B1/*At*CAD-B2 (85% identity), *AtCAD-C*/*AtCAD-D* (75% identity), and *At*CAD-E /*At*CAD-F (98% identity). In this family, *At*CAD-G is the most distant protein when compared with the others and shares less than 50% identity with the closest groups. When CADs previously identified and studied in other plant species were taken into consideration, phylogenetic analysis based on amino acid sequence comparison showed that Arabidopsis CADs are divided into four subfamilies (Fig. 1). Interestingly, in most cases, at least one CAD previously identified in other plants is present in each of the Arabidopsis subgroups. *At*CAD-1, -E, and -F make up a subfamily with *Msa*CAD-1 (Brill et al.,

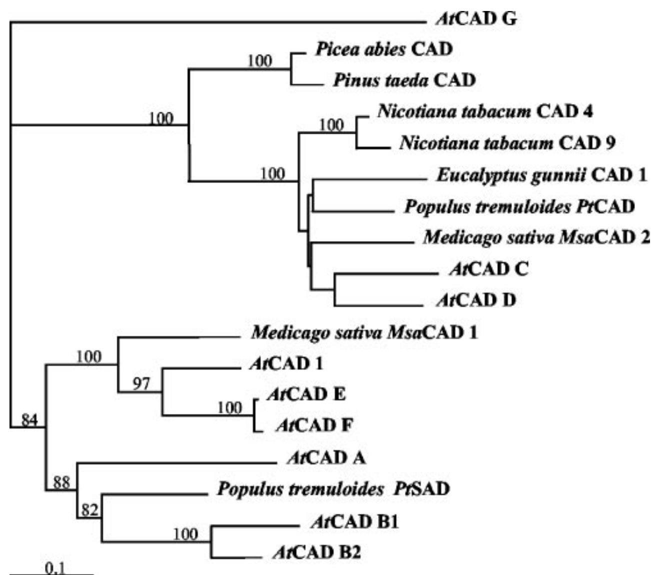

Figure 1. Relationship between the *At*CAD family and sequences of biochemically characterized enzymes of other plant species. The phylogenetic tree was built by neighbor-joining distance using a Kimura matrix (PHYLIP program, Phylogeny Inference Package, version 3.57c, Department of Genetics, University of Washington, Seattle) after alignment of amino acid sequence with Bioedit and ClustalW. Line lengths indicate the relative distances between nodes. Bootstrap values > 50% of 100 replications are shown for all branches. Accession number of proteins used to build the tree are: *At*CAD 1 (NP_195643), *At*CAD A (NP_195510), *At*CAD B1 (CAA48027), *At*CAD B2 (NP_195512), *At*CAD C (NP_188576), *At*CAD D (NP_195149), *At*CAD E (NP_179765), *At*CAD F (NP_179780), *At*CAD G (NP_177412), *E. gunii* CAD1 (Q42726), alfalfa *Msa*CAD1 (AAC35846), alfalfa *Msa*CAD2 (AAC35845), tobacco CAD4 (P30359), tobacco CAD9 (P30360), *Picea abies* CAD (Q08350), *Pinus taeda* CAD (P41637), *Populus tremuloides Pt*CAD (AAF43140), and *P. tremuloides Pt*Sad (AAK58693).

1999). *At*CAD-B1, *At*CAD-B2, and *At*CAD-A belong to the same subfamily as *Pt*SAD (Li et al., 2001). *AtCAD-C* and *AtCAD-D* belong to the well-characterized CAD group. *At*CAD-G constitutes by itself another subfamily, and only a few expressed sequence tags (ESTs) from other plant species match its cDNA (data not shown). We focused our attention on *AtCAD-C* and *AtCAD-D* due to their high similarity to poplar and tobacco CAD encoding the known CAD proteins and to the high number of ESTs (corresponding to these amino acid sequences) found in cDNA libraries (Goujon et al., 2003a). They both share 73% similarity with the first characterized poplar CAD (Van Doorsselaere et al., 1995). The tobacco, poplar, and alfalfa CADs have been shown to be involved in constitutive lignification (Halpin et al., 1994; Baucher et al., 1996, 1999) using antisense strategies. To determine the potential roles of *AtCAD-C* and *-D* in constitutive lignification, we have analyzed their expression profiles using northern hybridization, RFLP-reverse transcriptase (RT)-PCR, promoter β-glucuronidase (GUS) fusion, and western analysis. Null mutants corresponding to each of these genes were identified, and the consequences of the mutations on lignin content and structure were determined.

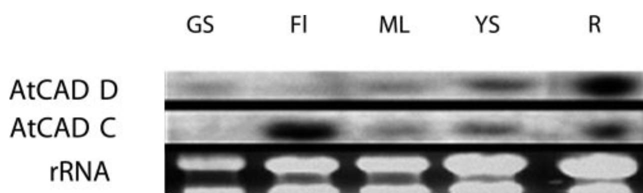

Figure 2. Northern-blot analysis of total RNA (20 μg lane^{-1}) from green siliques (GS), flowers (Fl), matures leaves without the central vein (ML), young stems (YS), and roots (R) hybridized with *AtCAD-C* and *AtCAD-D*. Total RNA loading is illustrated by ethidium bromide staining of ribosomal RNA in the bottom panel.

Transcription Pattern of *AtCAD-C* and *AtCAD-D*

Northern and RFLP-RT-PCR Analysis

The number of ESTs corresponding to *AtCAD-C* (25) and *AtCAD-D* (20) in the databases of GenBank + EMBL + DDBJ sequences were quite similar. However, it must be noted that no ESTs originating from a floral stem cDNA library were available, making it impossible to estimate their relative abundance in this highly lignified tissue. Northern-blot analyses were performed using different tissues. *AtCAD-D* was most strongly expressed in roots, less actively in young stems, and at relatively low levels in leaves, siliques, and flowers. The expression profile of *AtCAD-C* was similar to that of *AtCAD-D* but with stronger expression in flowers and no detectable expression in siliques.

Further analysis using semiquantitative RFLP-RT-PCR experiment revealed differences in the level of

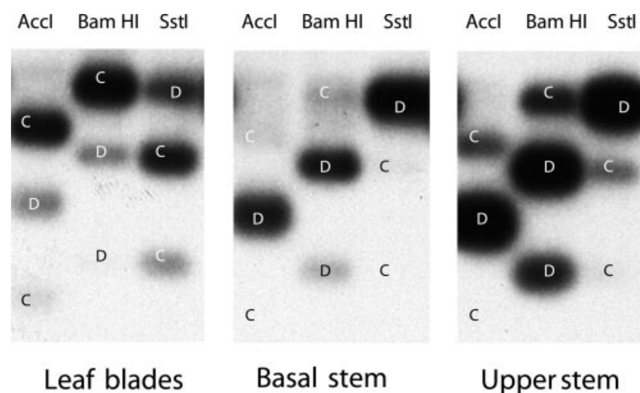

Figure 3. Restriction products of semiquantitative RFLP-RT-PCR conducted on leaf blades, basal, and upper stems of wild-type Arabidopsis using common primers for both *AtCAD-C* and *-D* cDNA. PCR products were digested by *Acc*I, *Bam*HI, and *Sst*I and then transferred onto DNA membranes before hybridization with an equal mix of the specific probes. Fragments corresponding to *AtCAD-C* are identified with the letter "C." Fragments corresponding to *AtCAD-D* are identified with the letter "D."

expression of these two *CAD* genes in leaf blades and stem parts (Fig. 3). *AtCAD-D* transcripts were more abundant in stem tissue than in leaf tissue, and were most abundant in upper stem tissues. *AtCAD-C* appeared to be most strongly expressed in leaf tissue, less strongly in upper tissues, and only at low levels in basal stem tissue.

