Molecular characterization of a cellulose synthase gene (AaxmCesA1) isolated from an Acacia auriculiformis \( \times \) Acacia mangium hybrid

Seok Yien Christina Yong · Ratnam Wickneswari

Published online: 23 August 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract Cellulose is the major component of plant cell walls, providing mechanical strength to the structural framework of plants. In association with lignin, hemicellulose, protein and pectin, cellulose forms the strong yet flexible bio-composite tissue of wood. Wood formation is an essential biological process and is of significant importance to the cellulosic private sector industry. Cellulose synthase genes encode the catalytic subunits of a large protein complex responsible for the biogenesis of cellulose in higher plants. The hybrid Acacia auriculiformis \( \times \) Acacia mangium represents an important source of tree cellulose for forest-based product manufacturing, with enormous economic potential. In this work, we isolate the first cellulose synthase gene, designated AaxmCesA1, from this species. The isolated full-length AaxmCesA1 cDNA encodes a polypeptide of 1,064 amino acids. Sequence analyses revealed that AaxmCesA1 cDNA possesses the key motif characteristics of a CesA protein. AaxmCesA1 shares more than 75 % amino acid sequence identity with CesA proteins from other plant species. Subsequently, the full-length AaxmCesA1 gene of 7,389 bp with partial regulatory and 13 intron regions was also isolated. Relative gene expression analysis by quantitative PCR in different tissues of the Acacia hybrid, suggests the involvement of the AaxmCesA1 gene in primary cell wall synthesis of rapidly dividing young root cells. Similarity analyses using Blast algorithms also suggests a role in primary cell wall deposition in the Acacia hybrid. Southern analysis predicts that AaxmCesA1 is a member of a multigene family with at least two isoforms in the genome of the Acacia hybrid.

Keywords Acacia hybrid · Cellulose synthase gene · Conserved motif · Relative gene expression · Isoform · Promoter region

Introduction

Acacia is a genus of shrubs and trees consisting of approximately 1,400 species found worldwide, with more than 900 species native to Australia. Acacia species spread around the tropical-warm–temperate regions of both hemispheres, including Southern Asia, Africa and the Americas (Maslin et al. 2003). Many Australia tropical acacias have substantial commercial importance for the pulp and timber industries, particularly in Southeast Asia. The main Acacia species with high commercial values are Acacia mangium, A. auriculiformis, A. crassarcarpa and the more recently emerging A. auriculiformis \( \times \) A. mangium hybrid (McDonald et al. 2001). In Malaysia, the first natural hybridization of Acacia auriculiformis \( \times \) A. mangium hybrid (hereafter Acacia hybrid) was observed in an Acacia plantation in Ulu Kukut, Sabah in 1971 (Tham 1976). Better stem straightness compared to A. mangium was found in the hybrid; the hybrid also inherits better self-pruning ability, stem roundness and disease resistance than A. auriculiformis. The hybrid also possesses intermediate characteristics between the two parental Acacia species (Bowen 1981) and appears to have higher cellulose and lower lignin contents, which give better pulp yields than both parental trees (Yamada et al. 1990).

Cellulose is the major component in wood, which is present universally in plant cell walls and provides mechanical strength to the plant. Cellulose accounts for about 20 %
and 50% of the primary and secondary cell walls in higher plants, respectively. Membrane-bound cellulose synthase enzyme complexes (CelS) are believed to play a key role in the biosynthesis of cellulose (Saxena and Brown 2005). These complexes are discernible as hexameric rosettes under freeze-fracture electron microscopy and consist of 36 cellulose synthase (CesA) proteins per rosette (Doblin et al. 2002; Saxena and Brown 2005; Somerville 2006). Since the first isolation of a plant CesA gene from a cotton tree 16 years ago, many similar attempts have been made in numerous plants with the aim of understanding the role of CesA genes and the mechanisms involved during cellulose biosynthesis. Previous studies revealed that the catalytic subunits of CelS in many plant species are encoded by a family of genes or isoforms (Richmond and Somerville 2000; Djerbi et al. 2005). Isoforms are alternative forms of a gene generated through splicing (Wang et al. 2012) and are found commonly in many plant species. Isoforms with differential expression patterns in different types of tissues throughout different plant developmental stages have also been reported (Hayashi et al. 2005; Joshi et al. 2004; Nairn and Haselkorn 2005; Roberts et al. 2004; Taylor et al. 2003). Mutant analyses conducted on ten CesA genes from Arabidopsis thaliana (AtCesAs) also revealed a number of glycosyltransferases encoded by these AtCesAs that function in cellulose biosynthesis in primary and secondary cell walls (Burn et al. 2002; Taylor et al. 1999; Turner and Somerville 1997). A study on Eucalyptus had also identified several CesAs associated with primary and secondary cell wall formations (Ranik and Myburg 2006). Cellulose biogenesis is a complex process, which still requires ongoing study to comprehend completely its pathways and mechanisms. Even with the completion of the Arabidopsis thaliana genome sequencing project, researchers still struggle to understand the biogenesis of cellulose in A. thaliana, let alone other plants that are less well studied.

