J Wood Sci (2012) 58:89–103
DOI 10.1007/s10086-011-1234-6

ORIGINAL ARTICLE

Collection of expressed genes from the transition zone of *Cryptomeria japonica* in the dormant season

Kazumasa Yoshida · Norihiro Futamura · Mitsuru Nishiguchi

Received: 19 July 2011 / Accepted: 21 October 2011 / Published online: 12 January 2012

Abstract Heartwood affects the utility of wood because it differs in some properties compared to sapwood. To regulate heartwood formation, its mechanism must be elucidated. However, the molecular basis underlying heartwood formation remains largely unknown. To obtain clues to understand the mechanism at a molecular level, we collected expressed sequence tags (ESTs) from the transition zone (TZ) of *Cryptomeria japonica* D. Don in November, in which heartwood formation is considered to proceed. A total of 1029 ESTs were assembled into 744 unique sequences (103 clusters and 641 singletons). Putative functions were assigned to 291 nuclear-encoded sequences, and they were grouped into 21 categories according to the eukaryotic orthologous groups functional classification. We selected 20 genes for enzymes or proteins, then examined their expression patterns among different organs. The expression levels of nine genes were higher in November than in June in the TZ. The genes encode two enzymes in glycolysis, invertase, methionine adenosyltransferase, glutathione transferase, the lipid transfer protein, Bet v 1 allergen, the dehydrin and the function-unknown protein. This study has provided the first large-scale EST information from the TZ of conifers, which will be useful for understanding the physiological processes in the TZ at a molecular level.

Keywords *Cryptomeria japonica* · Transition zone · EST · Gene expression · Heartwood formation

Introduction

Heartwood formation is a unique feature of woody plants. Heartwood is important both for a tree's life and wood usage. The roles of heartwood are considered to be optimizing sapwood volumes, recycling nutrients, and providing mechanical support with durability [1, 2]. For the usage of wood, heartwood has positive or negative effect on wood properties due to the presence of extractives [3]. To utilize trees more efficiently as wood sources, it is crucial to understand the mechanism of heartwood formation toward regulation of the formation. Heartwood formation has been studied from anatomical, chemical and biochemical perspectives [3–5]. However, why and how heartwood is formed remains largely elusive. It has been proposed that heartwood formation is an actively regulated physiological process [3, 5–7] and it is regarded as the final stage of wood differentiation [8]. The early stage is mature xylem formation, which includes cell division from vascular cambium, cell expansion, cell wall thickening and lignifications [1]. For this stage, expressed genes were collected and changes in gene expression were investigated on a large scale for several tree species belonging to *Pinus* [9–12], *Populus* [13–16], *Eucalyptus* [17–20] and *Acacia* [21]. In contrast, there have been only a few large-scale gene analyses for heartwood formation. Studies have been conducted on broad-leaved trees, *Robinia pseudoacacia* [22, 23] and *Juglans nigra* [24, 25], but there have been no reports on conifers. The sugi tree (*Cryptomeria japonica* D. Don) has sharply defined heartwood, and usually also a readily recognized transition zone (TZ) between sapwood

Part of this work was presented at the 23rd IUFRO World Congress, Seoul, Republic of Korea, August 2010, and the 61st Annual Meeting of the Japan Wood Research Society, Kyoto, March 2011.

K. Yoshida (✉) · N. Futamura · M. Nishiguchi
Department of Molecular and Cell Biology, Forestry and Forest Products Research Institute, Tsukuba, Ibaraki 305-8687, Japan
e-mail: ykazu@ffpri.affrc.go.jp

and heartwood [26]. Thus, *C. japonica* is a good sample material to study events having occurred during heartwood formation. To obtain clues to elucidate the mechanism of heartwood formation at a molecular level, we collected expressed sequence tags (ESTs) from the TZ in November, in which heartwood formation is considered to proceed. The ESTs were assembled and the resulting sequences were functionally categorized. Furthermore, the expression of genes selected as having potential involvement in heartwood formation was quantified in various organs.

Materials and methods

Plant materials

A 19-year-old *C. japonica* tree that grew in the nursery of the Forestry and Forest Products Research Institute (Tsukuba) was felled on November 28, 2003. A trunk was excised from a height of 1.2 m above ground. A disk 60-mm thick was cut from the bottom end of the log. The transition zone (2-year annual rings) was isolated from the disk, cut into small pieces, frozen in liquid nitrogen, and stored at −80°C. To examine the gene expression, the following organs were collected from one individual: TZ and sapwood (June 5 and November 26, 2008), inner bark (June 25, 2008), leaf buds (May 8, 2008), needles (June 25, 2008), pollen (March 21, 2008), male and female strobili (February 8, 2008), and young cones (June 25, 2008).

Construction of a cDNA library

Total RNA was prepared from the TZ collected in November according to the method previously reported [27]. cDNA was synthesized from 1.4 μg of total RNA using the Creator SMART cDNA library construction kit (Clontech, California, USA). The cDNAs were cloned into the pDNR-LIB vector (Clontech) and transformed into *Escherichia coli* strain DH5α.

Sequencing and data analysis

Sequences of cDNA inserts were determined with an ABI 3730 DNA analyzer (Applied Biosystems, California, USA). Raw sequence data were processed by ABI base caller with quality values. Low-quality sequences (quality score <20 at 750 bp) were discarded. Vector and adaptor sequences were trimmed. Trimmed sequences (≥100 bp) were assembled using Sequencher 4.1.2 (Gene Codes, Michigan, USA) with the parameters of minimum overlap = 40, minimum match = 95%. Sequences of ribosomal RNA (rRNA), chloroplast or mitochondrial DNA were identified by the BLASTN search against *Populus* rRNA sequences (accession numbers AF174629, AF206999, AF479118, AJ006440), *Arabidopsis* mitochondrial genome (NC001284) and *Populus* chloroplast genome (http://genome.ornl.gov/poplar_chloroplast/), and these were removed. Remaining sequences were searched locally against the eukaryotic orthologous groups protein databases [28] (KOG, ftp://ftp.ncbi.nih.gov/pub/COG/KOG/) and against the Universal Protein Resource [29] release 13.3 (UniProt, http://www.uniprot.org/) using the BLASTX program, respectively. The KOG comprises three databases containing orthologous proteins from at least three out of seven eukaryotic species, proteins from two species, and species-specific proteins. Sequences with an expectation (*E*) value of $<10^{-5}$ were considered to have significant homology, and were classified following the KOG functional classification. The sequences of ESTs have been submitted to the DNA Data Bank of Japan (DDBJ) under accession numbers DC882454–DC883482.

Gene expression analysis

Total RNA was isolated from the TZ and sapwood according to the method previously reported [27]. From the inner bark, leaf buds, needles, male strobili, female strobili and young cones, RNA was prepared as described by Futamura et al. [30]. RNA was also extracted according to Sone et al. [31] from pollen. The RNA was treated with RQ1 RNase-free DNase (Promega, Wisconsin, USA) to remove contaminant DNA before reverse transcription. The quality of the RNA was evaluated with the Agilent 2100 bioanalyzer (Agilent Technologies, California, USA). First-strand cDNA was synthesized from 100 ng of RNA using an AffinityScript QPCR cDNA synthesis kit (Stratagene, California, USA) with a mixture of oligo(dT) and random primers according to the manufacturer's instructions. The resultant cDNA solution was diluted 50 times with water and used as a template. The real-time quantitative polymerase chain reaction (qPCR) was performed with Mx3000P (Stratagene) using Brilliant II SYBR green QPCR master mix (Stratagene). The reaction mixture for PCR was composed of 5 μl of the diluted first-strand cDNA (equivalent to 0.5 ng of total RNA), 10 μl of 2× Brilliant II SYBR green QPCR master mix (Stratagene) and 0.2 μM gene-specific primers in a total volume of 20 μl. The primer sequences are listed in Table 1. Amplification was performed with an initial polymerase activation step at 95°C for 10 min followed by 40 cycles of denaturation at 95°C for 30 s, annealing at 58°C for 1 min, and an extension at 72°C for 30 s. To examine the specific amplification of target genes, melting curves were obtained after the amplification. Correct amplification was further