GUS Analysis of pAtCAD D*::GUS Lines and* pAtCAD C*::GUS Lines*

Analysis of the GUS pattern in several plants expressing this reporter gene under the control of the *AtCAD-D* (*pAtCAD D*::GUS lines) or the *AtCAD-C* (*pAtCAD C*::GUS lines) promoters confirmed results obtained with mRNA analysis. This comparative expression profiling allowed a more detailed study of the tissue specificity of these two genes (Fig. 4). GUS staining was observed in stems, leaves, flowers, siliques, and roots in both types of lines. However, leaves in *pAtCAD C*::GUS lines presented an overall GUS staining (Fig. 4, b and d), whereas staining in *pAtCAD D*::GUS lines was restricted to vascular tissues and hydathodes (Fig. 4, a and c). GUS staining was quite similar in roots of both lines but appeared sooner and more intensely in roots of *pAtCAD C*::GUS lines compared with *pAtCAD D*::GUS lines, both at seedling (data not shown) and mature stages (Fig. 4, e and f). In flowers, both lines showed stronger GUS activity after overnight incubation (Fig. 4, k and l), but monitoring allowed observation of a quicker staining of sepals and filaments of stamens (not of anthers) in *pAtCAD C*::GUS lines (Fig. 4n) when compared with *pAtCAD D*::GUS lines (Fig. 4m) after 1 hour incubation time.

In stems, GUS staining of both constructs was closely related to lignin deposition in xylem bundles vessels and evenly in interfascicular fibers (Fig. 4, g and h). No staining was observed in the pith. When sections of both line-types were subjected to the second method for staining (see "Materials and Methods"), GUS staining was observed at the proximity of the bundle cambium region that gives rise to xylem elements and also in the region of interfascicular cambium where the interfascicular elements originated and staining was localized in cells undergoing lignification within fascicular elements (Fig. 4, i and j). This method warranted that staining in fibers

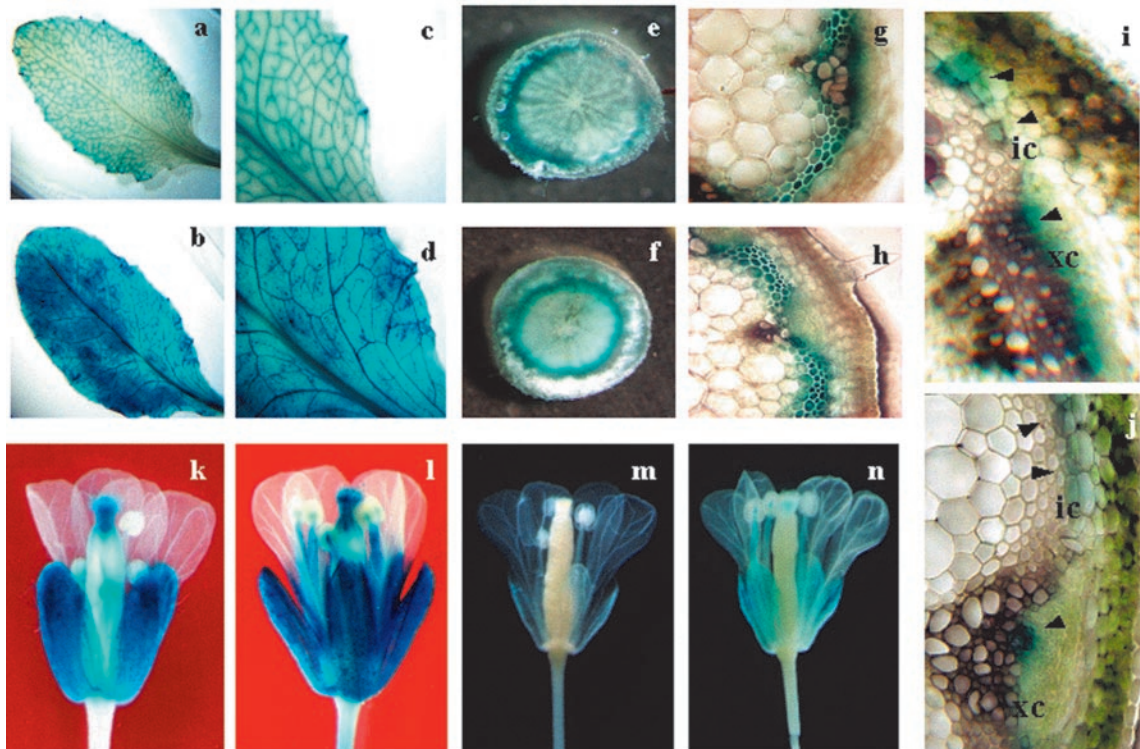

Figure 4. GUS assays performed on 8-week-old transgenic lines harboring the *AtCAD-D* and *AtCAD-C* promoters fused to the *uidA* gene (pCAMBIA 1391 xb). A distinct pattern of GUS expression between the two different constructs is observed in different parts of the plants (*pAtCAD D*::GUS lines: a, c, e, g, i, k, and m; *pAtCAD C*::GUS lines: b, d, f, h, j, l, and n). Organs tested are leaves (a–d), flowers (k–n), roots (e and f) and stems (g–j). Leaves were incubated overnight. Flowers were incubated overnight (k and l) or for 1 h (m and n). e and f, GUS assays of transverse sections of mature roots (magnification ×25) using method 2. g and h, GUS assays of transverse sections of mature stems (magnification ×50) using method 1. i and j, GUS assays of transverse sections of mature stems (magnification ×100) using method 2. Arrows show GUS staining in the cells close to the xylem cambium (xc) and interfascicular cambium (ic).

or in xylem was not due to diffusion of GUS products. We also noticed that in *pAtCAD C*::GUS lines, staining was higher in the interfascicular region than in the xylem vessels (Fig. 4h). This zonal and tissue-specific staining difference was not observed in *pAtCAD D*::GUS lines (Fig. 4g).

Isolation of *Atcad-C* and *Atcad-D* Mutants

To more precisely define the role of each *CAD* gene, we have identified mutant lines in the Versailles T-DNA insertion collection. One line (named *Atcad-D*) with a T-DNA insertion within the *AtCAD-D* gene was identified by reverse genetics. A second line containing a T-DNA insertion in the *AtCAD-C* gene (named *Atcad-C*) was identified by systematic border sequencing. The segregation of progenies of these lines, germinated on selective medium containing kanamycin, allowed us to infer that only one *nptII* insertion locus was present in each line. Hybridization experiments performed on digested genomic DNA from the mutant lines using radiolabeled DNA probes corresponding to the right and left borders of T-DNA confirmed the presence of a unique T-DNA insertion in each mutant (data not shown). Flanking regions of each T-DNA border were sequenced, and the site of the insertion was localized in the second and third intron for *AtCAD-C* and *AtCAD-D*, respectively (Fig. 5). No important deletions in the vicinity of either insertion were observed, demonstrating that only the *CAD* genes were targeted.

Homozygous lines for each insertion were obtained, and the impact of the T-DNA insertion on mRNA expression was determined by RT-PCR experiments on total RNA of each mutant using specific primers. Absence of mRNA signal for the specific *CAD* genes was confirmed for each mutant (data not shown). No visual phenotypes were observed when these mutant lines were grown in greenhouse conditions.

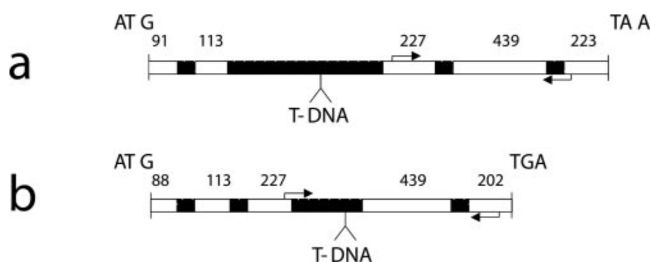

Figure 5. Diagrammatic representation of both *AtCAD-C* (a) and *AtCAD-D* (b) genes and localization of the T-DNA insertions in each mutant line. A single insertion occurred in each line. The figure indicates the number of nucleotides in each exon (represented by white rectangles) and positions of the primers (arrows) used to verify the absence of mRNA expression of either *AtCAD* gene.

CAD Activities and Changes in AtCAD Protein Quantity in *Atcad-C* and *Atcad-D* Mutants

Because no transcript from either of mutated genes was detected by RT-PCR analysis of the mutants, we performed western analysis in parallel with coniferyl alcohol dehydrogenase (conAD) or sinapyl alcohol dehydrogenase (sinAD) activities (Fig. 6).

