Bioenerg. Res.
DOI 10.1007/s12155-010-9080-8

Collection and Analysis of Expressed Sequence Tags Derived from Laser Capture Microdissected Switchgrass (*Panicum virgatum* L. Alamo) Vascular Tissues

Avinash C. Srivastava · Karuppaiah Palanichelvam · Junying Ma · Jarrod Steele · Elison B. Blancaflor · Yuhong Tang

Abstract Switchgrass is a perennial C4 grass that thrives in a wide range of North American habitats and is an emerging crop for the production of lignocellulosic biofuels. Lignin is an integral component of secondary plant cell walls that provides structural rigidity to the cell wall but it interferes with the conversion of cellulose to fermentable sugars by preventing chemical access to cellulose. Thus, one strategy for improving production of cellulosic ethanol is the down-regulation of lignin in plants. To achieve this goal, it is important to understand the molecular processes involved in vascular tissue development, lignification and secondary wall synthesis. Since active lignification occurs in the vascular system of the plant, we refined a protocol for isolating vascular tissues using laser-capture microdissection (LCM) in an effort to identify transcripts of switchgrass involved in lignification and secondary cell wall synthesis. ESTs (5,734) were sequenced from the *cDNA* libraries derived from laser microdissected vascular tissues. These Sanger sequences converged into 2,766 unigenes with an average length of 652 bp. Gene ontology of the unigenes indicated that 11% of the sequences were lignin and cell wall related. Several transcription factors involved in lignin and secondary cell wall synthesis and sugar- or vesicle-mediated transporters were also present in this EST data set. In situ hybridization of seven representative genes confirmed the preferential expression of five genes in the vascular tissues. Comparison of our switchgrass vascular tissue derived ESTs with that of other plant species validated our LCM approach. Furthermore, our switchgrass vascular tissue ESTs revealed additional lignin and cell wall related genes that were not present in other existing switchgrass EST collections. Inventory of the switchgrass vascular tissue ESTs presented here provides an important genomic resource for mining genes to reduce recalcitrance in this important bioenergy crop.

Keywords Expressed sequence tags · Laser-capture microdissection · Lignin · Secondary cell wall · Switchgrass · Vascular tissue

Introduction

Biofuel is defined as energy derived from any biological carbon source. Biofuels exist as bioether, biodiesel, biogas, syngas or bioethanol. Among these forms of biofuels, bioethanol, which is obtained from the conversion of carbon-based feedstock, can be produced from a variety of crops such as sugar cane, *Miscanthus*, sorghum, switchgrass, barley, *Hibiscus cannabinus*, potato, cassava, sunflower, maize, wheat, and cotton. Among these crops, research on switchgrass (*Panicum virgatum*) as a bioenergy crop, has increased dramatically in recent years [2, 38]. Switchgrass is a warm season perennial grass with immense biomass [23, 35] and is one of the dominant species of the central North American tallgrass prairie. Switchgrass is a

Electronic supplementary material The online version of this article (doi:10.1007/s12155-010-9080-8) contains supplementary material, which is available to authorized users.

A. C. Srivastava · K. Palanichelvam · J. Ma · J. Steele · E. B. Blancaflor · Y. Tang (✉)
Plant Biology Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
e-mail: ytang@noble.org

A. C. Srivastava · J. Ma · E. B. Blancaflor · Y. Tang
The BioEnergy Science Center (BESC) of U.S. Department of Energy, Ardmore, OK, USA

Published online: 14 February 2010

highly adaptable plant as evident from its broad habitat range, which spans south of latitude 55°N from Saskatchewan to Nova Scotia in Canada, and south over most of the United States east of the Rocky Mountains [36]. Because of its versatility and minimal fertilizer input needed to sustain its growth, switchgrass is poised to become a major crop plant for cellulosic ethanol production [5, 27].

Lignin, a network of phenylpropanoid subunits, is a prominent component of the plant secondary cell wall. The biosynthesis of lignin reinforces the structural integrity of the cell wall by cross-linking existing carbohydrate polymers leading to the formation of a considerably strengthened secondary cell wall in mature plant tissues [37]. Although a lignin-reinforced secondary wall is needed for maintaining normal tissue function necessary to sustain plant development, the presence of lignin compounds is a major limitation for efficient biomass conversion [37]. Therefore, lignin biosynthetic enzymes are viable targets for the genetic modification of plants for efficient cellulosic ethanol production. In fact, a recent proof-of-concept study in alfalfa indicated that down-regulation of genes that encode six different lignin biosynthetic pathway enzymes could reduce or eliminate the need for chemical pretreatment in the production of fermentable sugars. Alfalfa plants down regulated in these lignin biosynthetic genes yielded nearly twice as much sugar from cell walls compared to wild-type plants during biomass conversion [10, 37]. Such results are encouraging as they open up the possibility of reducing or eliminating acid pretreatment of plant biomass and make the fermentation step more efficient without acid residues in the reaction mixture.

Although efforts for down-regulation of lignin in switchgrass are currently underway, we still know very little about the basic biology of this emerging bioenergy crop. A key resource that could be instrumental in defining new targets for reducing recalcitrance in switchgrass would be a list of genes preferentially expressed in vascular bundles where active lignification and secondary cell wall modification occur. Laser-capture microdissection (LCM) provides a way to generate this resource because it allows RNA to be collected from specific tissue types [14, 30]. Indeed, LCM has been used to isolate *Arabidopsis* vascular tissue from flower stalks [19], rice phloem cells [1], and maize vascular bundles [29] leading to the identification of genes that function in vascular-bundle-related processes. Here, we describe the generation of ESTs from *cDNA* libraries created from LCM isolated vascular bundles of the switchgrass cultivar Alamo. The validity of our vascular tissue isolation was confirmed by the identification of switchgrass ESTs that were highly similar to genes previously identified in *Plantago* [31], maize [29], and *Cannabis* [6] using LCM or other approaches targeting vascular tissues. We verified expression of seven representative genes by in situ hybridization, five of which displayed strong expression in the vascular tissues. Furthermore, comparing the vascular bundle ESTs reported here with the ESTs derived from callus, crown and whole seedlings [38], and other sources revealed additional switchgrass genes involved in cell wall and lignin biosynthesis. This report represents a significant step toward the generation of an inventory of genes specifically expressed in vascular tissues in switchgrass that could be mined for reducing recalcitrance in this important bioenergy crop.

Methods

Plant Materials, Growth Condition, and Fixation for LCM

P. virgatum cultivar Alamo seeds were placed in 9-cm Petri dishes lined with moist Whatman filter paper. Seeds were incubated in the dark at 28°C for 4 days followed by a 4-day incubation in the light. Upon germination, seedlings were transplanted to soil (Sungrow Metromix 300 growing medium) and raised in a greenhouse at 28°C and 16/8-h day/night growth regimen, using supplemental lighting from halide lamps (250 mol photons m^{-2} s^{-1}). The second youngest internode from tillers of 5-month-old plants were fixed and subsequently used for LCM of vascular tissues. Fixation of the tissue and cryosectioning was performed essentially as described by Nakazono et al. [9] with minor modifications. Briefly, 5-mm segments of the second internode of the switchgrass plant were collected in ice cold fixative solution (75% (*v*/*v*) ethanol and 25% (*v*/*v*) acetic acid). This fixative was infiltrated into the tissues under vacuum (400 mm of Hg) on ice for 15 min. The vials containing plant samples were kept overnight at 4°C. To avoid tissue damage due to ice crystal formation during cryosectioning, fixed tissues were washed with 10% (*w*/*v*) sucrose, which was prepared with diethyl pyrocarbonate (DEPC)-treated PBS buffer (137 mM NaCl, 8.01 mM Na_2HPO_4, 2.68 mM KCl, and 1.47 mM KH_2PO_4, pH 7.3). The sucrose solution was infiltrated into the tissues under vacuum (400 mm of Hg) on ice for 15 min. The vials were swirled on a rotator at 4°C for 1 h and subsequently the 10% sucrose solution was exchanged with 15% (*w*/*v*) sucrose (in DEPC-treated PBS buffer). The vacuum infiltration/swirl step was repeated one more time and samples were incubated for an additional 24 h and stored at −20°C.

Cryosectioning and Dehydration of Sections

A Leica CM 1850 cryostat (Leica Microsystems Inc, Bannockburn, IL) was used for obtaining cross sections of switchgrass. After incubation in sucrose, switchgrass tissues were transferred to 4°C before embedding the tissues in

TissueTek OCT solution (Sakura Finetek, Torrance, CA). After securing the tissues in OCT solution, they were flash frozen in liquid nitrogen, and transferred to the cryostat for sectioning. We found that the optimum cross section thickness for our LCM tissue was 15 µm. Sections were mounted on membrane-coated slides (1 mm PEN membrane glass slides PALM Microlaser Technologies, Germany). To overcome the hydrophobic nature of the membrane, the membrane slides were irradiated with UV light at 254 nm for 30 min prior to mounting the sections. This procedure facilitated adherence of the frozen sections onto the slide. We also discovered that immediately after transferring sections onto slides, treatment with 100% (*v*/*v*) ethanol helped in flattening the sections and allowed sections to attach more firmly to the slides. This step also facilitated the removal of any remaining moisture and OCT solution from the sections. Sections were covered with 9-cm Petri dishes and kept at 4°C prior to LCM.

LCM Procedure

A Zeiss P.A.L.M. LCM (PALM Microlaser Technologies, Germany) was used for isolation of switchgrass vascular bundles into collection tubes for RNA preparation. A pulsed nitrogen laser (337 nm) was used for catapulting auto-circled vascular bundles (Fig. 1). Distance of AutoLPC (Laser Pressure Catapulting) shots was 20 µm, laser power was 92 mW, and laser pulse duration was 3 nanoseconds. Each LPC was estimated to capture ~15–20 cells from the vascular tissues. These captured cells were multiplied by the number of LPC used in the capture experiment to estimate the number of cells captured. Robot-Manipulator was used to pick the laser catapulted samples into specific adhesive opaque caps (PALM adhesive caps, PALM Microlaser Technologies, Germany). Around 100–150 vascular bundles were catapulted into each tube for RNA isolation.

RNA Extraction and Analysis

RNA was extracted from laser catapulted samples using RNeasy Micro Kit (Qiagen, Valencia, CA) according to the manufacturer's protocol with slight modification. RNA quality and quantity was estimated using RNA 6000 Nanochip on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). For analyzing low RNA quantity, RNA 6000 Pico chip was used (Fig. 1).