Table 1 Sequences of primers used for gene expression analysis

Annotation (sequence ID)	Forward primer (5′–3′)	Reverse primer (5′–3′)
Bet v 1 allergen (Cluster0012)	CTTGTGGAAGGAGGAATGC	GACCATGATGGTTTCAGCAG
Cysteine protease (TZ10B08)	CTGCGTATGGATATGCAGTG	CGTCAATCGAAACGACGTTG
Cytosolic thioredoxin (Cluster0013)	GATGTGACTGCAGAGTGG	CACGATTTCAAGGTAGCATGC
Dehydrin (Cluster0014)	CGGTCATGGCGAACAC	CGAACAAAACATTGCAGTTACAAG
Dormancy-associated protein (Cluster0011)	AGTACCATCGTCTCCCTC	CACCACTGTAAAGCCAGTC
Fructose-bisphosphate aldolase (TZ08D11)	CAAAGGTGTGCTACCTATTAC	TGTTCATGGCGTTCAAGTTC
Function-unknown protein (Cluster0010)	CACATGATCAAGGACACGTAG	GCGTAAGGTGAGGTTCAG
GDSL-motif lipase (TZ05F04)	TGAGTGATGGCATGCATCTG	GGTTCATTGAAGTCACTTGGC
Glucose-6-phosphate dehydrogenase (TZ02H09)	TTGAGTCCAGTGGAGGAC	CCATCCATTAGCAGAAGATGAAG
Glutathione transferase (TZ02A05)	GCAAATGAATCCCGTGCAC	CTTATCTGGAGAAGGAAGGC
Glyceraldehyde-3-phosphate dehydrogenase (Cluster0099)	GAGGTTGGTGCAGAGTATG	GTCAAATCCACAACTGATACATCTG
3-Hydroxy-3-methylglutaryl-CoA reductase (TZ02F02)	CTTGGCAAGAATAGTGGCAG	GAAACTAACCCCATGGTAGAG
Invertase (TZ11C11)	CTTGTTCAACAATGCCACTACAAC	CCAGAGCTCACATGCAAG
Lipid transfer protein (Cluster0009)	GTGCTGTAATGGCGTTAAGTC	CGTGTTGCAGTTAATCGAGG
Methionine adenosyltransferase (Cluster0070)	GTCTTTGGCAAAATCACCACC	GGCTTGATCACATGCTCC
Oleosin (Cluster0002)	CTGCAAGCCATGTGAAGG	GAACTGATCAAAACTACTCTGTAC
Phosphoglycerate mutase (TZ07G04)	GGTTCAAGTACATCAAGGAATGC	GGCAGGTATTGATCACTAGC
Pyruvate dehydrogenase E1 β subunit (TZ07D11)	GCAGGCTATTGATCACATC	GCAATCCACGAGCATCTTC
Sucrose-phosphate synthase (TZ09B07)	CACAATGTAGGATGAGACACC	CACCAAACTGGAGCAAAGAAAG
Splicing factor 3B subunit (Cluster0017)	GCAGAGAATGAGTCTATTGGAAG	CTCCTCATTGTTACACATCCATTG
Translation initiation factor	GTCAGATCTAGACGTTCAGATTC	GTCCTTCTTCACAATTCCAGC

verified by agarose gel electrophoresis and by the DNA sequencing of PCR products. The relative quantity of target mRNA was normalized using the gene for a translation initiation factor [27] (BJ936692) as an internal standard. The mean values and SD were calculated from the triplicate measurement.

Results

Collection of ESTs from the cDNA library of the TZ in November

In *C. japonica*, heartwood formation is considered to start in late summer and continue to winter in view of the cytological study [32]. In several previous studies, samples were collected in November to examine enzyme activities or expressed genes in relation to heartwood formation [8, 22, 23, 33–35]. Therefore, we constructed a cDNA library using the RNA prepared from the TZ collected in November. Over 1000 clones were sequenced and a total of 1029 ESTs were generated with an average length of 575 bp. After sequence assembly, 641 ESTs were found to be singletons, while the other 388 ESTs were assembled into 103 clusters containing from two to 44 ESTs.

Functional classification of ESTs

Sequences originated from chloroplast or mitochondria genomes, and those of rDNA were removed from the 744 unique sequences (singletons and clusters). The remaining 676 sequences were searched against the KOG using the BLASTX program. The sequences that showed significant similarity (an E value of $<10^{-5}$) with those in the database were annotated and assigned KOG functional classes. Annotation was given to 284 sequences. The sequences with an E value of $\geq 10^{-5}$, those assigned to "function unknown" or "unnamed proteins", and those that were not assigned were further searched against the UniProt. As a result, seven additional sequences were annotated. Finally, 291 sequences were annotated to known sequences with putative functions, 120 were similar to known genes whose functions were unknown and 265 had no similarity to sequences in the databases. The 291 annotated and 120 unknown-function sequences were categorized according to the KOG functional classification, and grouped into 22 categories (Fig. 1).

Sequences with annotation are listed in Table 2. The number of predicted proteins was fewer than that of sequences because sequences annotated as the same protein were represented by a sequence that had the lower E value.

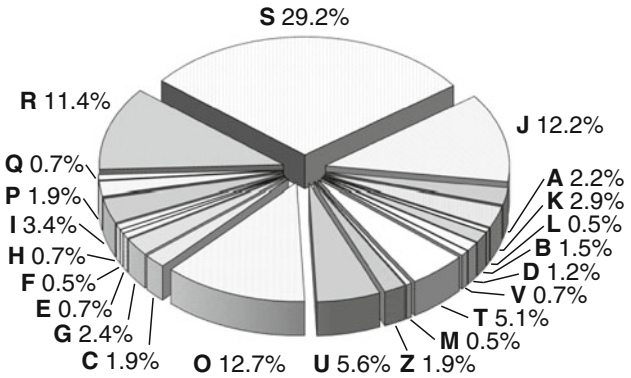

Fig. 1 Functional classification of unique transcripts in the transition zone in November. A total of 411 non-redundant sequences were assigned to the KOG functional category. Designations of functional categories: A, RNA processing and modification; B, chromatin structure and dynamics; C, energy production and conversion; D, cell cycle control, cell division, chromosome partitioning; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; J, translation, ribosomal structure and biogenesis; K, transcription; L, replication, recombination and repair; M, cell wall/membrane/envelope biogenesis; O, posttranslational modification, protein turnover, chaperones; P, inorganic ion transport and metabolism; Q, secondary metabolite biosynthesis, transport and catabolism; T, signal transduction mechanisms; U, intracellular trafficking, secretion and vesicular transport; V, defense mechanisms; Z, cytoskeleton; R, general functional prediction only; S, function unknown

"Posttranslational modification, protein turnover, chaperones (category symbol O)" was the largest category with putative function, which included 52 sequences encoding 33 different proteins (Table 2). A sequence encoding cytosolic thioredoxin, which reduces the disulfide bonds of target proteins and contributes to cell redox homeostasis, was most abundant in this category. The second largest category was "translation, ribosomal structure and biogenesis (J)." This consisted of 50 sequences encoding 42 independent proteins. Forty of the 50 sequences encoded various kinds of ribosomal proteins. Excluding "general function prediction only (R)", the next larger one was "intracellular trafficking, secretion, and vesicular transport (U)" which contain 22 sequences. This was followed by "signal transduction mechanisms (T)", "lipid transport and metabolism (I)", "transcription (K)", "RNA processing and modification (A)", and "carbohydrate transport and metabolism (G)". The remaining categories contained fewer than 10 sequences.

Clustering of ESTs revealed highly expressed transcripts. Except for rRNA encoded by nuclear, chloroplast or mitochondrial genomes, clusters containing five ESTs and more are listed in Table 3. The most abundant transcript encodes an oleosin that constitutes the surface layers of oil bodies, subcellular particles in cells [36]. Clusters encoding a lipid transfer protein (LTP), a function-unknown protein, a dormancy-associated protein, Bet v 1 allergen and a cytosolic thioredoxin included seven ESTs, respectively. Dehydrins were encoded by Cluster0014 containing six ESTs and Cluster0019 containing five ESTs.

Expression patterns of selected genes in various organs

It has been reported that the activities of several enzymes were increased in the TZ chiefly in the dormant season [33–35, 37]. Glycolysis and TCA cycle are fundamental metabolic pathways which produce energy and provide materials for secondary metabolites. Based on these findings, we selected genes for enzymes expected to be associated with heartwood formation from the ESTs collected in the present study. Namely, genes for sucrose-phosphate synthase, invertase (β-fructofuranosidase), fructose-bisphosphate aldolase, glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate mutase, a pyruvate dehydrogenase E1 β subunit, glucose-6-phosphate dehydrogenase, GDSL-motif lipase, cysteine protease, methionine adenosyltransferase (*S*-adenosylmethionine synthetase), glutathione transferase and 3-hydroxy-3-methylglutaryl-CoA reductase (HMGR). In addition, eight abundantly found transcripts containing more than five ESTs were included in the examination. The genes encode an oleosin, a LTP, a function-unknown protein encoded by Cluster0010, a dormancy-associated protein, Bet v 1 allergen, a cytosolic thioredoxin, a dehydrin 4 encoded by Cluster0014, and a splicing factor 3B subunit (Table 3). To estimate whether those enzymes and proteins participate in heartwood formation, the expression levels of the genes in the TZ in the dormant season (November) were compared with those in the TZ in early summer (June) and in several other organs.

