Tree Physiology 27, 1–9
© 2007 Heron Publishing—Victoria, Canada

Expressed sequence tags from *Cryptomeria japonica* sapwood during the drying process

KAZUMASA YOSHIDA,[1,2] MITSURU NISHIGUCHI,[1] NORIHIRO FUTAMURA[1] and TOKIHIKO NANJO[1]

[1] *Department of Molecular and Cell Biology, Forestry and Forest Products Research Institute, Tsukuba, Ibaraki 305-8687, Japan*

[2] *Corresponding author (ykazu@ffpri.affrc.go.jp)*

Received October 24, 2005; accepted April 1, 2006; published online October 2, 2006

Summary Secondary metabolites called norlignans are produced in the xylem of *Cryptomeria japonica* D. Don. Several norlignans have roles in the defense of sapwood against microbial invasion and in the coloration of heartwood. The biosynthetic pathway of norlignans is largely unknown. Norlignans have been reported to accumulate in the sapwood during the drying of *C. japonica* logs. To search for genes encoding enzymes that catalyze the synthesis of norlignans, we carried out suppression subtractive hybridization using the fresh sapwood of a felled log and the drying sapwood in which a norlignan, agatharesinol, accumulated. A total of 1050 expressed sequence tags were obtained from the subtracted cDNA library, and these were assembled into 146 contigs and 361 singletons. Of these 507 unique sequences, 263 were functionally classified into 12 categories. "Metabolism" was the largest category, with 23% (61) of classified sequences. Twenty-six sequences that encode 16 enzymes were assigned to "secondary metabolism." Expression analysis of 15 genes related to "secondary metabolism" revealed that 12 of these genes had transcripts that were induced during the sapwood drying process. Of the 12 genes, 10 encoded enzymes that use aromatic compounds as substrates. In addition, 58 sequences representing 22 defense-related proteins were found. Our subtraction library should be a useful source for isolating genes encoding proteins involved in secondary metabolism including norlignan biosynthesis and defense in *C. japonica* xylem.

Keywords: EST, RT-PCR, suppression subtractive hybridization.

Introduction

Some conifers and herbaceous plants contain norlignans—a group of secondary metabolites with a diphenylpentane skeleton (Umezawa 2000; Figure 1). Several norlignans in sugi trees (*Cryptomeria japonica* D. Don, Taxodiaceae) are thought to be involved in the defense of sapwood against microbial invasion and in the formation of the reddish color of heartwood. (*E*)-Hinokiresinol is produced in discolored sapwood of *C. japonica* formed after invasion by the sugi bark borer, *Semanotus japonicus*. It has antifungal activity against *Fusarium oxysporum, F. solani* and *Macrophoma sugi* (Yamada et al. 1988). Sequirin-C and cryptoresinol are associated with coloration of the heartwood (Takahashi 1981). Some norlignans from herbs are used in folk medicine in Asian and African countries (Nicoletti et al. 1992). Despite the physiological and economic importance of norlignans, their biosynthesis pathway is largely unknown.

(*Z*)-Hinokiresinol (nyasol) and (*E*)-hinokiresinol were shown to be synthesized from 4-coumaryl alcohol and 4-coumaroyl CoA, and from 4-coumaryl 4-coumarate, respectively, by enzyme preparations from fungal-elicited asparagus cultured cells or *C. japonica* cultured cells (Suzuki et al. 2002,

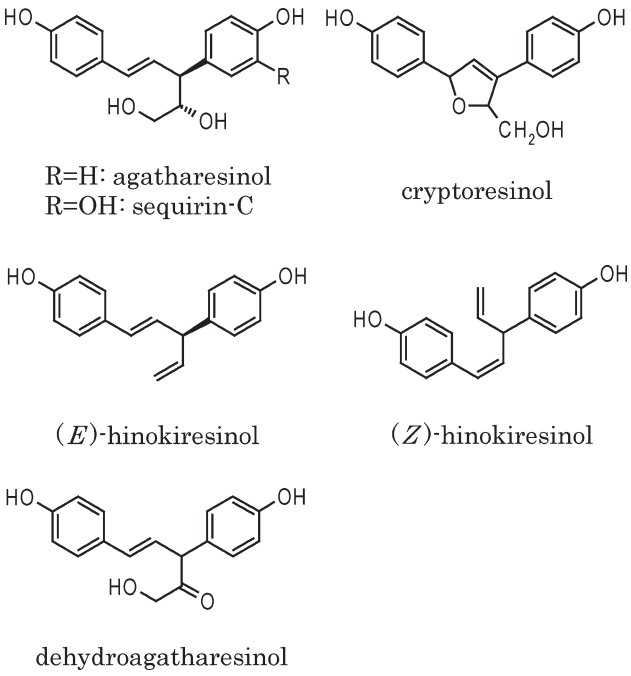

Figure 1. Structures of norlignans referred to in this paper. Dehydroagatharesinol is a putative compound in the proposed biosynthetic pathway of agatharesinol.

Downloaded from http://treephys.oxfordjournals.org/ by guest on October 30, 2011

2004). However, little is known about the biosynthetic pathway of other norlignans in conifers. Norlignans are thought to be synthesized mainly in autumn and winter in the transition zone between sapwood and heartwood during heartwood formation (Nobuchi et al. 1982, Taylor et al. 2002).

Agatharesinol, a major heartwood norlignan, accumulates in sapwood during the drying process of *C. japonica* logs (Ohashi et al. 1988, 1990, Yoshida et al. 2004). An inhibitor of phenylalanine ammonia lyase has been shown to prevent synthesis of agatharesinol in *C. japonica* logs (Ohashi et al. 1991). In addition, ^{2}H-labeled L-phenylalanine and ^{13}C-labeled *trans*-cinnamic acid absorbed into sapwood sticks of *C. japonica* were incorporated into agatharesinol during drying (Imai et al. 2006). These results strongly suggest that the accumulation of agatharesinol during the drying of sapwood is a result of de novo sysnthesis.

