Plant Physiology and Biochemistry 50 (2012) 44–53

Contents lists available at SciVerse ScienceDirect

Plant Physiology and Biochemistry

journal homepage: www.elsevier.com/locate/plaphy

Research article

Identification of water stress genes in *Pinus pinaster* Ait. by controlled progressive stress and suppression-subtractive hybridization

Pedro Perdiguero [a,b], Carmen Collada [a,b], María del Carmen Barbero [a,b], Gloria García Casado [c], María Teresa Cervera [a,b,d], Álvaro Soto [a,b,*]

[a] GENFOR Grupo de investigación en Genética y Fisiología Forestal, Universidad Politécnica de Madrid, E-28040 Madrid, Spain
[b] Unidad Mixta de Genómica y Ecofisiología Forestal, INIA/UPM, Madrid, Spain
[c] Centro Nacional de Biotecnología, CNB, Madrid, Spain
[d] Departamento de Ecología y Genética Forestal, CIFOR-INIA, Madrid, Spain

ARTICLE INFO

Article history:
Received 18 July 2011
Accepted 30 September 2011
Available online 8 October 2011

Keywords:
Pinus pinaster
Drought stress response
SSH
Microarray
Polyethylene glycol

ABSTRACT

Climate change is a major challenge particularly for forest tree species, which will have to face the severe alterations of environmental conditions with their current genetic pool. Thus, an understanding of their adaptive responses is of the utmost interest.

In this work we have selected *Pinus pinaster* as a model species. This pine is one of the most important conifers (for which molecular tools and knowledge are far more scarce than for angiosperms) in the Mediterranean Basin, which is characterised in all foreseen scenarios as one of the regions most drastically affected by climate change, mainly because of increasing temperature and, particularly, by increasing drought.

We have induced a controlled, increasing water stress by adding PEG to a hydroponic culture. We have generated a subtractive library, with the aim of identifying the genes induced by this stress and have searched for the most reliable expressional candidate genes, based on their overexpression during water stress, as revealed by microarray analysis and confirmed by RT-PCR.

We have selected a set of 67 candidate genes belonging to different functional groups that will be useful molecular tools for further studies on drought stress responses, adaptation, and population genomics in conifers, as well as in breeding programs.

1. Introduction

Climate change is a highly topical question nowadays. In view of the current climatic forecasts for the next hundred years, the need to study the adaptive responses of living organisms is broadly acknowledged among the scientific community. This is particularly relevant for tree species, which are sessile individuals and thus cannot flee from adverse conditions. Moreover, because of their longevity, similar to the predicted time-frame for climate change, they will have to face these perturbations with the same genetic makeup they currently possess.

Conifers, which represent approximately 34% of the world's forests (up to 61%, including mixed forest) and 60% of plantations for wood production [1], display peculiar characteristics that make the study of their genetic adaptations difficult. For instance, they usually have huge genomes, with a high percentage of repeated sequences and pseudogenes whose functions are not well known. As an illustration, whereas the genomes of *Arabidopsis thaliana* and *Populus trichocarpa* are approximately 150 and 550 Mbp long, respectively, the *Pinus pinaster* genome is about 30,900 Mbp, 70–75% of which is made up of highly repeated sequences [2]. Furthermore, angiosperms diverged from gymnosperms more than 300 million years ago. Thus, to a large extent, the knowledge and molecular tools developed for the former are not fully applicable to the latter. For these reasons, the selection of candidate genes for the study of adaptation in gymnosperms based solely on their homology with angiosperm genes, without further confirmation of their participation in the stress response, is not fully reliable.

The aim of this study was to identify and select for genes involved in the response to water stress in conifers. We used the maritime pine (*P. pinaster* Ait.), which is one of the most common conifer species in the western Mediterranean basin, as a model

* Corresponding author. GENFOR, G.I. Genética y Fisiología Forestal, Universidad Politécnica de Madrid, Ciudad Universitaria s/n, E-28040 Madrid, Spain. Tel.: +34 91 336 63 92; fax: +34 91 336 55 56.
E-mail address: alvaro.soto.deviana@upm.es (Á. Soto).

0981-9428/$ – see front matter © 2011 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.plaphy.2011.09.022

species. According to current models, this region will be dramatically affected by climate change, with an increase of 3.0–3.5 °C in the annual average temperature and a decrease of 10–20% in the annual rainfall, with increasingly long and more intense periods of drought [3]. Mediterranean species face an additional difficulty because the most favourable physiological period in terms of light and temperature, coincides with the period of maximum water restriction.

Notwithstanding its relatively small geographical range, the maritime pine has a large ecological amplitude, particularly in relation to rainfall regimes and water availability. Several works have revealed striking differences among maritime pine populations in growth and survival under drought-induced stress, and different strategies to face drought ([4,5] and references therein). These reasons make the maritime pine a uniquely suitable model species for the analysis of adaptation to drought-induced stress.

