

Víctor Chano, Unai López de Heredia, Carmen Collada and Álvaro Soto

Transcriptomic analysis of juvenile wood formation during the growing season in *Pinus canariensis*

Abstract: A noticeable proportion of low transcribed genes involved in wood formation in conifers may have been missed in previous transcriptomic studies. This could be the case for genes related to less abundant cell types, such as axial parenchyma and resin ducts, and genes related to juvenile wood. In this study, two normalized libraries have been obtained from the cambial zone of young individuals of *Pinus canariensis* C. Sm. ex DC, a species in which such cells are comparatively abundant. These two libraries cover earlywood (EW) and latewood (LW) differentiation, and reads have been *de novo* meta-assembled into one transcriptome. A high number of previously undescribed genes have been found. The transcriptional profiles during the growing season have been analyzed and several noticeable differences with respect to previous studies have been found. This work contributes to a more complete picture of wood formation in conifers. The genes and their transcription profiles described here provide a useful molecular tool for further studies focused on relevant developmental issues, such as wound response and the formation of traumatic wood, re-sprouting, etc., presumably related to those cells.

Keywords: cDNA libraries, earlywood (EW), genetic, juvenile wood, latewood (LW), microarrays, *Pinus canariensis*, transcriptome, wood formation

Introduction

In the last two decades, several works have approached the transcriptome basis of wood formation. Most of these works have been performed in angiosperm species (Sterky et al. 1998; Paux et al. 2004), including studies in the non-woody model *Arabidopsis thaliana* (L.) Heynh. (Zhang et al. 2011). Nevertheless, conifer and angiosperm woods display important differences, which support the need for specific studies on each group.

Unfortunately, gymnosperms, and more specifically conifers, display certain characteristics, which hamper transcriptomic and genomic analysis. They have huge genomes (10 000–40 000 Mbp), usually with a high percentage of repeated sequences and pseudo genes (Kovach et al. 2010; Lorenz et al. 2012; Nystedt et al. 2013). Additionally, a much lower number of sequences are available on public databases for conifers than for angiosperms, decreasing the proportion of successful annotations in these studies (López de Heredia and Vázquez-Poletti 2016).

Nevertheless, several studies have addressed wood formation in conifers. For instance, genes involved in wood formation have been identified in *Pinus pinaster* Ait. (Le Provost et al. 2003), *Pinus taeda* L. (Kirst et al. 2003) and *Pinus radiata* D. Don (Li et al. 2010). More recently, Raheison et al. (2015) and Lamara et al. (2016) described networks of co-expressed genes involved in wood formation in *Picea glauca* (Moench) Voss, and Mishima et al. (2014) published a transcriptomic analysis of wood formation in a Taxodiaceae, *Cryptomeria japonica* (Thunb. ex L.f.) D. Don. Other works have focused on the identification of genes involved in reaction wood formation (e.g. Allona et al. 1998; Yamashita et al. 2008), characterized by tracheids with thick and strongly lignified secondary walls and by a low proportion of parenchymatic cells or have used molecular markers to search for differences in wood quality associated with population genetics

(Sandak et al. 2015). Many of these works had mature wood in focus. However, as recently reported by Li et al. (2017) in *Larix kaempferi* (Lamb.) Carr., wood transcriptome varies along the life of the tree, so transcripts related to juvenile wood formation could have been missed in those studies.

This work reports a transcriptome analysis of juvenile wood formation in *Pinus canariensis* C. Sm. ex DC. This is not a common model species for this purpose, although it has noteworthy characteristics: it shows a remarkable healing capacity (recently described in López et al. 2015; Chano et al. 2015) and produces abundant resinous heartwood, which prevents putrefaction and reaches high market values, due to its aesthetic characteristics and its durability even under outdoor conditions (Climent et al. 1998; Taylor et al. 2002). A special feature of its secondary xylem, i.e. the comparatively abundant presence of axial parenchyma (Climent et al. 1998) may have been missed in previous studies in other conifer species (even in Pinaceae). Two cDNA libraries from the *P. canariensis* C. Sm. ex DC cambial zone have been prepared and analyzed, in order to cover the transcriptional profiles both of earlywood (EW) and latewood (LW) differentiation, aiming at obtaining a more complete picture of wood formation in conifers.

Materials and methods

Glossary:

cDNA	Double-stranded DNA synthesized from single-stranded RNA
contig	An overlapped DNA segment obtained during reads alignment, representing a consensus region of DNA
CT	Cycle threshold is the number of cycles required in qRT-PCR for the fluorescent signal to be differentiated from the background level
DEGs	Differentially expressed genes, inferred as statistically significant
delta-delta-CT	The delta-delta-CT is used to obtain normalized relative expression, assuming equal primer efficiency of the target gene and the reference gene
EST	Expressed sequence tag, partial cDNA sequence used to identify gene transcripts
GC content	Percentage of the sum of guanines and cytosines in a DNA or RNA sequence
GO	Gene ontology terms are assigned by the Gene Ontology Consortium to identify gene functions
L50	Corresponds to the smallest number of contigs whose length sum results in 50% of transcriptome total length
N50	Represents the length of the shortest contig at 50% of the transcriptome length distribution
PCR	Polymerase chain reaction, a technique for DNA or cDNA fragment amplification

qRT-PCR	Quantitative real-time polymerase chain reaction is a technique based on PCR for monitoring gene expression
singleton	Short reads not assembled to larger contigs
WGCNA	Weighted gene co-expression network analysis

Plant material: Three-year-old unrelated Canary Island pines grown in a nursery at the UPM facilities were investigated. Trees were grown in 3:1 (v/v) peat:vermiculite, in 650 ml cone-containers first and, after the first year, they were grown outdoor without watering. At the collecting dates, the trees were approximately 150 cm high, with a diameter at breast height of 2–3 cm. Three different unrelated trees were harvested independently at each sampling date. The stems were debarked and the cambial zone and differentiating xylem were scrapped with a sterile scalpel. These samples were immediately frozen in liquid nitrogen and stored at -80°C for further RNA extraction. For the construction of the libraries, samples were collected on 5 May and 12 July. To evaluate gene expression during the growing season, samples were collected on 7 March, 6 April, 6 May, 6 June, 6 July and 5 September. A 2-cm long portion of the stem of each plant was reserved for anatomical observation.

Anatomical observation: Cross-sections 20 μm thick were obtained from stem portions with a Leica SM2400 microtome (Leica Biosystems, Nussloch, Germany). Sections were treated with sodium hypochlorite, washed with distilled water 2–3 times, and then stained for 2 min with 1% safranin (v/v) and for 1 min with 1% alcian blue (w/v), washed with distilled water and dehydrated with an ethanol series. Finally, sections were mounted on slides for bright-field microscopy.

RNA isolation: Total RNA was isolated from differentiating xylem samples stored at -80°C , while the CTAB-LiCl precipitation method (Chang et al. 1993) was applied, and it was purified with the RNeasy Plant Mini Kit (Qiagen, CA, USA). The quantity of total RNA for each sample was measured with a Nanodrop model ND-1000 (Thermo Scientific, MA, USA), and RNA quality was checked by the Experion Bioanalyzer (Bio-Rad, CA, USA).

Library construction and transcriptome assembly: Two RNA pools (corresponding to spring and summer) were obtained by mixing equal amounts of RNA from each of the three plants collected at the sampling date. SMART approach (Zhu et al. 2001) was used for cDNA synthesis from each RNA pool. Amplified cDNA was purified with a QIAquick PCR Purification KIT (Qiagen, CA, USA) and concentrated by ethanol precipitation. DNA pellets were diluted with MilliQ water to a final cDNA concentration of 50 $\text{ng } \mu\text{l}^{-1}$. These cDNA samples were then normalized via the DSN normalization approach (Zhulidov et al. 2004).

Sequencing was performed by means of a Roche 454 Genome Sequencer FLX Titanium System (Roche/454 Life Sciences, CT, USA) by the life sequencing service (Universidad de Valencia, Spain). Read files were pre-processed to improve read quality via sequential PRINSEQ-lite v.0.20.4 (Schmieder and Edwards 2011) and SnoWhite v.1.1.4 (Barker et al. 2010), a cleaning pipeline for cDNA sequences that is based on SeqClean (<http://sourceforge.net/projects/seqclean/>) and trims polyA/T. Reads obtained were deposited in the National Center of Biotechnology Information (NCBI) Sequence Read Archive

(SRA) database (<http://www.ncbi.nlm.nih.gov/sra>; accession no SRP075611, BioProject PRJNA322445).

Assembly was done by Newbler v.2.5 (Roche/454 Life Sciences, CT, USA), based on the overlap layout consensus (OLC) paradigm, which first identifies all pairs of reads that show good overlap and then generates graphs where every read corresponds to one node, and there is an edge between any pair of overlapping reads. Preliminary *de novo* assemblies were pooled into single multi-fasta files, and meta-assembled with CAP3 (Huang and Madan 1999), using default parameters, to remove redundancies and achieve longer contigs. The quality of the final meta-assembled transcriptome was assessed quantitatively by computing the length of the contigs, the GC content and the N50 and L50 parameters using Quast v.2.3 (Gurevich et al. 2013) and rnaQUAST v.1.2.0. (Bushmanova et al. 2016). Additionally, quality of the transcriptome was also checked by aligning the contigs with available transcriptomes from other *Pinus* species. This transcriptome shotgun assembly project has been deposited at DDBJ/EMBL/GenBank under the accession GESU00000000 v01.

Functional annotation: The transcriptome assemblies were down-streamed as inputs for the local version of BlastX (Altschul et al. 1990). The contigs were launched against the Viridiplantae section of the RefSeq database (NCBI), with an *e-value* threshold of 10^{-5} and limited to a maximum of 50 hits. The output file was imported into Blast2GO v2.7.2 (Conesa et al. 2005), in order to provide a comprehensive functional annotation and sequences analysis. A statistical assessment of annotation was performed by the Mapping, Annotation, and GO-slim tools implemented in Blast2GO. InterPro annotation was also performed in order to retrieve motif/domain information and to improve annotation. Enzyme codes (EC) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were also obtained through Blast2GO by the direct mapping of GOs to the enzyme code equivalents.

Microarray analysis: The contig library was filtered removing isoforms and sequences shorter than 300 bp. The unigenes served for the design of a one-color 180 K microarray (Agilent, USA). Furthermore, contigs from other cDNA libraries of *Pinus pinaster* and *Pinus pinea*, as well as ESTs and sequences of the loblolly pine from the Pine Gene Index Database (http://www.mgel.msstate.edu/dna_libs.htm) were included. Probes included in the microarray are provided in Supplementary Information Table S1.

Three biological replicates (RNA from three different trees) from each sampling point were independently hybridized following the One Color Microarray-Based Gene Expression Analysis Protocol (Agilent Technologies, Palo Alto, CA, USA). Arrays were scanned at a 3-mm resolution on Agilent DNA Microarrays Scanner (G2565BA, Agilent Technologies), and the images were analyzed with Feature Extraction software (Agilent Technologies). Background correction and normalization of expression data were performed using LIMMA (Smyth and Speed 2003). For local background correction and normalization, the methods *normexp* and *loess* in LIMMA were applied, respectively (Smyth and Speed 2003). To have similar distribution across arrays and to achieve consistency among arrays, log-ratio values were scaled as a scale estimator for the median-absolute-value (Smyth and Speed 2003). Hybridizations and first statistical analysis were performed by the Genomics Facility at Centro Nacional de Biotecnología, Madrid. Data were deposited in the NCBI Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo> accession no GSE82217).

Normalized data were analyzed with Babelomics v.4.2 (Medina et al. 2010), performing a *maSigPro* (Time/Dosage Series) analysis, a methodology designed for the identification of differentially expressed genes (DEGs) in a time-course experiment. *maSigPro* is based on a regression modeling approach and reports expression changes considering the whole expression profile. The following options were selected for the analysis: (1) polynomial degree=4, which determines regression complexity; (2) FDR adjustment; (3) significance level of 0.05; and (4) k-means clustering for cluster method (k=6). Later on, an enrichment analysis of GO terms was performed for each of the six clusters compared to the complete transcriptome obtained from libraries, by means of the Fisher's exact test of the Gossip (Blüthgen et al. 2005) package implemented in Blast2GO, with an FDR value of 0.05.