Despite being cultivated for its timber, as well as for the pulp and paper industries, not much is known about the genetics underlying cellulose biosynthesis in the Acacia hybrid. Here, we report the first isolation and characterization of a cellulose synthase gene from the commercially valuable Acacia hybrid with the aim of understanding the deposition of cellulose in its primary and secondary cell walls.

Materials and Methods

Plant Material and Total RNA Preparation

Young leaves, a mixture of phloem and xylem (hereafter referred to as inner bark), flowers and greenish seedpod tissues were collected from a 7-year-old Acacia hybrid tree (AaHyF113) planted in Plot W, Plant Biotechnology Laboratory, Universiti Kebangsaan Malaysia. Young root tissue was obtained from tissue-cultured Acacia hybrid (Clone M5) provided by the Forest Research Institute Malaysia. Tissues were immersed in liquid nitrogen immediately upon collection. Total RNA was extracted from the five tissues using a RNeasy Plant Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer’s recommendations with minor modification, where 1% polyethylene glycol 6000 (PEG 6000) was added to the protocol to enhance total RNA yield. Additional on-column DNase digestions were performed twice using RNase free DNase I (Promega, Madison, WI). The concentration and quality of the total RNA isolated was estimated from the ratio of the absorbance measured at 260/280 nm, and by gel electrophoresis. cDNA was synthesized using Superscript II Reverse Transcriptase (Invitrogen, Carlsbad, CA) and cDNA quality was assessed by end point reverse transcription PCR using gene specific primers (GSP).

Relative Gene Expression Analysis by Quantitative PCR

GSP for AaxmCesA1 and actin genes were designed based on the 3' EST sequences identified from our Acacia hybrid EST dataset (Yong et al. 2011), using Primer Premier 5 (http://www.PremierBiosoft.com) and Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3) software, respectively. The primer sequences for CesA and actin genes were: CesAForw1: 5' CATACCCCGGTGGTGCATCC, CesARerv1: 5' CATCAACATGCTCATTTCATTG and ActinForw: 5' GGTAACATTGTCCTCTGTGGT, ActinRerv1: 5' CATCG TATTCTGCTCTCGA. The relative expression level of AaxmCesA1 gene was investigated in the inner bark, young seedpod, leaf, young root and flower tissues, with the inner bark tissue as calibrator sample. PCR reactions were carried out using SYBR Green Supermix (Bio-Rad, Hercules, CA) in a iCycler iQ5 Real-Time PCR Detection System (Bio-Rad). PCR was initiated with a pre-denaturation step at 95 °C for 4 min, followed by 60 cycles of denaturation, annealing and extension at 95 °C/15 s, 55–60 °C/30 s, 72 °C/15 s, respectively, then a final extension at 72 °C for 4 min. All standard dilutions, controls and samples were performed in four replicates. Standard curves for AaxmCesA1 and actin were constructed using 5-fold serial dilution of the RT product synthesized from inner bark tissue. The 10μL quantitative PCR (qPCR) reactions contained a final concentration of 1x SYBR Green Supermix, 0.2 μm of each primer and 10 ng template cDNA. Melt curve analysis was carried out from 55 °C to 95 °C at the end of the amplification to confirm that the individual qPCR product corresponded to a single homogeneous DNA species. qPCR products were also sequenced directly using ABI 3700 (Applied Biosystems, Foster City, CA) at Macrogen Corporation, Korea, to ensure amplification specificity.
Isolation of Full-Length *AaxmCesA1* cDNA

Total RNA was extracted from the leaf samples of a single *Acacia* hybrid individual plant (HyAa/F1 113) using an RNeasy Plant Mini Kit (Qiagen). The full-length *AaxmCesA1* cDNA was isolated using a BD SMART™ RACE cDNA amplification kit (Clontech, Palo Alto, CA) according to manufacturer’s instructions. RACE primers were designed using Primer Premier 5 software. The primers of RamA1: 5’ CCTTCCGAGCAGACCCTTCA, RCesA1-1Rb: 5’ GCCTGCTAAGGAAAATCTCAATGGACC and NRA11R: 5’ GCAATAGGCACCACAGACTCAAAGG were used for 5’ RACE PCR; FAMA1: 5’ GGTGGCATTCCTCCCTCAAC was used for 3’ RACE PCR. Each RACE-PCR reaction containing 2.5 μL first strand cDNA, 1x UPM A Mix, 0.2 μM GSP and 41.5μL master mix was performed using the Mastercycler PCR machine (Eppendorf, Hauppauge, NY). PCR was performed using a touch down cycling profile starting with 5 cycles of amplification at 94 °C/30s, 72 °C/3 min, followed by another 5 cycles at 94 °C/30s, 70 °C/30s, 72 °C/3 min, then 25 cycles at 94 °C/30s, 68 °C/30s, 72 °C/4 min, and finally an extension at 72 °C for 16 min. PCR products were gel purified and sequenced from both ends at Macrogen Corporation, Korea.