Genomic studies of the response to drought have been launched in different pine species. For instance, Heath et al. [6] and Watkinson et al. [7] performed a preliminary analysis of gene expression during drought-induced stress in *Pinus taeda*. These authors used microarrays that included cDNAs obtained from pre-existing libraries from the xylem, male cones and shoot tips, but did not attempt to exhaustively identify genes induced by drought-induced stress.

In addition, the comparison of EST libraries generated from the roots of loblolly pines during drought-induced stress and drought recovery as well as from well-watered roots led to the identification of 24 transcripts that were significantly induced by drought [8]. Very recently, this research group used a microarray to analyse the expression of 25,848 genes obtained from different libraries in the roots of *P. taeda* after seven days of withholding water and after two days of recovery. They identified up to 2445 genes that were responsive to drought-induced stress in this organism [9].

Sathyan et al. [10] generated an EST library from the roots of *Pinus halepensis* during drought-induced stress, and a preliminary expression analysis using a macroarray technique detected 156 transcripts that were likely to be induced by drought. This result was confirmed by qRT-PCR for 14 genes.

In *P. pinaster*, Dubos et al. [11,12] used the cDNA–AFLP technique to identify 48 differentially expressed fragments in the needles and roots of seedlings subjected to a mild water stress, checking their expression by reverse northern blot. Recently, 13 *P. pinaster* and *P. taeda* candidate genes selected from these works were evaluated for genetic variation in *P. pinaster* to detect patterns of selection and adaptation to local climatic conditions [13].

In this work we report the identification of 67 candidate genes (including 37 that had not been described in previous water stress studies), obtained from an SSH library from plants subjected to a progressive, controlled water deficit, and that were selected based on their expression patterns during this stress, analysed by microarrays and confirmed by RT-PCR. This set of genes constitutes a valuable molecular tool for further studies on adaptation to drought-induced stress in gymnosperms with important implications for breeding programs.

2. Results and discussion

2.1. Construction and differential screening of a subtracted cDNA library

A cDNA library enriched in genes that are induced by water stress was constructed, with the aim of identifying and characterising genes involved in the drought response in *P. pinaster* Ait. Our work differs from previous analyses of the water stress response in pines in terms of experimental design and methodology. First, we applied a progressive and more severe stress that is more similar to the first steps of natural drought-induced stress, with the aim of detecting genes involved in the different response pathways initiated by different levels of stress, as suggested by Watkinson et al. (2003). Although the inhibition of transcription shares importance

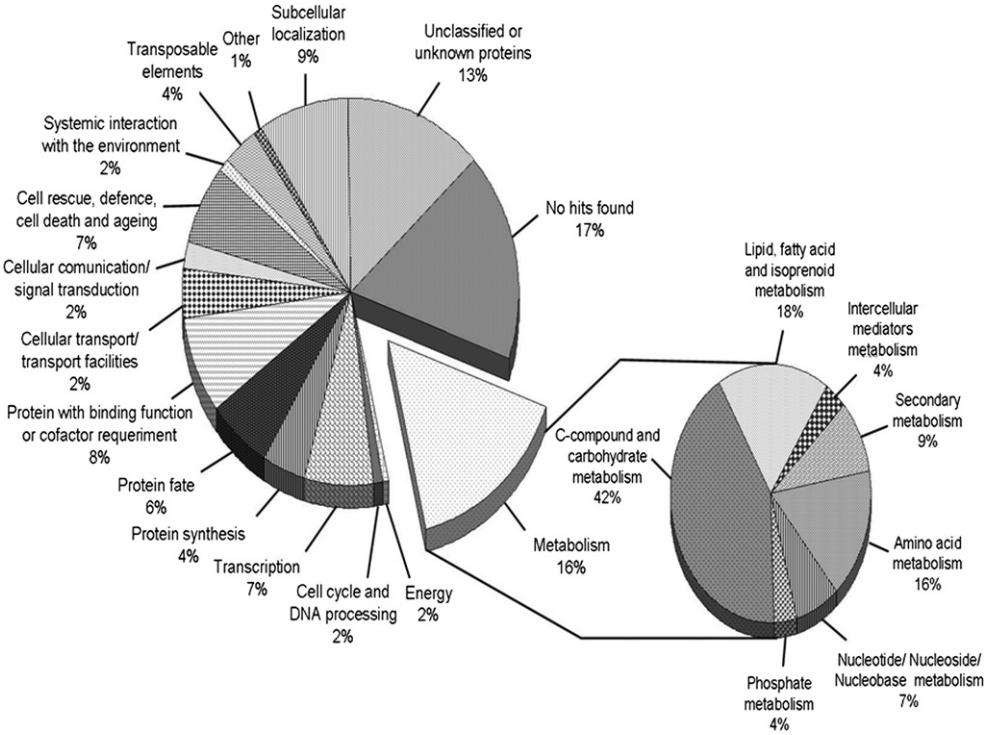

Fig. 1. Funtional distribution of genes obtained from the SSH library. In total, 351 unique ESTs were grouped according to MIPS functional categories for *Arabidopsis thaliana*. The percentage of gene transcripts in each group is listed.