Among the 20 selected genes, the expression levels of nine genes were significantly higher in the TZ in November than in June when the mean values were compared by Student's *t* test at $P < 0.05$. The genes were those encoding invertase, glyceraldehyde-3-phosphate dehydrogenase, phosphoglycerate mutase, methionine adenosyltransferase, glutathione transferase, the LTP, Bet v 1 allergen, the dehydrin 4 and the function-unknown protein (Fig. 2a–i). Conversely, the expression of genes for fructose-bisphosphate aldolase, the pyruvate dehydrogenase E1 β subunit, sucrose-phosphate synthase, cysteine protease, the oleosin and the dormancy-associated protein was lower in November ($P < 0.05$) (Fig. 2j–o). The expression levels of genes for glucose-6-phosphate dehydrogenase, GDSL-motif lipase, HMGR, the cytosolic thioredoxin and the splicing factor 3B subunit remained unchanged between November and June (Fig. 2p–t).

Table 2 Functional classification of genes expressed in TZ of *C. japonica* in November

Sequence ID	Locus code/ accession no.	*E* value	Annotation	Number of EST
Information storage and processing				
Translation, ribosomal structure and biogenesis (J)				
Cluster0018	At5g54760	1.18E−35	Translation initiation factor 1 (eIF-1/SUI1)	5
Cluster0031	At2g25210	8.19E−15	60S ribosomal protein L39	4
Cluster0033	At5g52650	8.17E−42	40S ribosomal protein S10	3
Cluster0038	At5g60390	4.89E−123	Translation elongation factor 1A/Tu	3
Cluster0054	At5g39850	1.05E−85	40S ribosomal protein S4	2
Cluster0068	At3g04400	1.60E−74	60S ribosomal protein L14/L17/L23	2
Cluster0076	At3g61110	4.38E−43	40S ribosomal protein S27	2
Cluster0086	At1g36240	3.32E−50	60S ribosomal protein L30	2
Cluster0091	At5g36230	3.86E−86	eIF4-gamma/eIF5/eIF2-epsilon domain-containing protein	2
Cluster0094	At3g55280	7.42E−42	60S ribosomal protein L23	2
TZ09G05	At1g36730	4.53E−63	Translation initiation factor 5 (eIF-5)	
TZ02G05	At1g17220	1.02E−121	Mitochondrial translation initiation factor 2	
TZ06H08	At1g57720	3.25E−87	Translation elongation factor 1B-gamma	
TZ10F05	At1g06220	3.04E−94	Protein similar to splicing factor Snu114	
TZ11F05	At1g25350	7.26E−57	Glutaminyl-tRNA synthetase	
TZ08D09	At3g48930	1.45E−70	40S ribosomal protein S11	
TZ01C03	At4g00100	6.49E−38	40S ribosomal protein S13	
TZ10E11	At3g52580	5.18E−58	40S ribosomal protein S14	
TZ01G07	At2g09990	3.81E−68	40S ribosomal protein S16	
TZ12B09	At5g04800	9.25E−56	40S ribosomal protein S17	
TZ11G09	At3g02080	1.41E−59	40S ribosomal protein S19	
TZ08D06	At5g60670	4.09E−36	40S ribosomal protein S2	
TZ10G01	At5g27700	5.31E−34	40S ribosomal protein S21	
TZ02A06	At5g02960	2.79E−73	40S ribosomal protein S23	
TZ02G07	At3g04920	1.26E−52	40S ribosomal protein S24	
TZ08E02	At4g39200	4.26E−29	40S ribosomal protein S25	
TZ07E05	At2g40590	2.04E−29	40S ribosomal protein S26	
TZ05G10	At3g44010	5.44E−25	40S ribosomal protein S29	
TZ07C08	At5g06360	7.39E−43	40S ribosomal protein S8e family protein	
TZ12A11	At5g24510	1.46E−18	60S acidic ribosomal protein P1	
TZ11F10	At2g27530	1.23E−93	60S ribosomal protein L10a	
TZ06A07	At3g07110	5.79E−97	60S ribosomal protein L13a	
TZ11B07	At4g27090	1.13E−47	60S ribosomal protein L14	
TZ03A12	At1g70600	1.22E−71	60S ribosomal protein L15/L27	
TZ07C12	At3g05590	4.31E−71	60S ribosomal protein L18	
TZ09F11	At1g02780	2.58E−75	60S ribosomal protein L19	
TZ04C03	At3g05560	1.11E−29	60S ribosomal protein L22	
TZ07F12	At3g49910	4.61E−42	60S ribosomal protein L26	
TZ05H11	At3g22230	1.26E−55	60S ribosomal protein L27	
TZ04A08	At1g26880	7.80E−42	60S ribosomal protein L34	
TZ01B02	At3g53740	8.63E−36	60S ribosomal protein L36	
TZ07C07	At3g59540	2.41E−30	60S ribosomal protein L38	
RNA processing and modification (A)				
Cluster0017	C7J5Z7	6.00 E−45	Splicing factor 3B subunit (*Oryza sativa*)	6
TZ11C12	At2g47420	9.67E−103	Ribosomal RNA adenine dimethylase	

Table 2 continued

Sequence ID	Locus code/ accession no.	*E* value	Annotation	Number of EST
TZ01H12	At5g16750	2.60E−82	WD40-repeat-containing subunit of the 18S rRNA processing complex	
TZ01G08	At3g62840	6.61E−47	Small nuclear ribonucleoprotein Sm core protein	
TZ01F05	At2g18740	1.33E−39	Small nuclear ribonucleoprotein E	
TZ09G03	At2g23930	3.95E−32	Small nuclear ribonucleoprotein G	
TZ12A07	At2g20490	6.48E−13	Nucleolar RNA-binding protein Nop10p	
TZ04A03	At5g51300	9.61E−40	Splicing factor 1 (RNA recognition motif superfamily)	
TZ07A11	At5g61220	8.05E−19	Complex 1 family protein/LVR family protein	
Transcription (K)				
TZ05B05	At5g09920	2.07E−47	RNA polymerase II subunit 4	
TZ02A11	At3g16980	5.65E−58	RNA polymerase II subunit 9	
TZ11G08	At5g19510	8.13E−56	Elongation factor 1 beta/delta chain	
TZ02E12	At3g58680	1.36E−54	Transcription factor MBF1	
TZ04B05	At4g39410	1.42E−36	WRKY superfamily transcription factors	
TZ03D11	At4g00990	3.56E−36	Transcription factor containing jumonji (jmjC) domain	
TZ02B06	At3g19080	1.31E−31	RNA polymerase I transcription factor UAF	
TZ03G07	At4g17500	3.37E−21	AP2 domain transcription factor	
TZ11E11	At3g26790	1.94E−22	APETALA2 and related transcription factors	
TZ06C11	At1g58110	1.73E−20	bZIP family transcription factor	
TZ09E12	At4g03250	5.41E−10	Homeobox-leucine zipper family protein	
TZ01G04	At1g79220	2.36E−07	Mitochondrial transcription termination factor family protein	
Replication, recombination and repair (L)				
TZ03E08	At3g26410	3.07E−62	Putative RNA methylase	
TZ09A12	At2g13840	1.30E−54	PHP domain-containing protein	
Chromatin structure and dynamics (B)				
Cluster0100	At5g59970	4.95E−40	Histone H4	2
TZ07D07	At5g54640	1.43E−42	Histone H2A	
TZ10D08	At1g07790	4.00E−43	Histone H2B	
Cellular processes and signaling				
Cell cycle control, cell division, chromosome partitioning (D)				
TZ06B10	At5g08290	3.15E−77	Mitosis protein DIM1	
TZ03H06	At5g21010	7.48E−47	Speckle-type poxvirus and zinc finger protein	
TZ10A07	At3g05540	5.32E−42	Microtubule-binding protein	
TZ11A09	At5g03460	5.58E−37	M-phase inducer phosphatase	
TZ12A10	At3g54670	3.64E−13	Cohesin (structural maintenance of chromosome protein 1)	
Defense mechanisms (V)				
TZ07F10	At1g65690	7.74E−41	Harpin (bacterial protein)-induced protein	
TZ11E07	At3g52180	7.13E−35	Plant-specific protein phosphatase	
TZ11D06	At3g25510	2.15E−16	Disease resistance protein (TIR-NBS-LRR class)	
Signal transduction mechanisms (T)				
Cluster0022	At3g43810	7.27E−76	Calmodulin	5
Cluster0045	Hs4502481	5.18E−16	Peripheral-type benzodiazepine receptor (*Homo sapiens*)	3
Cluster0074	At1g08650	6.22E−68	Ca^{2+}/calmodulin dependent protein kinase	2
TZ12A06	At5g01020	2.76E−63	Serine/threonine protein kinase	
TZ07F09	At4g03080	3.47E−88	Serine/threonine specific protein phosphatase	
TZ12A01	At1g07380	1.20E−61	Ceramidase family protein	
TZ07A10	At1g11280	4.28E−41	OTU (ovarian tumor)-like cysteine protease	
TZ08E03	At3g25570	2.54E−35	*S*-adenosylmethionine decarboxylase	