A weighted gene co-expression network analysis was performed by means of the WGCNA R package v1.51 (Langfelder and Horvath 2008) with the relative gene expression values for each biological replicate and sampling points. To identify expression modules in WGCNA, a soft threshold power value of 10 was applied and the networks were constructed for connected genes with $r \geq 0.90$. Finally, the co-expression network was visualized by Cytoscape v. 3.5.1 (Shannon et al. 2003).

qRT-PCR: The expression patterns of eight genes, covering the clusters obtained from the microarray analysis, were confirmed by qRT-PCR using RNA from one of the genotypes as a biological replicate in the microarrays. First-strand cDNA synthesis was performed by SuperScript™ III reverse transcriptase (Invitrogen, USA) following the manufacturer's instructions, with 4 µg of total RNA and random hexamers. Gene-specific primers were designed for the selected genes via the Primer3 software (Untergasser et al. 2012), with a m.p. between 60 and 65°C, and producing amplicons between 80 and 120 bp (Table 1). qRT-PCR was performed in a CFX96™ Real-Time PCR Detection System (Biorad, USA), by means of the SsoFast™ EVAgreen® Supermix (Biorad, USA), according to the manufacturer's protocol, and following the standard thermal profile: 95°C for 3 min, 40 cycles of 95°C for 10 s and 60°C for 10 s. To compare data from different qRT-PCR runs, the CT values were normalized by the R18S as housekeeping gene. The expression ratios were then obtained by the delta-delta-CT method corrected for the PCR efficiency for each gene (Pfaffl 2001).

Results and discussion

De novo transcriptome assembly

Two normalized cDNA libraries were obtained from the differentiating xylem, i.e. (1) in spring, when the growth rate is at its maximum and xylematic cells produced by periclinal divisions of cambial cells mostly differentiate into tracheids; and (2) in mid-summer, when the growth rate decreases and eventually stops, and a higher proportion of resin ducts and their associated axial parenchyma are produced.

Sequencing of this cDNA produced 458 498 reads in spring and 474 393 in summer. These two separate libraries were meta-assembled into a single transcriptome, removing redundancies, and 44 849 contigs and

Table 1: Primers used for qRT-PCR verification.

Contig name	Oligo name	Description	Orient- ation	Len	Tm	GC%	Sequence (5'-3')
Congit05410	Pc_05410_PECTINESTERASE_F1	Pectinesterase 2-like protein	For	20	63.1	55	GTACTCTCGCACGGTCTTCA
	Pc_05410_PECTINESTERASE_R1		Rev	20	62.5	45	ATAATAAGCGTCCCAACG
Contig12050	Pc_12050_MYB_F1	MYB46-like protein	For	20	62.8	45	ATCCCAACATGGAAGAAGC
	Pc_12050_MYB_R1		Rev	20	63.7	50	CTGCATCACCATCACACTCA
Contig06513	Pc_06513_ATHB15_F1	ATHB15-like protein	For	20	62.9	55	CCTTGAGGGAAGACGAAGAG
	Pc_06513_ATHB15_R1		Rev	20	62.7	45	AGCTCAGAATTCCTTTGCA
Contig02274	Pc_02274_NAC_F1	NAC74-like protein	For	20	64.2	50	ACCTTTCTGCTCGAAGCTCA
	Pc_02274_NAC_R1		Rev	20	63.9	45	TCAATGGCAGGTGTTGGTTA
Contig04961	Pc_04961_BEL1_F1	BEL1-like protein	For	21	64.2	48	TGAAACTTCAGCACCTCATCC
	Pc_04961_BEL1_R1		Rev	20	63.9	50	AGCTCTGGCTTAGGCAACAA
Contig03225	Pc_03225_EXPANSIN_F1	α -expansin-like protein	For	20	62.8	45	AAGCGGAGCTGATTCTTGAT
	Pc_03225_EXPANSIN_R1		Rev	20	63.1	60	CTCAGAGCCACAGAGACGAG
Contig04867	Pc_04867_WRKY_F1	WRKY4-like protein	For	20	63.8	50	CCTGTGCTAGCCATTTGTGA
	Pc_04867_WRKY_R1		Rev	20	63.7	50	TCAATGGCAGGTGTTGGTTA
Contig06476	Pc_06476_CCoAOMT_F1	CCoAOMT-like protein	For	20	64.0	50	GATTGAACAACCGAGGTGCT
	Pc_06476_CCoAOMT_R1		Rev	20	63.6	45	TGCAACACCTGAATCCAAC
Housekeeping	Ri18S_FW	18S ribosomal	For	19	62.4	53	GCGAAAGCATTTGCCAAGG
Housekeeping	Ri18S_RV		Rev	21	62.4	48	ATTCCTGGTCGGCATCGTTA

For, forward; Rev, reverse.

103 764 singletons (i.e. short reads that are not assembled into contigs). The size distribution of the contigs and singletons are presented in Figure 1a. The total length of the final transcriptome was 25 836 041 bp and the largest contig had a length of 3518 bp. The GC content was 42.15% and the N50 and L50 values were 1101 and 8755 bp, respectively (Table 2). This transcriptome is considerably larger than other ones related to wood formation in conifers and available in databases, such as those of *Pinus pinaster* Ait. (Le Provost et al. 2003), *Pinus taeda* L. (Kirst et al. 2003) or *Pinus radiata* D. Don (Li et al. 2010).

To check the quality of this *de novo* transcriptome, contigs were aligned with available transcriptomes from two closely related species, *Pinus pinaster* Ait. and *Pinus halepensis* Mill. From the total of *Pinus canariensis* C. Sm. ex DC contigs, 92.7%, and 79.3% were aligned, respectively, to the *P. pinaster* Ait. and *P. halepensis* Mill. unigenes.

Functional annotation

The transcriptome assemblies were aligned to the Viridiplantae section of the RefSeq database (NCBI). The analysis yielded 27 576 contigs (61.5%) and 28 542 singletons (27.5%) matching records of the RefSeq database.

The output file from the BlastX analysis was imported into Blast2GO (Conesa et al. 2005), and a statistical assessment of annotation, including InterProScan, was performed in order to retrieve motif/domain information and

merge this to the actual annotation. GO terms were assigned to 22 227 contigs (49.6% of the total) and 20 743 singletons (20.0% of the total). A total of 172 365 GO terms were obtained, which are distributed among different levels for the biological process (BP; 92 988GOs, 54%), molecular function (MF; 32 923 GOs, 19%) and cellular component (CC; 46 454 GOs, 27%) categories. The main level for BP and MF was level 3 (31 423 and 11 914 GO terms, respectively), while it was level 8 for CC (23 680 GOs). Again, this level distribution was consistent with those obtained for *Pinus halepensis* Mill. (Pinosio et al. 2014), *Pinus contorta* Douglas ex Loudon (Parchman et al. 2010), or *Pinus radiata* D. Don (Li et al. 2009). Moreover, Figure 1b shows the most represented GO terms in a direct GO count for the three categories, highlighting *cellular process*, *biosynthetic process*, *response to stress*, *cellular component organization* and *biological process* for BP, *nucleotide binding and kinase activity* for MF, or *membrane*, *cell wall* and *nucleus* for CC. The complete annotation of the 44 849 contigs is available in Supplementary Information Table S2. Conversely, no homology was found in the public databases for up to 17 273 contigs.

Transcription profiles during the growing season

Samples of the cambial zone of *Pinus canariensis* C. Sm. ex DC collected at six dates from March to September (Figure

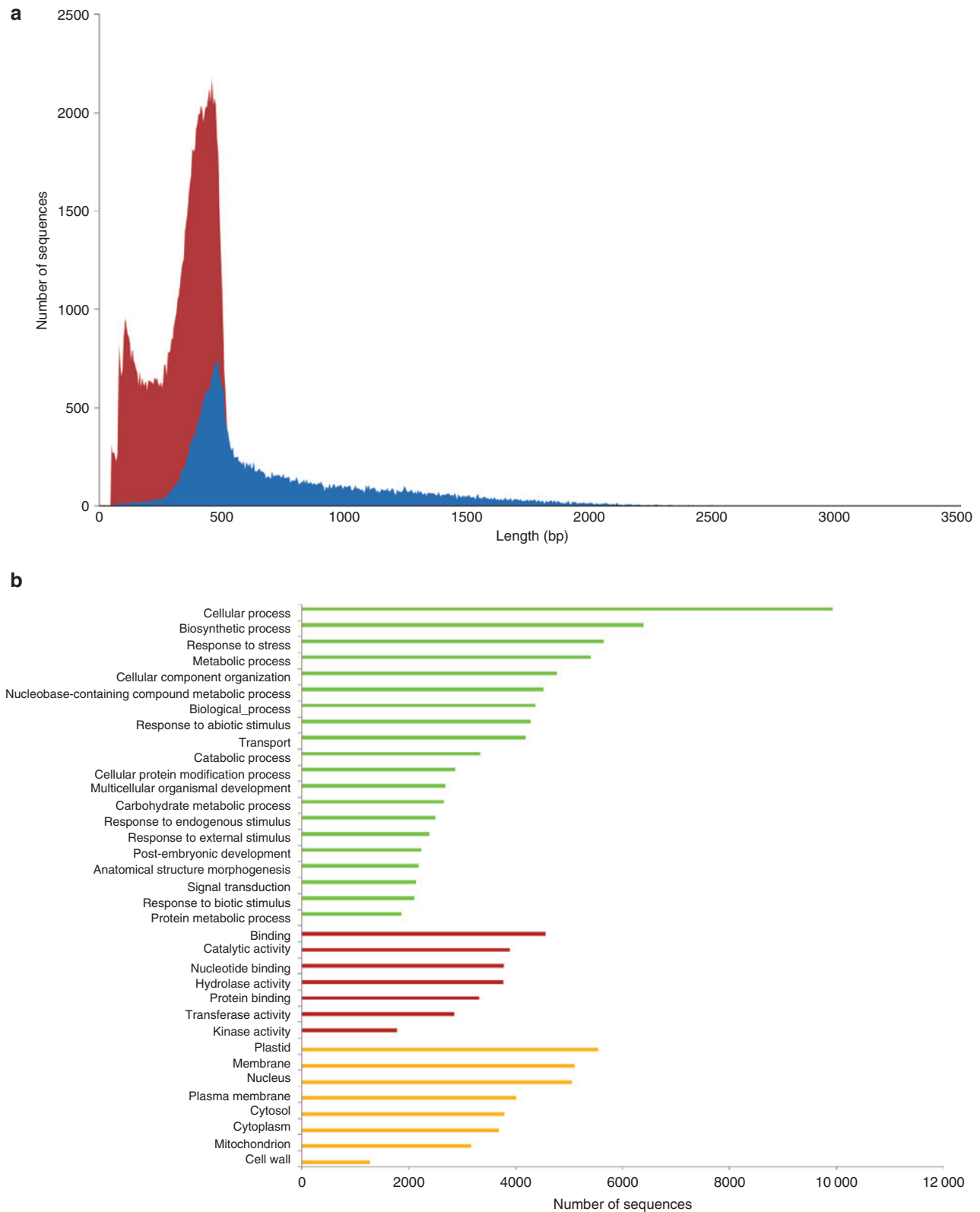


Figure 1: (a) Size distribution of contigs and singletons achieved after meta-assembly with CAP3. The vertical line indicates the number of sequences against their respective length in bp (red graph = singletons; blue graph = contigs). (b) Distribution of the most abundant GO terms assigned to the *Pinus canariensis* C. Sm. ex DC transcriptome during xylogenesis (green bars: biological process; red bars: molecular function; yellow bars: cellular component).