By utilizing the synthesis of agatharesinol during sapwood drying as an experimental system, we intended to isolate cDNAs encoding enzymes for norlignan biosynthesis by applying a suppression subtractive hybridization to the sapwood when norlignans are generated. The expressed sequence tags obtained were functionally categorized, and expression of genes related to secondary metabolism was examined by reverse transcription-polymerase chain reaction.

Materials and methods

Plant material

A 15-year-old *C. japonica* tree growing in the nursery of the Forestry and Forest Products Research Institute (Tsukuba) was felled in May 2001. A trunk section 0.8 m long was excised at a height of 1 m above ground and stored at 20–25 °C. A disk 10-mm thick was cut from the bottom end of the log and discarded preceding the collection of 60-mm-thick sample disks on Days 0, 10, 20, 41 and 70 post-harvest. The inner sapwood was isolated from the disks, cut into small pieces, frozen in liquid nitrogen, and stored at –80 °C.

RNA isolation

Total RNA was prepared from the inner sapwood according to Bugos et al. (1995) with the following modifications. The concentration of 2-mercaptoethanol in the homogenization buffer was increased to 288 mM, the amount of starting material was reduced to 5 g or less, and the volume of all solutions was reduced to half that of the original method. Before precipitation with isopropanol, the aqueous phase was extracted with an equal volume of chloroform.

Suppression subtractive hybridization (SSH)

The cDNA was synthesized from 1 µg of total RNA from the sapwood on Days 0 and 41, respectively, using the SMART PCR cDNA synthesis kit (BD Biosciences Clontech, Palo Alto, CA). The SSH was performed with the PCR-select cDNA subtraction kit (BD Biosciences Clontech) using the cDNA from the sapwood on Day 41 as a tester and that on Day 0 as a driver. The subtracted cDNAs were cloned into the pGEM-T vector (Promega, Madison, WI) and transformed into *Escherichia coli* strain DH10B.

Sequencing and data analysis

Sequences of subtracted cDNA inserts were determined with an ABI 3730 DNA analyzer (Applied Biosystems, Foster City, CA) at Shimadzu Biotech (Kyoto, Japan). 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, Ann Arbor, MI) with the parameters of minimum overlap = 40, minimum match = 95%. Sequences of rRNA, and chloroplast and mitochondrial DNA were identified by the BLASTN search against *Populus* rRNA sequences (AF174629, AF206999, AF479118, AJ006440), *Arabidopsis* mitochondrial genome (NC001284) and *Populus* chloroplast genome (http://genome.ornl.gov/poplar_chloroplast/), and these were removed. The remaining sequences were searched locally against the Munich Information Center for Protein Sequences (MIPS), the *A. thaliana* Genome Database (MATDB) (download website, ftp://ftpmips.gsf.de/cress/) and the Protein Information Resource (PIR) non-redundant reference protein database (NREF) (download website, ftp://ftp.pir.georgetown.edu/pir_databases/nref/) by BLASTX. Sequences with an expectation (*E*)-value of $<10^{-5}$ were considered to have significant homology. Resulting sequences were classified manually following the MIPS functional catalogue (FunCat) (http://mips.gsf.de/projects/funcat). The EST sequences have been submitted to DNA Data Bank of Japan (DDBJ) under Accession Numbers BB940565 to BB941614.

RT-PCR

Before reverse transcription, total RNA was treated with RQ1 RNase-free DNase (Promega) to remove contaminating DNA. First-strand cDNA was synthesized from 1 µg of RNA using SuperScript III first-strand synthesis (Invitrogen, Carlsbad, CA) with an oligo (dT) primer. The reaction mixture for PCR comprised 1 µl of first-strand cDNA, 12.5 µl of GoTaq green master mix (Promega) and 0.5 µM gene-specific primers in a total volume of 25 µl. The primers are listed in Table 1. Amplification involved an initial denaturation step at 95 °C for 90 s followed by 23 to 30 cycles of denaturation at 95 °C for 30 s, annealing at 58 °C for 30 s, and extension at 72 °C for 40 s. Expression of the gene for translation initiation factor was used as a control. The PCR products were separated on 0.7% agarose gels and stained with ethidium bromide.

Results

Collection of ESTs from the subtracted cDNA library

Norlignans were not detected in the sapwood of *C. japonica* on the day of felling. During the drying of *C. japonica* logs,

Downloaded from http://treephys.oxfordjournals.org/ by guest on October 30, 2011

Table 1. Primers used for RT-PCR analysis. Abbreviation: TIF = translation initiation factor.