Table 2 continued

Sequence ID	Locus code/ accession no.	*E* value	Annotation	Number of EST
TZ11A06	At2g35050	3.04E−34	Protein kinase family protein	
TZ03A09	At2g23070	8.41E−31	Casein kinase II alpha chain	
TZ05H06	At3g11930	2.25E−20	Universal stress protein family protein	
TZ08H05	At1g63700	6.67E−08	MEKK and related serine/threonine protein kinases	
TZ08A05	Q6ESH1	4.00E−06	Putative MAP3K α 1 protein kinase (*Oryza sativa*)	
Cell wall/membrane/envelope biogenesis (M)				
TZ09B07	At1g04920	2.96E−66	Sucrose-phosphate synthase (glycosyltransferase)	
TZ09A03	At2g33430	4.97E−65	Plastid developmental protein DAG	
Cytoskeleton (Z)				
Cluster0042	At5g56600	1.27E−54	Profilin3 (actin monomer-binding protein)	3
Cluster0072	At5g09810	2.68E−81	Actin and related proteins (ACT2, ACT7)	2
TZ12B06	At5g12250	1.03E−72	β Tubulin (TUB6)	
TZ06B09	At2g47500	7.26E−64	Kinesin (microtubule motor)	
TZ11H07	At1g01750	1.07E−53	Actin depolymerizing factor (ADF11)	
TZ01A11	At2g46600	3.54E−31	Ca^{2+}binding protein (centrin/caltractin), EF-Hand superfamily protein	
TZ06C03	At1g54560	1.09E−11	Myosin-like proteins (myosin class V heavy chain)	
Intracellular trafficking, secretion, and vesicular transport (U)				
Cluster0049	At3g62290	3.34E−99	GTP-binding ADP-ribosylation factor Arf1	2
TZ03E07	At4g29160	8.13E−75	Protein involved in vesicle-mediated transport (SNF7.1)	
TZ03D09	At1g56590	1.14E−11	Clathrin-associated protein medium chain	
TZ12C01	At3g13772	1.39E−54	Endosomal membrane proteins (EMP70)	
TZ01H06	At3g20920	2.08E−40	Membrane component of ER protein translocation complex	
TZ06A11	At4g32150	5.54E−76	Synaptobrevin/VAMP-like protein (v-SNARE)	
TZ02F11	At3g23660	1.66E−93	Vesicle coat complex COPII, subunit SEC23	
TZ04A06	At3g27325	1.63E−24	Negative regulator of COPII vesicle formation	
TZ08F09	At4g30600	3.91E−110	Ran GTPase, small soluble GTP-binding protein	
TZ03G03	At2g30060	7.08E−41	Ran (GTPase)-binding protein 1b (RanBP1b)	
TZ11B09	At2g47970	2.56E−99	NPL4 (nuclear protein localization 4) family protein	
TZ07D02	At1g27970	7.98E−25	Nuclear transport factor	2
TZ12B01	At2g37410	1.02E−49	Mitochondrial import inner membrane translocase, subunit TIM17	
TZ10H12	At3g46560	4.84E−34	Mitochondrial import inner membrane translocase, subunit TIM9	
TZ04F09	At5g43970	4.64E−21	Mitochondrial outer membrane translocase complex, subunit TOM22	
TZ10A03	At1g73030	5.40E−55	Vacuolar assembly/sorting protein (VPS46.2)	
TZ03G04	At3g20170	7.61E−14	Armadillo repeat protein VAC8 required for vacuole fusion	
Posttranslational modification, protein turnover, chaperones (O)				
Cluster0013	At3g51030	5.15E−36	Cytosolic thioredoxin	7
Cluster0020	At1g48130	4.02E−76	1-Cysteine peroxiredoxin family of antioxidants	5
Cluster0029	At1g65980	4.78E−67	Alkyl hydroperoxide reductase/peroxiredoxin	4
Cluster0052	At2g21130	9.26E−75	Cyclophilin type peptidyl-prolyl cis–trans isomerase	2
Cluster0058	At3g12490	1.92E−54	Cysteine proteinase inhibitor B (cystatin B)	2
Cluster0064	At3g12580	1.08E−121	Molecular chaperones HSP70	2
Cluster0101	At3g44110	6.79E−121	Molecular chaperone (DnaJ superfamily)	2
TZ05H07	At4g04910	1.09E−64	AAA^{+}-type ATPase	
TZ07H07	At2g34560	1.02E−29	Katanin (AAA^{+}-type ATPase)	
TZ01H10	At5g45390	1.25E−66	ATP-dependent Clp protease, proteolytic subunit	
TZ01E05	At5g40280	7.71E−48	β Subunit of farnesyltransferase	

Table 2 continued

Sequence ID	Locus code/ accession no.	*E* value	Annotation	Number of EST
TZ09C11	At1g56340	3.94E−106	Calreticulin	
TZ10B08	At2g21430	2.25E−84	Papain family cysteine protease	
TZ08G08	At5g40200	5.49E−35	Serine protease	
TZ05G03	At1g06200	1.32E−55	Mitochondrial inner membrane protease, subunit IMP2	
TZ12A09	At5g51750	1.14E−56	Subtilisin-like proteinase	
TZ07F06	At1g72280	1.76E−65	ER membrane-associated oxidoreductin involved in disulfide bond formation	
TZ12A12	At1g12390	1.43E−42	ER vesicle integral membrane protein	
TZ05A07	At5g16400	5.45E−49	Chloroplastic thioredoxin	
TZ12B11	At5g40370	3.41E−36	Glutaredoxin and related proteins	
TZ02A05	At1g78370	7.37E−63	Glutathione transferase	
TZ08A10	At1g31340	1.32E−63	Ubiquitin-like proteins	
TZ02G11	At4g37880	2.00E−86	Predicted E3 ubiquitin ligase	
TZ01E02	At5g53300	3.69E−81	Ubiquitin-protein ligase	
TZ09A07	At5g42190	9.75E−23	SCF ubiquitin ligase, Skp1 component	
TZ08H06	At5g66140	1.23E−107	20S proteasome, α5 subunit	
TZ12A02	At4g31300	1.37E−44	20S proteasome, β subunit PBA1	
TZ03D05	At5g40580	3.19E−15	20S proteasome, β subunit PBB2	
TZ02E10	At5g23540	1.61E−111	26S proteasome regulatory complex, subunit RPN11	
TZ05F09	At5g05780	3.92E−138	26S proteasome regulatory complex, subunit RPN8	
TZ03F03	At3g53970	8.27E−18	Proteasome formation inhibitor PI31	
TZ12B07	At2g04030	3.29E−35	Molecular chaperone (HSP90 family)	
TZ12B05	At4g12400	3.69E−81	Molecular co-chaperone STI1	
Metabolism				
Energy production and conversion (C)				
Cluster0056	At2g34420	3.01E−111	Chlorophyll *A*/*B* binding protein	2
TZ05D04	At1g49380	2.29E−97	Cytochrome *c* biogenesis protein family	
TZ07D11	At5g50850	4.21E−95	Pyruvate dehydrogenase E1 β subunit	
TZ10H04	At1g64200	2.66E−74	Vacuolar H^+ATPase V1 sector, subunit E	
TZ02G06	At4g22220	3.73E−61	Iron binding protein involved in Fe–S cluster formation	
TZ07A09	At1g79750	1.53E−46	$NADP^+$-dependent malic enzyme	
TZ10E07	At1g23740	9.75E−18	Zinc-binding oxidoreductase	
Carbohydrate transport and metabolism (G)				
Cluster0035	At3g53420	1.57E−123	Aquaporin (major intrinsic protein family)	3
Cluster0099	At1g13440	9.48E−118	Glyceraldehyde-3-phosphate dehydrogenase	2
TZ07G04	At1g09780	1.17E−120	Phosphoglycerate mutase	
TZ08D11	At2g36460	3.01E−120	Fructose-bisphosphate aldolase	
TZ02H09	At1g24280	3.20E−65	Glucose-6-phosphate dehydrogenase	
TZ11C11	At1g12240	1.67E−43	Invertase (β-fructofuranosidase)	
TZ10D04	At5g54800	2.93E−41	Glucose-6-phosphate/phosphate and phosphoenolpyruvate/ phosphate antiporter	
TZ07G07	At5g39790	1.78E−25	Protein involved in Snf1 protein kinase complex assembly	
TZ01H08	At5g09600	2.15E−09	Succinate dehydrogenase, cytochrome b subunit	
Amino acid transport and metabolism (E)				
TZ11G07	At4g38220	5.19E−71	Aminoacylase ACY1 and related metalloexopeptidases	
TZ06H01	At4g36760	3.29E−65	Xaa-Pro aminopeptidase	
TZ12B08	At5g40780	6.08E−41	Amino acid transporters	