2) were hybridized against a microarray including a set of 16 455 unigenes from the contig library (removing isoforms and sequences shorter than 200 bp), in order to examine

the transcription patterns during the growing season. Correlation among biological replicates is presented in Supplementary Information Figure S1.

Time-course analysis revealed that 3302 of the contigs included in the microarray (ca. 18%) showed significant differences in their transcription levels during the growing season, being considered as DEGs. According to their transcription patterns, they were clustered in six main trends (Figure 3). Most of the DEGs correspond to two opposite profiles. First, 43.8% of the DEGs have shown an increase in transcript level during spring, which corresponds to cluster 1 (1226 DEGs; 37.1%) and cluster 6 (222 DEGs; 6.7%), highlighting a decrease in the summer and an upturn at September for cluster 6, less prominent

Table 2: Summary of transcriptome quality assessment.

No. of transcripts	44 849
No. of transcripts > 500 bps	26 377
No. of transcripts > 1000 bps	10 565
Total length (bp)	25 836 041
Largest contig (bp)	3518
GC content (%)	42.15
N50 contig size (bps)	1101
L50 (contig no.)	8755

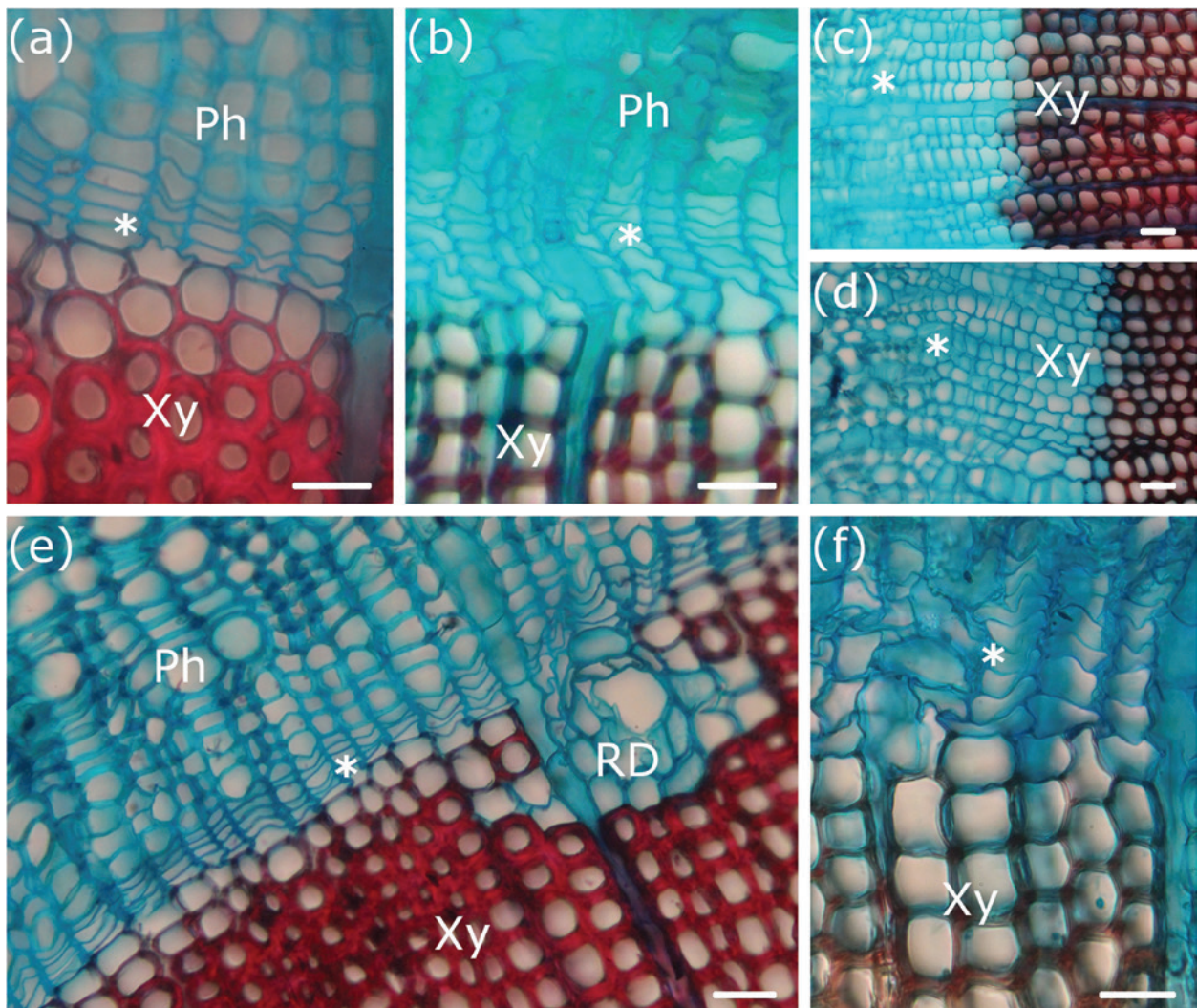


Figure 2: Micrographs in bright-field microscopy of the cambial zone in *Pinus canariensis* during the growing season (bars = 50 nm; asterisk = cambial zone; Ph = phloem; Xy = xylem; RD = resin duct).

a–f: March, April, May, June, July and September, respectively. Notice the highest cambial activity in May and June. Several rows of new tracheids, not completely lignified yet, can be observed. Conversely, cambial activity decreases in July and resin ducts and smaller tracheids, with thicker walls, are forming.

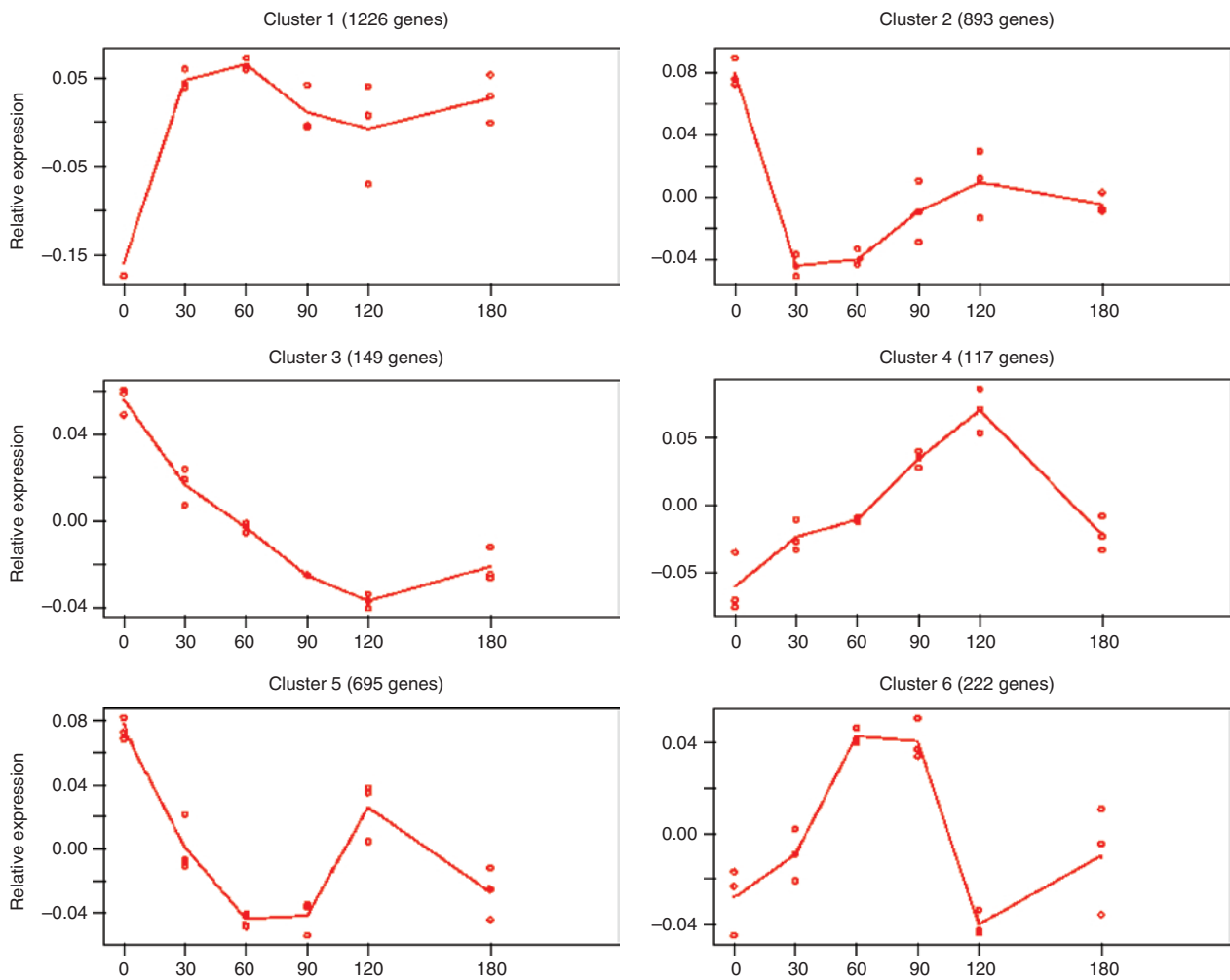


Figure 3: K-means clustering ($k=6$) of differentially expressed genes, representing the main tendencies of gene expression profiles during xylogenesis in *Pinus canariensis* C. Sm. ex DC after maSigPro analysis.

The Y-axis represents the relative expression values and the X-axis represents time in days. The number of genes in each cluster is indicated.

in cluster 1. Conversely, clusters 2, 4 and 5 showed low transcript levels during mid-spring (April and May) and a local maximum in the summer. Thus, cluster 2 (893 DEGs; 27.0%) and cluster 5 (695 DEGs; 21.1%) showed a decreasing transcript level during spring (abruptly from March to April for cluster 2, but progressively from March to May–June for cluster 5) and a new upturn in July for both clusters, while cluster 4 showed a continuous increase of transcript level from March to July, turning down in September. Finally, cluster 3 (149 DEGs; 4.5%) showed a progressive decrease from March to July–September. A complete list of detected DEGs is provided in Supplementary Information Table S3. Additionally, qRT-PCR analysis of the transcription profiles of eight genes, covering the main functional groups and expressional clusters, were performed. High Pearson's correlation values (>80%) were obtained between microarray and RT-PCR

expression profiles for most genes, thus validating the microarray expression patterns (Figure 4).