**Table 1** Gene specific primers (GSP) used in intron isolation

| Primer pair | GSP Primer sequence (5’—3’) | Annealing temperature (°C) | Product size (bp) |
|-------------|-----------------------------|---------------------------|------------------|
| 1 | FamCes | GGTGGCATTCCTCCCTCAAC | 64 | 897 |
| RamCes | CCTTCCGAGCAGACCCTTCA | 64 |
| 2 | CesA1-F2 | GAAGAGGGCAGCAAAAAGAAC | 64 | 648 |
| CesA1-R2 | GGGAGCACACAGTAAGCAATC | 60 |
| 3 | A1WF8 | CCAATCTGTGCGAACTAC | 48.3 | 1,050 |
| A1WF8b | CCTCCACTATGACCTAAG | 43.7 |
| 4 | A1WF9 | GGAAGGGGATGACGATG | 52.8 | 1,013 |
| AIWR9 | GCAATAGGCACCACAG | 52.7 |
| 5 | A1WF10 | GGGCCAAAGGCCCCCTGA | 55.9 | 1,782 |
| AIWR10 | GCCAACGGAACGACCT | 56.5 |
| 6 | A1WF11 | GAAAGGGTCTCCTCACA | 36.7 | 504 |
| AIWR11 | TAAACAACAGATAAGG | 36.6 |
| 7 | A1WF17 | TTGGTTATGGTTTCTCGTG | 49.2 | 1,534 |
| AIWR17 | TGAGTTGAAAGGGTGA | 49.2 |
| 8 | CES1In1F | CTGGGGAATGTGCGGTTG | 65 | 1,000 |
| CES1In1R | TTCTTACAACAGGACACAAAACGGTGTA | 65 |
| 9 | CES1In2F | CATACCCAGTGTGCTGCC | 56 | 1,420 |
| AIWR9 | TCACTCCCTACGTCATCCC | 56 |
of each adaptor and gene specific primer and 2 U Hot start Taq, and amplification used a hot-start touch-down PCR cycling profile. Amplification started with a pre-denaturation at 96 °C for 3 min, followed by 7 cycles of amplification at (96 °C/15 s, 72 °C/3 min), then another 20 cycles at (96 °C/15 s, 68 °C/3 min) and ended with a final extension at 68 °C for 15 min. Gel-purified secondary PCR product was sequenced from both ends using ABI PRISM 3730XL DNA analyzer (Applied Biosystems) at Research BioLabs Technologies, Singapore.

Sequence and Structure Analyses

The full-length AaxmCesA1 cDNA sequence isolated was subjected to similarity search using Blast (http://www.ncbi.nlm.nih.gov). A total of 24 CesA protein sequences of the closest homologues were retrieved from the NCBI database (http://www.ncbi.nlm.nih.gov) for phylogenetics analysis. The accession numbers for these sequences are: AAT66940.1, XP 003522623.1, AEK31219.1, XP 002515536.1, XP 002324291.1, AAY60847.1, AAY78952.3, ACT16001.1, AAY43218.1, NP 194967.1, AAO25536.1, NP 001054788.1, NP 001105574.1, AAP97497.1, AAF89963.1, XP 002880684.1, NP 180124.1, NP 001051648.1, NP 001105621.1, ADZ16121.1, ADZ16119.1, XP003540527.1, AAT66941.1, and AAY43219.2. In addition to these 24 sequences, selected CesA protein sequences from Arabidopsis (NP 195645.1, NP201279.1, NP 196136.1, Q84JAJ6, Q8L778, NP 197244.1 and NP 567564 .1), Hordeum vulgare (AY483150, AY483152, AY4831555), Oryza sativa (BAD30574), Zea mays (BAD30574), Gossypium hirsutum (U58283), Bambusa oldhamii (AAY43221.1) and Eucalyptus grandis (U58283, AAY60843.1, AAY60845.1, AAY60846.1 and AAY60848.1) were included. Multiple sequences alignment was performed using ClustalW2 (http://www.ebi.ac.uk/Tools/es/cgi-bin/clustalw2) and phylogenetic analysis was carried out using MEGA Version 5 (Tamura et al. 2011). The unrooted phylogenetic tree was constructed using the neighbor-joining method based on distance matrices with 1,000 bootstrap replicates and the Jones-Taylor-Thornton model.