Construction of *cDNA* Libraries

The vascular bundle specific *cDNA* libraries were constructed using Creator™ Smart™ *cDNA* Library Construction Kit (Clontech Laboratories, Inc, Mountain View, CA) as per manufacturer's protocol. The *cDNA* library was prepared from 500 ng total RNA, which was pooled from total RNA isolated from vascular tissues of the leaf sheath and second internode of switchgrass. The library was constructed in the pDNR-LIB vector (Clontech). First-strand *cDNA* was constructed using MMLV Reverse Transcriptase and amplified by PCR using 5′ PCR primers and CDS III/3′ PCR primer (supplied by Clontech) on a PTC-200 Peltier Thermal Cycler (MJ Research, Watertown, MA). *cDNA* amplification was done using first-strand *cDNA* as a template and the same set of primers in the above-mentioned thermal cycler with the following programs: 95°C for 1 min followed by 26 cycles (95°C for 15 s, 68°C 6 min). The resulting *cDNA* was digested with

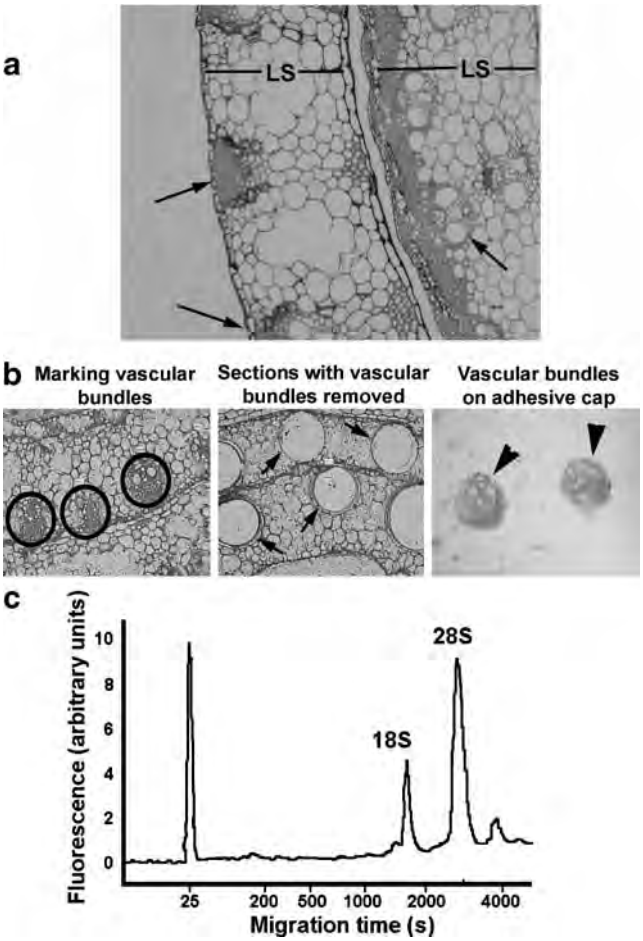

Fig. 1 LCM mediated isolation of vascular tissues from *P. virgatum* Alamo. **a** A 15-µm section of the second internode from the top of the tiller of a 5-month-old switchgrass plant. Note that a cross section from the second internode typically consists of two layers of leaf sheaths (*LS*) with prominent vascular bundles on the abaxial side (*arrows*). **b** Marking of vascular tissues (*black circles*, *left panel*) using the Robot-Manipulator feature of the LCM. After catapulting, sections without the vascular tissues (*arrows*, *center panel*) remain on the glass slides. A pair of isolated vascular bundles (*arrows*) on the cover of a collection tube (*right panel*). **c** Qualitative and quantitative analysis of total RNA isolated from the vascular tissues using a Bioanalyzer

Sfil enzyme and fractionated with CHROMA SPIN-400 column. The top three fractionations were combined and ligated into pDNR-LIB and transformed into Eletromax DH10B (Invitrogen, Carlsbad, CA) *E. coli* cells.

Automated Sequencing Procedure

The libraries were plated and individual colonies were randomly picked and arrayed into 384-well plates and grown in 150 µl liquid TB media for 20 h at 470 rpm in the HiGro microwell plate growth system (Genomic Solutions, Ann Arbor, MI). Plasmid DNA isolation was performed by standard procedures using Solutions I, II, and III using the Biomek FX^P (Beckman Coulter, Fullerton, CA). Sequencing was conducted using BigDye Terminator v3.1 (Applied Biosystem, Foster City, CA) standard procedures and with the following primer sequence: 5′-TTATCAGTCGACGG TACCGGACAT-3′. Sequencing reactions were purified using Agencourt® CleanSEQ® Dye-Terminator Removal System (Beckman Coulter, Beverly, MA) following the manufacturer's protocol using the Biomek FX^P (Beckman Coulter). Purified sequencing samples were analyzed using a 3730 DNA Analyzer (Applied Biosystems).

Sequencing Assembly and Data Annotation

Raw EST sequences were trimmed for vector and adapter/linker sequences and sequence reads were assembled into contigs using an in-house tool "EST Pipeline". For gene annotation, both contigs and singlets were searched against various sequence databases, including NCBI NR (http://blast.ncbi.nlm.nih.gov/Blast.cgi), TAIR (http://www.arabidopsis.org/) and PlantDB (http://www.plantgdb.org/cgi-bin/blast/PlantGDBblast) using BlastX or tBLASTX. Various metabolic pathways were identified using the web site of KEGG Metabolic Pathways (http://fire2.scl.genome.ad.jp/kegg/metabolism.html).

To determine the percentage of novel switchgrass transcripts present in our EST data, the unigene set was used to search against existing switchgrass ESTs downloaded from GenBank (Oct. 15, 2009) as well as DFCI Switchgrass Gene Index version 1.0-PaviGI (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=switchgrass) using Blastn. An *E* value cutoff of 1.0E−5 was used to determine homology because an *E* score larger than this value gives very short (<30 bp) or low similarity alignment.

Tissue Preparation and In Situ Hybridization

For in situ hybridization, the second internode of switchgrass was used. Tissue harvesting, fixation, sectioning and mounting on gelatin coated slides (Lab Scientific Inc., Livingston, NJ) were done as described above for LCM procedure. Slides were baked on a slide warmer at 42°C overnight before proceeding with the prehybridization steps. Tissue sections were next prehybridized in a prehybridization solution (47% formamide (*v*/*v*), 2× SSC, 50 mM phosphate buffer (pH 7.0), 1× Denhardt's solution, 10% dextran sulfate, 50 mM DTT, 250 µg ml^{-1} yeast tRNA (Sigma), 5 µg/ml polydeoxyadenylic acid, and 500 µg/ml denatured and sheared salmon sperm DNA for 2 h at 37°C. Tissue sections were hybridized overnight at 38°C in a humid chamber using a solution with the same composition as prehybridization solution but it also contained digoxigenin-labeled probe (50 pM). Oligonucleotide (45–50 mer) probes were synthesized by tailing with the DIG Oligonucleotide Tailing Kit according to the manufacturer's protocol (Roche Applied Science, Indianapolis, IN). For probe design, we usually choose an antisense fragment of 50 bp corresponding to the gene of interest and sense fragments were used as negative controls (Supplementary Table 2). Subsequently, tissue sections were washed in 2× SSC, 1× SSC, and 0.25× SSC at room temperature for 2×15 min each. For staining, sections were washed with buffer 1 [100 mM Tris–HCl (pH 7.5), 150 mM NaCl] for 2×10 min and incubated in blocking solution [buffer 1 containing 0.1% Triton X-100 and 2% normal sheep serum (Sigma)] for 30 min. The blocking solution was then decanted and sections were incubated with buffer 1 containing 0.1% Triton X-100, 1% normal sheep serum, and a various dilutions of sheep anti-DIG-alkaline phosphatase (Fab fragments) for 2 h in a humid chamber. For optimal detection, we used several sections (from the same sample) with different dilutions of the antibody (1:100; 1:200, and 1:500). Finally, sections were washed with buffer 1 for 2×10 min, then incubated for 10 min with buffer 2 [100 mM Tris–HCl (pH 9.5), 100 mM NaCl, 50 mM $MgCl_2$] and for immunological detection, each section was covered and incubated with approximately 100 µl color solution [(*NBT*/*BCIP* in 100 mM Tris–HCl (pH 9.5), 100 mM NaCl, 50 mM $MgCl_2$, 1 mM (2.4 mg/10 ml) levamisole (Sigma)] in a humid chamber for 2–24 h in the dark. Color reaction was stopped by incubating the slides in buffer 3 [10 mM Tris–HCl (pH 8.1), 1 mM EDTA]. Progression of the color reaction was terminated by washing the slides for 5 min in H_2O, followed by air drying in the dark at room temperature. Slides were next covered with cover-slips for microscopic analyses using Paramount (Fisher Scientific) as mounting media.

Photographs of sections were captured using an inverted Nikon TE300 compound microscope equipped with differential interference contrast optics. Still images of the sections were acquired using a Nikon DXM 1200 camera running on ACT-1 software (Nikon Instruments, Melville, NY).

Results

Optimization of LCM for Switchgrass Internodes

Although there are several examples of LCM applications in plants [22, 29], the methods for tissue structural preservation and extraction of quality RNA have to be optimized for each plant species. For example, it has been shown that optimal structural preservation and extraction of RNA from root cells of *Medicago truncatula* colonized with mycorrhizal fungi involved embedding in low-melting-point Steedman's wax [18]. Here, we tried three different post-fixation and embedding techniques for preparing switchgrass internode sections for LCM. We first tried isolating RNA from vibratome-sectioned internodes. Although this approach yielded the highest quality RNA, the thickness of the sections (ca. 50–70 μm) proved problematic because of poor section adherence to membrane-coated glass slides, and incomplete cutting of vascular bundles using the LCM microscope (data not shown). We also tried embedding switchgrass internodes in paraffin [22] and low-melting-point Steedman's wax [18]. Although wax embedding resulted in good structural preservation, we were unable to consistently obtain good quality RNA for *cDNA* library construction (data not shown). The method we decided to follow involved cryosectioning [29] as it provided a good compromise between internode structure that allowed us to efficiently enrich for vascular tissue with adequate RNA quality for *cDNA* library construction (Fig. 1).