Sequence ID	Forward primer sequence	Reverse primer sequence
Cj02B02	5′-GGAAAGACAGAAGCGATCG-3′	5′-CTAACACTAGACAAGGTTACATAG-3′
Cj02F02	5′-CAGTGAAATGATGGACGAGG-3′	5′-CCCTGATCTTGTGATCCAG-3′
Cj03A03	5′-GCAACCAGTGTGCAACTTAG-3′	5′-CACATATTCGTCGTTGACATCG-3′
Cj05E02	5′-TGGCACCCTTGAAATCAGAG-3′	5′-CACTAACACCACCATACATGC-3′
Cj07H11	5′-GGTGAAGTCCATTCTGCC-3′	5′-CGGTTCTTAACACAGTGAAAGC-3′
Cj09E12	5′-CGAGAAGCTCATGGTACAG-3′	5′-GATGGCTTGTTCTAAGCATCTG-3′
Cj10B09	5′-TGCACACTTGGAGGCTTC-3′	5′-CTGCTTTTGGCATTGTCAAACC-3′
Cj11C06	5′-CTCCATGACGCTAAGCTTG-3′	5′-GTACTACCTTCCATAGAAGAAACG-3′
Cj12A17	5′-CACTGTTGAATTCGGAACAATCC-3′	5′-GCAGACTGAATCGTTTCCAG-3′
Contig0019	5′-CCAAGATGATGAGGTTCCTG-3′	5′-GATAAGTAGATAACTCATGACCAGC-3′
Contig0030	5′-GTGCTGCCTCTGTTGAG-3′	5′-GTTGCGTTATAGCCGTCAATG-3′
Contig0089	5′-CTCACGCTAAACTCCTCAAC-3′	5′-TCTGCAACGCTCCCTTTC-3′
Contig0092	5′-CTCAATCCTTACATGCAGAGAATAG-3′	5′-CTCCTTGACTATTGCTTGCAAG-3′
Contig0098	5′-AATACAGCTGATGCACGAGC-3′	5′-CATGAGATTGGTGTTGGAGG-3′
Contig0276	5′-GTTAGTGGGAACTCAAAGATG-3′	5′-GTTTCCAATGCGTTCTTTGAGG-3′
TIF	5′-GTCAGATCTAGACGTTCAGATTC-3′	5′-GTCCTTCTTCACAATTCCAGC-3′

agatharesinol accumulated in the sapwood. The agatharesinol content significantly increased from Day 20 to Day 41 (Yoshida et al. 2004). The subtracted cDNA library was constructed using cDNA from the sapwood on Day 0 as a driver and on Day 41 as a tester. Over 1000 clones were picked and sequenced to generate 1050 ESTs with a mean length of 352 bp. After sequence assembly, 361 ESTs were singletons, whereas the other 689 ESTs clustered into 146 contigs containing from 2 to 61 ESTs.

Functional classification of ESTs

Three sequences that originated from chloroplast or mitochondrial genomes were removed from the 507 unique sequences (singleton and contig consensus). The remaining 504 sequences were annotated by BLASTX with an *E*-value of $< 10^{-5}$. MATDB was chosen first for annotation because genes in MATDB have their own FunCat numbers based on the MIPS FunCat scheme; thus sequences homologous to *Arabidopsis* genes could be readily functionally classified. Annotation was given to 150 sequences with similarity to known function proteins in MATDB. The sequences with an *E*-value of $\geq 10^{-5}$, and those assigned to "classification not yet clear-cut" and "unclassified proteins," were further searched against PIR-NREF, and 54 additional sequences were annotated. Finally, 204 sequences were annotated to known sequences with putative functions, 59 were similar to known genes but had unknown functions and 241 had no similarity to sequences in the databases. The 263 sequences were categorized according to the MIPS FunCat scheme, and functionally grouped into 12 categories (Figure 2).

"Metabolism" was the largest putative function category, and included 61 sequences that encoded 41 different proteins (Table 2). The number of predicted proteins was fewer than the number of sequences, because sequences annotated as the same protein were grouped, and the group was represented by the sequence with the lowest *E*-value. Putative proteins encoded by the abundant ESTs were monophenol monooxygenase, leucoanthocyanidin dioxygenase, glutaredoxin and β-galactosidase. A major subcategory was "secondary metabolism," which contained 26 sequences encoding 16 enzymes. Of the enzymes, 11 were oxidoreductases, two were transferases, two were lyases, and one was a hydrolase.

The second largest category was "defense." It included 58 sequences representing 22 proteins (Table 3). This group contained a large number of ESTs assigned to germin-like proteins (GLPs). The ESTs annotated as GLPs amounted to 84, accounting for 8.0% of entire ESTs. The ESTs encoding β-1,3-glucanase, class I chitinase, pathogenesis-related (PR) protein PR-1 and PR-5 were also abundant.

Other sequences were assigned to nine categories with putative function (Table 4). A large proportion of protein-synthesis-related ESTs encoded ribosomal proteins. The ESTs for the translation initiation factor Sui1 homolog were rich in this category. Energy-related ESTs included an abundance of 1,4-benzoquinone reductase. The "protein fate" category was represented by peptidylprolyl *cis-trans* isomerase, cysteine proteinase and proteins of the ubiquitin system. Several sequences

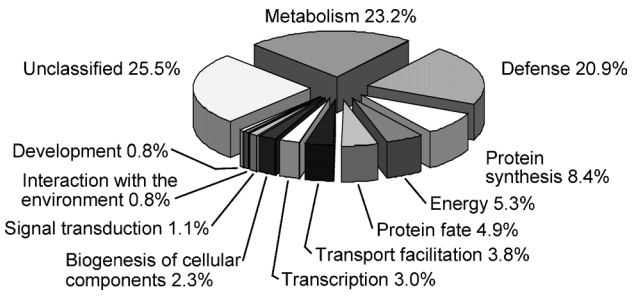

Figure 2. Functional classification of unique sequences with significant homology with the sequences in MATDB and PIR-NREF according to the MIPS classification scheme.

Downloaded from http://treephys.oxfordjournals.org/ by guest on October 30, 2011

Table 2. Sequences assigned to the "metabolism" category. Annotations for sequences without species names originated from *Arabidopsis thaliana*.