Considering the high amino acid homology between tobacco CAD and AtCAD-C and -D proteins, we carried out western-blot analysis using antibodies directed against the tobacco CAD. Long migration on acrylamide gel allowed identification of two proteins at the apparent molecular mass of 44 and 42 kD in protein extracts originating from the basal and upper parts of stems, siliques, and roots of the wild type. Although one of these bands was absent in flowers, one additional band at 36 kD was observed in this organ. Probably due to low abundance of these proteins and the greater abundance of proteins such as Rubisco, it was not possible to characterize extracts of leaf blades.

The 44-kD band was absent in stems of *Atcad-D*, suggesting that this band corresponds to the *AtCAD-D* protein. This band was clearly prominent in siliques as shown on Figure 6d. Similarly, the 42-kD band was absent in *Atcad-C*, and this likely corresponds to the AtCAD-C protein. This signal was less intense in the whole stem (Fig. 6, a and b), confirming the RFLP-RT-PCR analysis on wild-type plants. In contrast, this band was prominent in flowers of wild type confirming northern analysis, whereas no signal assigned to *AtCAD-D* (Fig. 6e) was detectable with these analyses.

Expression profiling of both genes was complemented by assays of CAD activities in both mutants. The conAD and sinAD activities were reduced in organs of both mutants confirming CAD biochemical functions of the corresponding proteins. These activities were more drastically reduced in *Atcad-D* than in *Atcad-C* except in flowers, confirming the predominance of *AtCAD-C* in this plant part (Fig. 6). Predominance of sinAD activity observed in stems of wild type was completely abolished within stems of *Atcad-D* with a 12-fold reduction, and conAD activity was 5-fold reduced. The conAD and sinAD activities were also reduced in *Atcad-C*, albeit less drastically (conAD and sinAD activities were reduced by 1.3- and 3-fold, respectively, in stems of this mutant). The sinAD activity was too low in roots of both mutants and wild type to be characterized with confidence. It is interesting to note that conAD and sinAD activities were not modified significantly in siliques of *Atcad-C* but were highly reduced in *AtCAD-D* in accordance with northern-blot hybridization and western analyses (Figs. 2 and 6d). Once again, CAD activities were not significant enough in leaf blades to be characterized in each mutant.

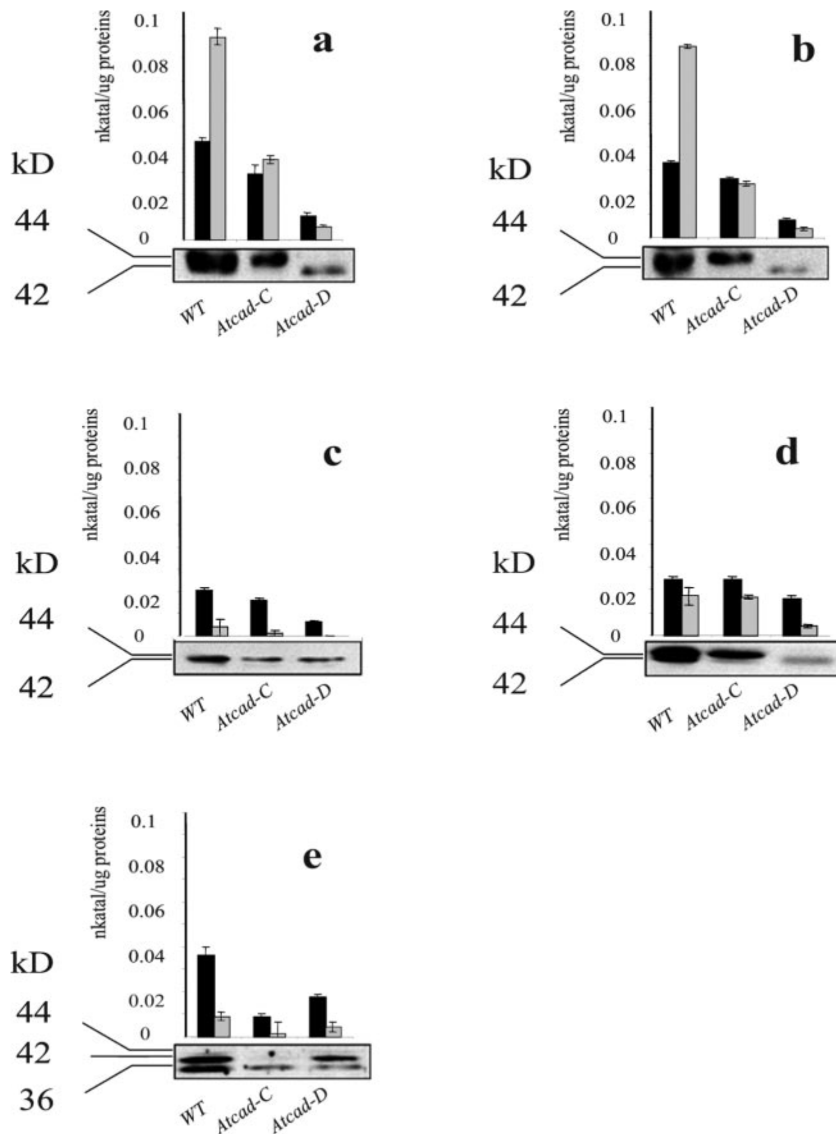

Figure 6. CAD activities and western-blot analysis in different tissues from wild type, *Atcad-C* and *Atcad-D* plants. Gray histograms represent sinAD activity, whereas the black histograms represent conAD activity. Bars = SEs of the mean of three assays. Crude extracts from different parts of wild type (WT), *Atcad-C*, and *Atcad-D* plants were assayed by immunoblotting using anti-tobacco CAD antibodies. Analyses were performed on the basal part of stems (a), upper part of stems (b), roots (c), siliques (d), and flowers (e). The apparent molecular masses of *At*CAD-D (44 kD), *At*CAD-C (42 kD), and an unknown protein (36 kD) are indicated.

Lignin Modification in Mutants

The histochemical analysis of lignified stems using the Wiesner (phloroglucinol-HCl) reagent or the Maüle reagent did not reveal any perturbation of lignification between the control and the *Atcad-C* or *Atcad-D* mutants (data not shown). The lignin content of extract-free floral stems was determined by the Klason standard method (Dence, 1992), which systematically includes the removal of extractives before analysis and the correction for ash content, if any. The wild-type and mutant lines displayed similar amount of extractives and negligible ash levels in the Klason lignin fraction. As already observed for other Arabidopsis lines (Goujon et al., 2003b), we found that the Klason lignin level of the extract-free stems displayed substantial variation between culture replication. Nevertheless, four different replications (comprising a total of 10 and four repetitions for *Atcad-D* and *Atcad-C*, respectively) revealed lower Klason lignin values for *Atcad-C* (14.44 ± 0.46) and *Atcad-D* (14.23 ± 0.29) stems when compared with the control line (15.20 ± 0.29). An ANOVA followed by least squares means tests determined that this difference was significant ($P = 0.08$) only for *Atcad-D*.

The *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin-derived monomers released by thioacidolysis of wild-type and homozygous mutant lines were analyzed by gas chromatography-mass spectrometry (GC-MS). The data reported in Table I for one replication series were confirmed by other replications. The thioacidolysis yield, when expressed on the basis of the Klason lignin content of the extract-free stems, did not clearly discriminate the control and mutant lines. In contrast, the proportion of the H, G, and S monomers revealed that *Atcad-D* lignins systematically released a lower proportion of thioacidolysis S

Table I. *Thioacidolysis analysis of wild-type and mutant lines*

The data are means (±SE) of lignin-derived monomers recovered from extract-free floral stems or from roots of mutant and control lines grown in the same culture series. All the data are from duplicate experiments, except for the *Atcad-D* root sample (single assay).