Our sectioning and RNA quality evaluations also prompted us to select the second internode from the top of 5-month-old switchgrass plants for vascular tissue isolation. Cross sections from this part of the plant typically consisted of two layers of leaf sheaths and a hollow inner core. Vascular bundles were mostly found close to the abaxial side of the leaf sheaths (Fig. 1a). The vascular bundles were marked using the Robot-Manipulator feature of the LCM microscope and the cut vascular bundles were catapulted into the caps of collection tubes for subsequent RNA extraction. For extracting vascular bundles, we drew circles with the same diameter to mark vascular bundles in the section that had roughly the same size as this facilitated rapid material collection (Fig. 1b). Smaller diameter circles were drawn for smaller vascular bundles to minimize contaminating parenchyma cells. We found that RNA isolated from vascular bundles from the leaf sheaths of the second internode yielded the best quality RNA for *cDNA* library construction (Fig. 1c). The highly lignified vascular tissues from the more mature internodes were more difficult to cut and catapult using LCM. Furthermore, the abundance of dead tissues in the more mature internodes yielded less RNA.

Identification of Genes Expressed in Switchgrass Vascular Tissues

A total of 5,734 partial DNA sequences (ESTs), were obtained from switchgrass vascular bundles and submitted to Genbank (Accessions GR875377 to GR881110). The consensus sequences derived from the EST sequences are attached to Supplemental File 1. The 5,734 ESTs that were expressed in vascular tissues represent 2,766 unigenes, including 151 ESTs that did not have any annotation information. The genes that were expressed from our switchgrass vascular tissues and for which functions could be predicted are listed in Table 1 and Supplemental Table 1. They were grouped into the following categories according to their functions and each of these categories are described below: (1) cell wall biosynthesis; (2) lignin biosynthesis; (3) transcription control and regulatory genes implicated in lignin synthesis; (4) C1 metabolism; (5) transcription control and other regulatory genes; (6) transport facilitators, channels, and pumps; (7) stress and defense; (8) signal transduction; (9) cellular response to hormones, growth and development; (10) cytoskeleton; (11) energy metabolism; (12) protein synthesis and (13) uncharacterized genes. The genes for category 1 through 4 are listed in Table 1 and categories 5 through 12 are shown in the accompanying Supplemental Table 1.

Distributions of these genes based on their functions indicate that cell wall and lignin biosynthesis genes together represented 11% of the vascular-bundle-derived unigenes. Genes encoding regulatory proteins comprised 12% of the ESTs. Other important categories of genes also made up a significant portion of the data set: transporters (7%), stress and defense (6%), energy and carbon metabolism (7%), signal transduction (7%), cytoskeleton (4%), hormones, growth and development (3%), and protein synthesis (9%; Supplemental Table 1; Fig. 2). In addition, ESTs classified as "other metabolism" and "uncharacterized" comprised 29% and 5% of the switchgrass library, respectively (Fig. 2).

On the basis of the classical lignin biosynthetic pathway that begins with phenylalanine ammonia lyase (PAL), transcripts of eight genes involved in lignin synthesis were found in the switchgrass vascular bundle library including PAL, 4-coumarate:CoA ligase (4-CL), cinnamoyl-CoA reductase (CCR), *O*-methyltransferase (OMT), cytochrome P450 monooxygenase, cinnamyl-alcohol dehydrogenase (CAD), peroxidase and laccase (Lac) transcripts [10, 11, 13, 34]. Overall, seven contigs (seven ESTs) were annotated as PAL representing 3 different genes, seven contigs (nine ESTs) as 4-CL, three contigs (three ESTs) as CAD, three contigs (three ESTs) as CCR, ten contigs (15 ESTs) as cytochrome P450 family protein, four contigs (four ESTs) as Lac, 15 contigs (27 ESTs) as peroxidase and one contig

Table 1 List of transcripts for cell wall and lignin-related genes derived from switchgrass vascular tissues

Cell wall biosynthesis

Contig ID	Contig length	EST No.	Target ID	Target description	*E* value
B01028, B00016	643	2	BAB84371	1,3-beta-glucan synthase component-like	2.0E−15
B01191	610	1	AT1G71170.1	6-phosphogluconate dehydrogenase NAD-binding domain	7.0E−07
A01546	704	1	ABG73467	6-phosphogluconolactonase	9.0E−75
A02059	416	1	NP_174282.2	Acetyltransferase-related	1.0E−70
A00036	563	18	NP_001050536	Acyl-CoA binding (ACBP)	3.0E−111
A00780, B01647	437	2	AT1G61050.1	Alpha 1,4-glycosyltransferase family protein	1.0E−01
A00374	459	2	Q9FXT4	Alpha-D-galactoside galactohydrolase	2.0E−20
B01538	380	1	AT3G46970.1	Alpha-glucan phosphorylase 2 (PHS2)	4.0E−01
B01091	722	1	AT1G24320.1	Alpha-glucosidase, putative	6.0E−75
B00969	513	1	BAC84349	Alpha-L-fucosidase 2	7.0E−10
B00984	233	1	AT3G62220.1	Arabinogalactan (AGP16)	9.6E−01
B00916	272	1	AT2G14890	Arabinogalactan-protein 9	6.0E−01
B02472A	423	1	AT4G40090	Arabinogalactan-protein 3	1.0E−02
B01306	493	1	NP_568718.1	Auxin F-Box protein 5	2.0E−73
A00898, A01608	395	2	BAD67750	Beta 1,3 glucan synthase	1.0E−16
A01615	877	1	CAJ47422	Beta-1,2-xylosyltransferase	5.0E−73
B02142	446	1	NP_187498.1	Beta-alanine-pyruvate aminotransferase	3.0E−73
A01960	395	1	AAK56125	Beta-expansin 2	2.0E−09
A01267	866	1	NP_001105643	Beta-expansin 7	2.0E−79
B01202	909	1	AT1G22650.1	Beta-fructofuranosidase, putative/invertase	0.0E+00
A00476	985	2	P49235	Beta-glucosidase (Beta-D-glucoside glucohydrolase)	7.0E−144
A01749	935	1	NP_001058010	Branch; Core-2/glycosyltransferase family 14 protein	1.0E−28
A00883	172	1	NP_001104946	Brittle stalk-2-like protein 3	4.0E−23
A00172	1985	4	NP_001052692	Calnexin 1 (CNX1)	2.0E−163
A01320, B01626	344	1	AY104712.1	Cell wall structural protein precursor	7.0E−35
A00243	673	3	NP_001104954	Cellulose synthase-1	8.0E−36
A01925	796	1	NP_001104954	Cellulose synthase-2	1.0E−37
A00723	960	1	NP_001059303	Cellulose synthase 2	8.0E−179
B00881	131	1	AT5G05170.1	Cellulose synthase 3	2.0E−01
A01842	489	1	AT5G44030.1	Cellulose synthase 4 (CESA4); IRX5	3.0E−17
A00804	959	1	NP_001104955	Cellulose synthase 5	1.0E−135
A01797	882	1	NP_001104955	Cellulose synthase 5	2.0E−162
A00197	701	4	NP_001104958	Cellulose synthase 8 (CESA8); IRX1	7.0E−72
A02064	213	1	AT4G18780	Cellulose synthase 8 (CESA8); IRX1	1.0E−31
A00142	1001	4	AT5G17420.1	Cellulose synthase catalytic subunit 7; IRX3	3.0E−71
A01984	461	1	AT1G23480.3	Cellulose synthase-like A3	1.0E−01
B00296	136	4	AT4G31590.1	Cellulose synthase-like C5	2.0E−01
A01365	472	1	NP_180869.1	Cellulose synthase-like D1	3.0E−52
A00659	421	2	NP_174497.1	Cellulose synthase-like D6	3.0E−50
B02289	386	1	AT5G60950.1	COBRA-like protein precursor 5 (COBL5)	2.0E−01
B00848	189	1	AT1G09790.1	COBRA-like protein precursor 6 (COBL6)	3.8E−01
A00118, B02131, B02128	662	3	AT3G22120.1	Cell wall–plasma membrane linker protein (CWLP)	2.0E−17
A02046	887	1	NP_001059165	Defective glycosylation 1 (DGL1)	2.0E−97
A01387	707	1	NP_001061329	Endo-1,3;1,4-beta-D-glucanase precursor	3.0E−55
A00400	631	2	BAA94257	Endo-1,4-beta-glucanase Cel1	2.0E−14
A00785, A00959	585	1	EAY98250	Exoglucanase precursor	3.0E−62

Table 1 (continued)