Table 2 continued

Sequence ID	Locus code/ accession no.	*E* value	Annotation	Number of EST
Nucleotide transport and metabolism (F)				
TZ08G12	At4g09320	2.91E−70	Nucleoside diphosphate kinase	
TZ05D02	At3g12670	1.47E−36	CTP synthase (UTP-ammonia lyase)	
Coenzyme transport and metabolism (H)				
Cluster0070	At3g17390	3.71E−129	Methionine transferase	2
Lipid transport and metabolism (I)				
Cluster0002	At3g18570	4.39E−17	Oleosin	12
Cluster0009	At3g51600	3.81E−12	Lipid transfer protein type 1	7
Cluster0075	At5g02100	2.02E−88	Oxysterol-binding protein	2
TZ01A04	At2g46210	1.33E−90	Delta 6-fatty acid desaturase/delta-8 sphingolipid desaturase	
TZ05F04	At3g11210	1.34E−60	GDSL-motif lipase	
TZ12B02	At5g43280	4.68E−59	Enoyl-CoA isomerase	
TZ02H05	At2g32260	5.93E−45	Phosphorylcholine transferase/cholinephosphate cytidylyltransferase	
TZ11B06	At4g10950	1.76E−38	GDSL-motif lipase/hydrolase family protein	
TZ02F02	At2g17370	1.16E−29	3-Hydroxy-3-methylglutaryl-CoA reductase	
TZ11H09	Hs11464975	2.68E−19	Fatty acyl-CoA elongase/polyunsaturated fatty acid specific elongation enzyme (*Homo sapiens*)	
Inorganic ion transport and metabolism (P)				
Cluster0087	At1g71050	1.26E−37	Copper chaperone	2
TZ07B01	At5g57490	2.54E−78	Porin-type voltage-dependent anion channel	
TZ12A04	At1g10830	1.43E−49	Sodium symporter-related	
TZ08B12	At1g55730	3.81E−42	Ca^{2+}/H^{+} antiporter (CAX2 family)	
TZ10H07	At2g19110	5.08E−29	Cation transport ATPase	
TZ09G09	At1g08830	5.94E−38	Cu^{2+}/Zn^{2+} superoxide dismutase CSD1	
Secondary metabolites biosynthesis, transport and catabolism (Q)				
Cluster0060	At3g50660	1.88E−26	Steroid 22α-hydroxylase (CYP90B1)	2
TZ07F02	At4g22010	2.19E−90	Multicopper oxidases	
TZ06E07	At1g77120	8.01E−30	Alcohol dehydrogenase, class III	
General function prediction only (R)				
Cluster0012	At1g24020	9.76E−07	Bet v 1 allergen family protein	7
Cluster0014	A9XKN0	2.00E−27	Dehydrin 4 (*Picea abies*)	6
Cluster0019	Q8H0M6	5.00E−11	Dehydrin 4 (*Pinus sylvestris*)	5
Cluster0023	At1g71695	2.20E−71	Peroxidase 12 (PER12)	4
Cluster0037	At3g05890	6.46E−19	Cold-inducible protein	3
Cluster0069	At1g20690	3.78E−20	HMG box-containing protein	2
Cluster0073	At5g60790	2.81E−58	Predicted transporter (ABC superfamily)	2
TZ03A02	At5g11900	1.19E−20	Eukaryotic translation initiation factor SUI1 family protein	
TZ11H04	At2g41060	2.54E−50	RNA-binding (RRM/RBD/RNP motifs) family protein	
TZ12B03	At4g16830	7.77E−11	Nuclear RNA-binding protein	
TZ08B02	At3g03710	8.64E−16	Predicted RNA-binding polyribonucleotide nucleotidyltransferase	
TZ02F05	At3g08640	5.28E−49	Alphavirus core protein family	
TZ09G11	At1g51200	1.25E−35	Zinc finger (AN1-like) family protein	
TZ08E11	At5g54470	5.55E−28	Zinc finger (B-box type) family protein	
TZ01F02	Q10R22	5.00E−22	Zinc finger, RING-type domain-containing protein (*Oryza sativa*)	
TZ07D10	At1g21570	2.44E−17	Zinc finger (CCCH-type) family protein	
TZ06F01	At1g70150	2.67E−08	Zinc finger (MYND type) family protein	
TZ12A08	At2g37020	2.05E−22	Transposable element gene; copia-like retrotransposon family	
TZ09F05	At5g48050	4.95E−07	Similar to retrotransposon protein	

Table 2 continued

Sequence ID	Locus code/ accession no.	*E* value	Annotation	Number of EST
TZ01B11	At3g05880	1.33E−13	Stress responsive protein	
TZ10B12	At1g32370	5.83E−14	Basic protein involved in tobamovirus multiplication in planta	
TZ06D09	At3g05370	2.21E−14	Receptor like protein containing leucine-rich repeat	
TZ01E01	At2g39940	1.35E−43	Protein containing Leu-rich repeats and a degenerate F-box motif.	
TZ03F10	At5g46170	1.00E−22	F-box family protein	
TZ09D12	At5g09390	1.77E−10	Protein containing polyproline-binding GYF domain	
TZ07A06	At1g20575	2.80E−33	Dolichyl-phosphate β-D-mannosyltransferase	
TZ08C08	At3g51800	1.78E−18	Metallopeptidase M24 family protein	
TZ03H12	At3g61050	1.05E−94	Ca^{2+} dependent lipid-binding protein	
TZ12A03	At5g45370	4.62E−47	Nodulin-related/integral membrane family protein	
TZ10F11	At2g16980	5.75E−41	Tetracycline transporter	
TZ02D04	At2g41380	1.24E−49	Methyltransferase	
TZ06D10	At4g35160	9.12E−25	*O*-Methyltransferase family 2 protein	
TZ04G03	At3g51450	1.12E−47	Strictosidine synthase family protein (alkaloid synthase)	
TZ12C02	At3g51680	6.89E−38	Short-chain dehydrogenase/reductase family protein	

A capital letter in parentheses after the category name indicates a category symbol corresponding to Fig. 1

Annotations for sequences without species names originated from *Arabidopsis thaliana*

Table 3 Abundant transcripts found in the transition zone of *C. japonica* in November

Sequence ID	Annotation	Locus code/UniProt entry	Number of EST
Cluster0002	Oleosin	At3g18570	12
Cluster0009	Lipid transfer protein	At3g51600	7
Cluster0010	Function-unknown protein (*Caenorhabditis elegans*)	Q9NES7	7
Cluster0011	Dormancy-associated protein	At2g33830	7
Cluster0012	Bet v 1 allergen	At1g24020	7
Cluster0013	Cytosolic thioredoxin	At3g51030	7
Cluster0014	Dehydrin 4 (*Picea abies*)	A9XKN0	6
Cluster0017	Splicing factor 3B subunit (*Oryza sativa*)	C7J5Z7	6
Cluster0018	eIF-1	At5g54760	5
Cluster0019	Dehydrin 4 (*Pinus sylvestris*)	Q8H0M6	5
Cluster0020	1-Cysteine peroxiredoxin	At1g48130	5
Cluster0022	Calmodulin	At3g43810	5

Clusters containing five ESTs and more are listed. Clusters for rRNA encoded by nuclear, chloroplast or mitochondrial genome were excluded

Annotations for sequences without species names originated from *Arabidopsis thaliana*

Discussion

To improve understanding of the heartwood formation process, we collected ESTs from the TZ of *C. japonica* in the dormant season when heartwood formation is considered to be underway. The ESTs were assembled into non-redundant sequences, and annotated according to the KOG functional classification. Sequences included in the two most abundant categories, "posttranslational modification, protein turnover, chaperones (O)" and "translation, ribosomal structure and biogenesis (J)", occupied a quarter of the classified sequences (Fig. 1). In ESTs previously collected from different organs of woody plants and functionally classified using the KOG [30, 38, 39], categories "O" and "J" were ranked in the top four in terms of abundance. Thus, it appears that the translational regulation of protein synthesis is maintained in the TZ in November as well as in other organs. Meanwhile, sequences that did not match known sequences in the KOG and UniProt databases accounted for 39% of all unique sequences. This