These patterns are consistent with anatomical observations (Figure 2). The beginning of EW development starts in April through periclinal divisions, and increase noticeably during spring. The growth rate decreases during the summer and resin ducts are formed, setting the period of LW development. Interestingly, the growing season upturns again in September. This growing pattern is not unusual in Mediterranean environments. In northern latitudes, summer is favorable for tree growth, with long days and warm temperatures, so that growth continues during spring and summer, while LW is formed at the beginning of autumn (Li et al. 2010). On the other hand, the drought in Mediterranean summer compromises growth, which can even stop in harsh summers. Wood ring analysis proves that a second growth may occur in

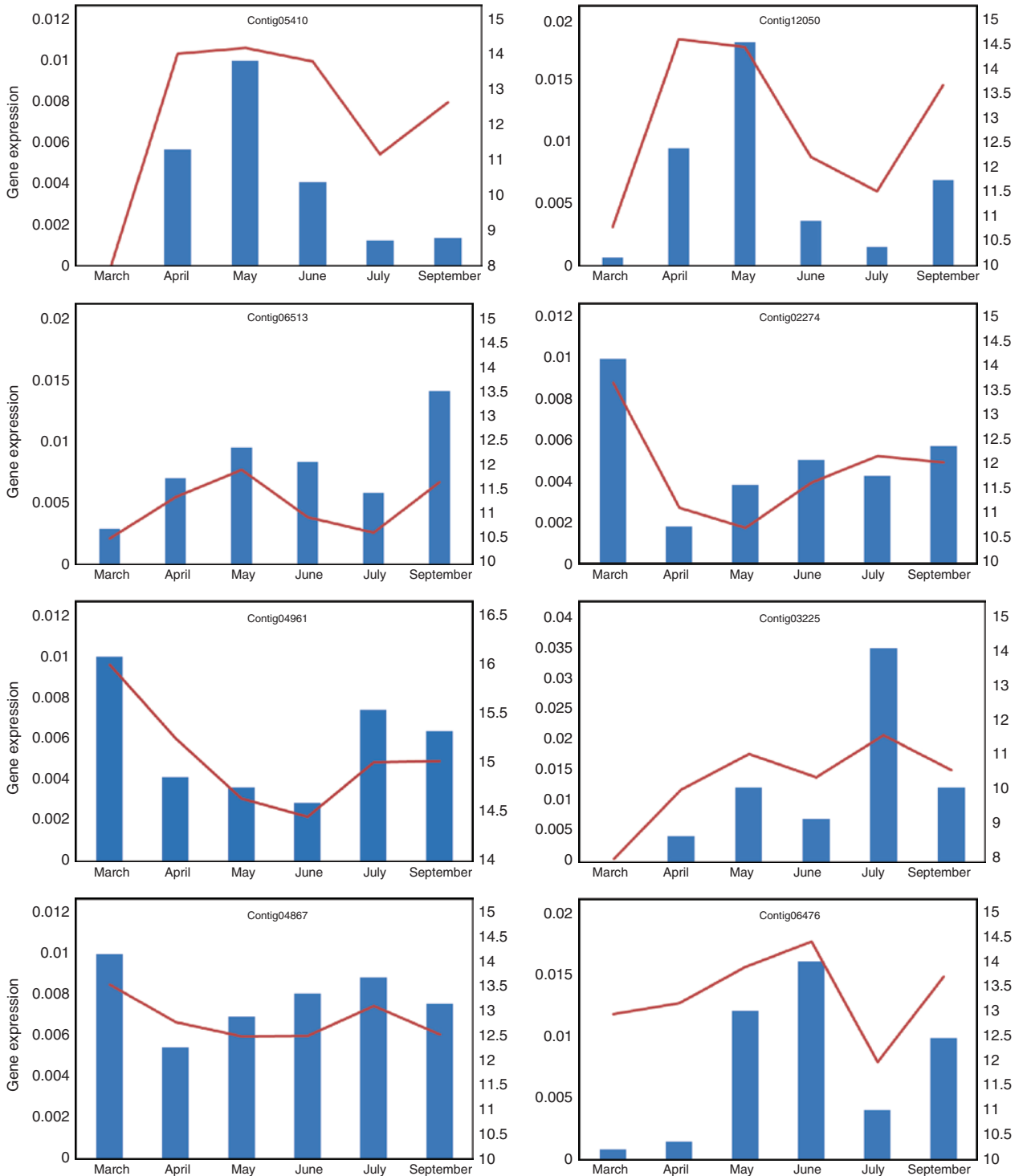


Figure 4: qRT-PCR validation of microarray transcription profiles. X-axis: time in days. Y-axis left: normalized gene expression values of selected DEGs for qRT-PCR (bars). Y-axis right: absolute expression values for microarray experiments (continuous line).

years of mild autumns, with mild temperatures and sufficient rainfall (Figure 5). This effect is also known as polycyclism, and can be observed in leaf formation and

shoot development too (Verdú and Climent 2007; Bobinac et al. 2012). Polycyclism has been described for several Mediterranean conifer species such as *Pinus pinaster* Ait.

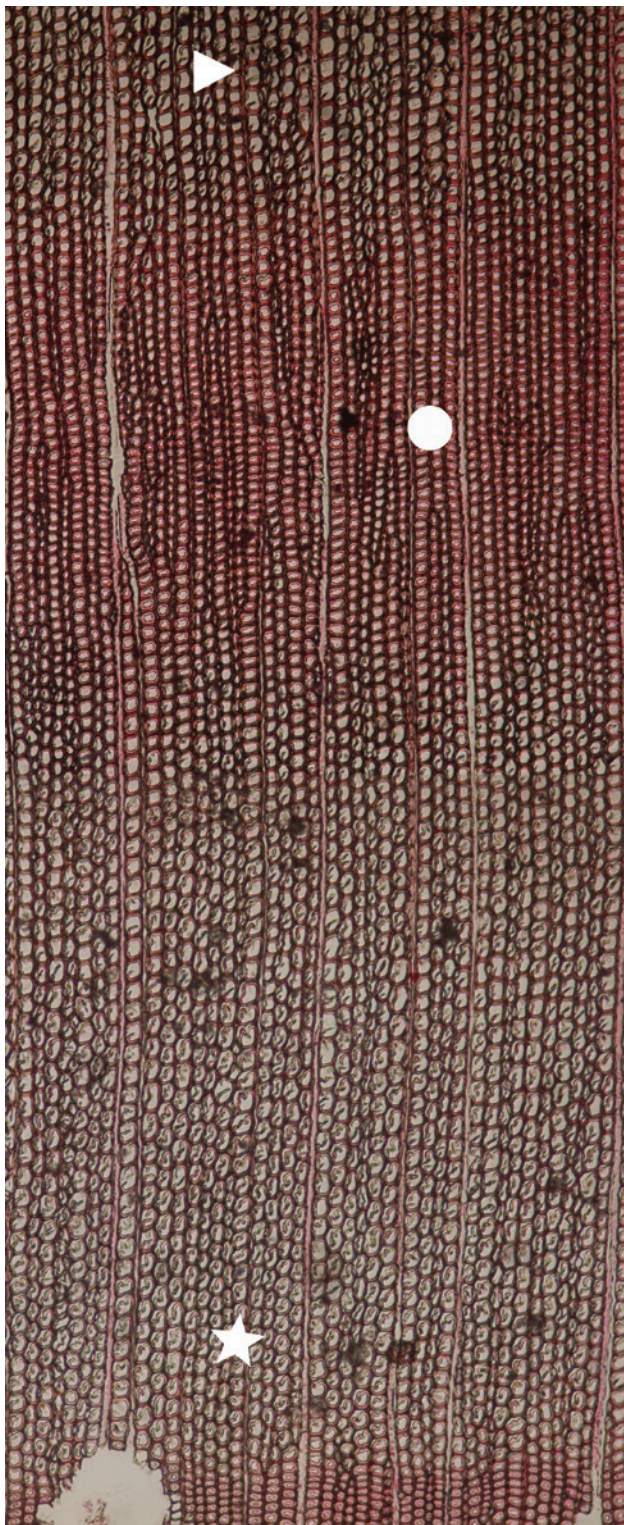


Figure 5: Picture of a *Pinus canariensis* C. SM. ex DC wood ring. At the bottom, earlywood formed at the beginning of the growing season can be observed (star). During the summer, the growth rate decreases and smaller tracheids with thicker walls are produced (dot). During autumn, a new increase in cambial activity takes place again producing wider tracheids with thinner walls (arrowhead).

(Vieira 2013), *Pinus pinea* L. (Mutke et al. 2013), or *Pinus halepensis* Mill. (Liphshitz et al. 1984; de Luis et al. 2007). Interestingly, higher variance has been observed among biological replicates for transcript values in the summer time samples for most genes, which is consistent with the individual variability in response to unfavorable summer conditions (drought, high temperature).

DEGs involved in EW development

As expected, most of the DEGs with a transcription maximum in April–May (clusters 1 and 6) are related with cell division and cell wall formation processes (Table 3). Thus, an enrichment analysis of GO terms reveals a significant abundance in these clusters of GO terms related to biosynthetic process, cell growth, biogenesis, anatomical structure, morphogenesis and development.

The most abundant DEGs of this group are those putatively involved in cell wall matrix formation and modification, such as genes for coding proteins of the cellulose synthase superfamily (contigs 05245, 06866 and FLPpisotig04143), KORRIGAN endoglucanase-like proteins (contigs 10173, 18777 and FLPpisotig01381), and several genes coding for the Carbohydrate Active enZymes (CAZymes; <http://www.cazy.org>; Park et al. 2010a). These are regulators of the principal biosynthesis pathways of hemicelluloses, such as glycosyl transferases and hydrolases (GTs and GHs), along with pectate lyases (PLs). PLs are involved in the degradation of pectin, a principal component of the primary cell walls during the secondary cell wall development, allowing intrusive growth of tracheids (Marín-Rodríguez et al. 2002). Expansins, which are regulators of cell growth (Cosgrove et al. 2002) and are related to tracheid elongation, are also present in these clusters. Additionally, some fasciclin like-arabinogalactan proteins associated with xylem development (Andersson-Gunnerås et al. 2006) were also found in cluster 1.

Another important group of DEGs present in clusters 1 and 6 comprises genes involved in the phenylpropanoid pathway and lignin biosynthesis, which play a key role in the development of tracheids. These genes have been well-characterized in plant models such as *Arabidopsis thaliana* (L.) Heynh., maize, rice, alfalfa or tobacco, as well as in forest species such as *Populus* L. and conifers like *Pinus* L. or *Picea* Link (reviewed by Vanholme et al. 2010). Phenylalanine ammonia-lyase (PAL), cinnamate-4-hydroxylase (C4H), hydroxycinnamoyl:CoA shikimate/quinatehydroxy cinnamoyl transferase (HCT), caffeic acid O-methyltransferase (COMT), caffeoyl-CoA

Table 3: Most relevant differentially expressed genes (DEGs) presumably involved in earlywood development.