Conserved domains were detected using Conserved Domain software (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), ClustalW2 and Interproscan software (http://www.ebi.ac.uk/Tools/cgi /iprscan.cgi). Promoter sequence was analyzed using Prediction of Plant Promoter (http://www.softberry.com/berry.html) and Neural Network Prediction (http://www.fruitly.org/seq_tools/promoter.html) software. The cDNA, introns and promoter sequences amplified from RACE-PCR, primer walking and genome walking respectively, were analyzed using CAP contig assembly (Huang and Madan 1999) and ClustalW2 to predict the AaxmCesA1 full-length gene sequence. The software “Finding 5’, internal and 3’ coding exons” (http://linux.softberry.com/cgi-bin/progrmas/gfind/fex.pl) was applied to predict exon–intron boundaries. The full-length gene sequence was also compared with the full-length gene sequences of other species using ClustalW2 to predict intron and exon positions.

Gene Isoform Analysis

Genomic DNA was extracted from young leaves of an Acacia hybrid individual plant (AaHy113F1) using the CTAB method (Doyle and Doyle 1990). Genomic DNA (15 μg) was digested with two restriction enzymes (BsrBI and BstZ17I; New England Biolabs) at 37 °C for 18 h. The phenol-chloroform purified digested DNA was electrophor- esed on a 0.8 % agarose gel in 1x TAE buffer at 25 V for 16 h and transferred to Hybond N filter membrane (Amersham, Little Chalfont, UK) by downward capillary transfer (Chomczynski 1992) in alkaline buffer (Sambrook and Russell 2001). The blot was hybridized with a 1.6 kb AaxmCesA probe labeled with 32P deoxycytidine 5′-Triphosphate (dCTP) disodium salt using Megaprime DNA Labelling Systems (Amersham), and unincorporated probe was filtered on a Sephadex G-75 column. Hybridization and washing were carried out according to standard procedures described in Sambrook and Russell (2001).

Results

Relative Gene Expression Analysis

The expression levels of AaxmCesA1 in young root, inner bark, young leaf, greenish seedpod and flower tissues were investigated using qPCR. Melt curve analysis showed a single peak for the amplification of AaxmCesA1 gene denoting PCR amplification specificity. Direct sequencing of the qPCR products also confirmed the originality of the amplified products. The highest level of transcript abundance was observed in young root tissue whereas the lowest was evident in inner bark tissue. The expression levels of the AaxmCesA1 gene in the Acacia hybrid were found in a decreasing pattern from leaf, seed pod to flower (Fig. 1). However, statistical analysis using independent t-test at 95 % confidence interval shows that the differences in the expression levels between inner bark tissue and all other tissues studied were not statistically significant with the exception of young root tissue.

Full-length AaxmCesA1 cDNA

The isolated full-length cDNA of the AaxmCesA1 gene is 3,835 bp long, which includes the 3’ UTR (437 bp) and 5’
UTR (212 bp) regions. The predicted open reading frame is 1,062 residues in length and encodes a protein of 146 kDa with a predicted pI of 8.84. Multiple alignment analysis performed on the deduced amino acid sequences of the Acacia hybrid and CesA protein sequences of other species revealed a high degree of amino acid similarity throughout most of the sequences length, diverging mostly at the N and C-terminal ends and hypervariable regions (HVR). Although the various CesA proteins from different species varied in their degree of sequence similarity, a conserved motif (D, D, D, QXXRW) was found in all sequences analyzed (Fig. 2). The first, second, third D and QXXRW were found at positions 710, 758, 891 and 929 residues of the translated protein sequence of the Acacia hybrid, respectively.

Two HVRs were also predicted from the deduced amino acid sequence of AaxmCesA1. HVRI was found before the first aspartic acid in Domain A and HVRII between the second aspartic acid in Domain A and third aspartic acid in Domain B (Fig. 3). HVRI was predicted at positions 201–330, and HVRII at positions 760–830. Other domains, including pfam cellulose_synt, which was identified in the region between residues 394 and 1126, glycosyltransferase family_2 at residues 882–1020, COG1215 (glycosyltransferase) at residues 882–1048 and bcsA at residues 880–1011, were found in the deduced amino acid sequence of the Acacia hybrid. Interpro Scan performed on the full-length AaxmCesA1 cDNA identified additional domains, namely zinc fingers (RING/FYVE/PHD-type) and transmembrane regions. Two putative zinc finger domains were predicted at residues 82–160 and 161–200 in the Acacia hybrid amino acid sequence. Two transmembrane domains were found in the amino terminal region and another six in the carboxy terminal region. The first two transmembrane domains were located before the beginning of the globular domain A at positions 361–381 and 390–410. The other six transmembrane domains were located just after globular domain B at positions 933–953, 963–983, 1002–1020, 1035–1053, 1116–1138 and 1249–1269.