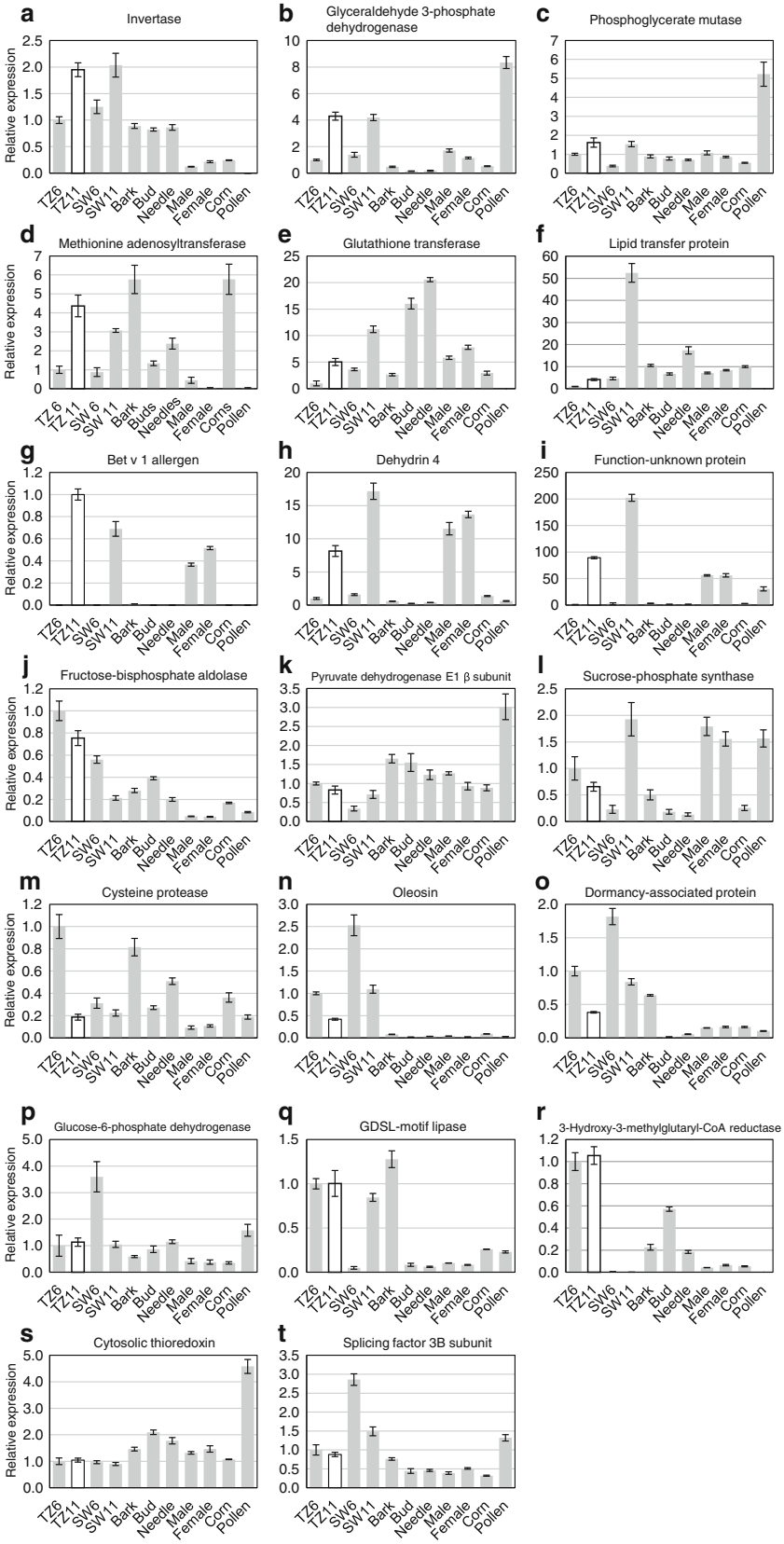

Fig. 2 Expression patterns of selected genes in the various organs of *C. japonica*. The values were normalized to the transcript levels of a gene for the translation initiation factor. *White bars* represent the relative gene expression levels in TZ in November. *Gray bars* show those in the other organs. *Error bars* indicate SD of three technical replicates. Expression levels were expressed as a ratio relative to the level of TZ6 except for the Bet v 1 allergen gene because its expression level in TZ6 was below the limit of detection. TZ6, transition zone (June); TZ11, transition zone (November); SW6, sapwood (June); SW11, sapwood (November); Bark, inner bark (June); Bud, leaf buds (May); Needle, needles (May); Male, male strobili (February); Female, female strobili (February); Corn, young corns (June); Pollen, pollen (March)

proportion is higher than that of *C. japonica* male strobili (30%) [30], suggesting that the expression of genes specific to the TZ would occur.

We examined the expression levels of the selected 20 genes, focusing on the differences between June and November in the TZ (Fig. 2). As a result, nine genes were up-regulated in the TZ in November as compared to June. The expression level of an invertase gene was approximately twice as high in November compared to in June (Fig. 2a). Among the three types (vacuolar, cell wall bound, and neutral) of invertases, which have been distinguished based on their subcellular localization, pH optima, etc. [40], the deduced invertase in this study was assumed to be classified as a vacuolar type. Invertases hydrolyze sucrose into glucose and fructose, and vacuolar types are involved in supplying these hexoses, not only for nutrients but also for osmoregulation [40]. The resulting hexoses in the TZ may function as osmoregulators in addition to entering glycolysis and/or the pentose phosphate cycle. The increased activity of a neutral invertase has been reported in the TZ of *Robinia pseudoacacia* in the winter season [34]. Two genes for enzymes in the glycolytic pathway, glyceraldehyde-3-phosphate dehydrogenase and phosphoglycerate mutase, were also up-regulated in November (Fig. 2b, c). The expression of these genes may contribute to the production of energy for heartwood formation, and the supply of material for the biosynthesis of extractives through the catalysis of sugars. The expression level of a gene for fructose-bisphosphate aldolase, which catalyzes an earlier step in the glycolytic pathway, was maintained in TZ in November (Fig. 2j).

Methionine adenosyltransferase is responsible for the synthesis of *S*-adenosylmethionine (SAM). Secondary metabolites synthesized during heartwood formation often possess methyl groups in their structures. In methylation reactions, SAM is utilized as a primary methyl donor. However, the major heartwood extractives of *C. japonica*, norlignans and diterpenes, are rarely methylated [41]. SAM is also the precursor for the synthesis of ethylene. In *Pinus radiata*, ethylene was produced in the TZ, and its quantity was larger than that from sapwood during the winter period [42]. In addition, increased ethylene production was reported for the TZ of *Eucalyptus tereticornis* [43] and *Juglans nigra* [44] early in the dormant season. Thus, increased demands for SAM might lead to the expression of a gene for methionine adenosyltransferase in the TZ in November (Fig. 2d).

A gene for a tau-class glutathione transferase (GST) was up-regulated in November (Fig. 2e). The tau-class GSTs are specific to plants, and are involved in several biological processes such as the detoxification of xenobiotics, reduction of oxidative stress, and regulation of flavonoid biosynthesis and trafficking [45–47]. A tau-class GST from barley (*Hordeum vulgare*) was induced in leaves during senescence and in response to low temperature [48]. In maize, a tau-class GST serves as a carrier protein for the vacuolar sequestration of anthocyanins in carnels [49, 50]. Thus, it is possible that the GST in the TZ functions to eliminate oxidative stress and/or transport secondary metabolites during heartwood formation.

Lipid transfer proteins have been isolated from a variety of plants. LTPs are capable of transferring lipids between membranes in vitro [51]. The physiological roles of LTPs are obscure, but they appear to be involved in antimicrobial defense, cuticle biosynthesis, cell wall loosening, anther development, etc., depending on isoforms [51]. The expression of a gene for an LTP increased in the TZ in November compared to in June (Fig. 2f). The LTP examined in this study shows no amino acid sequence similarity to an LTP isolated from *C. japonica* pollen [52], while sharing high sequence similarity (60% identity) with an LTP isolated from *Cycas revoluta* seeds, which was reported to possess weak antimicrobial activity [53]. The role of LTPs in the TZ is still unclear at present. The significantly higher level of gene expression in the sapwood in November implies an important role of the LTP in it.

Bet v 1 is the major pollen allergen of *Betula verrucosa*, and belongs to a family of plant pathogenesis-related protein 10 (PR-10). PR-10 proteins have been found in many plant species but their physiological roles are still obscure [54]. It has been reported that Bet v 1 is able to bind several biological molecules such as fatty acids, flavonoids and cytokinins [55], and Bet v 1 appears to function as the carrier of brassinosteroids [56]. The Bet v 1-like protein found in this study has significant amino acid sequence similarity (53–55% identitiy) to PR-10 proteins from *Pinus monticola* needles, which accumulated not only in response to wounding but also during winter [57]. It could be speculated that the Bet v 1-like protein play a functional role in the defense or developmental regulation of the TZ.

Dehydrins are known to be produced in many plant species in response to environmental stresses [58, 59]. Dehydrin proteins or transcripts have been found from buds, leaves, inner bark, fruits and xylem [59]. The general roles of dehydrins are thought to be chaperones or cryoprotectants. The expression of a dehydrin 4 gene increased in the TZ in November compared to that in June (Fig. 2h). Although functions of dehydrins in the TZ remain unclear, they may play a role in adaptation to winter cold and/or lower moisture content.