Cluster	Contig ID	BlastX Description	Hit ACC	eValue
Cell-wall matrix development and/or carbohydrate metabolism				
1	Contig00492	glucomannan 4-beta-mannosyltransferase 9-like	XP_006851749	0
1	Contig00603	beta-xylosidase alpha-l-arabinofuranosidase 2-like	NP_001146416	6.55E-121
1	Contig00713	mannan endo-beta-mannosidase 6-like	XP_004249033	5.18E-169
1	Contig01025	fasciclin-like arabinogalactan protein 10-like	XP_002320736	3.25E-105
1	Contig01628	mannose-1-phosphate guanylyltransferase 1-like	XP_008441398	0
1	Contig01691	probable beta-xylosyltransferase irx10l-like	XP_002511079	0
1	Contig01735	probable beta-xylosyltransferase irx14-like	XP_006850475	3.65E-91
1	Contig01867	probable polygalacturonase	XP_006840361	0
1	Contig02153	probable beta-galactosyltransferase 11-like	XP_008791474	1.51E-118
1	Contig02909	mannose-1-phosphate guanyltransferase alpha-like	XP_002281959	2.55E-111
1	Contig03225	expansin alpha	XP_002315043	8.72E-125
1	Contig05066	probable pectate lyase 15-like	XP_006282594	0
1	Contig05245	probable cellulose synthase a catalytic subunit 3	XP_002283406	4.80E-177
1	Contig05410	pectinesterase 2-like	XP_003521940	3.37E-143
1	Contig05424	probable xyloglucan endotransglucosylase hydrolase protein 8-like	XP_002518528	2.91E-119
1	Contig06866	cellulose synthase-like protein d3-like	XP_006478929	9.89E-115
1	Contig07097	probable beta-d-xylosidase 7-like	XP_002535159	3.14E-92
1	Contig08356	udp-glycosyltransferase 85a2-like	XP_008789349	7.55E-82
1	Contig09025	xyloglucan endotransglucosylase hydrolase protein 9-like	XP_002522701	1.27E-92
1	Contig09397	alpha-l-arabinofuranosidase 1-like	XP_007028421	0
1	Contig09434	fasciclin-like arabinogalactan protein 17-like	XP_002309262	2.66E-67
1	Contig09435	fasciclin-like arabinogalactan protein 17-like	XP_004249393	5.99E-92
1	Contig09907	probable xyloglucan endotransglucosylase hydrolase protein 32	XP_002523709	4.46E-80
1	Contig10110	xyloglucan galactosyltransferase katamari1-like	NP_001147481	2.12E-106
1	Contig10173	endoglucanase 24-like	XP_006487371	2.15E-124
1	Contig10178	beta-xylosidase alpha-l-arabinofuranosidase 2-like	XP_006855294	1.32E-71
1	Contig11436	probable polygalacturonase non-catalytic subunit jp650-like	XP_007153633	2.34E-43
1	Contig11975	probable beta-xylosyltransferase irx9h	XP_006852709	1.55E-131
1	Contig12072	glucomannan 4-beta-mannosyltransferase 9-like	XP_007220072	0
1	Contig12089	mannan endo-beta-mannosidase 7-like	XP_006841534	5.26E-179
1	Contig12414	expansin-b3-like	XP_007043563	1.01E-84
1	Contig12926	xylosyltransferase 1-like	XP_004309761	1.14E-20
1	Contig13281	probable pectinesterase 68-like	XP_007039841	8.74E-68
1	Contig13724	probable xyloglucan endotransglucosylase hydrolase protein 32-like	XP_008805707	8.06E-98
1	Contig13778	probable polygalacturonase at1g80170-like	XP_006840713	3.72E-56
1	Contig15337	xylosyltransferase 1-like	NP_198815	3.15E-14
1	Contig15778	udp-glycosyltransferase family protein	XP_004306713	2.55E-71
1	Contig18777	endoglucanase 24-like	XP_006849270	8.04E-82
1	Contig18811	expansin alpha	NP_001281040	1.04E-56
1	Contig18941	probable xyloglucan endotransglucosylase hydrolase protein 32	XP_006424466	2.05E-57
1	Contig21041	scopoletin glucosyltransferase-like	XP_006379195	2.97E-50
1	Contig21066	polygalacturonase qrt3-like	XP_003554934	9.14E-61
1	Contig21204	alpha-expansin 8	XP_002520863	1.99E-81
1	Contig22029	probable pectate lyase 15-like	XP_002265729	1.12E-39
1	Contig23968	mannan endo-beta-mannosidase 7-like	XP_006340651	1.49E-10
1	FLPpcontig01227	probable xyloglucan endotransglucosylase hydrolase protein 8	XP_002518528	5.08E-119
1	FLPpcontig04143	cellulose synthase-like protein d3	NP_001169669	6.35E-160
1	FLPpisotig01381	endoglucanase 17-like	XP_006487371	5.49E-47
1	FLPpisotig01642	beta-xylosidase alpha-l-arabinofuranosidase 2-like	NP_001146416	4.19E-109
1	Ppisotig03412	beta-xylosidase alpha-l-arabinofuranosidase 2-like	XP_006855294	4.17E-80
1	Ppisotig09772	alpha-l-arabinofuranosidase 1-like	XP_010904719	0
1	Ppisotig14238	fasciclin-like arabinogalactan protein 16	NP_850253	2.92E-30
1	Ppnisotig01389	mannan endo-beta-mannosidase 7-like	XP_008797528	1.13E-74
1	Ppnisotig07937	probable beta-xylosyltransferase irx9h	XP_011626438	1.72E-49
1	Ppnisotig12100	mannan endo-beta-mannosidase 7	XP_007046870	3.76E-23
6	Contig01745	gdp-mannose dehydratase 1-like	XP_001774456	4.81E-180

Table 3 (continued)

Cluster	Contig ID	BlastX Description	Hit ACC	eValue
6	Contig22956	xylosyltransferase 1-like	XP_007222964	9.26E-61
6	Ppisotig06462	probable xyloglucan glycosyltransferase 5	XP_006845528	2.24E-19
6	Ppnisotig01254	probable xyloglucan glycosyltransferase 6	XP_002960657	7.08E-76
Phenylpropanoid biosynthesis pathway				
1	Contig00537	caffeoyl- o-methyltransferase	XP_006856484	1.48E-157
1	Contig01861	cinnamoyl- reductase 1-like	XP_008228791	7.59E-174
1	Contig05567	trans-cinnamate 4-monooxygenase-like	XP_008792249	4.55E-166
1	Contig12190	caffeic acid 3-o-methyltransferase-like	XP_006829216	5.53E-163
1	Contig20935	shikimate o-hydroxycinnamoyltransferase-like	XP_002303858	1.06E-135
1	Contig23672	caffeic acid 3-o-methyltransferase-like	XP_004494877	5.02E-43
6	Contig06476	caffeoyl- o-methyltransferase	XP_001766967	1.01E-50
6	Ppisotig09705	phenylalanine ammonia-lyase	XP_010246007	0
Transcriptional regulation of meristem activity				
1	Contig01410	transcription factor bhlh130-like	XP_007205257	8.41E-49
1	Contig01913	transcription factor myb44-like	XP_003611666	6.58E-75
1	Contig02039	gata transcription factor 12-like	XP_006372032	5.89E-46
1	Contig02356	exordium like 3	XP_007038432	6.16E-121
1	Contig03423	floricaula leafy homolog	XP_002966714	7.09E-69
1	Contig03715	protein argonaute 4-like	XP_002308843	4.80E-19
1	Contig04966	transcription factor bhlh144-like	XP_007018175	3.83E-10
1	Contig06513	homeobox-leucine zipper protein athb-15-like	XP_006853643	1.37E-170
1	Contig06617	homeobox protein knotted-1-like 3-like isoform x2	XP_006483430	4.70E-121
1	Contig06813	wuschel-related homeobox 4-like	XP_006305551	3.78E-35
1	Contig10039	transcription factor bhlh61-like	XP_006486517	8.59E-07
1	Contig12050	transcription factor myb46-like	XP_007132019	1.10E-67
1	Contig12421	transcription factor bhlh63-like	XP_006878569	5.93E-69
1	Contig14178	clavata3 esr-related 12 family protein	XP_008785816	4.38E-08
1	Contig14511	homeobox-leucine zipper protein hat5-like	XP_007151094	6.18E-16
1	Contig18903	transcription factor bhlh155-like	XP_001763347	2.03E-114
1	Contig20476	exordium like 2	XP_002285759	5.88E-89
1	Contig21473	nac domain transcriptional regulator superfamily	XP_006852355	2.09E-10
1	Contig22505	transcription factor bhlh62-like	XP_006848450	4.74E-65
1	FLPpisotig02169	transcription factor myb46-like	XP_007211763	2.57E-70
1	Ppisotig11335	myb-related protein 308-like	XP_006854612	9.56E-71
1	Ppisotig16776	transcription factor myb24-like	XP_009795062	2.76E-40
1	Ppisotig19583	floricaula leafy homolog	XP_001762054	5.32E-51
1	Ppnisotig10775	scarecrow-like protein 28	XP_006477403	2.54E-14
6	Contig02513	gata transcription factor 9-like	XP_002319169	1.32E-38

O-methyltransferase (CCoAOMT), cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) can be found in clusters 1 and 6.

Among the transcription factors grouped in these clusters are several members of the MYB family, presumably related to the expression of monolignols and phenylpropanoids biosynthesis genes mentioned above, together with NAC factors such as NAC7 and NAC8 (Duval et al. 2014; Lamara et al. 2016). Contig 01913 encodes for a putative MYB44, which was found to be expressed in *Arabidopsis thaliana* (L.) Heynh. in response to aphid attacks (Liu et al. 2010). For instance, MYB46 is involved in the regulation of secondary wall biosynthesis and its

repression has been reported to cause a drastic reduction in secondary wall thickening in *A. thaliana* (L.) Heynh. (Zhong et al. 2007). Conversely, overexpression of MYB308, also found in cluster 1, inhibits the production of phenolic compounds in *Antirrhinum majus* L., which are needed for lignification of the secondary cell wall (Tamagnone et al. 1998). Also, a putative MYB24, which belongs to the R2R3-MYB subgroup and is involved in stamen filament development in *A. thaliana* (L.) Heynh. (Cheng et al. 2009), can be seen in cluster 1.

Other transcription factors, belonging to the homeodomain leucine-zipper (HD-Zip) family reach a transcript maximum in spring. An ortholog of HAT5 (contig 14511)

Table 4: Most relevant differentially expressed genes (DEGs) presumably involved in latewood development.

Cluster	Contig ID	BlastX Description	Hit ACC	eValue
Cell-wall matrix development and/or carbohydrate metabolism				
2	Contig10181	expansin-a4-like	XP_003544930	4.43E-95
2	Contig12274	galactinol-sucrose galactosyltransferase-like	XP_004232319	9.53E-143
2	Contig21865	galactinol-sucrose galactosyltransferase-like	NP_001275531	7.90E-57
4	Contig01916	xyloglucan glycosyltransferase 4-like	XP_006845528	1.63E-101
4	Contig06122	probable xyloglucan endotransglucosylase hydrolase protein 8-like	NP_001241628	9.09E-96
4	Contig08266	probable xyloglucan endotransglucosylase hydrolase protein 23	XP_002273742	1.46E-104
4	Contig12879	udp-glycosyltransferase superfamily protein	XP_006850431	1.22E-97
4	Contig14263	mannan endo-beta-mannosidase 7-like	XP_006841534	6.55E-20
4	Contig15857	cellulose synthase-like protein d3	NP_001169669	4.04E-92
4	Contig17013	probable xyloglucan endotransglucosylase hydrolase protein 23	NP_001149692	2.66E-13
4	Ppisotig00266	probable xyloglucan endotransglucosylase hydrolase protein 23	XP_012447657	3.64E-75
5	Ppnisotig11614	probable sucrose-phosphate synthase 3	XP_008794597	7.64E-105
Phenylpropanoid biosynthesis pathway				
2	Contig18416	cinnamoyl-reductase 1-like	XP_006847422	1.20E-154
Transcriptional regulation of meristem activity				
2	Contig02274	nac domain-containing protein 74-like	XP_008778272	1.62E-72
2	Contig02951	gata transcription factor 9-like	XP_001778541	8.26E-14
2	Contig04889	wuschel-related homeobox 8-like	XP_001777634	2.93E-43
2	Contig05450	transcription factor bhlh96-like	XP_006829211	8.27E-84
2	Contig05803	homeobox protein knotted-1-like 3-like	XP_008802010	3.28E-116
2	Contig06152	exordium like 2	XP_007042914	1.68E-109
2	Contig11838	homeobox protein knotted-1-like 2-like	XP_002285521	5.26E-99
2	Contig13895	nac transcription factor 29-like	XP_001771652	6.28E-27
2	Contig16468	scarecrow-like protein 6-like	XP_006440333	9.19E-15
2	Contig19392	nac domain-containing protein 78-like	XP_006855883	2.86E-18
2	Contig21263	transcription factor bhlh121-like	XP_006859207	3.87E-86
2	Contig23724	myb family transcription factor family protein	XP_008784286	4.36E-10
2	Ppnisotig00673	nac transcription factor 29-like	XP_012440667	3.44E-13
4	Contig01331	myb-like protein x-like isoform x2	XP_006380820	2.40E-09
4	Contig02773	topless-related protein 4-like isoform x1	XP_002960294	1.90E-135
4	Contig04552	bzip domain class transcription factor	XP_008813601	4.81E-133
4	Contig09007	exordium like 2	XP_007018740	4.46E-112
4	Ppisotig11957	protein exordium-like 2	XP_009760896	1.38E-91
5	Contig01876	myb-like protein x-like	XP_007034235	4.30E-26
5	Contig04867	probable wrky transcription factor 4	XP_008227105	6.96E-84
5	Contig05634	nac domain-containing protein 2-like	XP_008792531	3.49E-86
5	Contig12675	myb-related protein 3r-1-like isoform x1	XP_008800229	1.34E-19
5	Contig20447	bel1-like homeodomain protein 1	XP_002275098	2.74E-53
5	Contig22932	homeobox-leucine zipper protein athb-15-like	XP_001762557	3.65E-43
5	FLPpisotig03105	bel1-like homeodomain protein 1	XP_010247184	4.56E-58
5	Ppnisotig08121	myb-related transcription partner of profilin-like	XP_001751813	1.64E-34
Defense and stress related genes				
2	Contig04363	wound-responsive family protein isoform 1	XP_001781658	2.12E-40
2	Contig20190	pathogenesis-related homeodomain	XP_002269263	2.93E-11
2	Contig22601	probable disease resistance protein at4g33300-like	XP_008377462	1.01E-20
2	Ppnisotig04165	late embryogenesis abundant protein d-34-like	XP_002965692	1.84E-64
2	Ppnisotig09289	pathogenesis-related homeodomain protein isoform x1	XP_009349156	1.46E-43
2	Ppnisotig10670	protein enhanced disease resistance 2-like	XP_013458134	3.73E-76
2	Ppnisotig10692	late embryogenesis abundant protein family protein	XP_011628884	1.58E-51
3	Ppnisotig11932	11 kda late embryogenesis abundant protein	XP_010275352	1.02E-06
4	Contig40036	peroxidase 55	XP_006399923	5.33E-08
5	Contig00602	defensin ec-amp-d2-like	XP_006647552	1.21E-13
5	Contig01697	heat shock 70 kda protein 8-like	XP_008375010	1.42E-150
5	Contig01723	heat stress transcription factor a-1-like	XP_008800101	6.86E-15