Sequence ID	Annotation	*E*-value	Accession No.	No. of ESTs
Amino acid metabolism				
Contig0068	*S*-Adenosylmethionine synthetase-like protein	8.50E-30	At3g17390	3
Contig0016	Hydroxyacylglutathione hydrolase	4.58E-10	At3g10850	2
Cj10D12	4-Hydroxyphenylpyruvate dioxygenase	3.59E-28	At1g06570	
Cj08A04	Acetylornithine transaminase-like protein	5.48E-20	At1g80600	
Cj06F08	Homocysteine S-methyltransferase	7.22E-30	At3g63250	
Nitrogen and sulfur metabolism				
Cj06F06	Nitrilase-like protein	5.75E-17	At5g12040	
Nucleotide metabolism				
Contig0008	Putative glutaredoxin (*Oryza sativa*)	2.08E-11	NF01802982	9
Cj07B06	Dihydroorotate dehydrogenase-like-protein	8.56E-10	At3g17810	
Cj04B07	Putative tetrahydrofolate synthase	1.36E-25	At3g12290	
C-compound and carbohydrate metabolism				
Contig0007	β-Galactosidase-like protein	2.50E-20	At2g28470	6
Contig0035	Glucan endo-1,3-alpha-glucosidase (*Schizosaccharomyces pombe*)	7.48E-13	NF01763982	4
Contig0226	Carbonic anhydrase (CAH1)	8.44E-12	At1g08080	3
Contig0258	Formate dehydrogenase	2.93E-43	At5g14780	2
Contig0177	Short-chain type dehydrogenase/reductase (*Picea abies*)	6.27E-43	NF00538246	2
Contig0087	Endoxyloglucan transferase	4.17E-10	At1g32170	2
Cj12A16	Aldose 1-epimerase-like protein	1.63E-27	At3g17940	
Cj04A12	β-Fructofuranosidase	1.00E-27	At1g55120	
Cj05C10	β-Glucosidase-like protein	2.00E-29	At1g61820	
Cj01C11	*myo*-Inositol-1-phosphate synthase-like protein	2.60E-65	At5g10170	
Lipid, fatty-acid and isoprenoid metabolism				
Contig0014	Allene oxide synthase	1.64E-21	At5g42650	4
Cj04F06	*E*-α-Bisabolene synthase (*Abies grandis*)	4.94E-06	NF00378556	
Cj10B06	γ-Humulene synthase (*Abies grandis*)	4.97E-12	NF00378562	
Cj03F12	Steroid 22-alpha-hydroxylase	1.23E-20	At3g50660	
Metabolism of vitamins, cofactors, and prosthetic groups				
Cj03H04	GTP cyclohydrolase II	1.60E-38	At5g64300	
Cj04B01	Monodehydroascorbate reductase (NADH) -like protein	8.92E-18	At5g03630	
Secondary metabolism				
Contig0032	Monophenol monooxygenase (*Trifolium pratense*)	1.84E-07	NF01195790	36
Contig0019	Leucoanthocyanidin dioxygenase-like protein	5.67E-29	At1g77330	18
Contig0092	Cytochrome P450-like protein	2.94E-22	At4g15350	3
Contig0089	Alcohol dehydrogenase (ATA1)	2.56E-25	At3g42960	3
Contig0276	Mandelonitrile lyase-like protein	1.82E-25	At1g72970	2
Cj07H11	Reticuline oxidase-like protein	3.00E-27	At1g26400	
Cj03A03	Catechol oxidase-like protein (*Vicia faba*)	2.01E-10	NF00733990	
Cj12A17	Tyrosine 3-monooxygenase-like protein (*Anguilla anguilla*)	1.70E-11	NF00061765	
Cj11C06	*trans*-Cinnamate 4-monooxygenase-like protein	4.37E-19	At2g30490	
Cj05E02	4,5-DOPA dioxygenase extradiol-like protein	1.85E-35	NF00796464	
Cj10B09	Flavonoid 3′-monooxygenase-like protein	7.04E-29	At4g31940	
Cj02B02	Isochorismatase-like protein	6.48E-25	At5g23230	
Cj02F02	Glutathione transferase-like protein	6.67E-30	At2g29450	
Cj09E12	Hydroxylase/oxygenase (CTF2B)	2.11E-15	At2g29720	
Cj11G10	Caffeic acid *O*-methyltransferase, putative	4.51E-19	At5g54160	
Cj01A05	1-Aminocyclopropane-1-carboxylate synthase-like protein	1.00E-08	At3g61510	

encoding adenosine nucleotide translocators were prominent in the "transport facilitation" category. The "transcription" category included an AP2 domain containing protein along with bZIP and WRKY transcription factors. The category "biogenesis of cellular components" comprised tubulin, actin, histone H2A and an expansin-like protein. A senescence-associated protein was the only protein assigned to the "development" category.

Downloaded from http://treephys.oxfordjournals.org/ by guest on October 30, 2011

Table 3. Sequences assigned to the "defense" category. Annotations for sequences without species names originated from *Arabidopsis thaliana*.