The predicted protein encoded by Cluster0010 showed little similarity to proteins whose functions are estimated, thus this was referred to as a function-unknown protein. The protein has a lower degree of similarity to dehydrins, and possesses the lysine-rich 15 amino acid sequence (the K-segment) that is characteristic to dehydrins [58].

The expression pattern of this gene in the various organs resembled that of the dehydrin 4 encoded by Cluster0014 (Fig. 2h), except that the expression level in both the TZ and sapwood in June was extremely low (Fig. 2i). From these findings, we assume that the Cluster0010 encodes a dehydrin-related protein and the protein could have similar functions with the dehydrins mentioned above especially in the winter season in the TZ.

In the xylem of *C. japonica*, secondary metabolites called norlignans are produced especially during heartwood formation [60]. Previously, we isolated ESTs for enzymes, which are presumed to be involved in biosynthesis of norlignans, from the drying sapwood of *C. japonica* [27]. These ESTs, however, were not found in the present study. One possible reason is that the number of ESTs collected from the TZ was insufficient for including transcripts from the candidate genes for norlignan biosynthesis. The cDNA library in this study was not normalized, while the previous library was constructed by suppression subtractive hybridization, and thus it contained predominantly expressed genes in the drying sapwood accumulating a norlignan. Another possibility is that the genes isolated in the previous study are actually not involved in the biosynthesis of norlignans. This seems unlikely because very few sequences were contained in the "secondary metabolite biosynthesis, transport and catabolism (Q)" category (Fig. 1; Table 2).

Heartwood formation has been considered as a form of programmed cell death (PCD) [7, 8]. A key factor in PCD is proteases [61]. PCD in animal cells are known to be executed by caspases, which belong to a class of specific cysteine proteases. In plants, functionally similar proteases, including papain-type cysteine proteases, are involved in some cases of PCD [62, 63]. We found an EST predicted to encode a papain-type cysteine protease. The predicted protease showed significant sequence similarity to those isolated from the leaves and petals undergoing PCD during senescence [64–67], suggesting that the protease may be involved in PCD in the TZ as heartwood formation is thought to be a senescence process. The expression level of the corresponding gene for the protease in the TZ was about 5 times higher in June than that in November (Fig. 2m). Taking into consideration this expression pattern and a lack of putative norlignan biosynthetic genes, the gene expression associated with heartwood formation might be initiated before the onset of cytological changes observed in late summer [32], and the expression of genes for enzymes related to heartwood formation ceased in November.

To further understand the mechanism of heartwood formation at a molecular level, the annual changes in gene expression in the TZ must be investigated with simultaneous observation of cytological changes in parenchyma cells, and analysis of extractives in the TZ.

Conclusions

We collected 1029 ESTs from a cDNA library constructed from the TZ of *C. japonica* in November. The ESTs were assembled into 744 unique sequences. Putative functions were assigned to 291 nuclear-encoded sequences, and they were grouped into 21 functional categories. We also revealed that the expression levels of nine genes for enzymes involved in glycolysis, invertase, methionine adenosyltransferase, glutathione transferase, a LTP, Bet v 1 allergen, a dehydrin and a function-unknown protein were higher in November than in June in the TZ. These genes may play roles in maintaining the TZ function and/or forming heartwood. This study has provided the first large-scale EST information from the TZ of conifers, which will be useful for understanding the physiological processes in the TZ at a molecular level.

References

1. Plomion C, Leprovost G, Stokes A (2001) Wood formation in trees. Plant Physiol 127(4):1513–1523
2. Taylor AM, Gartner BL, Morrell JJ (2002) Heartwood formation and natural durability: a review. Wood Fiber Sci 34(4):587–611
3. Hillis WE (1999) The formation of heartwood and its extractives. An overview. In: Romeo JT (ed) Phytochemicals in human health protection nutrition and plant defense. Kluwer/Plenum Publishers, New York, pp 215–253
4. Bamber RK, Fukazawa K (1985) Sapwood and heartwood: a review. For Abstr 46(9):567–580
5. Higuchi T (1997) Regulation of heartwood formation biochemistry and molecular biology of wood. Springer, Berlin, pp 299–307
6. Bamber RK (1976) Heartwood, its function and formation. Wood Sci Technol 10(1):1–8
7. Spicer R (2005) Senescence in secondary xylem: heartwood formation as an active development program. In: Holbrook NM, Zwieniecki MA (eds) Vascular transport in plants. Elsevier, Burlington, pp 457–475
8. Magel EA (2000) Biochemistry and physiology of heartwood formation. In: Savidge RA, Barnett JR, Napier R (eds) Cell and molecular biology of wood formation. BIOS Scientific Publishers, Oxford, pp 363–376
9. Allona I, Quinn M, Shoop E, Swope K, St Cyr S, Carlis J, Riedl J, Retzel E, Campbell MM, Sederoff R, Whetten RW (1998) Analysis of xylem formation in pine by cDNA sequencing. Proc Natl Acad Sci 95(16):9693–9698
10. Kirst M, Johnson AF, Baucom C, Ulrich E, Hubbard K, Staggs R, Paule C, Retzel E, Whetten R, Sederoff R (2003) Apparent homology of expressed genes from wood-forming tissues of loblolly pine (*Pinus taeda* L.) with *Arabidopsis thaliana*. Pro Natl Acad Sci 100(12):7383–7388
11. Yang SH, van Zyl L, No EG, Loopstra CA (2004) Microarray analysis of genes preferentially expressed in differentiating xylem of loblolly pine (*Pinus taeda*). Plant Sci 166(5):1185–1195
12. Egertsdotter U, van Zyl LM, MacKay J, Peter G, Kirst M, Clark C, Whetten R, Sederoff R (2004) Gene expression during formation of earlywood and latewood in loblolly pine: expression profiles of 350 genes. Plant Biol 6(6):654–663