Table 4 (continued)

Cluster	Contig ID	BlastX Description	Hit ACC	eValue
5	Contig01763	heat shock protein binding	XP_004494547	2.88E-41
5	Contig03816	heat shock protein 83-like	XP_008801780	0
5	Contig03998	defensin ec-amp-d2-like	XP_006647552	1.58E-17
5	Contig05454	disease resistance protein	XP_006844417	3.09E-154
5	Contig09453	pathogenesis-related protein 5-like	XP_006847064	7.78E-12
5	Contig14467	heat shock 70 kda mitochondrial-like	XP_006471142	4.62E-17
5	Contig17617	defensin ec-amp-d2-like	XP_006647552	1.24E-17
5	Contig19071	heat shock protein 83-like	XP_006660235	1.89E-66
5	Contig20815	heat shock 70 kda mitochondrial-like	XP_001763121	3.91E-126
5	Contig23170	heat shock protein 83-like	XP_006290250	1.59E-42

was found in cluster 1, a HD-Zip class I member proposed to be related with abiotic stress response in *Arabidopsis thaliana* (Johannesson et al. 2003), but not related with meristem activity up to now. Different members of the HD-ZIP III subfamily group in these clusters. For instance, a putative ortholog of CORONA/ATHB15 factor has been found in this group; interestingly, this gene has been described in *Arabidopsis thaliana* (L.) Heynh. as a negative regulator of vascular tissue development under the control of miRNA165/166 (Ong and Wickneswari 2012). On the contrary, an miRNAresistant POPCORONA has been suggested to be involved in the upregulation of cellulose biosynthesis-related genes, but also with the down-regulation of lignin deposition-related genes in poplar (Du et al. 2011).

Finally, other genes coding for proteins involved in meristem activity have been detected as DEGs in these clusters. For instance, a putative WOX4-like protein was upregulated at the beginning of the growing season. WOX4 belongs to the WUSCHEL-related HOMEBOX (WOX) family, which is involved, for instance, in the regulation of proliferation from stem cell niches in root and shoot meristems after embryogenesis (Haecker et al. 2004), together with CLAVATA (CLV) (Miwa et al. 2009), and in the differentiation in the organizing center of the apical shoot (Mayer et al. 1998). WOX4 is involved in procambial and cambial growth, with function in vascular bundle development (Ji et al. 2010), which is consistent with our results.

It is noteworthy that up to 245 non-annotated contigs from *Pinus canariensis* C. Sm. ex DC are included in these clusters. Among them, 66 contigs are significantly co-expressed with genes involved in the development of the cellulosic matrix and in the synthesis of lignin, according to WGCNA analysis (Figure S2a and Table S4). However, they were not detected in the rather exhaustive analysis of the co-expression networks of NAC and MYB transcription factors during wood formation recently published in *Picea glauca* (Moench) Voss (Raherison et al. 2015; Lamara et al.

2016). They could correspond to genes expressed exclusively in *P. canariensis* or they could have been missed in previous works.

Another interesting point is that a considerable number of DEGs overexpressed during EW formation show high homology with genes reported to be involved in LW formation in other pine species. This is the case for 337 and 255 DEGs with *Pinus taeda* L. (Kirst et al. 2003) and *Pinus radiata* D. Don (Li et al. 2010) LW genes, respectively. Among them, genes related to proliferation and cell expansion, such as EXORDIUM or an expansin-like protein (contig 03225), and genes coding for proteins directly involved in cell wall formation, such as C4H (contig 05567), a CCoAMT (contig 00537) or a HCT (contig 20935), involved in lignin biosynthesis, a Cesa-like protein (contig 05245) or a KORRIGAN endoglucanase-like protein (contig 18777) or several CAZymes, can be found. This discrepancy is also found for several transcription factors, such as two bHLH-like transcription factors (contigs 04966 and 10039), a GATA-like transcription factor (contig 06513), a MYB-like transcription factor (contig 12050), a WOX4-like homeobox protein (contig 06813) or an ortholog of CORONA/ATHB15 (contig 06513).

DEGs involved in LW development

No homology was found for up to 287 of the DEGs showing a local maximum of transcript level in the summer (clusters 2, 4 and 5; Table 4). For the annotated DEGs, functional groups similar to those of EW were detected. Thus, an important group of these DEGs are involved in cell wall formation. For example, a putative sucrose phosphate synthase gene is included in cluster 5, as well as genes coding for expansins and genes involved in hemicelluloses (mainly xyloglucans) biosynthesis were included in cluster 2. This is consistent with the induction of

xyloglucan endotransglycosylases/hydrolases in dormant meristems of *Cryptomeria japonica* (Thunb. ex L.f.) D. Don (Mishima et al. 2014). In the same way, a CesA-like gene and several genes putatively coding for CAZymes, such as XET-like proteins, are present in cluster 4. DEGs related with the late steps of the lignin biosynthesis pathway, such as cinnamoyl-CoA reductase (CCR) also appear in these clusters. Considering the higher lignification degree of LW compared to EW, a greater abundance of DEGs related to lignin synthesis could have been expected in this group. Our results could reflect a special characteristic of *Pinus canariensis* C. Sm. ex DC juvenile wood. Additionally, the restrictiveness for the identification of DEGs can account for this result. In the same way, the presence of highly lignified compression wood, especially in the EW, could also hamper the detection of lignin genes as differentially overexpressed during LW formation. Although the trees did not appear noticeably inclined or distorted, this effect cannot be completely discarded; all the more, a small proportion of compression wood was detectable in certain sections (see, for instance, Figure 2a).

Among the transcription factors included in this group, some putative members of the MYB family and several members of the NAC family are detectable. The NAC (NAM/ATAC/CUC) family is supposed to be involved in multiple roles of plant development, as well as in responses to abiotic (drought, salinity, etc.) or biotic stress (reviewed by Hu et al. 2010). Recently, numerous NAC domains were reported to play crucial roles in xylogenesis, fiber development and wood formation in vascular plants, including conifers such as spruce (Duval et al. 2014; Raheison et al. 2015; Lamara et al. 2016). Transcripts encoding for NAC proteins were found in clusters 2 and 5. NAC78 (cluster 2) is putatively related to growth and development functionalities (Park et al. 2010b). Also in cluster 2, NAC29 has been reported to regulate CesA expression via MYB transcription factors in rice, and the interaction between NAC29/31 with SLR1 (SLENDER RICE 1) causes the inhibition of this regulatory pathway (Huang et al. 2015). A contig encoding for NAC2 was found as well in cluster 5. In *Capsicum annuum* L., a CaNAC2 belonging to the NAC2 subgroup was induced by cold, salt and abscisic acid treatments, suggesting its involvement in the response to abiotic stress (Guo et al. 2015). Also in this group, a DEG coding for a putative NAC74, contig 02274, which has not been associated to any functionality yet, was found.

As in cluster 1, a putative CORONA/ATHB15 coding gene is present in cluster 5. Consistently, Du et al. (2011) and Ong and Wickneswari (2012) reported this gene as being a negative regulator of xylem development in poplar and acacia, respectively. In a similar way, among

the WOX family, a WOX8-like gene was detected in cluster 2. WOX8 has been proposed in *Arabidopsis thaliana* (L.) Heynh. to have a specific role in the very early stages of embryo development, while it was not detected in postembryonic phases (Haecker et al. 2004). Palovaara et al. (2010) reported similar results for *Picea abies* (L.) H. Karst., where the homolog PaWOX8/9 is involved in embryo formation, differentiation and patterning. Analogously, this *Pinus canariensis* C. Sm. ex DC ortholog could be involved in the adjustments previous to cell proliferation, according to its transcription profile: it showed maximum transcript levels in March and July, prior to major meristem activity corresponding to the two periods of seasonal growth described in this work.

The three-amino-acid-loop-extension (TALE) homeodomain proteins are involved in the control of meristem activity (Traas and Vernoux 2002) and are classified into two family groups: the KNOTTED-like homeodomain (KNOX) and the BEL1-like homeodomain (BELL) subfamilies. Two transcripts coding for putative BEL1-like proteins (contigs 20447 and FLPpisotig03105) were grouped in cluster 5. A third member appears in cluster 3 (contig 04961), which shows a profile opposed to that of cluster 4. Other transcription factors found in cluster 4 were bZIP-like and zinc-finger protein motifs, as well as TOPLESS-like proteins, reported to repress root-promoting genes in the top half of the embryo in *Arabidopsis thaliana* (L.) Heynh. (Szemenyei et al. 2008).

EXORDIUM (EXO) is another relevant gene found in this work, previously described in proliferating cells (Farrar et al. 2003). First, it was suggested to be a negative regulator of cell division control and meristem maintenance. Further experiment revealed that EXO gene expression is brassinosteroid-dependent (Coll-Garcia et al. 2004) and mediates growth through cell expansion promoted by brassinosteroids (Schröder et al. 2009). A putative EXO-like gene is seen in cluster 2 and two more in cluster 4. However, two additional EXO-like proteins are present in cluster 1.

Several of the DEGs with a maximum transcript level in summer are putatively related to abiotic stresses, mainly drought, which is typical under Mediterranean summertime conditions. For instance, an ortholog gene encoding a putative WRKY4 was found in cluster 5. This gene has been reported to be induced by stress conditions in *Arabidopsis thaliana* (L.) Heynh. (Lai et al. 2008). Furthermore, many NAC transcription factors, mainly present in clusters 2 and 5, are also supposed to be involved in stress response.

The presence of stress-related genes in the LW clusters is consistent with the results of Mishima et al. (2014),

who described the abundance of “defense mechanism genes” in the “cessation of growth clusters” obtained for the cambial zone and differentiating xylem in *Cryptomeria japonica* (Thunb. ex L.f.) D. Don. LW differentiation is partly a consequence of a lower availability of water and resources at the end of the growing season, which detracts cell development compared to conditions for EW cell development in spring. Actually, lignin deposition and carbohydrate compounds involved in cell wall formation differ in EW and LW cells (Antonova et al. 2014). LW is also a constitutive, inner barrier for pathogen expansion. Therefore, it is not surprising that genes involved in stress response show high expression values during LW development.

WGCNA revealed another network of co-expressed LW DEGs including contigs related to ubiquitination and degradation of proteins and RNA, as well as two non-annotated contigs (Figure S2b). Among them, the following genes were found: one coding for an ubiquitin regulatory X (UBX) domain-containing protein, which is presumably involved in proteolysis (Vale 2000), one coding for an ubiquitin carboxyl-terminal hydrolase 13-like (UBP13), another one regulating protein degradation via lysosome and proteasome (Glickman and Ciechanover 2002) and a U-box domain containing protein 44-like (PUB44), which is presumably involved in ubiquitin ligation (Azevedo et al. 2001).