Sequence ID	Annotation	*E*-value	Accession No.	No. of ESTs
Disease, virulence and defense				
Contig0001	Germin-like protein	2.13E-65	At5g39190	61
Contig0002	β-1,3-Glucanase	9.14E-39	At3g57270	14
Contig0088	Putative class I chitinase (*Cryptomeria japonica*)	1.42E-60	NF01565751	12
Contig0009	PR protein 5 precursor, putative	6.62E-30	At1g19320	9
Contig0096	PR protein 1, putative	4.47E-50	At3g19690	8
Contig0293	Acidic endo-chitinase	4.22E-62	At5g24090	2
Contig0316	Antimicrobial peptide 4 (*Pinus sylvestris*)	6.16E-30	NF01555571	2
Contig0189	Osmotin precursor-like protein	2.60E-50	At4g11650	2
Contig0266	Metallothionein (*Xerophyta humilis*)	4.41E-12	NF01836177	2
Cj05C11	Putative PR-10 protein (*Picea glauca*)	5.09E-37	NF00538391	
Cj02C09	NtPRp 27 (PR-17 protein) (*Nicotiana tabacum*)	6.41E-41	NF00216846	
Cj02G05	Class IV chitinase	1.87E-44	At3g54420	
Cj03F10	Peroxidase-like protein	2.92E-37	At2g41480	
Cj09G01	Activated disease resistance (ADR) 1	4.41E-15	At1g33560	
Cj07D04	Putative dirigent protein	1.58E-08	At1g64160	
Cj09A02	Disease resistance-like protein	9.38E-10	At3g53240	
Cj12A48	Nematode resistance protein, putative	1.54E-26	At2g40000	
CJ09E02	Putative disease resistance response protein	3.18E-08	At5g42500	
Cj01D09	Putative hypersensitive response protein	1.74E-07	At3g11650	
Cj05B02	Glutathione peroxidase (*Zantedeschia aethiopica*)	2.25E-06	NF00224441	
Cj01E02	Putative DnaJ protein	9.15E-08	At2g17880	
Cj10D04	Low-temperature and salt-responsive protein	2.12E-20	At3g05880	

Expression patterns of genes of secondary metabolism

To determine whether the isolated ESTs encoding enzymes involved in secondary metabolism participate in the biosynthesis of norlignans, the expression patterns of genes for those ESTs were examined during sapwood drying. Although the 1,4-benzoquinone reductase-like protein did not belong to the "metabolism" category, it was included in this examination because it is associated with the metabolism of aromatic compounds. The 15 genes for which primers were successfully designed were subjected to RT-PCR analysis (Figure 3). These genes were divided into three groups according to their expression patterns. The first group contained four genes encoding reticuline oxidase-like protein, catechol oxidase-like protein, mandelonitrile lyase-like protein and monophenol monooxygenase, which were not expressed on Days 0 and 10, but were expressed on Days 20 and 41. In the second group, genes were not expressed on Day 0, but were expressed on Days 10, 20 and 41. This group included eight genes for tyrosine 3-monooxygenase-like protein, *trans*-cinnamate 4-monooxygenase-like protein, 1,4-benzoquinone reductase-like protein, 4,5-DOPA dioxigenase extradiol-like protein, flavonoid 3′-monooxygenase-like protein, leucoanthocyanidin dioxygenase-like protein, cytochrome P450-like protein and alcohol dehydrogenase. The third group contained three genes encoding isochorismatase-like protein, glutathione transferase-like protein and hydroxylase/oxygenase, which were expressed throughout the period. All 10 genes involved in the metabolism of aromatic compounds were induced during the drying process (Figure 3).

Discussion

To increase our understanding of the norlignan biosynthetic process in *C. japonica*, we applied SSH to drying sapwood in which agatharesinol was synthesized and analyzed the subtracted cDNAs. The "metabolism" and "defense" categories included 59% of the sequences with putative functions. Sequences that did not match any known sequence in MIPSMATDB and PIR-NREF accounted for 48% of the total sequences. This percentage was larger than for the ESTs collected from inner bark (31%), seed cones (41.9%) and pollen cones (19.0%) of *C. japonica* (Ujino-Ihara et al. 2000, 2003). It is suggested that the length and quality of DNA sequences are correlated with the ability to identify similar sequences in public databases (Allona et al. 1998). The mean sequence length of ESTs was 352 bp for this study, whereas it was 450 bp for inner bark, 495 bp for seed cones and 564 bp for pollen cones. Because the sequence quality of our ESTs was high (quality score ≥ 20 at 750 bp), a significant proportion of no-hit sequences may be attributed to the short sequence length owing to digestion of cDNA with the restriction enzyme *Rsa* I in the SSH procedure.

The abundance of ESTs in the "defense" category may be related to the experimental method, which included cutting the trunk wood and leaving the log under non-aseptic conditions. The sample presumably suffered from wounding, dehydration and biotic stresses, thus the living cells in the cut log that might have expressed defense-related genes to protect the trunk wood against these stresses, resulting in many transcripts related to defense in the subtraction library. Annotated se-

Downloaded from http://treephys.oxfordjournals.org/ by guest on October 30, 2011

Table 4. Sequences assigned to functional categories other than "metabolism" and "defense." Annotations for sequences without species names originated from *Arabidopsis thaliana*.