Line	Wild type	*Atcad-C*	*Atcad-D*
		μmole g^{-1} Klason lignin	
Extract-free floral stems			
Total yield of main monomers	1,145 ± 32	1,209 ± 50	1,022 ± 16
H	1.9 ± 0.1	2.1 ± 0.1	0.8 ± 0.1
G	69.8 ± 0.1	72.1 ± 0.2	77.7 ± 0.3
S	28.3 ± 0.1	25.8 ± 0.2	21.5 ± 0.3
Whole dried roots			
H	3.6 ± 0.1	2.9 ± 0.1	3.0
G	88.4 ± 0.3	89.3 ± 0.2	92.5
S	8.0 ± 0.1	7.8 ± 0.2	4.5

main monomers (S-CHSEt-CHSET-CH_2SEt) than the wild-type or *Atcad-C* homologous samples. This result indicates that the *AtCAD-D* mutation induces some perturbations in the formation of sinapyl alcohol, the precursor of lignin syringyl-glycerol units (S-CHOH-CHOAr-CH_2OH). Thioacidolysis of whole dried roots confirmed this specific trait of *Atcad-D* lignins: Although the proportion of S monomer released by root lignins in all three lines was found particularly low, the lowest level of S monomer was obtained from *Atcad-D* sample (Table I).

Recent studies with appropriate model compounds revealed that, when incorporated into lignins by peroxidasic oxidation, the main ether linkage modes of coniferaldehyde and sinapaldehyde were at C4OH and at Cβ, respectively. When subjected to thioacidolysis, the former coniferaldehyde end group structures gave rise to G-CHR-CHR-CH_2R (R = SEt). In contrast, the latter sinapaldehyde structures linked at their Cβ carbon gave rise to two very specific indene isomers that have been authenticated recently (Kim et al., 2002). These G and S marker compounds, which are diagnostic for coniferaldehyde end groups and for sinapaldehyde units linked at Cβ, were quantified by GC-MS. Their proportions, relative to the main G and S conventional monomers were determined for six replications of wild-type and *Atcad-D* samples and are reported in Table II. Although the comparison of wild-type and *Atcad-C* samples with regard to structural traits did not reveal any difference (data not shown), we could see that sinapaldehyde units linked at Cβ were systematically incorporated into *Atcad-D* stem lignins at a level that was more than 10-fold that observed in the control line. Similar results were observed for the root samples (data not shown). In contrast, the proportion of coniferaldehyde end groups was not found to be significantly different in the control and mutant lines. These thioacidolysis data are consistent with the fact that Wiesner staining did not discriminate between the control and mutant lines. We recently established that coniferaldehyde end groups react positively with the phloroglucinol reagent, whereas the reaction is negative with hydroxy-cinnamaldehyde units linked at Cβ. In other words, the Wiesner staining reaction is not appropriate to reveal the incorporation of sinapaldehyde into lignins, as recently reported for CAD-deficient poplars (Kim et al., 2002).

Whereas the major alteration induced by the *AtCAD-D* mutation was the incorporation of sinapaldehyde units in lignins, differences in other structural traits were observed that were reminiscent of the traits reported for CAD-deficient poplars (Lapierre et al., 1999). The proportion of G lignin units with free

Table II. *Proportions (mol %) of coniferaldehyde end groups (G-CH = CH-CHO linked at C4OH) relative to conventional G units (G-CHOH-CHOAr-CH_2OH) and of sinapaldehyde units linked at Cβ (S-CH = COAr-CHO) relative to conventional S units (S-CHOH-CHOAr-CH_2OH), as determined by GC-MS of their specific thioacidolysis markers within stem of wild type and Atcad-D*

Replication	Coniferaldehyde End Groups/G		Sinapaldehyde Units Linked at Cβ/S	
	Wild type	Atcad-D	Wild type	Atcad-D
			mol %	
1	0.68	0.77	0.16	2.76
2	0.42	0.55	0.05	1.39
3	0.69	0.68	0.14	1.79
4	0.75	0.96	0.15	1.93
5	0.70	0.89	0.18	2.69
6	0.62	0.66	0.21	2.26
Mean ± SD	0.64 ± 0.12	0.75 ± 0.15	0.15± 0.05	2.14± 0.53

phenolic groups was observed to be higher in *Atcad-D* lignins than in the control ones (percentage of G units with free OH per 100 β-O-4 linked G units: 14.33 ± 0.04 in wild type versus 15.56 ± 0.03 in the *AtCAD-D* homologous sample). In addition, although the amount of syringaldehyde released from the cell walls by mild alkaline hydrolysis (1 M NaOH, 24 h, room temperature) was very low, it nevertheless discriminated the wild-type and *Atcad-D* stems, the latter providing about 3 times more syringaldehyde than the control (120 versus 40 ng g^{-1} cell wall).

DISCUSSION

AtCAD-C and *AtCAD-D* Belong to a Small Multigene Family in Arabidopsis

Different studies have highlighted that *CAD* genes, which have been relatively well studied, could be present in more than one copy in several plant species, except in conifers such as the loblolly pine (MacKay et al., 1995). Tavares et al. (2000) used the uncompleted sequence of the Arabidopsis genome to detail the structure of the *CAD* gene family along with three other multigene families. After reexamining the entire Arabidopsis genome, in the present study, we found that the *AtCAD* family may include a ninth gene that we named *AtCAD-G*.

Proteins involved in CAD activity associated with lignification were previously thought to act on three different cinnamaldehydes (coniferaldehyde, sinapaldehyde, and *p*-coumaraldehyde; for review, see Baucher et al., 1998), but differences in substrate specificity of paralogs have led to the hypothesis that CAD polymorphism could play a role in the control of lignin heterogeneity (Hawkins and Boudet, 1994). To date, the *At*CAD family in Arabidopsis is the only complete CAD family to be described. According to our analysis, this model plant seems to contain genes encoding CAD-like proteins similar to those previously characterized in other species, thus confirming Arabidopsis as a relevant model for an extensive study of this gene family. Some paralogs (*At*CAD-1, *At*CAD-E, and -F) are close to *Msa*CAD1, a gene known to be wound inducible in alfalfa, where the corresponding protein is active on a range of cinnamyl, benzyl, and aliphatic aldehyde substrates (Brill et al., 1999). The *At*CAD-B paralogs corresponding to ELI-3 proteins have been studied previously and were originally identified as part of the defense response in parsley (*Petroselinum crispum*) and in Arabidopsis (Kiedrowski et al., 1992). Williamson et al. (1995) and Somssich et al. (1996) demonstrated that ELI-3 was neither a CAD nor a malate dehydrogenase but rather a benzyl alcohol dehydrogenase, which accepts various benzaldehyde substrates. In our phylogenetic analysis, these proteins (*At*CAD-A, *At*CAD-B1, and -B2) fall within the same cluster as *Pt*SAD, which has been recently identified and characterized in poplar (Li et al., 2001). *Pt*SAD was shown to be highly specific for sinapaldehyde and was proposed to be responsible for S-unit deposition in lignins of poplar fibers. Lignins of Arabidopsis stems are composed of approximately 25% of S units; therefore, the presence of SAD paralogs could be predicted.

*At*CAD-C and *At*CAD-D belong to the same subfamily as *E. gunii* CAD, *Pt*CAD, and *Msa*CAD-2. These proteins correspond to some of the best characterized CAD enzymes. Both substrates, sinapaldehyde and coniferaldehyde, were accepted by these CAD proteins (for review, see Baucher et al., 1998; Mellerowicz et al., 2001). However, the recent study of Li et al. (2001) has shown that *Pt*CAD displays a higher affinity for coniferaldehyde than for sinapaldehyde in poplar. The *AtCAD-C* and *AtCAD-D* genes display a high degree of similarity (84%) and a conserved genome structure in terms of number and position of introns (Fig. 5) as discussed by Tavares et al. (2000). These two genes could probably be the result of a duplication of a common ancestor, a relatively frequent occurrence in Arabidopsis (Arabidopsis Genome Initiative, 2000) and, in particular, in the *CAD* gene family, as shown in this study. These duplicated genes may have acquired differential biological functions through evolution and finding the role of either gene in lignin metabolism is a challenging goal.