13. Sterky F, Regan S, Karlsson J, Hertzberg M, Rohde A, Holmberg A, Amini B, Bhalerao R, Larsson M, Villarroel R, Van Montagu M, Sandberg G, Olsson O, Teeri TT, Boerjan W, Gustafsson P, Uhlen M, Sundberg B, Lundeberg J (1998) Gene discovery in the wood-forming tissues of poplar: analysis of 5, 692 expressed sequence tags. Proc Natl Acad Sci 95(22):13330–13335
14. Hertzberg M, Aspeborg H, Schrader J, Andersson A, Erlandsson R, Blomqvist K, Bhalerao R, Uhlen M, Teeri TT, Lundeberg J, Sundberg B, Nilsson P, Sandberg G (2001) A transcriptional roadmap to wood formation. Proc Natl Acad Sci 98(25):14732–14737
15. Israelsson M, Eriksson ME, Hertzberg M, Aspeborg H, Nilsson P, Moritz T (2003) Changes in gene expression in the wood-forming tissue of transgenic hybrid aspen with increased secondary growth. Plant Mol Biol 52(4):893–903
16. Déjardin A, Leplé JC, Lesage-Descauses MC, Costa G, Pilate G (2004) Expressed sequence tags from poplar wood tissues—a comparative analysis from multiple libraries. Plant Biol 6(1): 55–64
17. Paux E, Tamasloukht M, Ladouce N, Sivadon P, Grima-Pettenati J (2004) Identification of genes preferentially expressed during wood formation in *Eucalyptus*. Plant Mol Biol 55(2):263–280
18. Foucart C, Paux E, Ladouce N, San-Clemente H, Grima-Pettenati J, Sivadon P (2006) Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*. New Phytol 170(4):739–752
19. Ranik M, Creux NM, Myburg AA (2006) Within-tree transcriptome profiling in wood-forming tissues of a fast-growing *Eucalyptus* tree. Tree Physiol 26(3):365–375
20. Rengel D, Clemente HS, Servant F, Ladouce N, Paux E, Wincker P, Couloux A, Sivadon P, Grima-Pettenati J (2009) A new genomic resource dedicated to wood formation in *Eucalyptus*. BMC Plant Biol 9:36
21. Suzuki S, Suda K, Sakurai N, Ogata Y, Hattori T, Suzuki H, Shibata D, Umezawa T (2010) Analysis of expressed sequence tags in developing secondary xylem and shoot of *Acacia mangium*. J Wood Sci 57(1):40–46
22. Yang JM, Park S, Kamdem DP, Keathley DE, Retzel E, Paule C, Kapur V, Han KH (2003) Novel gene expression profiles define the metabolic and physiological processes characteristic of wood and its extractive formation in a hardwood tree species, *Robinia pseudoacacia*. Plant Mol Biol 52(5):935–956
23. Yang JM, Kamdem DP, Keathley DE, Han KH (2004) Seasonal changes in gene expression at the sapwood-heartwood transition zone of black locust (*Robinia pseudoacacia*) revealed by cDNA microarray analysis. Tree Physiol 24(4):461–472
24. Beritognolo I, Magel E, Abdel-Latif A, Charpentier JP, Jay-Allemand C, Breton C (2002) Expression of genes encoding chalcone synthase, flavanone 3-hydroxylase and dihydroflavonol 4-reductase correlates with flavanol accumulation during heartwood formation in *Juglans nigra*. Tree Physiol 22(5):291–300
25. Huang ZL, Tsai CJ, Harding SA, Meilan R, Woeste K (2010) A cross-species transcriptional profile analysis of heartwood formation in black walnut. Plant Mol Biol Rep 28(2):222–230
26. Kurotori S (1954) Studies on the "White rings" of Sugi (1) (in Japanese). J Jpn For Soc 36(1):15–19
27. Yoshida K, Nishiguchi M, Futamura N, Nanjo T (2007) Expressed sequence tags from *Cryptomeria japonica* sapwood during the drying process. Tree Physiol 27(1):1–9
28. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA (2003) The COG database: an updated version includes eukaryotes. BMC Bioinform 4:41
29. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang HZ, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LSL (2005) The universal protein resource (UniProt). Nucleic Acids Res 33:D154–D159
30. Futamura N, Totoki Y, Toyoda A, Igasaki T, Nanjo T, Seki M, Sakaki Y, Mari A, Shinozaki K, Shinohara K (2008) Characterization of expressed sequence tags from a full-length enriched cDNA library of *Cryptomeria japonica* male strobili. BMC Genomics 9:383
31. Sone T, Komiyama N, Shimizu K, Kusakabe T, Morikubo K, Kino K (1994) Cloning and sequencing of cDNA coding for *Cry j* I, a major allergen of Japanese cedar pollen. Biochem Biophys Res Commun 199(2):619–625
32. Nobuchi T, Kuroda K, Iwata R, Harada H (1982) Cytological study of the seasonal features of heartwood formation of Sugi (*Cryptomeria japonica* D. Don). Mokuzai Gakkaishi 28(11):669–676
33. Hillinger C, Höll W, Ziegler H (1996) Lipids and lipolytic enzymes in the trunkwood of *Robinia pseudoacacia* L. during heartwood formation II. Radial distribution of lipases and phospholipases. Trees 10(6):376–381
34. Hauch S, Magel E (1998) Extractable activities and protein content of sucrose-phosphate synthase, sucrose synthase and neutral invertase in trunk tissues of *Robinia pseudoacacia* L. are related to cambial wood production and heartwood formation. Planta 207(2):266–274
35. Magel EA, Hillinger C, Wagner T, Höll W (2001) Oxidative pentose phosphate pathway and pyridine nucleotides in relation to heartwood formation in *Robinia pseudoacacia* L. Phytochemistry 57(7):1061–1068
36. Hsieh K, Huang AHC (2004) Endoplasmic reticulum, oleosins, and oils in seeds and tapetum cells. Plant Physiol 136(3):3427–3434
37. Shain L, Mackay JFG (1973) Seasonal fluctuation in respiration of aging xylem in relation to heartwood formation in *Pinus radiata*. Can J Bot 51(4):737–741
38. Nanjo T, Futamura N, Nishiguchi M, Igasaki T, Shinozaki K, Shinohara K (2004) Characterization of full-length enriched expressed sequence tags of stress-treated poplar leaves. Plant Cell Physiol 45(12):1738–1748
39. Futamura N, Ujino-Ihara T, Nishiguchi M, Kanamori H, Yoshimura K, Sakaguchi M, Shinohara K (2006) Analysis of expressed sequence tags from *Cryptomeria japonica* pollen reveals novel pollen-specific transcripts. Tree Physiol 26(12):1517–1528
40. Roitsch T, González MC (2004) Function and regulation of plant invertases: sweet sensations. Trends Plant Sci 9(12):606–613
41. Umezawa T (2000) Chemistry of extractives. In: Hon DNS, Shiraisi N (eds) Wood and cellulosic chemistry, 2nd edn. Marcel Dekker, New York, pp 213–241
42. Shain L, Hillis WE (1973) Ethylene production in xylem of *Pinus radiata* in relation to heartwood formation. Can J Bot 51(7):1331–1335
43. Nelson ND (1978) Xylem ethylene, phenol-oxidizing enzymes, and nitrogen and heartwood formation in walnut and cherry. Can J Bot 56(6):626–634
44. Hillis WE (1975) Ethylene and extraneous material formation in woody tissues. Phytochem 14(12):2559–2562
45. Dixon DP, Lapthorn A, Edwards R (2002) Plant glutathione transferases. Genome Biol 3:reviews3004
46. Frova C (2003) The plant glutathione transferase gene family: genomic structure, functions, expression and evolution. Physiol Plant 119(4):469–479
47. Basantani M, Srivastava A (2007) Plant glutathione transferases—a decade falls short. Can J Bot 85(5):443–456
48. Kunieda T, Fujiwara T, Amano T, Shioi Y (2005) Molecular cloning and characterization of a senescence-induced tau-class glutathione *S*-transferase from barley leaves. Plant Cell Physiol 46(9):1540–1548

49. Alfenito MR, Souer E, Goodman CD, Buell R, Mol J, Koes R, Walbot V (1998) Functional complementation of anthocyanin sequestration in the vacuole by widely divergent glutathione *S*-transferases. Plant Cell 10(7):1135–1149
50. Zhao J, Dixon RA (2010) The 'ins' and 'outs' of flavonoid transport. Trends Plant Sci 15(2):72–80
51. Yeats TH, Rose JKC (2008) The biochemistry and biology of extracellular plant lipid-transfer proteins (LTPs). Protein Sci 17(2):191–198
52. Ibrahim ARN, Kawamoto S, Nishimura M, Pak S, Aki T, Diaz-Perales A, Salcedo G, Asturias JA, Hayashi T, Ono K (2010) A new lipid transfer protein homolog identified as an IgE-binding antigen from Japanese cedar pollen. Biosci Biotechnol Biochem 74(3):504–509
53. Yokoyama S, Kato K, Koba A, Minami Y, Watanabe K, Yagi F (2008) Purification, characterization, and sequencing of antimicrobial peptides, Cy-AMP1, Cy-AMP2, and Cy-AMP3, from the Cycad (*Cycas revoluta*) seeds. Peptides 29(12):2110–2117
54. Liu J–J, Ekramoddoullah AKM (2006) The family 10 of plant pathogenesis-related proteins: their structure, regulation, and function in response to biotic and abiotic stresses. Physiol Mol Plant Pathol 68(1–3):3–13
55. Mogensen JE, Wimmer R, Larsen JN, Spangfort MD, Otzen DE (2002) The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands. J Biol Chem 277(26): 23684–23692
56. Markovic-Housley Z, Degano M, Lamba D, von Roepenack-Lahaye E, Clemens S, Susani M, Ferreira F, Scheiner O, Breiteneder H (2003) Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier. J Mol Biol 325(1):123–133
57. Liu J–J, Ekramoddoullah AKM, Yu X (2003) Differential expression of multiple PR10 proteins in western white pine following wounding, fungal infection and cold-hardening. Physiol Plant 119(4):544–553
58. Campbell SA, Close TJ (1997) Dehydrins: genes, proteins, and associations with phenotypic traits. New Phytol 137(1):61–74
59. Kosová K, Vítámvás P, Prášil IT (2007) The role of dehydrins in plant response to cold. Biol Plant 51(4):601–617
60. Suzuki S, Umezawa T (2007) Biosynthesis of lignans and norlignans. J Wood Sci 53(4):273–284
61. Solomon M, Belenghi B, Delledonne M, Menachem E, Levine A (1999) The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. Plant Cell 11(3):431–444
62. Martinez M, Diaz I (2008) The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. BMC Evol Biol 8:198
63. Woltering EJ (2010) Death proteases: alive and kicking. Trends Plant Sci 15(4):185–188
64. Eason JR, Ryan DJ, Pinkney TT, O'Donoghue EM (2002) Programmed cell death during flower senescence: isolation and characterization of cysteine proteinases from *Sandersonia aurantiaca*. Funct Plant Biol 29(2):1055–1064
65. Scharrenberg C, Falk J, Quast S, Haussuehl K, Humbeck K, Krupinska K (2003) Isolation of senescence-related cDNAs from flag leaves of field grown barley plants. Physiol Plantarum 118(2):278–288
66. Yamada T, Ichimura K, Kanekatsu M, van Doorn WG (2007) Gene expression in opening and senescing petals of morning glory (*Ipomoea nil*) flowers. Plant Cell Rep 26(6):823–835
67. Chen HJ, Su CT, Lin CH, Huang GJ, Lin YH (2010) Expression of sweet potato cysteine protease SPCP2 altered developmental characteristics and stress responses in transgenic *Arabidopsis* plants. J Plant Physiol 167(10):838–847