It is noteworthy that a low number of genes related to resin production were detected. Several genes involved in terpene biosynthesis are included in the transcriptome obtained from the libraries, such as abietadiene synthase, geranylgeranyl phosphate synthase, 1-deoxy-d-xylulose-5-phosphate reductoisomerase, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase and 3-hydroxy-3-methylglutaryl-CoA reductases. Only three contigs with this latter annotation were differentially overexpressed during LW formation. This may be partially explained by the construction of libraries, as those samples were collected in May and July. At this later point, resin ducts are differentiating, but they are not completely filled with resin. Other genes involved in resin production and secretion would be expressed later on and, therefore, are not included in the microarray, which was intended for xylem cell differentiation analysis.

As reported above, in the EW DEGs' section, a substantial number of *Pinus canariensis* C. Sm. ex DC LW DEGs were detected to be homologous to genes involved in EW formation in other pine species, such as *Pinus taeda* L. (181 DEGs) or *Pinus radiata* D. Don (90 DEGs). In this group, two different EXORDIUM-like genes (contigs 06152 and 09007), involved in cell proliferation, can be

seen, as well as genes related to cell wall biosynthesis and lignin deposition, such as CAZymes or a putative CCR-like gene (contig 18416). Several transcription factors are also in this group, namely a GATA-like (contig 02951) and a MYB-like transcription factor (contig 01331), as well as a putative KNOT-like (contig 05803) or a TIFY-8-like transcription factor (contig 11973), reported to act as a repressor of primary growth in roots of *Arabidopsis thaliana* (L.) Heynh. (Cuéllar Pérez et al. 2014). These discrepancies in the expression patterns among pine species could underlie the anatomic differences in their secondary xylem and their specific wood characteristics.

Finally, it is noteworthy that genes of the same family are found in clusters with opposite patterns. The abundance of duplications is typical for conifer genomes (Kovach et al. 2010; Lorenz et al. 2012), which can lead to the specialization of different closely related genes, as reported for WRKY transcription factors (Zhang and Wang 2005), terpene synthases (Trapp and Croteau 2001), etc. For instance, an ortholog of CORONA/ATHB15 is in cluster 1, but another ortholog is present in cluster 5, consistent with its role as an inhibitor of vascular tissue development in *Arabidopsis thaliana* (L.) Heynh. Another example is the presence of different genes putatively coding for expansins and hemicelluloses in EW- and LW-related DEGs, which would be consistent with differences between them not only in terms of cell expansion and wall thickness but also in cell wall composition throughout the growing season (Mellerowicz and Sundberg 2008).

Conclusions

To minimize the loss of less abundant transcripts, two normalized libraries were obtained from the cambial zone of young *Pinus canariensis* C. Sm. ex DC individuals, i.e. a species, in which axial parenchyma and resin ducts are comparatively abundant. Samples were collected in spring (during EW formation) and summer (when resin ducts and LW are formed). A high number of contigs without homology in the public databases have been identified. The transcription patterns during the growing season of approximately 16 500 unigenes, more than 300-bp long, have been analyzed. Three thousand three hundred and two genes showed significant variations in their transcript levels during this period. Approximately half of the genes displaying variable transcription levels during the growing season (DEGs) show a transcript level peak in spring, and are putatively involved in EW formation. The other DEGs show a peak during summer, when

xylem growth decreases and lignification begins. Most of the DEGs in both groups are involved in cell wall biosynthesis. Interestingly, some genes presumably related to abiotic stress appear among the LW DEGs, which is related to the summer drought typical of the Mediterranean climate. Expectedly, several genes with antagonistic effects are found in both EW and LW. The presence of closely related DEGs with contrasting and even opposite expression profiles suggests the specialization of diverse members of gene families, involved in the different composition of EW and LW. Of note, results for *P. canariensis* C. Sm ex DC show several discrepancies with those obtained for other species, even within the Pinaceae group. The differences may be partly species-specific. Additionally, no homology was found in Viridiplantae databases for more than 1000 DEGs identified in *P. canariensis* C. Sm. ex DC. Future studies on cell-type-specific expression patterns should give further insight on conifer wood formation. Moreover, the transcriptome reported here will constitute valuable information for further investigation of relevant issues such as the formation of traumatic wood in conifers.

Acknowledgments: We want to thank the two anonymous reviewers and the editor for their helpful comments on the manuscript. We also want to thank Dr. Pedro Perdiguero for his help in collecting samples for the transcriptome construction and for his comments and suggestions. We also thank Dr. J.L. Vázquez-Poletti for computational support in transcriptome assembly. This work was supported by the projects AGL2009-10606 (Spanish Ministry of Science and Innovation), SPIP2014-01093 (Spanish National Parks Agency, Ministry of Agriculture) and FP7-KBBE-2011-5 289841. VC had a predoctoral fellowship granted by the Spanish Ministry of Science and Innovation.

Author contribution

AS and CC designed and supervised the experiments and collected the samples for the construction of libraries. ULH performed the *de novo* assembly of the transcriptome. ULH and VC annotated the transcriptome. VC carried out the expression analysis. VC and AS drafted the manuscript.

References

- Allona, I., Quinn, M., Shoop, E., Swope, K., St Cyr, S., Carlis, J., Riedl, J., Retzel, E., Campbell, M.M., Sederoff, R., Whetten, R.W. (1998) Analysis of xylem formation in pine by cDNA sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 95:9693–9698.
- Altschul, S., Gish, W., Miller, W., Myers, E., Lipman, D. (1990) Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.
- Andersson-Gunnerås, S., Mellerowicz, E.J., Love, J., Segerman, B., Ohmiya, Y., Coutinho, P.M., Nilsson, P., Henrissat, B., Moritz, T., Sundberg, B. (2006) Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. *Plant J.* 45:144–165.
- Antonova, G.F., Varaksina, T.N., Zheleznicenko, T.V., Stasova, V.V. (2014) Lignin deposition during earlywood and latewood formation in Scots pine stems. *Wood Sci. Technol.* 48:919–936.
- Azevedo, C., Santos-Rosa, M.J., Shirasu, K. (2001) The U-box protein family in plants. *Trends Plant Sci.* 6:354–358.
- Barker, M.S., Dlugosch, K.M., Dinh, L., Sashikiran Challa, R., Kane, N.C., King, M.G., Rieseberg, L.H. (2010) Evopipes.net: Bioinformatic tools for ecological and evolutionary genomics. *Evol. Bioinform.* 2010:143–149.
- Blüthgen, N., Kiebas, S.M., Herzel, H. (2005) Inferring combinatorial regulation of transcription *in silico*. *Nucleic Acids Res.* 33:272–279.
- Bobinac, M., Batos, B., Miljković, D., Radulović, S. (2012) Polycyclism and phenological variability in the common oak (*Quercus robur* L.). *Arch. Biol. Sci.* 64:97–105.
- Bushmanova, E., Antipov, D., Lapidus, A., Suvorov, V., Prjibelski, A.D. (2016) rnaQUAST: a quality assessment tool for *de novo* transcriptome assemblies. *Bioinformatics* 32:2210–2212.
- Chang, S., Puryear, J., Cairney, J. (1993) A simple and efficient method for isolating RNA from pine trees. *Plant Mol. Biol. Rep.* 11:113–116.
- Chano, V., López, R., Pita, P., Collada, C., Soto, Á. (2015) Proliferation of axial parenchymatic xylem cells is a key step in wound closure of girdled stems in *Pinus canariensis*. *BMC Plant Biol.* 15:64.
- Cheng, H., Song, S., Xiao, L., Soo, H.M., Cheng, Z., Xie, D., Peng, J. (2009) Gibberellin acts through jasmonate to control the expression of MYB21, MYB24, and MYB57 to promote stamen filament growth in *Arabidopsis*. *PLoS Genet.* 5:20–22.
- Climent, J., Gil, L., Pardos, J.A. (1998) Xylem anatomical traits related to resinous heartwood formation in *Pinus canariensis* Sm. *Trees* 12:139–145.
- Coll-García, D., Mazuch, J., Altmann, T., Müssig, C. (2004) EXOR-DIUM regulates brassinosteroid-responsive genes. *FEBS Lett.* 563:82–86.
- Conesa, A., Götz, S., García-Gómez, J.M., Terol, J., Talón, M., Robles, M. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21:3674–3676.
- Cosgrove, D.J., Li, L.C., Cho, H.-T., Hoffmann-Benning, S., Moore, R.C., Blecker, D. (2002) The growing world of expansins. *Plant Cell Physiol.* 43:1436–1444.
- Cuellar Pérez, A., Nagels Durand, A., Vanden Bossche, R., De Clercq, R., Persiau, G., Van Wees, S.C.M., Pieterse, C.M.J., Gevaert, K., De Jaeger, G., Goossens, A., Pauwels, L. (2014) The non-JAZ TIFY protein TIFY8 from *Arabidopsis thaliana* is a transcriptional repressor. *PLoS One* 9:e84891.
- de Luis, M., Gričar, J., Cufar, K., Raventós, J. (2007) Seasonal dynamics of wood formation in *Pinus halepensis* from dry and semi-arid ecosystems in Spain. *IAWA J.* 28:389–404.