Comparative Expression Profiling of *AtCAD-C* and *AtCAD-D*. Commonality and Dissimilarity

Consulting EST databanks demonstrated that *AtCAD-C* and *AtCAD-D* were observed in different cDNA libraries obtained from seedlings, leaves, and roots. In this work, the expression profiles of *AtCAD-C* and *-D* were determined using several approaches. Although expression patterns for both genes seem similar in a first approach (northern analysis), organ specificity was shown using more extensive studies (semiquantitative RFLP-RT-PCR analyses on leaf blade and stem), allowing us to deduce that the *AtCAD-D* to *AtCAD-C* mRNA ratio is organ dependent. *AtCAD-D* is clearly the main protein in stem, albeit both mRNA transcripts were detected within this highly lignified tissue. In-depth analysis using *AtCAD* promoter-GUS fusion demonstrated that *AtCAD-C*, which is expressed in xylem elements and fibers, is also expressed at a high level in other tissues such as flowers and leaf parenchyma and, therefore, seems less regulated than *AtCAD-D*. Our GUS assay (method 2) indicated that CAD proteins are probably synthesized early in stems and roots close to the cambium when secondary development occurs for xylem and fiber formation. This expression close to the cambial zone has been observed previously in poplar by Hawkins et al. (1997).

The Null Mutants *Atcad-C* and *Atcad-D* Facilitate a Better Understanding of the Importance of Each CAD Protein

Because phylogenetic analysis and determination of the expression patterns of the two *CAD* genes did not resolve the respective abundance and importance of these proteins, we have characterized T-DNA insertion mutants corresponding to null mutants for these genes. Homozygous lines containing T-DNA insertions in each gene were obtained and characterized. Western experiments using an antiserum raised against a tobacco CAD (Halpin et al., 1994) allowed us to identify the proteins corresponding to *AtCAD-C* and *AtCAD-D*. This analysis confirmed that differences in the *AtCAD-D* to *AtCAD-C* mRNA ratio within tissues was also manifested at the protein level. In addition, a signal (36 kD) that could not be assigned to *AtCAD-C* or *-D* was found in flowers and at a lower level in leaves and stems. This 36-kD band suggests the possibility that at least one more *CAD* gene is expressed in these organs.

Atcad-D and *Atcad-C* show drastically significant reductions in conAD and sinAD activities when compared with the control plants. This result clearly confirmed the biochemical function of the corresponding genes. CAD activity assays of these mutants for substrate specificity (coniferyl and sinapyl alcohols) showed that *At*CAD-D is responsible for the main conAD and sinAD in vitro activities in stems even if *At*CAD-C is involved to a lower extent in these activities. The combination of CAD activities and western analyses shows clearly that *AtCAD-D* is unambiguously the main CAD protein in lignified tissues (stem) but not in other tissues such as flowers.

The *AtCAD-D* Mutation Has an Impact on Lignin Content and Structure

To evaluate the respective roles of *AtCAD-C* and *AtCAD-D* in constitutive lignification, lignin characteristics have been determined in stems and roots of both mutants and wild-type lines. A lower Klason lignin content was observed in *Atcad-D* in four different biological replications carried out at different times. This phenotype was observed to a lesser extent in *Atcad-C*, but the difference was not significantly different. Reduction of lignin content has been observed in the pine *cad* mutant (MacKay et al., 1997), in the maize *bm1* (Halpin et al., 1998), in *CAD* antisense young poplars with less than 5% residual conAD activity (C. Lapierre, unpublished data) and in older *CAD* antisense poplars with about 20% to 30% residual conAD activity (Lapierre et al., 1999; Pilate et al., 2002). However, reduced lignin content was not observed in alfalfa and tobacco *CAD* antisense plants (Halpin et al., 1994; Baucher et al., 1999). A lower proportion of β-O-4-linked syringyl-glycerol units has been repeatedly reported to occur in various *CAD*-deficient plants (for review, see Dixon et al., 2001). Therefore, the results obtained for the *Atcad-D* mutant are consistent with the tendency observed in lignins from *CAD*-deficient dicots. In addition, we have also established that this reduced proportion of "conventional" S-lignin units was accompanied by the incorporation of sinapaldehyde units linked at their Cβ carbon and unreactive toward the Wiesner reagent, a structural trait recently reported for *CAD*-deficient poplar lines (Kim et al., 2002). These results serve to reinforce the limitations of the histochemical tests for revealing lignin alterations. No modifications of lignin structure were observed in *Atcad-C*, suggesting a marginal if any role of *AtCAD-C* in lignification.

***AtCAD-D* Is Involved in Biosynthesis of G or S Lignin Precursors or Both?**

Drastic decrease of conAD and sinAD activities (20% and 6% of residual activities in wild type, respectively), accumulation of sinapaldehyde on one hand and total disappearance of the *AtCAD-D* protein in stems of *Atcad-D* on the other hand leads us to hypothesize that this protein is able to use both cinnamaldehydes but with a greater preference for sinapaldehyde. In contrast, the ability to reduce coniferyl alcohol remains relatively elevated in some organs of this mutant such as flowers and siliques. Deficiency in both activities could have a higher impact on S lignin biosynthesis because this lignin type is synthesized at the latter stage of cell wall formation (Donaldson, 2001), and this deposition may be dependent on the appropriate amount of G-lignin type. However, it must be noticed that coniferyl and sinapyl alcohols are still incorporated in high proportion in lignins of *Atcad-D* and that lignin quantity is only slightly reduced. These characteristics suggest that *AtCAD-D* is not the only CAD involved in the reduction of cinnamaldehydes. Some other *At*CADs could participate in their reduction, but evidently, these *At*CAD proteins are not able to compensate for the reduction of total CAD activities, sinapaldehyde accumulation, and Klason lignin decrease observed in the *Atcad-D* mutant.

Li et al. (2001) have suggested recently that coniferaldehyde and sinapaldehyde are respectively reduced by *Pt*CAD and *Pt*SAD proteins in aspen. They showed that *Pt*CAD was immunolocalized exclusively in xylem elements, whereas *Pt*SAD was conspicuous in phloem fiber cells. *AtCAD-D* and *-C* are highly similar to *Pt*CAD and belong to the same cluster of previously characterized CAD (see cladogram). However, our results regarding enzymatic activities and lignin structure clearly show that absence of *AtCAD-D* results in a decrease in both conAD and sinAD activities and induces impact on deposition of S lignins. Promoter fusion analysis suggested that corresponding *AtCAD-D* and *-C* genes may be expressed in xylem and fibers, although only

fibers contain S lignin. The ability to reduce sinapaldehyde to sinapyl alcohol in xylem (a G unit-enriched tissue) is consistent with the fact that sinapyl alcohol was synthesized in xylem bundles when ferulate-5-hydroxylase was overexpressed in transgenic Arabidopsis (Meyer et al., 1998; Sibout et al., 2002). These observations demonstrate that the absence of SAD-type protein in vessels of stems and roots likely would not be the limiting step for S-unit biosynthesis in these tissues, at least in Arabidopsis, and that other CADs could be involved in this processes. Therefore, we propose that *AtCAD-D* could participate in S-unit biosynthesis.

Absence of the *At*CAD-D protein in the mutant could certainly have indirect consequences. The reduction of S unit incorporation observed in its lignin could be due to a decreased activity of an *At*SAD protein as a consequence of coniferaldehyde accumulation. Li et al. (2001) have shown that the two substrate interactions (sinapaldehyde and coniferaldehyde) are of the competitive inhibition type for both *Pt*CAD and *Pt*SAD proteins. However, unlike in the pine *CAD* mutant and the *bm1* maize and tobacco antisense lines (Halpin et al., 1998; Ralph et al., 1998), we did not observe any increase of coniferaldehyde incorporation in lignins of *Atcad-D*.

An *AtSAD* gene involved in lignification has not been characterized until now in Arabidopsis, but characterization of null mutants for the other seven Arabidopsis *CAD* genes is under way and may allow us to get a clearer view of the last step of the synthesis of the monolignol monomers.