Cell wall biosynthesis

Contig ID	Contig length	EST No.	Target ID	Target description	*E* value
A01365A	301	1	AT1G21310	Extensin 3	5.9E−01
A00601, B01807, B01916	326	3	NP_172673.1	Extensin-like protein (ELP)	1.0E−63
A00234	590	1	NP_196257.2	Fasciclin-like protein FLA17	2.0E−108
A01231	143	1	AT4G01970.1	Galactinol-raffinose galactosyltransferase	2.0E−01
B02833	215	1	AT5G55120.1	Galactose-1-phosphate guanylyltransferase (VTC5)	2.0E−01
A01402A	120	1	AT3G14960.1	Galactosyltransferase family protein	4.0E−02
B01025	548	1	AT2G26100.1	Galactosyltransferase family protein	1.0E−30
B01434	529	1	AT1G27120.1	Galactosyltransferase family protein	1.0E−113
A01977, A01949	638	2	ABA99401	Galactosyltransferase family, putative	1.0E−90
A00808	1005	1	NP_001050360	Galacturonosyltransferase 10 (GAUT10/LGT4)	2.0E−117
B01204, B02078	308	3	AT4G38270.1	Galacturonosyltransferase 3 (GAUT3)	3.0E−01
B01870	154	1	AT4G02130.3	Galacturonosyltransferase-like 6 (GATL6)	7.4E−01
B00305	380	4	EU964859	Glucan endo-1,3-beta-glucosidase 4 precursor	2.0E−65
A01585	499	1	AAM19120	Glucan synthase	2.0E−65
B00953A	127	1	AT1G24100.1	Glucan synthase-like 3 (GSL03)	2.0E−01
B00489A	100	2	AT4G03550.1	Glucan synthase-like 5 (GSL05)	2.7E−01
B02051	245		AT4G03550.1	Glucan synthase-like 5 (GSL05)	2.0E−01
B00891A, B00941A	85	2	AT3G07160.1	Glucan synthase-like10 (GSL10)	8.0E−01
B01374A, B01836A, A01866	120	3	AT3G59100.1	Glucan synthase-like11 (GSL11)	3.0E−01
A00733	841	1	NP_001106058	Glucose-1-phosphate adenylyltransferase large subunit 3	2.0E−67
B02734	272	1	AT5G58210.4	Glycoprotein	2.4E−01
B02445	124	1	NP_199783.1	Glycoside hydrolase 9A1 (GH9A1)	4.0E−13
B02649	252	1	NP_177697.1	Glycoside hydrolase 9B7	8.0E−44
A01853	251	1	NP_177697.1	Glycoside hydrolase 9B7 (GH9B7)	8.0E−44
B00527	177	2	AT4G11050.1	Glycoside hydrolase 9C3 (GH9C3)	8.0E−01
B02735	378	1	AT3G54440.1	Glycoside hydrolase family 2	5.0E−71
A00354, B00474A, B00841	532	3	AT3G62110.1	Glycoside hydrolase family 28	6.4E−01
B00915, B01509	826	2	AT3G57790.1	Glycoside hydrolase family 28 /polygalacturonase family	2.0E−87
B01262, A01507	945	2	AT3G62110.1	Glycoside hydrolase family 28 /polygalacturonase family	e−103
A01987	422	1	AT2G04270.4	Glycoside hydrolase starch-binding domain	9.0E−06
A01775	984	1	EAZ08383	Glycoside hydrolase, family 17 protein	3.0E−157
A00112, A00592, A00763, A01962, B02726	1252	12	NP_001051013	Glycosyl hydrolase family 1 protein	2.0E−153
B01469	161	1	NP_195577.2	Glycosyl hydrolase family 10	1.0E−19
B00327A, B01407A	115	4	AT5G55180.1	Glycosyl hydrolase family 17 protein	1.0E−01
A02012, B00681A, B1042	825	4	EAY74993	Glycosyl hydrolase family 18 protein	5.0E−77
B01321A	170	1	AT5G04885.1	Glycosyl hydrolase family 3	2.0E−02
B01728, B01063	917	2	AT3G26720.1	Glycosyl hydrolase family 38 protein	2.0E−90
B01506	149	1	AT4G28320.1	Glycosyl hydrolase family 5	1.0E−01
B01431	518	1	AT5G15870.1	Glycosyl hydrolase family 81 protein	1.6E−01
B00791	283	2	AT5G14480.1	Glycosyl transferase family 17 protein	1.0E−01
A02048	120	1	AT3G48820.1	Glycosyl transferase family 29 protein	9.0E−01
A01502	592	1	NP_001056954	Glycosyl transferase, family 31 protein	4.0E−73
A01685	570	1	NP_001060535	Glycosyl transferase, family 4 protein	2.0E−34

Table 1 (continued)

Cell wall biosynthesis

Contig ID	Contig length	EST No.	Target ID	Target description	*E* value
A01805	980	1	EAZ02335	Glycosyl transferase, family 8 protein	1.0E−150
A01706	838	1	CAJ19325	Glycosyl transferase-like protein	8.0E−87
A01483	521	1	EU969907.1	Glycosyltransferase	5.0E−71
A01484	528	1	EU969907.1	Glycosyltransferase	5.0E−71
A01856	355	1	NP_001105849	Glycosyltransferase	2.0E−10
B00929	354	1	NP_001105751	Glycosyltransferase	9.0E−14
B01119	353	1	EU956619.1	Glycosyltransferase	2.0E−18
A00669	684	1	NP_001060977	Glycosyltransferase family 14 protein	3.0E−113
B01436A	327	1	AT2G28080.1	Glycosyltransferase family protein	1.9E−01
B01164	262	1	AT1G30620.2	HSR8/MUR4/UXE1 (MURUS 4)	3.0E−16
B00853	114	1	AT1G73480.1	Hydrolase, alpha/beta fold family protein	4.0E−01
A00305	596	1	AT5G58210.4	Hydroxyproline-rich glycoprotein	2.0E−01
A00375	526	2	EAZ08078	Impaired sucrose induction 1 (ISI1)	6.0E−03
A00324	461	2	AF272760	Kinesin-like protein16	5.0E−14
A00899	962	1	Q6YWK8	Mannan synthase 11 (Cellulose synthase-like protein A11)	1.0E−69
A02017	810	1	NP_001053800	Mannosyl-oligosaccharide 1,2-alpha-mannosidase	2.0E−66
B01339A	123	1	AT1G02145	Mannosyltransferase, putative	2.0E−01
A00465	764	2	CAO63647	Nucleotide-rhamnose synthase/epimerase-reductase	4.0E−54
B01254	938	1	AT1G62760.1	Pectin methylesterase inhibitor family	2.0E−56
B00101, B00102A	333	12	AT4G19420.2	Pectinacetylesterase family protein	8.0E−02
B00780	119	1	AT3G06830.1	Pectinacetylesterase family protein	6.1E+02
B01231A	167	1	AT3G62060.1	Pectinacetylesterase family protein	1.0E−01
B01744	166	1	AT3G05620.1	Pectinacetylesterase family protein	1.0E−01
A01238, A00553A	461	3	NP_001060382	Pectinacetylesterase, putative	4.0E−10
A01554	400	1	AT5G09760.1	Pectinacetylesterase, putative	1.0E−02
B00555A	146	2	AT2G43050.1	Pectinesterase	2.0E−01
B01827A	186	1	AT3G49220	Pectinesterase family protein	3.0E−01
A01311	330	1	NM_001154483	Peptide transporter (PTR2)	1.0E−23
B01376	768	1	AT5G08200.1	Peptidoglycan-binding LysM domain-containing protein	1.0E−04
A01275	499	1	AT2G47320.1	Peptidyl-prolyl *cis–trans* isomerase cyclophilin-type	3.0E−08
B01460	671	1	AT3G66654.3	Peptidyl-prolyl *cis–trans* isomerase cyclophilin-type	8.0E−33
B01135	900	1	AT1G26940.1	Peptidyl-prolyl *cis–trans* isomerase cyclophilin-type	E−105
A02035	218	1	EU957959	Permease	8.0E−173
B01576	557	1	X66422	PG gene for polygalacturonase	2.0E−19
B01246	644	1	AT3G28360.1	PGP16 (P-GLYCOPROTEIN 16)	1.5E−01
B01171	428	1	AT5G26570.2	Phospholucan wer dikinase (GWD3, OK1, PWD)	4.0E−13
B01723	175	1	AT4G33330.1	Plant glycogenin-like starch initiation protein 3 (PGSIP3)	5.0E−01
B00966	339	1	AT5G64370.1	Polygalacturonase inhibitor 2	4.0E−13
A01835, B01899	523	1	AT2G43860.1	Polygalacturonase, putative / pectinase	4.8E−02
B00597	300	1	BAD61522	Polygalacturonase-like	6.0E−16
B01207A	232	1	AT1G24170.1	Polygalacturonate 4-alpha-galacturonosyltransferase	1.0E−01
A01643	663	1	A2YNH4	Probable 6-phosphogluconolactonase 2 (6PGL 2)	2.0E−122
B01573, B00783	271	2	NP_200894.1	Rab GTPase homolog A1f	2.0E−39
A00729	399	1	AT5G15650.1	Reversibly glycosylated polypeptide 2 (RGP2)	3.0E−77
A01766	114	1	AT5G03650.1	Starch branching enzyme 2.2 (SBE2.2)	7.0E−01
B01636	451	1	AT3G01180.1	Starch synthase 2 (SS2)	7.2E−02

Table 1 (continued)

Cell wall biosynthesis					
Contig ID	Contig length	EST No.	Target ID	Target description	*E* value
B02259A, A01864, A01912A, B01065, A00805	135	5	AT3G22142.1	Structural constituent of cell wall	6.0E−02
B02601	160	1	AT5G20280.1	Sucrose phosphate synthase 1F (SPS1F)	2.0E−01
B02819	253	1	NP_197528.1	Sucrose phosphate synthase 1F (SPS1F)	3.0E−39
A01134	829	1	NP_001105411	Sucrose synthase	5.0E−73
A01317	807	1	NP_001105323	Sucrose synthase 1	8.0E−123
B01950	167	1	AT4G02280.1	Sucrose synthase 3 (SUS3); UDP-glycosyltransferase	2.0E−24
B02739	446	1	NP_566865.2	Sucrose synthase 4 (SUS4); UDP-glycosyltransferase	3.0E−85
B02618	323	1	NP_173134.2	Sugar binding transferase	5.0E−24
A01532	246	1	AT3G10370.1	Sugar dependent 6 (SDP6)	3.0E−01
A01344	864	1	ABG25862	Trehalase	1.0E−100
B02511	577	1	NP_172129.1	Trehalose-phosphatase	1.0E−25
B01566A	359	1	AT5G52560.1	UDP- sugar pyrophosphorylase (USP)	1.0E−02
B01639	457	1	NP_192834.1	UDP-D-galactose 4-epimerase 5 (UGE5)	1.0E−78
A00652	834	1	NP_001049561	UDP-D-glucuronate 4- epimerase	3.0E−37
A00227, A00392, A01291	1264	6	AT3G46440	UDP-glucuronate decarboxylase	0.0E+00
A00768	581	1	ABC67799	UDP-D-xylose epimerase 3	2.0E−18
B01108	1033	1	AT3G59360.2	UDP-galactose transporter 6 (UTR6)	E−168
A01127	395	1	NP_001063685	UDP-glucoronosyl/UDP-glucosyl transferase	1.0E−09
A01464	388	1	AT2G29750.1	UDP-glucoronosyl/UDP-glucosyl transferase family	6.0E−01
A01834	212	1	AT3G21800.1	UDP-glucoronosyl/UDP-glucosyl transferase family	4.0E−01
B00874, B02417, B00952	114	3	AT3G46670.1	UDP-glucoronosyl/UDP-glucosyl transferase family	4.0E−01
B00222A	167	5	AT1G78570.1	UDP-glucose 4,6-dehydratase	1.0E−01
A00788	967	1	BAC24803	UDP-glucose 4-epimerase	3.0E−81
A01869	614	1	NP_001066706	UDP-glucose 6-dehydrogenase	4.0E−39
B00185,A00196	2059	3	NP_001066705	UDP-glucose 6-dehydrogenase, putative	0.0E+00
A00103	857	2	AAO48422	UDP-glucose pyrophosphorylase	3.0E−97
A00483	964	2	ABW78938	UDP-glucose pyrophosphorylase	6.0E−127
A00104	1802	5	Q43772	UDP-glucose pyrophosphorylase (UGPase) (UDPGP)	8.0E−117
B01995, A00868A	375	2	AT2G15480.1	UDP-Glucosyl transferase 73B5 (UGT73B5)	1.4E+00
B00954A, B01100,	247	3	AT5G53460.1	UDP-Glucosyl transferase 74B1 (UGT74B1)	8.0E−02
B01117A	132	1	AT2G31750.1	UDP-Glucosyl transferase 74D1 (UGT74D1)	7.0E−01
A01015	575	1	AT1G05530.1	UDP-Glucosyl transferase 75B2 (UGT75B2, UGT2)	3.0E−01
B00743	487	2	AT3G62170.1	Vanguard 1 homolog 2 (VGDH2); pectinesterase	1.0E−01
B02061A	167	1	AT1G21210.1	Wall associated kinase (WAK4)	7.0E−01
B01856	553	1	AT4G25810.1	Xyloglucan endotransglycosylase 6 (XTR6)	1.0E−78
B01124A	116	1	AT1G74380.1	Xyloglucan xylosyl transferase 5 (XXT5)	3.0E−01
B00796	274	1	AT3G44990.1	Xyloglucan:xyloglucosyl transferase (XTR8)	2.0E−01
Lignin biosynthesis					
A01868	919	1	AAW65140	3-dehydroquinate dehydratase/shikimate 5-dehydrogenase	1.0E−111
A00268	644	3	NP_001105258	4-coumarate coenzyme A ligase	2.0E−36
A01170	505	1	ACA09448	4-coumarate:CoA ligase	2.0E−31
A01458	372	1	NP_175579.1	4-coumarate:CoA ligase 1 (4CL1)	5.0E−66
A00367, A00719	672	2	EU966506.1	4-coumarate:CoA ligase 2 (4CL2)	5.0E−123
B02410	365	1	AT1G65060.2	4-coumarate:CoA ligase 3 (4CL3)	1.0E−18