- Du, J., Miura, E., Robischon, M., Martinez, C., Groover, A.T. (2011) The *Populus* Class III HD ZIP transcription factor POPCORONA affects cell differentiation during secondary growth of woody stems. *PLoS One* 6:1–14.
- Duval, I., Lachance, D., Giguère, I., Bomal, C., Morency, M.J., Pelletier, G., Boyle, B., MacKay, J.J., Séguin, A. (2014) Large-scale screening of transcription factor-promoter interactions in spruce reveals a transcriptional network involved in vascular development. *J. Exp. Bot.* 65:2319–2333.
- Farrar, K., Evans, I.M., Topping, J.F., Souter, M.A., Nielsen, J.E., Lindsey, K. (2003) EXORDIUM – A gene expressed in proliferating cells and with a role in meristem function, identified by promoter trapping in *Arabidopsis*. *Plant J. Cell Mol. Biol.* 33:61–73.
- Glickman, M.H., Ciechanover, A. (2002) The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* 82:373–428.
- Guo, W.-L., Wang, S.-B., Chen, R.-G., Chen, B.-H., Du, X.-H., Yin, Y.-X., Gong, Z.-H., Zhang, Y.-Y. (2015) Characterization and expression profile of CaNAC2 pepper gene. *Front. Plant Sci.* 6:755.
- Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G. (2013) QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075.
- Haecker, A., Gross-Hardt, R., Geiges, B., Sarkar, A., Breuninger, H., Herrmann, M., Laux, T. (2004) Expression dynamics of WOX genes mark cell fate decisions during early embryonic patterning in *Arabidopsis thaliana*. *Development* 131:657–668.
- Huang, X., Madan, A. (1999) CAP3: a DNA sequence assembly program. *Genome Res.* 9:868–877.
- Huang, D., Wang, S., Zhang, B., Shang-Guan, K., Shi, Y., Zhang, D., Liu, X., Wu, K., Xu, Z., Fu, X., Zhou, Y. (2015) A gibberellin-mediated DELLA-NAC signaling cascade regulates cellulose synthesis in rice. *Plant Cell* 27:1681–1696.
- Hu, R., Qi, G., Kong, Y., Kong, D., Gao, Q., Zhou, G. (2010) Comprehensive analysis of NAC domain transcription factor gene family in *Populus trichocarpa*. *BMC Plant Biol.* 10:145.
- Ji, J., Strable, J., Shimizu, R., Koenig, D., Sinha, N., Scanlon, M.J. (2010) WOX4 promotes procambial development. *Plant Physiol.* 152:1346–1356.
- Johannesson, H., Wang, Y., Hanson, J., Engström, P. (2003) The *Arabidopsis thaliana* homeobox gene *ATHB5* is a potential regulator of abscisic acid responsiveness in developing seedlings. *Plant Mol. Biol.* 51:719–729.
- Kirst, M., Johnson, A.F., 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*. *Proc. Natl. Acad. Sci. U.S.A.* 100:7383–7388.
- Kovach, A., Wegrzyn, J.L., Parra, G., Holt, C., Bruening, G.E., Loopstra, C.A., Hartigan, J., Yandell, M., Langley, C.H., Korf, I., Neale, D.B. (2010) The *Pinus taeda* genome is characterized by diverse and highly diverged repetitive sequences. *BMC Genomics* 11:420.
- Lai, Z., Vinod, K., Zheng, Z., Fan, B., Chen, Z. (2008) Roles of *Arabidopsis* WRKY3 and WRKY4 transcription factors in plant responses to pathogens. *BMC Plant Biol.* 8:68.
- Lamara, M., Raheison, E., Lenz, P., Beaulieu, J., Bousquet, J., MacKay, J. (2016) Genetic architecture of wood properties based on association analysis and co-expression networks in white spruce. *New Phytol.* 210:240–255.
- Langfelder P, Horvath S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559.
- Le Provost, G., Paiva, J., Pot, D., Brach, J., Plomion, C. (2003) Seasonal variation in transcript accumulation in wood-forming tissues of maritime pine (*Pinus pinaster* Ait.) with emphasis on a cell wall glycine-rich protein. *Planta* 217:820–830.
- Li, X., Wu, H., Dillon, S., Southerton, S. (2009) Generation and analysis of expressed sequence tags from six developing xylem libraries in *Pinus radiata* D. Don. *BMC Genomics* 10:41.
- Li, X., Wu, H.X., Southerton, S.G. (2010) Seasonal reorganization of the xylem transcriptome at different tree ages reveals novel insights into wood formation in *Pinus radiata*. *New Phytol.* 187:764–776.
- Li, W.F., Yang, W.H., Zhang, S.G., Han, S.Y., Qi, L.W. (2017) Transcriptome analysis provides insights into wood formation during larch tree aging. *Tree Genet. Genomes* 13:19.
- Lipshitz, N., Lev-Yadun, S., Rosen, E., Waisel, Y. (1984) The annual rhythm of activity of the lateral meristems (cambium and phellogen) in *Pinus halepensis* Mill. and *Pinus pinea* L. *IAWA Bull.* 5:263–274.
- Liu, R., Lü, B., Wang, X., Zhang, C., Zhang, S., Qian, J., Chen, L., Shi, H., Dong, H. (2010) Thirty-seven transcription factor genes differentially respond to a harpin protein and affect resistance to the green peach aphid in *Arabidopsis*. *J. Biosci.* 35:435–450.
- López, R., Brossa, R., Gil, L., Pita, P. (2015) Stem girdling evidences a trade-off between cambial activity and sprouting and dramatically reduces plant transpiration due to feedback inhibition of photosynthesis and hormone signaling. *Front. Plant Sci.* 6:285.
- López de Heredia, U., Vázquez-Poletti, J.L. (2016) RNA-seq analysis in forest tree species: bioinformatic problems and solutions. *Tree Genet. Genomes* 12:30.
- Lorenz, W.W., Ayyampalayam, S., Bordeaux, J.M., Howe, G.T., Jermstad, K.D., Neale, D.B., Rogers, D.L., Dean, J.F.D. (2012) Conifer DBMagic: a database housing multiple de novo transcriptome assemblies for 12 diverse conifer species. *Tree Genet. Genomes* 8:1477–1485.
- Marín-Rodríguez, M.C., Orchard, J., Seymour, G.B. (2002) Pectate lyases, cell wall degradation and fruit softening. *J. Exp. Bot.* 53:2115–2119.
- Mayer, K., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., Laux, T. (1998) Role of WUSCHEL in regulating stem cell fate in the *Arabidopsis* shoot meristem. *Cell* 95:805–815.
- Medina, I., Carbonell, J., Pulido, L., Madeira, S.C., Goetz, S., Conesa, A., Tárraga, J., Pascual-Montano, A., Nogales-Cadenas, R., Santoyo, J., García, F., Marbà, M., Montaner, D., Dopazo, J. (2010) Babelomics: an integrative platform for the analysis of transcriptomics, proteomics and genomic data with advanced functional profiling. *Nucleic Acids Res.* 38(Web Server Issue):W210–213.
- Mellerowicz, E.J., Sundberg, B. (2008) Wood cell walls: biosynthesis, developmental dynamics and their implications for wood properties. *Curr. Opin. Plant Biol.* 11:293–300.
- Mishima, K., Fujiwara, T., Iki, T., Kuroda, K., Yamashita, K., Tamura, M., Fujisawa, Y., Watanabe, A. (2014) Transcriptome sequencing and profiling of expressed genes in cambial zone and

- differentiating xylem of Japanese cedar (*Cryptomeria japonica*). *BMC Genomics* 15:219.
- Miwa, H., Kinoshita, A., Fukuda, H., Sawa, S. (2009) Plant meristems: CLAVATA3/ESR-related signaling in the shoot apical meristem and the root apical meristem. *J. Plant Res.* 122:31–39.
- Mutke, S., Gordo, J., Khouja, M.L., Fady, B. (2013) Low genetic and high environmental diversity at adaptive traits in *Pinus pinea* from provenance tests in France and Spain. *Options Méditerranéennes* 79:72–79.
- Nystedt, B., Street, N.R., Wetterbom, A., Zuccolo, A., Lin, Y.-C., Scofield, D.G., Vezzi, F., Delhomme, N., Giacomello, S., Alexeyenko, A., Vicedomini, R., Sahlin, K., Sherwood, E., Elfstrand, M., Gramzow, L., Holmberg, K., Hällman, J., Keech, O., Klasson, L., Koriabine, M., Kucukoglu, M., Käller, M., Luthman, J., Lysholm, F., Niittylä, T., Olson, A., Rilakovic, N., Ritland, C., Rosselló, J. a, Sena, J., Svensson, T., Talavera-López, C., Theißen, G., Tuominen, H., Vanneste, K., Wu, Z.-Q., Zhang, B., Zerbe, P., Arvestad, L., Bhalerao, R., Bohlmann, J., Bousquet, J., Garcia Gil, R., Hvidsten, T.R., de Jong, P., MacKay, J., Morgante, M., Ritland, K., Sundberg, B., Thompson, S.L., Van de Peer, Y., Andersson, B., Nilsson, O., Ingvarsson, P.K., Lundeberg, J., Jansson, S. (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature* 497:579–584.
- Ong, S.S., Wickneswari, R. (2012) Characterization of microRNAs expressed during secondary wall biosynthesis in *Acacia mangium*. *PLoS One* 7:1–13.
- Palovaara, J., Hallberg, H., Stasolla, C., Hakman, I. (2010) Comparative expression pattern analysis of WUSCHEL-related homeobox 2 (WOX2) and WOX8/9 in developing seeds and somatic embryos of the gymnosperm *Picea abies*. *New Phytol.* 188:122–135.
- Parchman, T.L., Geist, K.S., Grahnen, J.A., Benkman, C.W., Buerkle, C.A. (2010) Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics* 11:180.
- Park, B.H., Karpinet, T.V., Syed, M.H., Leuze, M.R., Uberbacher, E.C. (2010a) CAZymes Analysis Toolkit (CAT): Web service for searching and analyzing carbohydrate-active enzymes in a newly sequenced organism using CAZy database. *Glycobiology* 20:1574–1584.
- Park, M.R., Yun, K.Y., Mohanty, B., Herath, V., Xu, F., Wijaya, E., Bajic, V.B., Yun, S.J., De Los Reyes, B.G. (2010b). Supra-optimal expression of the cold-regulated OsMyb4 transcription factor in transgenic rice changes the complexity of transcriptional network with major effects on stress tolerance and panicle development. *Plant Cell Environ.* 33:2209–2230.
- Paux, E., Tamasloukht, M.B., Ladouce, N., Sivadon, P., Grima-Pettenati, J. (2004) Identification of genes preferentially expressed during wood formation in *Eucalyptus*. *Plant Mol. Biol.* 55:263–280.
- Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29:e45.
- Pinosio, S., González-Martínez, S.C., Bagnoli, F., Cattonaro, F., Grivet, D., Marroni, F., Lorenzo, Z., Pausas, J.G., Verdú, M., Vendramin, G.G. (2014) First insights into the transcriptome and development of new genomic tools of a widespread circum-Mediterranean tree species, *Pinus halepensis* Mill. *Mol. Ecol. Resour.* 14:846–856.
- Raherison, E., Giguère, I., Caron, S., Lamara, M., MacKay, J. (2015) Modular organization of the white spruce (*Picea glauca*) transcriptome reveals functional organization and evolutionary signatures. *New Phytol.* 207:172–187.
- Sandak, J., Sandak, A., Cantini, C., Autino, A. (2015) Differences in wood properties of *Picea abies* L. Karst. in relation to site of provenance and population genetics. *Holzforschung* 69:358–397.
- Schmieder, R., Edwards, R. (2011) Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864.
- Schröder, F., Lisso, J., Lange, P., Müssig, C. (2009) The extracellular EXO protein mediates cell expansion in *Arabidopsis* leaves. *BMC Plant Biol.* 9:20.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13:2498–2504.
- Smyth, G.K., Speed, T. (2003) Normalization of cDNA microarray data. *Methods* 31:265–273.
- 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, T.T., Boerjan, W., Gustafsson, P., Uhlén, 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. U.S.A.* 95:13330–13335.
- Szemenyei, H., Hannon, M., Long, J.A. (2008) TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. *Science* 319:1384–1386.
- Tamagnone, L., Merida, A., Parr, A., Mackay, S., Culianez-Macia, F., Roberts, K., Martin, C. (1998) The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10:135–154.
- Taylor, A.M., Gartner, B.L., Morrell, J.J. (2002) Heartwood formation and natural durability – A review. *Wood Fiber Sci.* 34:587–611.
- Traas, J., Vernoux, T. (2002) The shoot apical meristem: the dynamics of a stable structure. *Philos. Trans. R. Soc. B-Biol. Sci.* 357:737–747.
- Trapp, S.C., Croteau, R.B. (2001) Genomic organization of plant terpene synthases and molecular evolutionary implications. *Genetics* 158:811–832.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., Rozen, S.G. (2012) Primer3-new capabilities and interfaces. *Nucleic Acids Res.* 40:e115.
- Vale, R.D. (2000) AAA proteins: lords of the ring. *J. Cell Biol.* 150:F13–F19.
- Vanholme, R., Demedts, B., Morreel, K., Ralph, J., Boerjan, W. (2010) Lignin biosynthesis and structure. *Plant Physiol.* 153:895–905.
- Verdú, M., Climent, J. (2007) Evolutionary correlations of polycyclic shoot growth in *Acer* (Sapindaceae). *Am. J. Bot.* 94:1316–1320.
- Vieira, J. (2013) Cambial activity and wood formation of maritime pine in a drought-prone environment: the effect of growth rate, size and climate. Doctoral Thesis. Universidade de Coimbra.
- Yamashita, S., Yoshida, M., Yamamoto, H., Okuyama, T. (2008) Screening genes that change expression during compres-

- sion wood formation in *Chamaecyparis obtusa*. *Tree Physiol.* 28:1331–1340.
- Zhang, Y., Wang, L. (2005) The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evol. Biol.* 5:1.
- Zhang, J., Elo, A., Helariutta, Y. (2011) *Arabidopsis* as a model for wood formation. *Curr. Opin. Biotechnol.* 22:293–299.
- Zhong, R., Richardson, E.A., Ye, Z.H. (2007) The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in *Arabidopsis*. *Plant Cell* 19:2776–2792.
- Zhu, Y., Machleder, E., Chenchik, A., Li, R., Siebert, P. (2001) Reverse transcriptase template switching: a SMART (TM) approach for full-length cDNA library construction. *Biotechniques* 30:892–897.
- Zhulidov, P.A., Bogdanova, E.A., Shcheglov, A.S., Vagner, L.L., Khaspekov, G.L., Kozhemyako, V.B., Matz, M. V, Meleshkevitch, E., Moroz, L.L., Lukyanov, S.A., Shagin, D.A. (2004) Simple cDNA normalization using kamchatka crab duplex-specific nuclease. *Nucleic Acids Res.* 32:e37.

Supplemental Material: The online version of this article (DOI: 10.1515/hf-2017-0014) offers supplementary material, available to authorized users.