CONCLUSION

This work aims to contribute to a better understanding of the lignin monomer pathway in the context of a small multigene family. The *Atcad-D* mutant, in which the corresponding gene is specifically expressed in lignified elements in wild-type plants, displayed structural modifications within its constitutive lignin. This phenotype is consistent with those observed in plants where CAD was down-regulated as a consequence of mutations or antisense strategies. Other *AtCAD* genes are not able to compensate the *Atcad-D* phenotype; therefore, *AtCAD-D* could be considered a major *CAD* gene for monolignol biosynthesis among the small *CAD* multigene family in Arabidopsis.

The role of *AtCAD-C* in constitutive lignification, despite its expression in lignified tissues, is less obvious because no major lignin structural modifications have been detected in the *Atcad-C* null mutant. However, expression of *AtCAD-C* is partly redundant to *AtCAD-D* (at the whole organ level), and the absence of *AtCAD-C* could be compensated by the *AtCAD-D* protein, especially if this step is not limiting for lignin biosynthesis (Anterola et al., 2002). Its basal expression level in many plant parts and its high expression in flowers suggested a role in other pathways (suberin or lignan biosynthesis) or in pathogen defense. Characterization of a double *Atcad-C*/*Atcad-D* null mutant and other *Atcad* mutants, which is underway, will be very useful for further studies and better understanding of the role of each *CAD* gene.

MATERIALS AND METHODS

Plant Material and Growth Conditions

The ecotype Wassilewskija was used in this work except for *Atcad-D* promoter cloning, where genomic DNA from the Columbia ecotype was used. Mutants were identified in the Arabidopsis T-DNA insertion collection of Versailles (Bouché and Bouchez, 2001). The transformed and wild-type Arabidopsis plants were grown together in the same greenhouse to ensure uniform environmental conditions. For stem sections and GUS assays, plants were grown in growth chambers at 23°C with 12 h of light for 5 weeks. Plants were grown in aeroponic conditions for root analyses.

DNA and RNA Analyses

Southern-Blot Hybridization

Genomic DNA was extracted from leaves of wild-type and *Atcad* mutant plants as described by Doyle and Doyle (1990). Southern blots were performed according to Sambrook et al. (1989) with Genescreen Plus membranes (NEN Research, Boston).

Reverse Genetic and Flanked Sequence Tag

DNA screening for the *Atcad-D* mutant was achieved using the following primers: 5′-CTACAAATTGCCTTTTCTTATCGAC-3′ and 5′- ATGCTCCCTATYAAGCTCCC-3′, using DNA pools of the Versailles collection of T-DNA lines. The *Atcad-C* mutant was selected using a systematic border sequencing program (http://flagdb-genoplante-info.infobiogen.fr) in the same collection of mutants.

Northern-Blot Hybridization

Total RNA was extracted from several Arabidopsis tissues (for leaves, the central vein was eliminated before freezing) and prepared as described by Verwoerd et al. (1989). RNA was then redissolved in diethyl pyrocarbonate-treated water, and RNA concentrations were determined by A_{260}. Equal amounts of total RNA (20 μg) were denatured with formamide/formaldehyde and fractionated on a 1.2% (w/v) agarose formaldehyde gel (Sambrook et al., 1989). Total RNA quality was confirmed by ribosomal RNA integrity observed after ethidium bromide staining of the gel. The gel was blotted using a capillary procedure (Sambrook et al., 1989) onto Genescreen plus membranes (NEN Research), and RNAs were cross-linked to membranes by UV radiation. Specific *AtCAD-C* and *AtCAD-D* radiolabeled probes synthesized using ESTs (GenBank accession no. Z34154) corresponding to *AtCAD*-D (Höfte et al., 1993) and from a *AtCAD*-C cDNA (GenBank accession no. T45746) provided by The Arabidopsis Information Resource (www.Arabidopsis.org) were used. *AtCAD-C* and *D* probe production was synthesized through PCR amplification using *AtCAD-C* and *AtCAD-D* shared primers (5′-GGATCAGATGTGAGCAAGTT-3′ and 5′-ATGCTCCCTATYAAGCTCCC-3′) with ^{32}P-dCTP in the reaction mix. Preliminary studies have shown no cross hybridization between *AtCAD-C* and *Atcad-D* probes when membranes were washed under stringent conditions. The 674-bp PCR product probes were purified with the QIAquick nucleotide removal kit (Qiagen USA, Valencia, CA). Membranes were hybridized as for Southern-blot hybridization experiments and were washed in 0.1% (w/v) SSC and 0.1% (w/v) SDS at 65°C.

Semiquantitative RLFP-RT PCR

An RT reaction was carried out on 10 μg of DNAse-treated (Promega, Madison, WI) total RNA in a 50-μL volume. Five microliters of the RT reactions was used as template for a semiquantitative PCR by using primers common to both *AtCAD-C* and *-D* genes as described before (Lurin and Jouanin, 1995). The PCR products were digested by restriction enzymes (*Acc*I, *Bam*HI, and *Sst*I), transferred onto Genescreen plus membranes, then hybridized with a mix of specific probes for each gene using the following primers: *AtCAD-C*fw, 5′-GCACGAGGTAGTAGGNGARGT-3′; *AtCAD-C*up, 5′-AAAGCCAACACTTCTTCNGTYTC-3′; *AtCAD-D*fw, 5′-GTGGGATCA-GATGTGAGCAA-3′; and *AtCAD-D*up, 5′-AACGCACATCGTTCTTCTCG-3′ using the cDNA previously described for northern-blot hybridization experiments as template.

Western-Blot Analysis and Enzyme Activities

Total protein extracts were obtained by homogenization of fresh tissues in 100 mM Tris-HCl (pH 7.5) containing 0.4% (w/v) polyvinylpolypyrrolidone, 0.5% (w/v) polyethylene glycol, and 15 μM β-mercaptoethanol, and quantified according to Bradford (1976).

Western Analysis

Protein samples (15 μg) were heated at 95°C for 5 min in Laemli buffer, cooled, and centrifuged briefly before loading on a 12% (w/v) acrylamide SDS-PAGE with a 10% (w/v) resolving gel using a Bio-Rad Protean II apparatus (Bio-Rad Laboratories, Hercules, CA) and run at 50 V for 2 or 4.5 h. Proteins were transferred onto a 0.45-μm nitrocellulose membrane (Amersham Biosciences Inc., Piscataway, NJ) using an electroblotting apparatus (Bio-Rad Laboratories). The polyclonal antibody raised against tobacco (*Nicotiana tabacum*) xylem CAD 2 (Halpin et al., 1994) was used at a 1:2,000 dilution (w/v). Blots were developed using the ECL western blotting analysis system (Amersham Biosciences).

Enzyme Activities

Crude extracts were assayed spectrophotometrically (Ultra microplate, Bio-Tek Instruments, Winooski, VT) for aromatic alcohol dehydrogenase activity by oxidation of coniferyl alcohol (conAD activity) or sinapyl alcohol (sinAD activity). Assays were carried out at 25°C for 30 min in 250 μL of 100 mM Tris-HCl (pH 8.8), NADP (20 mM), and 50 mM of coniferyl or sinapyl alcohols using a micro-ELISA plate. Twenty micrograms of total protein for stem extracts, 40 μg for roots and siliques, and 80 μg for flowers were used for these reactions. Formation of hydroxy-cinnamaldehydes was monitored at 400 nm using the following molar extinction coefficient: coniferaldehyde 2.10×10^4 M^{-1} cm^{-1} and sinapaldehyde 1.68×10^4 M^{-1} cm^{-1}, as described by Mitchell et al. (1994). An assay without NADP was used as a control. Resulting units are defined as the amount of activity that converts 1 nmol of hydroxy-cinnamyl alcohol into the corresponding aldehyde per second (1 nKatal) per microgram of crude protein extract.