Table 1 (continued)

Cell wall biosynthesis

Contig ID	Contig length	EST No.	Target ID	Target description	*E* value
B01348	362	1	AT3G21230.1	4-coumarate:CoA ligase 5 (4CL5)	8.0E−01
B01395, A01443	605	2	EU968435.1	Acetyltransferase 1-like	3.0E−46
A01961	307	1	NP_201454.1	Aldose 1-epimerase family protein	9.0E−52
A00674	701	1	AT3G25585	Aminoalcoholphosphotransferase 1 (AAPT1)	5.0E−130
A01394, B00903	899	2	NP_001046388	Aspartate aminotransferase 1 (ASP1)	7.0E−118
B01526	376	2	NP_187498.1	Beta-alanine-pyruvate aminotransferase	7.0E−67
A00623, B01892	494	2	AT2G33590	Cinnamoyl-CoA reductase	3.0E−69
B01641	552	1	NP_565557.1	Cinnamoyl-CoA reductase	2.0E−50
B00973, B01345	516	2	FJ554574.1	Cinnamyl-alcohol dehydrogenase 4 (CAD4)	1.0E−81
B01692	320	1	AT1G66800.1	Cinnamyl-alcohol dehydrogenase family	1.0E−47
B01740A	51	1	AT3G20080.1	CYP705A15 Cytochrome P450, family 705	2.0E−01
B02752	537	1	AY072299.1	Cytochrome P450 monooxygenase	8.0E−32
A00074	680	6	T02955	Cytochrome P450 monooxygenase	8.0E−51
A01210	732	1	AAL66770	Cytochrome P450 monooxygenase (CYP72A5)	3.0E−74
B01609	221	1	AT1G55940.1	Cytochrome P450, family 708 (CYP708A1)	1.0E−01
B00831	166	1	AT1G13110.1	Cytochrome P450, family 71 (CYP71B7)	1.7E−01
B01426	459	1	AT2G28860.1	Cytochrome P450, family 710 (CYP710A4)	1.5E−02
B00976, B01166A	537	2	AT2G34690.1	Cytochrome P450, family 724 (CYP724A1)	4.0E−19
B01158A	306	1	AT1G13710.1	Cytochrome P450, family 78 (CYP78A5)	6.0E−01
A00862	816	1	EU964079	Dirigent	6.0E−10
A01193	805	1	EU960469.1	Dirigent	2E−168
B01462	291	1	AT5G04330.1	Ferulate-5-hydroxylase	8.9E−02
A01346	824	1	EAY96614	Isoflavone reductase-like1	2.0E−51
A01973	112	1	AT3G09220.1	Laccase 7 (LAC7)	1.1E+01
B01922	201	1	AT5G01190.1	Laccase 10 (LAC10); copper ion binding	9.0E−18
B01520A, B00794	181	2	AT5G07130.1	Laccase 13 (LAC13); copper ion binding	1.0E−01
B00997	461	1	AT1G77530.1	O-methyltransferase family 2 protein	4.0E−14
A01201	698	1	NP_001065568	Peroxidase	2.0E−71
B02413	436	1	EU966352.1	Peroxidase 1 (PER1) (PRXR1)	6.0E−74
B00400	834	3	AT3G21770.1	Peroxidase 30 (PER30)	5.0E−97
A00383	576	2	AT5G66390.1	Peroxidase 72 (PER72) (PRXR8)	3.0E−38
A00203	1004	4	EAY73397	Peroxidase 72 precursor	2.0E−145
A01722, A01577	516	2	NP_001058379	Peroxidase family protein	4.0E−90
B00792A, B01158, B01387, B01762A, B01865, B00137, B00553A	160	13	AT4G37530.2	Peroxidase, putative	2.0E−01
A01065	437	1	NP_001151940.1	Peroxidase 52	1.0E−15
B02470	276	1	NP_001105334	Phenylalanine ammonia-lyase	1.0E−12
A01668	661	1	NP_001047482	Phenylalanine ammonia-lyase	9.0E−128
A01916	485	1	NP_001105334	Phenylalanine ammonia-lyase	4.0E−78
A01647	262	1	NP_190894.1	Phenylalanine ammonia-lyase 2 (PAL2)	7.0E−13
A01053	456	1	NP_001053324	Phenylalanine ammonia-lyase, putative	8.0E−19
A00431, B00851	677	2	NP_187645.1	Phenylalanine ammonia-lyase-4 (PAL4)	5.0E−126
B02005	223	1	AF239818.1	Protein kinase CK2 regulatory subunit CK2B3	7.0E−27
B00452, A01131	788	1	AT2G21940	Shikimate kinase domain-containing protein	1.0E−91
B01112	332	1	Q2TJK6	Sesquiterpene synthase	7.2E−01

Table 1 (continued)

Cell wall biosynthesis					
Contig ID	Contig length	EST No.	Target ID	Target description	*E* value
Transcription control and regulatory genes in lignin synthesis					
A00776	290	1	NP_850273.1	DNA-binding domain-containing protein	7.0E−37
A01553	278	1	NP_198106.1	Embryo defective 2473 (EMB2473)	1.0E−40
A01728	612	1	EAZ23030	Fragile fiber 3 (FRA3)	1.0E−07
A00347	488	2	NP_001049655	LIM domain-containing protein	3.0E−20
B02263	635	1	NP_172491.1	LIM1; transcription factor	5.0E−05
A01991	703	1	NP_001067462	LIM domain-binding protein.	2.0E−24
B01076	450	1	AT3G09230.1	MYB domain protein 1 (MYB1); DNA-binding	6.0E−01
B01433	287	1	AT3G01140.1	Myb domain protein 106 (MYB106)	2.3E−01
B02164A, B00899	253	2	AT3G62610.1	Myb domain protein 11 (MYB11)	5.0E−01
B00824	158	1	AT3G29020.2	Myb domain protein 110 (MYB110)	1.1E−01
A00330, B00707A	170	4	AT5G58850.1	Myb domain protein 119	1.1E−01
A00854A	217	1	AT2G36890.1	Myb domain protein 38 (MYB38/RAX2)	8.0E−01
B01568A	68	1	AT4G28110.1	Myb domain protein 41 (MYB41)	3.6E−01
B01332	140	1	AT1G16490.1	Myb domain protein 58	1.6E−01
B01151	1028	1	AT5G59780.3	Myb domain protein 59 (MYB59)	6.0E−66
B01937	327	1	AT2G17770.1	Myb family transcription factor	1.0E−03
B01416A	90	1	AT3G24120	Myb family transcription	1.2E+00
A00750	483	1	AT3G25790	Myb family transcription factor	1.5E−01
B02084	257	1	AT2G40970.1	Myb family transcription factor	6.0E−02
A01248	261	1	AT2G40970.1	Myb family transcription factor	9.7E−02
A01610	520	1	NP_001045011	Myb- MADS16/mpik39b	5.0E−37
A00861	138	1	AT4G34430.2	Myb- SWI3D, CHB3; transcription factor	3.0E−01
B01238A	119	1	AT1G32870.1	NAC domain-containing protein 13 (NAC013)	2.9E−01
B01421	354	1	AT3G04430.1	NAC domain-containing protein 49 (NAC049)	1.0E−01
B02654	233	1	NP_566375.1	NAC domain-containing protein 51	3.0E−16
A01858	892	1	NP_566375.1	NAC domain-containing protein 51/52 (NAC051/NAC052)	1.0E−172
A01488A	126	1	AT5G41090.1	NAC domain-containing protein 95 (NAC095)	6.0E−01
A00658	188	1	AT5G50820.1	NAC domain-containing protein 97 (NAC097)	5.5E−01
A00938	942	1	AT3G12390.1	Nascent polypeptide-associated complex alpha chain	8.0E−67
A00271	898	3	AT1G73230.1	Nascent polypeptide-associated complex (NAC) domain	5.0E−73
A01552, A01979	994	1	AT5G16360.1	NC domain-containing protein	1.0E−26
A01801	232	1	AT5G06370.1	NC domain-containing protein	3.0E−27
B02360	557	1	NP_177507.1	Protein kinase family protein	4.0E−87
B01111	256	1	NP_172630.1	RNA binding protein 45 (RBP45B)	3.0E−38
B02312	528	3	AT3G02150	Transcription factor	5.0E−01
B00585A	576	1	AT4G36160.1	Vascular-related NAC domain 2 (ANAC076/VND2)	2.0E−01
C1 metabolism					
A00084	1438	2	CAB40376	Adenosine kinase	1.0E−189
A01097, B02444	820	2	AT3G09820.2	Adenosine kinase 1 (ADK1)	E−104
A00176	1083	4	NP_001047479	Adenosine kinase 2 (ADK2)	1.0E−175
A00474, A00214, A01047	1038	7	NP_001067759	Adenylate kinase 1 (ADK1), putative	1.0E−121
A00120, A01795	1192	8	NP_001105217	Glyoxalase I	1.0E−165
A00532, A00293, A00519, A00564	951	9	AAL73979	Methionine synthase	9.0E−146
A01578	988	1	AAL33589	Methionine synthase	8.0E−150