Sequence ID	Annotation	*E*-value	Accession No.	No. of ESTs
Protein synthesis				
Contig0106	Translation initiation factor Sui1 homolog	7.75E-20	At5g54760	6
Contig0031	60S Ribosomal protein L31	7.12E-15	At5g56710	3
Contig0091	60S Ribosomal protein L23	2.24E-43	At2g33370	2
Cj09C10	60s Ribosomal like protein L37	2.49E-17	At1g15250	
Cj01D08	60S Ribosomal protein L23	6.57E-49	At3g04400	
Cj04H05	60S Ribosomal protein L30	8.10E-16	At2g36620	
Cj01E03	40S Ribosomal protein S7-like	5.37E-53	At1g48830	
Cj03G01	Ribosomal protein 48	2.95E-38	At3g22230	
Cj10A11	Ribosomal protein L11-like	2.56E-55	At4g18730	
Cj07H08	Ribosomal protein L13a-like	2.64E-24	At5g48760	
Cj06G01	Ribosomal protein L24-like	1.72E-08	At5g23535	
Cj06D12	Ribosomal protein L32	4.04E-27	At4g18100	
Cj03G08	Ribosomal protein L35a-like	1.61E-19	At1g07070	
Cj01D07	Ribosomal protein L39	6.23E-08	At2g25210	
Cj10E12	Ribosomal protein L5-like	5.04E-23	At3g25520	
Cj01F02	Ribosomal protein L8, putative	5.51E-17	At2g18020	
Cj01A07	Ribosomal protein S30 homolog	3.90E-22	At4g29390	
Cj06F07	Elongation factor 1-alpha	8.97E-22	At1g07940	
Cj11H08	Methionine aminopeptidase-like protein	9.70E-06	At2g44180	
Energy				
Contig0030	1,4-Benzoquinone reductase-like	1.49E-54	At5g54500	9
Contig0090	Phosphoenolpyruvate carboxykinase (ATP)-like protein	2.58E-63	At5g65690	3
Contig0075	Ubiquinol-cytochrome c reductase-like protein	1.10E-20	At5g25450	2
Cj06D03	Phosphoglycerate dehydrogenase-like protein	2.10E-15	At4g34200	
Cj01F10	Putative fructose/tagatose bisphosphate aldolase (*Oryza sativa*)	9.84E-06	NF02002445	
Cj07G05	Strong similarity to transaldolase	1.48E-29	At1g12230	
Cj0_G01	Malate dehydrogenase-like protein	2.75E-14	At5g25880	
Cj07G11	Cytochrome b5-like	4.06E-07	At2g32720	
Cj11E05	Chlorophyll a/b-binding protein CP29	1.90E-16	At2g40100	
Cj09B03	Putative thioredoxin	3.67E-51	At4g04950	
Cj10F08	ATP synthase beta chain (*Picea abies*)	8.05E-31	NF00538274	
Protein fate				
Contig0097	Peptidylprolyl *cis-trans* isomerase, cyclophilin	1.59E-54	At4g38740	3
Contig0278	Cysteine proteinase-like protein	8.74E-18	At5g43060	2
Cj02A01	Strong similarity to polyubiquitins	3.30E-24	At4g05050	
Cj09D07	Ubiquitin-conjugating enzyme E2	2.03E-26	At1g14400	
Cj02E07	Ubiquitin-like protein	2.85E-18	At3g45180	
Cj10C09	Polyubiquitin (ubq10)	2.55E-11	At1g31340	
Cj08A09	Proteinase inhibitor-like protein	6.91E-11	At2g38870	
Cj09F04	Serine-type carboxypeptidase like protein	6.28E-28	At1g15000	
Cj07D12	Calreticulin-like protein	1.12E-54	At1g08450	
Cj02B08	Luminal binding protein; heat shock cognate 71 kDa protein	2.03E-37	At5g28540	
Cj01B03	Prefoldin-like protein	2.23E-22	At3g22480	
Cj05G04	Putative ubiquitin/ribosomal protein CEP52	2.78E-19	At2g36170	
Transport facilitation				
Contig0138	Citrate binding protein (*Solanum tuberosum*)	5.17E-31	NF01158137	3
Contig0193	Adenosine nucleotide translocator	3.67E-57	At5g13490	2
Contig0183	Adenylate translocator	5.41E-11	At3g08580	2
Cj10A07	Putative adenine nucleotide translocase	4.09E-21	At1g15500	
Cj03C02	ADP, ATP carrier-like protein	2.78E-24	At4g28390	
Cj05E11	K+ channel, beta subunit, putative	2.93E-21	At1g04690	
Cj04H11	Putative hexose transporter	3.07E-57	At4g02050	
Cj01D11	Mitochondrion ATPase subunit 9 (*Emericella nidulans*)	1.28E-06	NF00176423	
Cj07C10	Putative ABC transporter	2.36E-15	At2g36380	
Cj05A09	OEP8-like protein	4.91E-22	At4g15800	

Continued on facing page.

Downloaded from http://treephys.oxfordjournals.org/ by guest on October 30, 2011

Table 4 continued. Sequences assigned to functional categories other than "metabolism" and "defense." Annotations for sequences without species names originated from *Arabidopsis thaliana*.

Sequence ID	Annotation	*E*-value	Accession No.	No. of ESTs
Transcription				
Contig0025	AP2 domain containing protein, putative	3.29E-10	At5g21960	5
Contig0187	Splicing factor 3B subunit 5-like protein (*Oryza sativa*)	2.29E-44	NF01667982	2
Cj07C12	bZIP transcription factor ATB2	9.20E-09	At4g34590	
Cj10G07	WRKY transcription factor 31 (WRKY31)	3.79E-12	At4g22070	
Cj04E11	Oligouridylate-specific hnRNP-like protein (RBP45c)	4.23E-11	At4g27000	
Cj08B01	Pirin-like protein	4.24E-09	At2g43120	
Cj05F01	Putative DNA-directed RNA polymerase subunit	1.45E-08	At4g07950	
Cj08E12	Putative small nuclear ribonucleoprotein G (*Oryza sativa*)	1.19E-08	NF00255159	
Biogenesis of cellular components				
Contig0308	Tubulin alpha-6 chain (TUA6)	5.27E-104	At4g14960	3
Cj02G06	Actin 2/7	5.29E-15	At5g09810	
Cj01A10	Histone H2A like-protein	2.99E-35	At1g54690	
Cj08G07	Expansin-like protein (EXLB1)	1.22E-18	At4g17030	
Signal transduction				
Cj04H10	Caltractin-like protein	5.00E-30	At2g46600	
Cj11C02	Ras-related small GTP-binding protein	3.23E-09	At3g16100	
Cj06E01	CBL-interacting protein kinase 20 (CIPK20)	2.14E-38	At5g45820	
Interaction with the environment				
Contig0099	AX1/AUX1-like permease	1.27E-51	At2g38120	
Cj05C02	Phytosulfokine peptide (*Cryptomeria japonica*)	9.74E-10	NF01284868	
Development				
Contig0026	Putative senescence-associated protein 5 (*Oryza sativa*)	1.08E-55	NF01260646	6

quences contained a considerable number of ESTs encoding GLPs, which are widely distributed in plants. Although the functions of most GLPs have yet to be clarified, several GLPs are involved in defense against fungal pathogens in barley and wheat, and against a virus in hot pepper plants (Wei et al. 1998, Schweizer et al. 1999, Park et al. 2004). The sequences encoding transcription factors involved in stress responses, such as an AP2 domain containing protein, bZIP and WRKY, were classified in the "transcription" category.