Promoter Cloning and uidA Fusion

Gene fusion products with the gene coding for GUS gene (*uidA*), under the control of *AtCAD-C* and *AtCAD-D* promoters, were constructed for monitoring expression of these genes in different plant parts and tissues. For both constructions, *Eco*RI *and Spe*I sites were inserted at the 5′ ends of the primers (underlined on the primer sequences) for cloning into pCAMBIA1391xb (Cambia, Canberra, Australia). The *AtCAD-C* promoter (1,762 bp) was cloned using Arabidopsis genomic DNA (ecotype Columbia) with the following oligonucleotides: 5′- *GAATTC*TGTTCATTGAGGCC-CAAGTATTTGTGTATT-3′ and 5′-*ACTAGT*CTTTTCTCCTGCTTCTACAC-TTCCCATTTC-3′. The *AtCAD-D* promoter (1,780 bp) was cloned using Arabidopsis genomic DNA (ecotype Wassilewskija) with the following oligonucleotides: 5′-G*GAATTC*GAAATTCTCCACTCGTAGCTCTTCGTTCTG-3′ and 5′- *ACTAGT*TTTCCTCTCTGCCTCCATTATTCCCATTTTTTGATG-3′. PCR products were cloned in pGEM-T Easy Vector (Promega) and sequenced. Promoter sequences were digested from pGEM-T Easy Vector with the appropriate enzyme and thus cloned in pCAMBIA1391xb according to standard methods (Sambrook et al., 1989).

Gus Staining

Entire leaves, flowers, and seedlings were harvested and immediately incubated in a 5-bromo-4-chloro-3-indolyl-β-D-GlcUA reaction medium as described by Jefferson et al. (1987) for 2 to 6 h depending on the rate of staining and were then dehydrated in 95% (v/v) ethanol. Two different sample preparations were used for stem sections. The first technique (method 1) consisted of cutting thin sections of stem sample (about 1 cm long), by hand or with a vibratome that had previously been subjected to staining as described for entire organs. In the second method (method 2), 1-cm stem samples were incubated in 0.5% (v/v) formaldehyde for exactly 1 min at room temperature immediately after harvesting. Sections of these samples were cut in 100 mM potassium phosphate buffer (pH 7.0) and then incubated in 5-bromo-4-chloro-3-indolyl-β-D-GlcUA reaction medium at 37°C for 15 min.

Arabidopsis Transformation

The binary vectors were introduced in the *Agrobacterium tumefaciens* strain C58pMP90 (Koncz and Schell, 1986) by electroporation. Plants were transformed by the flower infiltration protocol (Bechtold and Pelletier, 1998). T_1 transgenic plants were selected on Estelle and Somerville (1987) medium containing hygromycin (50 mg L^{-1}) or kanamycin (100 mg L^{-1}). T_2 plants were used for the GUS bioassays.

Lignin Analysis

Dried mature stems were collected after removal of leaves and siliques. Extract-free samples were prepared using a Soxhlet apparatus by sequentially extracting the ground material with toluene:ethanol (2:1 [v/v]), ethanol, and water. The determination of lignin content was carried out on the extract-free samples using the standard Klason procedure (Dence, 1992). The evaluation of lignin structure was carried out on whole plant material or on extract-free material, using the thioacidolysis procedure (Lapierre et al., 1995; 1999). The lignin-derived monomers were identified by GC-MS as their trimethyl-silylated derivatives. Low-molecular mass phenolics were analyzed as described by Lapierre et al. (1999).

Screening Databases, DNA Sequence Analysis, Protein Alignments, and Statistics

Databases were screened with BLAST algorithms (Altschul et al., 1990). DNA and protein alignments were carried out with GCG, BioEDIT, and ClustalW. PHYLIP was used for phylogenetic analysis. Statistical analysis of Klason lignin content was carried out using SAS (SAS Institute, Cary, NC).

ACKNOWLEDGMENTS

The authors thank Frédéric Légée for Klason lignin analysis, Hervé Ferry for plants cultivation in the greenhouse, Nicolas Feau for his helpful work in phylogenetic analyses, and Michéle Bernier-Cardou for statistical analyses. We are also grateful to Claire Halpin for providing antibodies and to Janice Cooke and Denis Lachance for suggestions and reading this manuscript.

Received January 27, 2003; returned for revision February 23, 2003; accepted March 20, 2003.

LITERATURE CITED

Altschul S, Gish W, Miller W, Myers W, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol **215:** 403–410

Anterola AM, Jeon JH, Davin LB, Lewis NG (2002) Transcriptional control of monolignol biosynthesis in *Pinus taeda*: factors affecting monolignol

ratios and carbon allocation in phenylpropanoid metabolism. J Biol Chem **277:** 18272–18280

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796–815

Baucher M, Bernard-Vailhe MA, Chabbert B, Besle JM, Opsomer C, Van Montagu M, Botterman J (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

Baucher M, Chabbert B, Pilate G, Van Doorsselaere J, Tollier MT, Petit-Conil M, Cornu D, Monties B, Van Montagu M, Inze D et al. (1996) Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. Plant Physiol **112:** 1479–1490

Baucher M, Monties B, Van Montagu M, Boerjan W (1998) Biosynthesis and genetic engineering of lignin. Crit Rev Plant Sci **17:** 125–197

Bechtold N, Pelletier G (1998) *In planta Agrobacterium*-mediated transformation of adult *Arabidopsis thaliana* plants by vacuum infiltration. Methods Mol Biol **82:** 259–266

Bouché N, Bouchez D (2001) *Arabidopsis* gene knockout: phenotypes wanted. Curr Opin Plant Biol **4:** 111–117

Boudet A (2000) Lignin and lignification: selected issues. Plant Physiol Biochem **38:** 81–96

Bradford M (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem **248:** 248–254

Brill EM, Abrahams S, Hayes CM, Jenkins CL, Watson JM (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

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

Dence C (1992) Lignin determination. *In* C Dence, S Lin, eds, Methods in Lignin Biochemistry. Springer Verlag, Berlin, pp 33–61

Dixon RA, Chen F, Guo D, Parvathi K (2001) The biosynthesis of monolignols: a "metabolic grid," or independent pathways to guaiacyl and syringyl units? Phytochemistry **57:** 1069–1084

Donaldson LA (2001) Lignification and lignin topochemistry: an ultrastructural view. Phytochemistry **57:** 859–873

Doyle J, Doyle J (1990) Isolation of plant DNA from fresh tissues. Focus **12:** 13–15

Estelle M, Somerville C (1987) Auxin-resistant mutants of *Arabidopsis thaliana* with an altered morphology. Mol Gen Genet **206:** 200–206

Franke R, Hemm MR, Denault JW, Ruegger MO, Humphreys JM, Chapple C (2002a) Changes in secondary metabolism and deposition of an unusual lignin in the *ref8* mutant of *Arabidopsis*. Plant J **30:** 47–59

Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C (2002b) The *Arabidopsis REF8* gene encodes the 3-hydroxylase of phenylpropanoid metabolism. Plant J **30:** 33–45

Goffner D, Van Doorsselaere J, Yahiaoui N, Samaj J, Grima-Pettenati J, Boudet AM (1998) A novel aromatic alcohol dehydrogenase in higher plants: molecular cloning and expression. Plant Mol Biol **36:** 755–765

Goujon T, Sibout R, Eudes A, MacKay J, Jouanin L (2003a) Genes involved in the biosynthesis of lignin precursors in *Arabidopsis thaliana*. Plant Physiol Biochem (in press)

Goujon T, Sibout R, Pollet B, Maba B, Nussaume L, Bechtold N, Lu F, Ralph J, Mila I, Lapierre C et al. (2003b) A new *Arabidopsis thaliana* mutant deficient in the expression of *O*-methyltransferase: I. Impact on lignins and on sinapoyl esters. Plant Mol Biol **51:** 973–989

Grima-Pettenati J, Feuillet C, Goffner D, Borderies G, Boudet AM (1993) Molecular cloning and expression of a *Eucalyptus gunnii* cDNA clone encoding cinnamyl alcohol dehydrogenase. Plant Mol Biol **21:** 1085–1095

Halpin C, Holt K, Chojecki J, Oliver D, Chabbert B, Monties B, Edwards K, Barakate A, Foxon GA (1998) Brown-midrib maize (*bm1*)-a mutation affecting the cinnamyl alcohol dehydrogenase gene. Plant J **14:** 545–553