Table 1 (continued)

Cell wall biosynthesis					
Contig ID	Contig length	EST No.	Target ID	Target description	*E* value
A01824	908	1	CAJ01714	Methionine synthase 2	5.0E−124
A01863	912	1	NP_001104947	Methylenetetrahydrofolate reductase	3.0E−163
A02005, A01607	706	2	NP_001060314	Phospho-2-dehydro-3-deoxyheptonate aldolase 1	3.0E−74
B01452	292	1	AT4G32840.1	Phosphofructokinase family protein	2.0E−01
B01491	348	1	AT1G23190.1	Phosphoglucomutase, putative	2.0E−01
A00707, A00978	804	2	CAA83914	Phosphoglycerate mutase	7.0E−201
B01060	192	1	AT2G21520.2	Phosphoglyceride transfer protein, putative	1.0E−02
B01162	1007	1	NP_001105368	Phosphohexose isomerase1	5.0E−153
A00749	667	1	P50299	*S*-adenosylmethionine synthetase 1	7.0E−34
A01927	871	1	NP_001054575	*S*-adenosylmethionine synthetase 1	4.0E−91
A00090	737	7	NP_001054575	*S*-adenosylmethionine synthetase 2 (MAT2/SAM-2)	2.0E−86
A00171	1511	3	NP_001054575	*S*-adenosylmethionine synthetase 2 (MAT2/SAM-2)	7.0E−217
A00247, A01314	948	3	ABA93502	Serine hydroxymethyltransferase	9.0E−91
A01584	1016	1	BAB89667	Transaldolase	3.0E−112

(1 EST) for *O*-methyltransferases (Table 1). In addition we also observed two ESTs for two dirigents with one containing the full-length sequence. This is important as a role of dirigent proteins in macromolecular lignin assembly has been postulated [11–13]. A single contig (1 EST) for isoflavone reductase was also present in our EST. The cytochrome P450 is a diverse group of heme-containing enzymes that catalyze a wide range of oxidative reactions and have been shown to participate in a variety of biochemical pathways, including biosynthesis of phenylpropanoids [9]. We found ten contigs for cytochrome P450 in our EST library

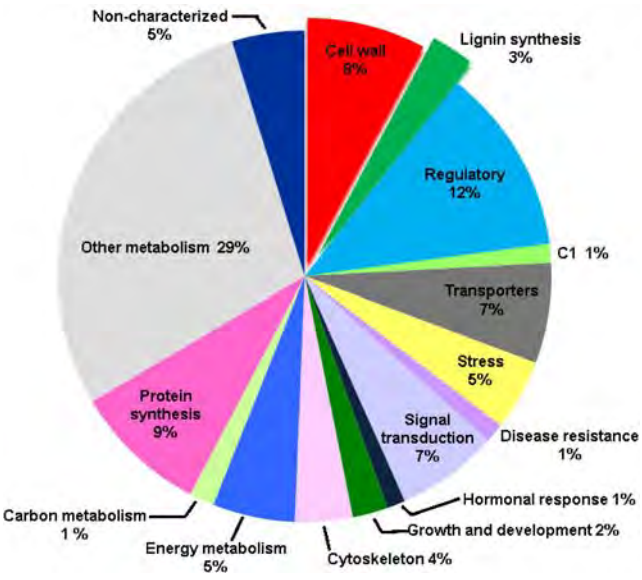

Fig. 2 Functional categories of switchgrass vascular tissue derived ESTs using plant gene ontology terms

representing different gene families. Peroxidases, which comprise a large gene family representing 138 distinct gene models in plants and involved in various hypersensitive responses, play pivotal roles in lignin biosynthesis [4, 25]. In our switchgrass vascular tissue EST library, we identified 15 contigs for peroxidases with three ESTs closely related to anionic peroxidase and tomato peroxidase (TPX1) both of which have been implicated in lignification [32]. Laccases are a large gene family with 17 members in Arabidopsis. Here, we recorded four contigs in our library representing three different members of this gene family.

Transcriptional control and regulatory genes were the most abundant ESTs in our switchgrass vascular bundle library, accounting for 12% of the ESTs. In *Arabidopsis*, which is the most well studied plant genome, about 6% of genes code for transcription factors [33]. Among these regulatory proteins (Table 1; supplementary Table 1), three groups of transcription factors namely Myb, NACs, and LIMs, which have been shown to participate in lignin biosynthesis [20, 40], comprised a significant portion of this category of genes among the ESTs (Table 1). It has been shown that the LIM1 protein acts as a potential transcription factor in lignin biosynthesis by binding to PAL-box motif that could block several steps in phenylpropanoid biosynthesis and also regulate the expression of other genes in the pathway that contain the PAL-box sequence in their promoter regions [20]. Our switchgrass vascular bundle library contained genes encoding three LIM-binding-domain proteins with all three representing different genes including LIM1. Similarly, 18 contigs (20 ESTs) for Myb transcription factors were present in this dataset. Some of these Myb transcription factors were shown

to be involved in the regulation of the phenylpropanoid pathway, such as Myb 1, 58 and 63 [40]. NAC transcription factors are another group of key regulators of secondary wall biosynthesis in *Arabidopsis* [40]. In our switchgrass vascular tissue library, eight contigs (ten ESTs) representing different NAC transcription factors were identified.

The most abundant EST related to cell wall biosynthesis in our library was annotated as glycosyl hydrolases and UDP-glucoronosyl/UDP-glucosyl transferase family. Glycosyl hydrolase families 1, 17, 18, 28 were prominent gene families in the switchgrass vascular bundle library. Presence of these glycosyl hydrolases could be important as recent reports suggest that the repertoires of plant glycosyl hydrolases, with their associated catalytic activities and polysaccharide binding affinities, may have valuable applications in modifying plant cell wall architecture and in the development and characterization of new bioenergy feedstocks [24]. A total of 14 contigs (26 ESTs) for cellulose synthase were also present and they represented seven different *cellulose synthase* genes. Nine glucan synthase-like contigs and 19 contigs for pectinacetylesterase family protein genes were also present (Table 1). In addition, a brittle stalk-2-like protein, calnexin 1, sucrose synthase (4 EST), UDP-glucose 6-dehydrogenase (three ESTs) and five ESTs of structural constituents of the cell wall were other prominent ESTs found in this vascular bundle library (Table 1).

Another group of genes of potential interest in cell wall modification were genes related to carbon 1 (C1) metabolism as this group of genes has been shown to provide the activated methyl groups for methylation of several lignin monomers. Some of these C1 metabolism genes have been shown to be highly expressed in vascular tissues [6]. In this regard, two genes encoding *methionine synthase* (*MS*) and *S-adenosyl methionine synthetase* (SAM-synthetase) that are involved in C1 metabolism were present in our vascular tissue library. Based on sequence similarity, six contigs (11 ESTs) encoded at least three different *methionine synthases* and four contigs (12 ESTs) encoding *SAM-synthetase* were found. Other genes most likely involved in C1 metabolism were *serine hydroxymethyltransferase*, *adenosine kinase*, and *adenylate kinase* (Table 1).

Comparison of Switchgrass Vascular Tissue ESTs with Previously Published Switchgrass ESTs

Previous major switchgrass EST sequencing efforts produced 11,990 and 61,585 ESTs [38, 39] from the northern lowland ecotype "Kanlow". To determine whether our vascular tissue data set yielded additional switchgrass genes, all existing switchgrass EST were downloaded from GenBank and used to build an EST database for local blast analysis. We also compared our switchgrass data set with Switchgrass Gene Index PaviGI. This dataset was built from 435,448 ESTs which included data reported by Tobias et al. [38]. Overall, 16% of our ESTs did not have any significant hits to the existing EST data. Out of 216 unigenes reported for cell wall synthesis, we found that 14% unigenes from our vascular tissue data set were not present in the existing switchgrass ESTs. This list included glycosyl hydrolases, cell wall protein precursors, sucrose synthases, and galactosyltransferase family proteins (Fig. 3; Supplemental File 2). We also found that 21% of the unigenes related to lignin biosynthesis reported in our study were absent from the aforementioned switchgrass database. This indicates that transcripts involved in secondary cell wall related processes were enriched in our data set. Also notable are additional ten transcription factors and eight transporter genes in our vascular tissue data set that were not present in previous switchgrass EST data sets (Fig. 3; Supplemental File 2).