Blast X analysis revealed significant matches between the deduced AaxmCesA1 protein sequence and the CesA protein sequences of A. mangium (AY643519.1), Eucalyptus spp (AEK31219.1; AAY60847.1), Bohmeria nivea (AY78952.1), Populus spp (XP_002308657.1; AA025536.1), Arabidopsis thaliana (NP_194967.1), Bambusa oldhamii (AAY43218.1; AAY43216.1) and Ricinus communis (XP_002515536.1), with E-value = 0 and amino acid identity of more than 70 %. Based on the predictions of Blast N analysis, the AaxmCesA1 cDNA shares 95 % nucleotide sequence similarity with Acacia mangium CesA1 (gb|AY643519.1) at E-value = 0. It also exhibits over 75 % nucleotide sequence similarity with Ricinus communis CesA6 (ref|XM002515490.1), Glycine max CesA1 (XP_003522623.1), Populus tremuloides CesA4 (gb|162181.1), Populus trichocarpa CesA (ref|XM_002308621.1), Eucalyptus camaldulensis CesA5 (gb|HQ864587.1) and Arabidopsis thaliana CesA5 (ref|NM_119393.2). The high level of similarities in the nucleotide and protein sequences of these CesA genes suggest an orthologous relationship among them. Blast analyses also suggested a possible role for the AaxmCes1 gene in primary cell wall formation of the Acacia hybrid as most of the matched CesA proteins with known function from other species are related to primary cell wall formation.

A total of 43 CesA amino acid sequences from other species, in plants ranging from dicots, monocots, angiosperms to gymnosperms, were retrieved from the NCBI website (http://www.ncbi.nlm.nih.gov) and compared with the AaxmCesA1 predicted protein sequence. Figure 4 shows an unrooted tree constructed based on neighbor-joining analysis for the deduced amino acid of AaxmCesA1 with the 43 protein sequences of other plant species. Although not all the CesA genes of these plants have yet been identified, a preliminary conclusion that can be drawn is that, as in most large gene families, orthologous genes are more similar than paralogous genes. Thus, many of the groups in this analysis contain members from plants of both monocot and dicot or angiosperm and gymnosperm lineages.

Full-Length AaxmCesA1 Gene

A total of nine primer pairs were used to isolate all the intronic regions of the AaxmCesA1 gene. The alignment and assembly of the full-length cDNA sequences obtained from RACE-PCR, and the sequences of introns amplified through primer walking, resulted in a partial gene sequence of 6,545 bp with the predicted introns. The gene structure was predicted based on multiple alignment analysis of the full-length AaxmCesA1 gene sequence with the CesA full-length gene sequences of A. thaliana CesA1 (NM119393.2) and CesA10 (128111.2) and full-length cDNA sequences of A. mangium (AY643519) and E. grandis (DQ014509.1).
Prediction of these 13 introns was also supported by “Finding 5′, internal and 3′ coding exons” software.

A 1-kb fragment was amplified successfully from *Acacia* hybrid genomic DNA digested with *HpaI* using the genome

![Springer](https://www.springer.com)
walking approach, which yielded a partial promoter region of 844 bp. Promoter Prediction and Neural Network Prediction software predicted several core promoter elements from the partial nucleotide sequence of the promoter region isolated in this study. The transcription start site (TSS) was found at position 840 bp. Three core elements of the promoter, namely TATA, CAAT and GC boxes were predicted at −32, −53 and −151 bp, respectively, upstream of the TSS of full-length AaxmCesA1 gene. A full-length AaxmCesA1 gene of 7,389 bp was then obtained by assembling the promoter region sequence with the partial promoter region sequence. The isolated AaxmCesA1 gene (ref|NC003075.7) located on chromosome 4 and the AtCesA10 gene located on chromosome 2 of the A. thaliana genome, with E-values of 3e−149 and 0, respectively.

Southern blot analysis using genomic DNA digested with restriction enzymes BstZ17I and HpaI revealed several hybridization bands (Fig. 5). Knowing that most CesA genes in plant species are larger than 4 kb, the hybridization pattern in the Southern analysis suggested that the CesA gene family might consist of up to four isoforms in the Acacia hybrid genome. However, two of these hybridization bands (indicated by arrows in Fig. 5) showed higher intensity than other bands in both digested genomic DNA samples. Based on this, we are confident that there are at least two AaxmCesA isoforms present in the genome of the Acacia hybrid. As the 1.6 kb AaxmCesA1 probe included the HVRII region in its design, sequence variation might cause failure of the probe to hybridize and detect all the different isoforms. This also reflects the possible high sequence variation in the HVRII region of the AaxmCesA genes of the Acacia hybrid but further investigation is necessary.