Halpin C, Knight ME, Foxon GA, Campbell MM, Boudet AM, Boon JJ, Chabbert B, Tollier MT, Schuch W (1994) Manipulation of lignin quality by down regulation of cinnamyl alcohol dehydrogenase. Plant J **6:** 339–350

Hawkins S, Samaj J, Lauvergeat V, Boudet A, Grima-Pettenati J (1997) Cinnamyl alcohol dehydrogenase: identification of new sites of promoter activity in transgenic poplar. Plant Physiol **113:** 321–325

Hawkins SW, Boudet AM (1994) Purification and characterization of cinnamyl alcohol dehydrogenase isoforms from the periderm of *Eucalyptus gunnii* Hook. Plant Physiol **104:** 75–84

Höfte H, Desprez T, Amselem J, Chiapello H, Rouzé P, Caboche M, Moisan A, Jourjon M-F, Charpenteau J-L, Berthomieu P et al. (1993) An inventory of 1152 expressed sequences tags obtained by partial sequencing of cDNAs from *Arabidopsis thaliana*. Plant J **4:** 1051–1061

Jefferson R, Kavanagh T, Bevan M (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. EMBO J **6:** 3901–3907

Jones L, Ennos AR, Turner SR (2001) Cloning and characterization of *irregular xylem4* (*irx4*): a severely lignin-deficient mutant of *Arabidopsis*. Plant J **26:** 205–216

Jouanin L, Goujon T, Sibout R, Pollet B, Mila I, Maba B, Ralph J, Petit-Conil M, Lapierre C (2001) Tuning lignin structure through silencing, restoring, or increasing caffeic acid *O*-methyltransferase activity: evaluation in poplar and *Arabidopsis*. *In* Proceedings of the ISWPC Meeting, Nice, pp 25–28

Kiedrowski S, Kawalleck P, Hahlbrock K, Somssich IE, Dangl JL (1992) Rapid activation of a novel plant defense gene is strictly dependent on the *Arabidopsis RPM1* disease resistance locus. EMBO J **11:** 4677–4684

Kim H, Ralph J, Lu F, Pilate G, Leplé J-C, Pollet B, Lapierre C (2002) Identification of the structure and origin of thioacidolysis marker compounds for cinnamyl alcohol dehydrogenase deficiency in angiosperms. J Biol Chem **277:** 47412–47419

Koncz C, Schell J (1986) The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. Mol Gen Genet **204:** 383–396

Lapierre C, Pollet B, MacKay JJ, Sederoff RR (2000) Lignin structure in a mutant pine deficient in cinnamyl alcohol dehydrogenase. J Agric Food Chem **48:** 2326–2331

Lapierre C, Pollet B, Petit-Conil M, Toval G, Romero J, Pilate G, Leple JC, Boerjan W, Ferret VV, De Nadai V et al. (1999) Structural alterations of lignins in transgenic poplars with depressed cinnamyl alcohol dehydrogenase or caffeic acid *O*-methyltransferase activity have an opposite impact on the efficiency of industrial kraft pulping. Plant Physiol **119:** 153–164

Lapierre C, Pollet B, Rolando R (1995) New insights into the molecular architecture of hardwood lignins by chemical degradation methods. Res Chem Intermed **21:** 397–412

Li L, Cheng XF, Leshkevich J, Umezawa T, Harding SA, Chiang VL (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

Lurin C, Jouanin L (1995) RFLP of RT-PCR products: application to the expression of CHS multigene family in poplar. Mol Breed **1:** 411–417

MacKay JJ, Liu W, Whetten R, Sederoff RR, O'Malley DM (1995) Genetic analysis of cinnamyl alcohol dehydrogenase in loblolly pine: single gene inheritance, molecular characterization and evolution. Mol Gen Genet **247:** 537–545

MacKay JJ, O'Malley DM, Presnell T, Booker FL, Campbell MM, Whetten RW, Sederoff RR (1997) Inheritance, gene expression, and lignin characterization in a mutant pine deficient in cinnamyl alcohol dehydrogenase. Proc Natl Acad Sci USA **94:** 8255–8260

Mansell RL, Gross GG, Stoeckigt J, Franke H, Zenk MH (1974) Purification and properties of cinnamyl alcohol dehydrogenase from higher plants involved in lignin biosynthesis. Phytochemistry **13:** 2427–2436

Mellerowicz EJ, Baucher M, Sundberg B, Boerjan W (2001) Unravelling cell wall formation in the woody dicot stem. Plant Mol Biol **47:** 239–274

Meyer K, Shirley AM, Cusumano JC, Bell-Lelong DA, Chapple C (1998) Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. Proc Natl Acad Sci USA **95:** 6619–6623

Mitchell HJ, Hall JL, Barber MS (1994) Elicitor-induced cinnamyl alcohol dehydrogenase activity in lignifying wheat (*Triticum aestivum* L.) leaves. Plant Physiol **104:** 551–556

Pilate G, Guiney E, Holt K, Petit-Conil M, Lapierre C, Leple JC, Pollet B, Mila I, Webster EA, Marstorp HG et al. (2002) Field and pulping performances of transgenic trees with altered lignification. Nat Biotechnol **20:** 607–612

Ralph J, Hatfield RD, Piquemal J, Yahiaoui N, Pean M, Lapierre C, Boudet AM (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

Ralph J, MacKay JJ, Hatfield RD, O'Malley DM, Whetten RW, Sederoff RR (1997) Abnormal lignin in a loblolly pine mutant. Science **277:** 235–239

Reiter W-D (1998) *Arabidopsis thaliana* as a model system to study synthesis, structure, and function of the plant cell wall. Plant Physiol Biochem **36:** 167–176

Ruegger M, Chapple C (2001) Mutations that reduce sinapoylmalate accumulation in *Arabidopsis thaliana* define loci with diverse roles in phenylpropanoid metabolism. Genetics **159:** 1741–1749

Sambrook J, Fritsch E, Maniatis T (1989) Molecular cloning: A Laboratory Manual, Ed 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Sibout R, Baucher M, Gatineau M, Van Doorsselaere J, Mila I, Pollet P, Maba B, Pilate G, Lapierre C, Boerjan W et al. (2002) Expression of a poplar cDNA encoding a ferulate-5-hydroxylase increases S lignin deposition in *Arabidopsis thaliana*. Plant Physiol Biochem **40:** 1087–1096

Somssich IE, Wernert P, Kiedrowski S, Hahlbrock K (1996) *Arabidopsis thaliana* defense-related protein ELI3 is an aromatic alcohol:NADP(+) oxidoreductase. Proc Natl Acad Sci USA **93:** 14199–14203

Tavares R, Aubourg S, Lecharny A, Kreis M (2000) Organization and structural evolution of four multigene families in *Arabidopsis thaliana*: *AtLCAD*, *AtLGT*, *AtMYST* and *AtHD-GL2*. Plant Mol Biol **42:** 703–717

Turner S, Taylor N, Jones L (2001) Mutations of secondary cell wall. Plant Mol Biol **47:** 209–219

Van Doorsselaere J, Baucher M, Feuillet C, Boudet AM, Van Montagu M, Inze D (1995) Isolation of cinnamyl alcohol dehydrogenase cDNAs from two important economic species: alfalfa and poplar. Demonstration of a high homology of the gene within angiosperms. Plant Physiol Biochem **33:** 105–109

Verwoerd T, Dekker B, Hoekema A (1989) A small-scale procedure for the rapid isolation of plant RNAs. Nucleic Acids Res **17:** 2362

Williamson JD, Stoop JM, Massel MO, Conkling MA, Pharr DM (1995) Sequence analysis of a mannitol dehydrogenase cDNA from plants reveals a function for the pathogenesis-related protein ELI3. Proc Natl Acad Sci USA **92:** 7148–7152

Wyrambik D, Grisebach H (1975) Purification and properties of isoenzymes of cinnamyl-alcohol dehydrogenase from soybean-cell-suspension cultures. Eur J Biochem **59:** 9–15