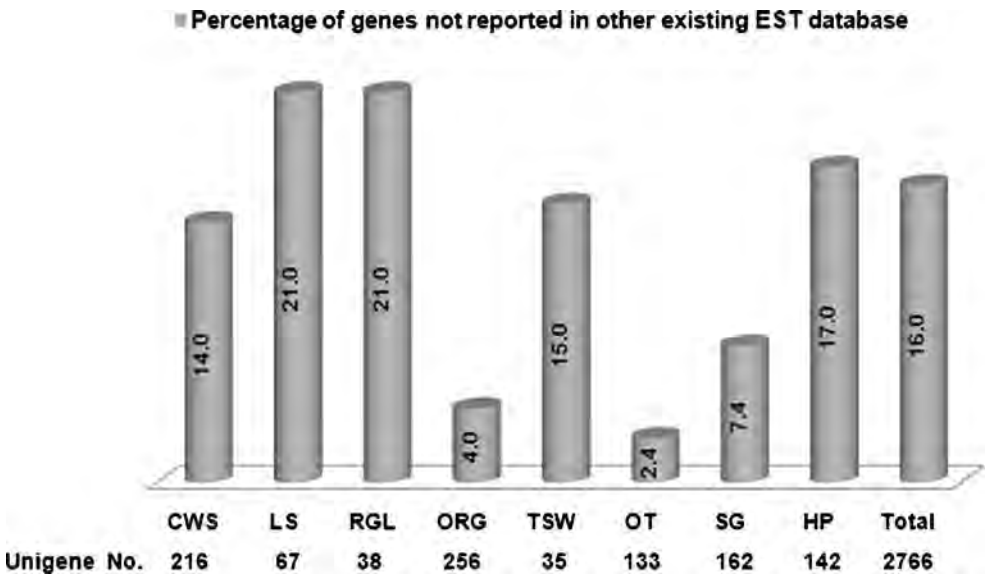

Fig. 3 Comparative blast analysis of unigenes derived from vascular tissues (this study) and other existing switchgrass ESTs revealed a significant number of transcripts not present in other studies. *CWS* Cell wall synthesis, *LS* lignin synthesis, *RGL* regulatory genes for lignin synthesis, *ORG* other regulatory genes, *TSW* transporters for cell wall, *OT* other transporters, *SG* signal transduction, *HP* hypothetical proteins and other metabolism

Validation of Vascular Tissue Specificity of ESTs by In Situ Hybridization

Validation of ESTs recovered from vascular tissues was carried out by in situ hybridization. Seven genes were selected for in situ hybridization and labeled antisense oligo probes were used to hybridize with the desired mRNA while labeled sense probes were used as negative controls. Antisense probes were selected based on conserved sequence of various representative genes. The genes we tested included *dehydrin*, *brittle stalk*, *MS*, *copper chaperon* (*CCH*), *dirigent*, a*nionic peroxidase* or *peroxidase 30* homolog (*Per30*), and *aspartate aminotransferase 1* (*ASP1*). Out of the seven genes tested, five genes namely *MS*, *CCH*, *ASP1*, *Per30*, and *dirigent* had strong expression in the vascular tissues (Fig. 4a–g). *Dehydrin* (Fig. 4h) and *brittle stalk 2* (data not shown) were expressed in both the vascular bundles and in other cells of the internode. Representative antisense controls for some genes are shown in Fig. 4i–l.

Discussion

The objective of this study was to generate an inventory of genes expressed in switchgrass vascular tissues. Although there have been extensive efforts in producing genomic resources for switchgrass [38], there is still a need to generate additional genomic tools using more targeted approaches such as LCM to further increase our understanding of the biology of this emerging bioenergy crop. Particular emphasis should be on cell wall and lignin biosynthesis because of the significance of these processes to plant recalcitrance. Since a close functional relationship has been established between lignin deposition and secondary cell wall biogenesis in vascular tissues [17, 21, 28], information regarding ESTs derived exclusively from vascular tissues should provide an important resource for switchgrass researchers interested in reducing recalcitrance.

To validate our EST list, we compared our switchgrass data set with vascular tissue EST data obtained from other plant species. A summary of this comparative analysis is provided in Table 2. It is worth noting that the list of genes represented in our switchgrass vascular tissue library was well represented in vascular-tissue-specific genes from other plant species. For example, Nakazono et al. [29] reported that 71 transcripts were enriched in LCM isolated vascular tissues of maize. Homologs of all 71 transcripts shown in that study were present in our switchgrass vascular bundle EST library (Table 2). Similarly, 3,247 independent mRNAs from *Plantago* vasculature were studied [31] and out of 90 highly expressed contigs (the 20 "non-characterized" genes are not included) in vascular

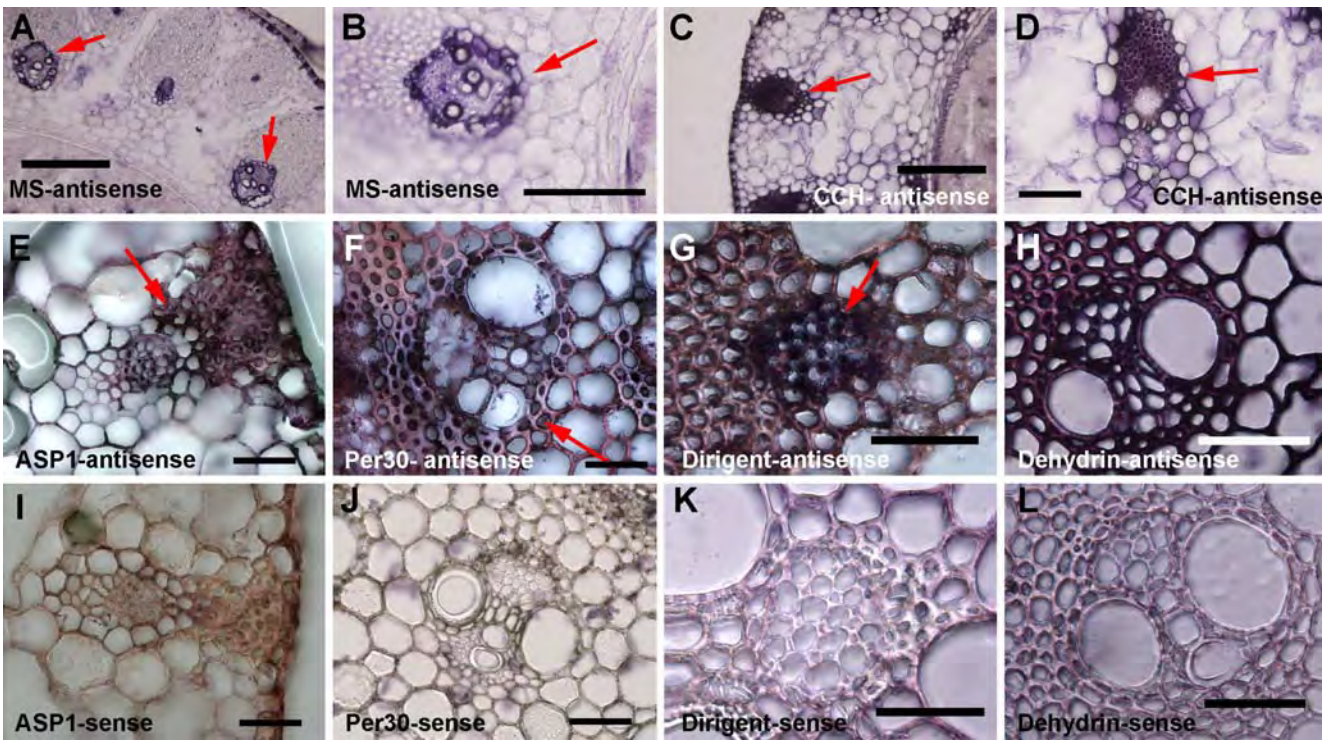

Fig. 4 In situ hybridization of selected genes in switchgrass cross sections. Antisense (*AS*) probes of *methionine synthase* (*MS*; **a**, **b**), *copper chaperon* (*CCH*; **c**, **d**), *aspartate aminotransferase 1* (*ASP1*; **e**) *peroxidase 30* (*Per30*; **f**), *dirigent* (**g**), and *dehydrin* (**h**) were detected in vascular tissues (*red arrows*). Representative sense controls are shown for *ASP1* (**i**), *Per30* (**j**), *dirigent* (**k**) and *dehydrin* (**l**). *Scale bars* 100 μm (**a**, **c**); 30 μm (**b**, **d**, **e**, **f**, **i**, **j**); 20 μm (**g**, **h**, **k**, **l**)

Table 2 List of genes identified by previous work of potential vascular specificity that were also identified in switchgrass LCM derived vascular tissue

Genes present in this library	Unigene ID	Deeken et al. (2008; *Arabidopsis*[a])	Pommerrenig et al. 2006; plaintain)	Nakazono et al. 2003; maize[a])	Broeck et al. 2007; *Cannabis*)
Cell wall biosynthesis					
Cellulose synthase (IRX3)	A00142			+	+
Endo-1,4-Beta-D-glucanase	A01387				+
Sucrose synthase	A01317				+
Lignin Biosynthesis					
Cinnamoyl-CoA reductase	B01641				+
4-coumarate coenzyme A ligase	A00367				+
Phenylalanine ammonia-lyase	B02470			+	+
Regulatory genes in lignin synthesis					
NAC domain-containing protein 82	A01875	+			
LIM1; transcription factor	B02263				
MYB transcription factor	B01151		+	+	
C1 metabolism					
Methionine synthase 2	A01824				+
S-adenosylmethionine synthetase 1	A01927				+
Adenosine kinase 2 (ADK2)	A00176			+	+
Transport facilitators					
Copper chaperone	A00270	+	+	+	+
Sulfate transporter protein	A01180			+	
Sugar transporter	A01522	+	+		
Abiotic and biotic stress					
Senescence-associated protein, putative	A00073	+			
Pathogen-related protein	B01016	+			
Glutathione transferase 9	A00434	+		+	
Signal transduction					
Calmodulin 2	A00119	+		+	+
Calcium-dependent protein kinase	B01454	+		+	+
1-phosphatidylinositol-3-phosphate 5-kinase	A01475	+			
Cytoskeleton					
Tubulin alpha-1 chain	A00028		+		+
Beta-expansin 7	A01267		+		
Actin depolymerizing factor	A00113		+		
Other					
MADS-domain transcription factor	A00391		+		
14-3-3 protein	A00221	+			
WRKY40 (WRKY DNA-binding protein 40)	A01702	+			

+ reported as vascular specific transcripts

[a] These transcripts were obtained using LCM

tissues, our switchgrass vascular tissue library had 75 common sequences. Recently phloem-based transcripts of *Arabidopsis* were enriched from phloem cells isolated using laser pressure capture microdissection (LMPC) or from phloem exudates [15]. One hundred thirteen phloem mobile signal transcripts were reported in the aforementioned study and more than 90% of these were seen in our switchgrass vascular tissue library. In *Cannabis sativa*, a total of 178 clones encoding 44 highly expressed proteins for lignin and cell wall synthesis were reported in vascular tissues. Our switchgrass EST dataset had more than 90% similarity with the *Cannabis* vascular transcript list (Table 2). Although we did not have an EST library from non-vascular tissue of switchgrass for comparison, the fact

that our gene inventory closely matched the gene lists from vascular tissues of other plant species strongly indicate that ESTs acquired using our LCM method truly represent genes that are expressed in switchgrass vascular tissues.