Discussion

Isolation of a candidate CesA gene is a fundamental step towards understanding the gene functions and mechanisms involved during cellulose biogenesis. Although mutant analyses conducted on Arabidopsis have provided significant information on gene function, we are still far from understanding the complete pathway of cellulose biosynthesis in plants. Furthermore, detailed information on the CesA gene family is available only from a limited number of plants, such as barley (Burton et al. 2004), Arabidopsis (Richmond 2000) and popular (Djerbi et al. 2004). In the current study, we have isolated and characterized the AaxmCesA1 gene from the Acacia auriculiformis x mangium hybrid for the first time.
study conducted on the CesA genes of *Oryza sativa* revealed that *OsCesA1, OsCesA3, OsCesA8* and *OsCesA4*, *OsCesA7, OsCesA9* are strongly co-expressed in tissue of typical primary and secondary cell walls respectively. These findings suggest that each set of these *OsCesA* subunits belongs to the same cellulose synthase complex, which catalyze similar cellulose biosynthesis (Wang et al. 2010; Tanaka et al. 2003). Different types of CesA isoforms have also been found to be co-expressed in barley; *HvCesA1, HvCesA2* and *HvCesA6* (Burton et al. 2004) and *Eucalyptus; EgCesA1, EgCesA2* and *EgCesA3* (Ranik and Myburg 2006). Thus, it is very likely that we will find other co-expressing CesA isoforms responsible for the biogenesis of cellulose in either primary or secondary cell wall in the *Acacia* hybrid.

Based on our Blast analysis, we propose a role in primary cell wall synthesis for the isolated *AaxmCesA1*, concuring with our relative gene expression study. Blast analyses revealed a high degree of amino acid sequence similarity between the *AaxmCesA1* isolated in this study and *GmCesA1, EgCesA5* (Ranik and Myburg 2006), *OsCesA1* (Wang et al. 2010), *BoCesA2* (Chen et al. 2010), *AtCesA1* (Desprez et al. 2007) and *AmCesA1*. The high degree of amino acid sequence similarity suggests that *AaxmCesA1* might be orthologous to *EgCesA5, AtCesA1, BoCesA2* and *OsCesA1*, which are all known to be associated with primary cell wall formation. Mutant analysis performed on *Arabidopsis* in past study showed that a single amino acid substitution in *AtCesA1* resulted in a marked reduction of crystalline cellulose synthesis, and the disassembly of cellulose synthase complexes (Arioli et al. 1998). This finding suggests that *AtCesA1* is involved in primary cell wall synthesis in *Arabidopsis*.

The length (3,835 bp) of the *AaxmCesA1* cDNA isolated in the current study is similar to that of *A. thaliana* (3–3.5 kb; Richmond 2000) and poplar (3–3.3 kb; Joshi et al. 2004). The D, D, QXXRW signature motif predicted from the deduced amino acid sequence of the *Acacia* hybrid is a strong predictor for family 2 of β-glycosyltransferases (Campbell et al. 1997). β-Glycosyltransferases can be categorized into processive or non-processive glycosyltransferases depending on the transfer of multiple or single sugar residue to an acceptor molecule. Processive glycosyltransferase is composed of two conserved domains (domains A and B), while non-processive glycosyltransferase contains only domain A (Saxena and Brown 1997). Based on this classification, the *AaxmCesA1* isolated in our study is postulated to belong to the processive class because it contains both domains A (consisting two conserved aspartic acid, D) and B (consisting one aspartic acid, D and the QXXRW sequence). The two aspartic acid residues in domain A, and a single aspartic acid residue and the sequence QXXRW in domain B are found to be strictly conserved in all the CesA genes analyzed here. This finding