Agatharesinol is synthesized from L-phenylalanine and *trans*-cinnamic acid in the sapwood of *C. japonica* during the drying process (Imai et al. 2006). However, the biosynthetic pathway after *trans*-cinnamic acid is unclear. In early reports, two possible pathways were proposed. One is the pathway via (*E*)-hinokiresinol (Beracierta and Whiting 1978). (*E*)-Hinokiresinol is a candidate because hydroxylation of a double bond in (*E*)-hinokiresinol can lead to the formation of agatharesinol. The alternative pathway is via dehydroagatharesinol (Erdtman and Harmatha 1979). Because compounds with quinone moieties are expected intermediates in both pathways, the 1,4-benzoquinone reductase found in the library may be involved in the biosynthesis of agatharesinol.

Some norlignans in *C. japonica* have the same carbon framework but with different numbers of hydroxyl groups on the one aromatic ring, such as agatharesinol and sequirin-C (Figure 1). It is postulated that norlignans with two hydroxyl groups could be made by hydroxylation of those with one hydroxyl group. Three sequences for putative enzymes (tyrosine 3-monooxygenase, *trans*-cinnamate 4-oxygenase and flavonoid 3′-monooxygenase) catalyzing hydroxylation of aromatic rings were isolated from the library. Twelve genes for enzymes, including those catalyzing hydroxylation and 1,4-benzoquinone reductase, were not expressed on Day 0, but they were up-regulated toward Day 41 (Figure 3). Similarly, although agatharesinol was barely detectable on Day 0, it appeared on Day 10 and sharply increased from Days 20 to 41 (Yoshida et al. 2004). The strong correlation between expression patterns of genes and changes in agatharesinol content suggests the involvement of enzymes in the de novo synthesis of agatharesinol and other norlignans; however, as a final proof, the enzyme(s) responsible for catalyzing the conversion step must be identified. The expression of 12 out of 15 genes coding for putative enzymes involved in secondary metabolism was induced during the drying process, indicating that cells in sapwood on Day 41 still possessed metabolic activity and that subtractive hybridization was effective in this experiment.

In conclusion, we collected over 1000 ESTs from a subtraction library constructed using fresh and drying sapwood, and demonstrated that a significant proportion of ESTs encoded proteins in the "metabolism" and "defense" categories, indicating that the library is a useful source for the isolation of genes related to "metabolism" and "defense." Future research will focus on the characterization of recombinant enzymes

Downloaded from http://treephys.oxfordjournals.org/ by guest on October 30, 2011

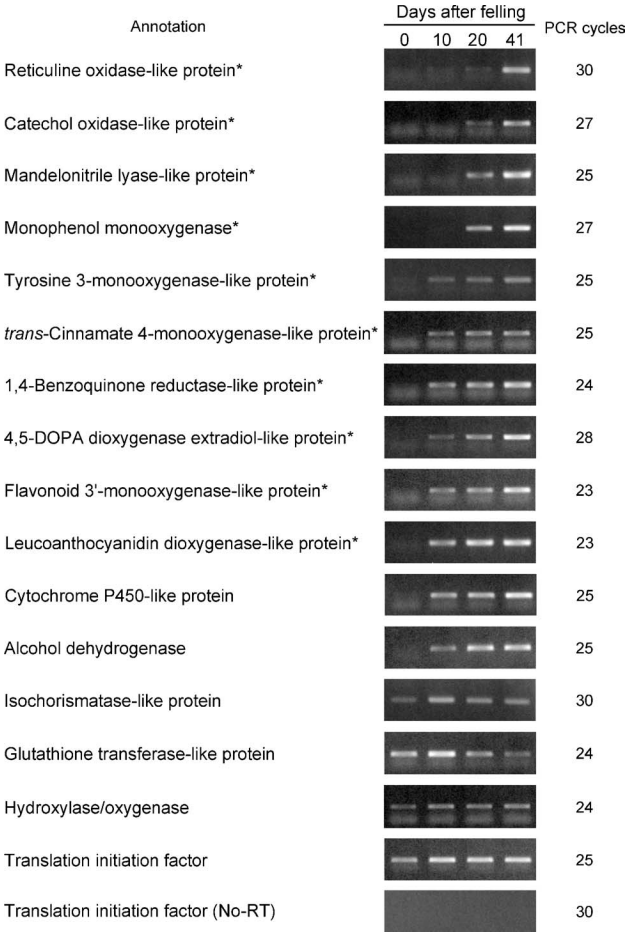

Figure 3. Expression patterns of selected genes during the drying process of *Cryptomeria japonica* sapwood. The gene expression of translation initiation factor was examined as an internal and no-reverse transcription (No-RT) control. Asterisks indicate the enzymes postulated to use aromatic compounds as substrates.

prepared from genes found by SSH to identify the enzymes involved in norlignan biosynthesis.

Acknowledgments

We thank Dr. Tokuko Ujino-Ihara for providing information on the EST sequences of *C. japonica*. This study was supported by a Grant-in-Aid for Scientific Research (No. 17580147) from the Japan Society for the Promotion of Science, and by a research grant (No. 200310) from the Forestry and Forest Products Research Institute.