To further verify the specificity of our switchgrass vascular tissue library, we examined the expression of seven genes from our list by in situ hybridization. Out of the seven genes selected, five genes had enhanced expression in vascular tissues compared to other tissues. The genes with strong expression in the vascular tissues were *CCH*, *MS*, *Per30*, *ASP1*, and *dirigent* (Fig. 4). Among these five genes, the vascular tissue expression of three genes was consistent with the expression patterns of homologous genes reported in other plant species. For instance, the CCH protein was shown to accumulate in sieve elements of *Arabidopsis* stems, leaves and phloem exudates [26] while immunolabeling indicated that anionic peroxidases, which closely matches the switchgrass *peroxidase 30* gene reported here, are expressed in the differentiating xylem of tomato and *Zucchini*, particularly during secondary wall formation [3, 4, 8]. Furthermore, subcellular localization of protein epitopes in *F. intermedia* stem sections demonstrated that dirigent is expressed in vascular cambium, and secondary xylem fibers [11, 13]. In this study, it appears that one the switchgrass dirigent gene we identified is dominantly expressed in the phloem tissue. We also found that not all the ESTs we tested by in situ hybridization are vascular-specific. For example, dehydrin and brittle stalk expression was not limited to the vascular region. But this is not surprising since in situ immunolocalization of dehydrin-like proteins was previously shown to be expressed in all tissues of the plant [7, 16].

When compared to previous switchgrass EST databases [38, 39], the unigene dataset generated from vascular tissues uncovered additional gene sequences that were not present in the larger switchgrass EST collections. We not only found additional sequences for cell wall and lignin-related genes but also new sequence information for transcription factors, sugar and ABC transporters, mobile signal proteins, and hypothetical proteins. Considering the fact that the switchgrass genome is still not fully sequenced, the vascular EST sequence data set we present here along with previously published ESTs of switchgrass should be very useful in deciphering the genomic structure of secondary cell wall formation for this important bioenergy crop.

In conclusion, we describe the generation of EST sequences from switchgrass vascular tissues forming 2,766 unigenes. This EST data set contained transcripts related to lignin and cell wall biosynthesis, and transcription factors that are known to regulate these pathways. These gene sequences provide baseline information for defining molecular targets for modifying lignin content not only in switchgrass but also in other monocotyledonous bioenergy crops. Furthermore, this gene list paves the way for more in-depth analysis of the tissue-specific expression patterns for thousands of other switchgrass genes that could be novel targets for reducing recalcitrance.

Acknowledgments We thank Drs Zengyu Wang and Jiyi Zhang for critical reading of the manuscript, and Dr. Ji He, for assistance with the EST data analysis. The research described in this paper was carried out as part of the BESC (The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science US Department of Energy) and also funded by the Samuel Roberts Noble Foundation. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of the authors expressed herein do not necessarily reflect those of the United States Government or any agency thereof.

References

1. Asano T, Masumura T, Kusano H, Kikuchi S, Kurita A, Shimata H et al (2002) Construction of a specialized *cDNA* library from plant cells isolated by laser capture microdissection: toward comprehensive analysis of the genes expressed in the rice phloem. Plant J 32:401–408
2. Balat M (2008) Global trends on the processing of bio-fuels. Intl J Green Energy 5:212–238
3. Botella MA, Amaya I, Valpuesta V (2000) A tomato peroxidase involved in the synthesis of lignin and suberin. Plant Physiol 122:1119–1127
4. Botella MA, Quesada MA, Kononowicz AK, Bressan RA, Pliego F, Hasegawa PM et al (1994) Characterization and in situ localization of a salt-induced tomato peroxidase mRNA. Plant Mol Biol 25:105–114
5. Bouton JH (2007) Molecular breeding of switchgrass for use as a biofuel crop. Curr Opin Genet Dev 17:553–558
6. Broeck HCvd, Maliepaard C, Ebskamp MJM, Toonen MAJ, Koops AJ (2008) Differential expression of genes involved in C1 metabolism and lignin biosynthesis in wooden core and bast tissues of fiber hemp (*Cannabis sativa* l.). Plant Sci 174:205–220
7. Carjuzaa P, Castellión M, Distéfano AJ, Del Vas M, Maldonado S (2008) Detection and subcellular localization of dehydrin-like proteins in quinoa (*Chenopodium quinoa* willd.) embryos. Protoplasma 233:149–156
8. Carpin S, Crevecoeur M, Greppin H, Penel C (1999) Molecular cloning and tissue-specific expression of an anionic peroxidase in *Zucchini*. Plant Physiol 120:799–810
9. Chapple C, Carpita N (1998) Plant cell walls as targets for biotechnology. Curr Opin Plant Biol 1:179–185
10. Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. Nature Biotechnol 25:759–761

11. Davin LB, Jourdes M, Patten AM, Kim KW, Vassão DG, Lewis NG (2008) Dissection of lignin macromolecular configuration and assembly: comparison to related biochemical processes in allyl/propenyl phenol and lignan biosynthesis. Nat Prod Rep 25 (6):1015–1090
12. Davin LB, Lewis NG (2000) Dirigent proteins and dirigent sites explain the mystery of specificity of radical precursor coupling in lignan and lignin biosynthesis. Plant Physiol 123:453–462
13. Davin LB, Lewis NG (2005) Lignin primary structures and dirigent sites. Curr Opin Biotechnol 16(4):407–415
14. Day A, Addi M, Kim W, David H, Bert F, Mesnage P et al (2005) ESTs from the fibre-bearing stem tissues of flax (*Linum usitatissimum* l.): expression analyses of sequences related to cell wall development. Plant Biol 7:23–32
15. Deeken R, Ache P, Kajahn I, Klinkenberg J, Bringmann G, Hedrich R (2008) Identification of *Arabidopsis thaliana* phloem RNAs provides a search criterion for phloem-based transcripts hidden in complex datasets of microarray experiments. Plant J 55:746–759
16. Dejun L, Chunhua Y, Xiaobing L, Guobiao J, Lihuang Z (2008) Sense and antisense osdof12 transcripts in rice. BMC Mol Biol 9:80
17. Dixon RA, Lamb CJ (1990) Molecular communication in interactions between plants and microbial pathogens. Ann Rev Plant Biol 41:339–367
18. Gomez SK, Harrison MJ (2009) Laser microdissection and its application to analyze gene expression in arbuscular mycorrhizal symbiosis. Pest Manag Sci 65:504–511
19. Ivashikina N, Deeken R, Ache P, Kranz E, Pommerrenig B, Sauer N et al (2003) Isolation of Atsuc2 promoter-gfp-marked companion cells for patch-clamp studies and expression profiling. Plant J 36:931–938
20. Kawaoka A, Ebinuma H (2001) Transcriptional control of lignin biosynthesis by tobacco LIM protein. Phytochem 57:1149–1157
21. Keller B, Templeton MD, Lamb CJ (1989) Specific localization of a plant cell wall glycine-rich protein in protoxylem cells of the vascular system. Proc Natl Acad Sci 86:1529–1533
22. Kerk NM, Ceserani T, Tausta SL, Sussex IM, Nelson TM (2003) Laser capture microdissection of cells from plant tissues. Plant Physiol 132:27–35
23. Keshwani DR, Cheng JJ (2009) Switchgrass for bioethanol and other value-added applications: a review. Bioresour Technol 100:1515–1523
24. Lopez-Casado G, Urbanowicz BR et al (2008) Plant glycosyl hydrolases and biofuels: a natural marriage. Curr Opin Plant Biol 11:329–337
25. Lucena MA, Romero-Aranda R, Mercado JA, Cuartero J, Valpuesta V, Quesada MA (2003) Structural and physiological changes in the roots of tomato plants over-expressing a basic peroxidase. Physiol Plant 118:422–429
26. Mira H, Martinez-Garcia F, Penarrubia L (2001) Evidence for the plant-specific intercellular transport of the *Arabidopsis* copper chaperone CCH. Plant J 25:521–528
27. Mulkey VR, Owens VN, Lee DK (2008) Management of warm-season grass mixtures for biomass production in South Dakota USA. Bioresour Technol 99:609–617
28. Müse G, Schindler T, Bergfeld R, Ruel K, Jacquet G, Lapierre C et al (1997) Structure and distribution of lignin in primary and secondary cell walls of maize coleoptiles analyzed by chemical and immunological probes. Planta 201:146–159
29. Nakazono M, Qiu F, Borsuk LA, Schnable PS (2003) Laser-capture microdissection, a tool for the global analysis of gene expression in specific plant cell types: identification of genes expressed differentially in epidermal cells or vascular tissues of maize. Plant Cell 15:583
30. Nelson T, Tausta SL, Gandotra N, Liu T (2006) Laser microdissection of plant tissue: what you see is what you get. Annu Rev Plant Biol 57:181–201
31. Pommerrenig B, Barth I, Niedermeier M, Kopp S, Schmid J, Dwyer RA et al (2006) Common plantain. A collection of expressed sequence tags from vascular tissue and a simple and efficient transformation method. Plant Physiol 142:1427
32. Quiroga M, Guerrero C, Botella MA, Barcelo A, Amaya I, Medina MI et al (2000) A tomato peroxidase involved in the synthesis of lignin and suberin. Plant Physiol 122:1119–1128
33. Riechmann JL, Heard J, Martin G, Reuber L, Jiang ZC et al (2000) *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. Sci 290:2105–2110
34. Rubin EM (2008) Genomics of cellulosic biofuels. Nature 454 (7206):841–845
35. Sarath G, Mitchell RB, Sattler SE, Funnell D, Pedersen JF, Graybosch RA et al (2008) Opportunities and roadblocks in utilizing forages and small grains for liquid fuels. J Ind Microbiol Biotech 35:343–354
36. Schmer MR, Vogel KP, Mitchell RB, Perrin RK (2008) Net energy of cellulosic ethanol from switchgrass. Proc Natl Acad Sci 105:464
37. Sticklen MB (2008) Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. Nat Rev Genet 9:433–443
38. Tobias CM, Sarath G, Twigg P, Lindquist E, Pangilinan J, Penning B et al (2008) Comparative genomics in switchgrass using 61,585 high-quality expressed sequence tags. The Pl Genome 1:111–124
39. Tobias CM, Twigg P, Hayden DM, Vogel KP, Mitchell RM, Lazo GR et al (2005) Analysis of expressed sequence tags and the identification of associated short tandem repeats in switchgrass. TAG Theor Appl Genetics 111:956–964
40. Zhong RQ, Lee CH, Zhou JL, McCarthy RL, Ye ZH (2008) A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. Plant Cell 20:2763–2782