Cellulose is the main component in cell wall formation and is found in all plant tissues, but the deposition of cellulose in different cell types varies (Delmer and Amor 1995). In our analysis, qPCR relative gene expression patterns revealed that *AaxmCesA1* is expressed in all the tissues investigated but at different levels. The higher levels of transcript abundance observed in young root tissue most probably demonstrates an active involvement of this *AaxmCesA1* in the primary cell wall formation of actively dividing young root tissue. Fagard et al. (2000) and Wang et al. (2001) also showed that some CesAs acted early in the process of root hair outgrowth and elongation. The varying expression patterns of this *AaxmCesA1* in the five tissues tested here suggested the possibility that more than one CesA is needed for the biosynthesis of cellulose in primary and secondary cell walls in the *Acacia* hybrid. Previous studies have proposed the involvement of at least three types of CesA isoforms namely, α1, α2 and β, in the spontaneous arrangement of CesAs in the cellulose synthase complex of higher plants (Ding and Himmel 2006; Doblin et al. 2002). In *Arabidopsis*, a cellulose synthase complex consisting of three unique subunits (*AtCesA1, AtCesA3* and *AtCesA6*) was required for primary cell wall synthesis (Persson et al. 2007). Samuga and Joshi (2002) also showed that two subunits viz., *PtrCesA1* and *PtrCesA2* are involved in secondary cell wall formation in *Populus tremuloides*. Another
Amino acid distances between conserved residues, and their position, in the deduced amino acid sequence of the *Acacia* hybrid

| Conserved residues | Amino acid distance between conserved residues in *CesA* of *Acacia* hybrid (residues) | Range of amino acid distance (residues) |
|--------------------|---------------------------------------------------------------------------------|--------------------------------------|
| D1−D2              | 48                                                                              | 46–120                               |
| D2−D1              | 133                                                                             | 92–138                               |
| D3−QXXRW           | 38                                                                              | 33–38                                 |

We also predicted eight transmembrane regions in the *AaxmCesA1* amino acid sequence; the distribution of these regions is similar to that of Richmond’s prediction in *A. thaliana*. According to Richmond (2000), all members of the cellulose synthase superfamily appear to be integral membrane proteins, with one or two transmembrane domains in the amino terminal region, and three to six transmembrane domains in the carboxy terminal region of the protein. This nature of all cellulose synthase superfamily members is crucial, as cellulose is a water-insoluble polymer. Thus, cellulose cannot be synthesized inside the cells but can only be deposited on the external surface of each plant cell. These transmembrane helices are believed to be involved in creating a channel whereby the glucan chain is secreted outside the cell to form new cell walls (Kurek et al. 2002). We also found that the intron–exon organization is highly conserved among most of the *CesA* genes analyzed in this study. We predicted 13 intron regions in the full-length *AaxmCesA1* gene, which is comparable to the arrangement found in *A. thaliana* (Richmond 2000) and poplar (Joshi et al. 2004). The highly conserved organization of the *CesA* gene structure across species might reflect its importance for gene functionality in plants.

Analysis of the *AaxmCesA1* partial promoter sequence of 840 bp predicted several core elements. These core elements consist of TATA, CAAT and GC boxes, which are generally found in the regulatory regions of promoters and are involved in the regulation of gene expression. Little is known about the transcriptional regulation of this gene family generally, and the *Acacia* hybrid in particular. A study conducted by Creux et al. (2008) demonstrated that *CesA* promoters from the *Eucalyptus* tree are functional in *Arabidopsis* despite the low sequence similarity between them. Others reported that co-expressed genes harbor conserved transcription factor binding sites (Harmer et al. 2000). Study of the promoter region of *AaxmCesA1* is important as it provides an avenue to regulate the expression of *AaxmCesA1* in the *Acacia* hybrid. We aim to isolate the complete promoter region of the *AaxmCesA1* and investigate the regulatory mechanism of this gene in our future efforts. Isolation and mutant analyses of the promoter could provide experimental proof for the hypothetical mechanisms orchestrating the transcriptional activity of this gene.

Past investigations conducted on many plant species, including poplar, maize and cotton, have demonstrated that *CesA* belongs to a multigene family. Desprez et al. (2007) have shown that at least three *CesA* isoforms are needed for the biosynthesis of cell walls. Different isoforms have been found to be co-expressed in many plant species including rice (Tanaka et al. 2003), *Arabidopsis* (Taylor et al. 2003), poplar (Joshi et al. 2004) and barley (Burton et al. 2004). Mutant analyses of *CesA1*, *CesA3* and *CesA6* showed defects in the primary cell walls of *Arabidopsis* (Arioli et
Therefore, we expect AAXM CesA to identify other family in the Acacia aspartic residues D1, D2, D3 enable the identification of AAXM CesA1. We conjectured that there are other CesA genes whereas the HVRII helps in differentiating the different isoforms. HVRII was designated as a class specific region (CSR) and has been used to isolate cDNAs encoding CesAs. This is because this region could discriminate each member of the CesA family efficiently (Liang and Joshi 2004; Ranik and Myburg 2006). Therefore, we expect to identify other AAXM CesAs and at the same time distinguish the different isoforms of the AAXM CesA gene family in the Acacia hybrid using this probe. In this current study, we failed to identify a third isoform. Nevertheless, this failure might possibly reflect high sequence variation in the HVRII region of the AAXM CesA genes of the Acacia hybrid. In a previous study conducted on aspen, Samuga and Joshi (2002) concluded that orthologous genes show a higher percentage of identity than paralogous genes. In their work, they found that CesA from the same species exhibits more amino acid sequence variation than between CesA of different species. This variation might explain why we did not detect a third isoform using this probe.