References

Allona, I., M. Quinn, E. Shoop et al. 1998. Analysis of xylem formation in pine by cDNA sequencing. Proc. Natl. Acad. Sci. USA 95:9693–9698.

Beracierta, A.P. and D.A. Whiting. 1978. Stereoselective total syntheses of the (±)-di-*O*-methyl ethers of agatharesinol, sequirin-A, and hinokiresinol, and of (±)-tri-*O*-methyl-sequirin-E, characteristic norlignans of coniferae. J. Chem. Soc. Perkin Trans. I 1257–1263.

Bugos, R.C., V.L. Chiang, X.-H. Zhang, E.R. Campbell, G.K. Podila and W.H. Campbell. 1995. RNA isolation from plant tissues recalcitrant to extraction in guanidine. Biotechniques 19:734–737.

Erdtman, H. and J. Harmatha. 1979. Phenolic and terpenoid heartwood constituents of *Libocedrus yateensis*. Phytochemistry 18: 1495–1500.

Imai, T., M. Nomura and K. Fukushima. 2006. Evidence for involvement of the phenylpropanoid pathway in the biosynthesis of the norlignan agatharesinol. J. Plant Physiol. 163:483–487.

Nicoletti, M., C. Galeffi, I. Messana and G.B. Marini-Bettolo. 1992. Hypoxidaceae. Medicinal uses and the norlignan constituents. J. Ethnopharmacol. 36:95–101.

Nobuchi, T., K. Kuroda, R. Iwata and H. Harada. 1982. Cytological study of the seasonal features of heartwood formation of Sugi (*Cryptomeria japonica* D. Don). Mokuzai Gakkaishi 28:669–676.

Ohashi, H., T. Goto, T. Imai and M. Yasue. 1988. Accumulation and distribution of extractives in the living stem of *Cryptomeria japonica*. Res. Bull. Fac. Agr. Gifu Univ. 53:301–314.

Ohashi, H., T. Imai, K. Yoshida and M. Yasue. 1990. Characterization of physiological functions of sapwood: fluctuation of extractives in the withering process of Japanese cedar sapwood. Holzforschung 44:79–86.

Ohashi, H., N. Kato, T. Imai and S. Kawai. 1991. Characterization of physiological functions of sapwood: fluctuation of heartwood extractives in the withering process of Japanese cedar sapwood fed an inhibitor of phenylalanine ammonia-lyase. Holzforschung 45: 245–252.

Park, C.-H., J.-M. An, Y.-C. Shin, K.-J. Kim, B.-J. Lee and K.-H. Paek. 2004. Molecular characterization of pepper germin-like protein as the novel PR-16 family of pathogenesis-related proteins isolated during the resistance response to viral and bacterial infection. Planta 219:797–806.

Schweizer, P., A. Christoffel and R. Dudler. 1999. Transient expression of members of the germin-like gene family in epidermal cells of wheat confers disease resistance. Plant J. 20:541–552.

Suzuki, S., T. Nakatsubo, T. Umezawa and M. Shimada. 2002. First in vitro norlignan formation with *Asparagus officinalis* enzyme preparation. Chem. Commun. 1088–1089.

Suzuki, S., M. Yamamura, M. Shimada and T. Umezawa. 2004. A heartwood norlignan, (*E*)-hinokiresinol, is formed from 4-coumaryl 4-coumarate by a *Cryptomeria japonica* enzyme preparation. Chem. Commun. 2838–2839.

Takahashi, K. 1981. Heartwood phenols and their significance to color in *Cryptomeria japonica* D. Don. Mokuzai Gakkaishi 27: 654–657.

Taylor, A.M., B.L. Gartner and J.J. Morrell. 2002. Heartwood formation and natural durability—a review. Wood Fiber Sci. 34: 587–611.

Ujino-Ihara, T., K. Yoshimura, Y. Ugawa, H. Yoshimaru, K. Nagasaka and Y. Tsumura. 2000. Expression analysis of ESTs derived from the inner bark of *Cryptomeria japonica*. Plant Mol. Biol. 43:451–457.

Ujino-Ihara, T., Y. Taguchi, K. Yoshimura and Y. Tsumura. 2003. Analysis of expressed sequence tags derived from developing seed and pollen cones of *Cryptomeria japonica*. Plant Biol. 5:600–607.

Umezawa, T. 2000. Chemistry of extractives. *In* Wood and Cellulosic Chemistry. 2nd Edn. Eds. D.N.-S. Hon and N. Shiraisi. Marcel Dekker, New York, pp 213–241.

Wei, Y., Z. Zhang, C.H. Andersen, E. Schmelzer, P.L. Gregersen, D.B. Collinge, V. Smedegaard-Petersen and H. Thordal-Christensen. 1998. An epidermis/papilla-specific oxalate oxidase-like protein in the defence response of barley attacked by the powdery mildew fungus. Plant Mol. Biol. 36:101–112.

Downloaded from http://treephys.oxfordjournals.org/ by guest on October 30, 2011

Yamada, T., H. Tamura and K. Mineo. 1988. The response of sugi (*Cryptomeria japonica* D. Don) sapwood to fungal invasion following attack by the sugi bark borer. Physiol. Mol. Plant Pathol. 33:429–442.

Yoshida, K., M. Hiraide, M. Nishiguchi, S. Hishiyama and A. Kato. 2004. A heartwood norlignan, agatharesinol, was generated in sapwood during withering of a Sugi (*Cryptomeria japonica* D. Don) log. Bull. For. For. Prod. Res. Inst. 3:25–28. Available at http://www.ffpri.affrc.go.jp/labs/kanko/390-4.pdf.

Downloaded from http://treephys.oxfordjournals.org/ by guest on October 30, 2011