Assessment of gene function through mutant analysis and gene expression studies are important. Although remarkable progress has been made in recent decades, especially in the model plant Arabidopsis, towards unraveling the process of cellulose biosynthesis, many unanswered questions still remain. While mutant analysis can reveal whether a gene is critical for a particular process, it cannot provide us with information on the exact pathways or mechanisms that take place during cellulose deposition. We still have no complete picture of how plants synthesize cellulose during primary and secondary cell wall formation. In the case of the Acacia hybrid, only the AAXM CesA gene has been isolated thus far. Thus, efforts to isolate and investigate all the AAXM CesA genes must continue in order to obtain more a comprehensive insight into the biogenesis of cellulose in the Acacia hybrid. Our current report represents the first step towards achieving this goal.

Acknowledgments This study was funded by the Ministry of Science, Technology and Innovation Malaysia under the Intensified Research Project Area Scheme (Priority Research Grant 01-02-02-0010-PR0003/03-01) and eScience (Project No. 02-01-02-SF0403).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Arioli T, Peng L, Betzner AS, Burn J, Witte W, Herth W, Camilleri C, Hofte H, Plazinski J, Birch R, Cork A, Glover J, Redmond J, Williamson RE (1998) Molecular analysis of cellulose biosynthesis in Arabidopsis. Science 279:717–720

Bowen MR (1981) Acacia mangium. A note on seed collection, handling and storage techniques including some experimental data and information on A. auriculiformis and the probable A. mangium x A. auriculiformis hybrid. Occasional Technical and Scientific Notes, Seed Series Number 3. FAO/UNDP-MAL/78/5009

Burn TE, Hocart CH, Birch RJ, Cork AC, Williamson RE (2002) Functional analysis of the cellulose synthase genes CesA1, CesA2 and CesA3 in Arabidopsis. Plant Physiol 129:797–807

Burton RA, Shirley NJ, King BJ, Harvey AJ, Fincher GB (2004) The CesA gene family of barley—quantitative analysis of transcripts reveals two groups of co-expressed genes. Plant Physiol 134:224–236

Campbell JA, Davies GJ, Bulone VV, Henriass S (1997) A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities. Biochem J 326:929–939

Chen CY, Hsieh MH, Yang CC, Lin CS, Wang YJ (2010) Analysis of the cellulose synthase genes associated with primary cell wall synthesis in Bambusa oldhamii. Phytochemistry 71:1270–1279

Chomczynski P (1992) One-hour downward alkaline capillary transfer for blotting of DNA and RNA. Anal Biochem 201:134–139

Creux NM, Ranik M, Berger DK, Myburg AA (2008) Comparative analysis of orthologous cellulose synthase promoters from Arabidopsis, Populus and Eucalyptus: evidence of conserved regulatory elements in angiosperm. New Phytol 179:722–737

Delpier DP, Amor Y (1995) Cellulose biosynthesis. Plant Cell 7:987–1000

Desprez T, Juraniec M, Crowell EF, Jouy H, Pochylowa Z, Parry F, Hofte H, Gonneau M, Vernhette S (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana. Proc Natl Acad Sci USA 104(39):15572–15577

Ding SY, Himmel ME (2006) The maize primary cell wall microfibril: a new model derived from direct visualization. J Agric Food Chem 54:597–606

Djerbi S, Lindskog M, Arvedst L, Sterky F, Teeri TT (2005) The genome sequence of black cottonwood (Populus trichocarpa) reveals 18 conserved cellulose synthase (CesA) genes. Planta 221:739–746

Djerbi S, Aspeborg H, Nilsson P, Sundberg B, Mellerowicz E, Blomqvist K, Teeri TT (2004) Identification and expression analysis of genes encoding putative cellulose synthase (CesA) in the hybrid aspen, Populus tremula (L.) x P. tremuloides (Michx). Cellulose 11:301–312

Doblin MS, Kurek I, Jacob-Wilk D, Delpier DP (2002) Cellulose biosynthesis in plants: from genes to rosettes. Plant Cell Physiol 43:1407–1420

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Fagard M, Desnos T, Desprez T, Goulet F, Regrefger G, Mouselle G, Maureen M, Rayon C, Vernhette S, Hofte H (2000) PRO-CUSTE1 encodes a cellulose synthase required for normal cell elongation specifically in roots and dark-grown hypocotyls of Arabidopsis. Plant Cell 12:2409–2423