Table 1
Functional classification of the 67 selected candidate genes induced during the treatment.

ID_sequence	Accession No.	Annotation
1. Metabolism		
1.1. *Amino Acid Metabolism*		
Ppter_DR_61	FR846597	Prephenate Dehydrogenase Family Protein
1.4. *Phosphate metabolism*		
Ppter_DR_227	FR846763	Pyrophosphorylase 4
1.5. *C-compound and carbohydrate metabolism*		
Ppter_DR_4	FR846540	Aldehyde Dehydrogenase 7B4
Ppter_DR_33	FR846569	Seed Imbibition 2
Ppter_DR_52	FR846588	Quitinase IV
Ppter_DR_99	FR846635	Glyceraldehyde-3-Phospate Dehydrogenase
Ppter_DR_101	FR846637	Glycosyltransferase Family 14 Protein
Ppter_DR_120	FR846656	Seed Imbibition 2
Ppter_DR_276	FR846812	Malate Synthase
Ppter_DR_305	FR846841	Beta-Galactosidase 3
1.6. *Lipid, fatty acid and isoprenoid metabolism*		
Ppter_DR_48	FR846584	Lipase_3
Ppter_DR_90	FR846626	Myo-Inositol-1-Phostpate Synthase 2
Ppter_DR_244	FR846780	Sugar-Dependent 1
Ppter_DR_278	FR846814	Cytochrome P450 Like_TPB
1.7. *Metabolism of intercellular mediators*		
Ppter_DR_180	FR846716	Lesion Initiation 2; Copropophyrogen Oxidase
1.20. *Secondary metabolism*		
Ppter_DR_77	FR846613	Acc Oxidase 2
Ppter_DR_274	FR846810	12-Oxophytodienoate Reductase
11. Transcription		
Ppter_DR_213	FR846749	NAC Domain Containing Protein 28
Ppter_DR_254	FR846790	AP2 Domain Containing Transcription Factor
Ppter_DR_296	FR846832	DEAD-Box Rna Helicase
Ppter_DR_315	FR846851	BEL1-Like Homeodomain 1
Ppter_DR_319	FR846855	AP2 Domain Containing Transcription Factor
14. Protein fate		
Ppter_DR_136	FR846672	DNAJ Heat Shock N-Terminal Domain Containing Protein
Ppter_DR_287	FR846823	DEGP Protease 9
16. Protein with binding function or cofactor requirement		
Ppter_DR_13	FR846549	Accelerated Cell Death 1
Ppter_DR_242	FR846778	Unusual Floral Organs
Ppter_DR_275	FR846811	(R)-Mandelonitrile Lyase, Putative/(R)-Oxynitrilase, Putative
20. Transport		
Ppter_DR_123	FR846659	Inositol Transporter 2
Ppter_DR_162	FR846698	Pleiotropic Drug Resistance 12
Ppter_DR_187	FR846723	Farnesylated Protein 3
Ppter_DR_239	FR846775	Peroxisomal Adenine Nucleotide Carrier 1
Ppter_DR_329	FR846865	General Control Non-Repressible 1
30. Cellular communication/signal transduction		
Ppter_DR_137	FR846673	WD 40 Repeat Family Protein
Ppter_DR_198	FR846734	CBL-Interacting Protein Kinase 20
32. Cell rescue, defence, cell death and ageing		
Ppter_DR_41	FR846577	Nonphotochemical Quenching
Ppter_DR_72	FR846608	Late Embryogenesis Abundant 14
Ppter_DR_106	FR846642	Tau Class Glutatione S-Transferase
Ppter_DR_152	FR846688	Heat Shock Protein 90.5
Ppter_DR_159	FR846695	Phenylalanine Ammonia-Lyase 2
Ppter_DR_235	FR846771	Heat Shock Protein 90
32. Systemic interaction with the environment		
Ppter_DR_262	FR846798	Aluminium Induced LP1
38. Transposables elements		
Ppter_DR_37	FR846573	Retrotransposon Protein
40. Cell fate		
Ppter_DR_22	FR846558	Tetraspanin 3
41. Plant development		
Ppter_DR_95	FR846631	Late Embryogenesis Abundant Protein, Putative
70. Subcellular localisation		
Ppter_DR_44	FR846580	Metallo-Beta-Lacmatasa
Ppter_DR_80	FR846616	Unknown
Ppter_DR_334	FR846870	Unknown
99. Unclassified or unknown proteins		
Ppter_DR_6	FR846542	Unknown
Ppter_DR_51	FR846587	Plasmodesmata-Located Protein 8
Ppter_DR_115	FR846651	Nodulin MtN3 Family Protein
Ppter_DR_122	FR846658	Unknown
Ppter_DR_160	FR846696	Unknown
Ppter_DR_163	FR846699	Nodulin MtN3 Family Protein
Ppter_DR_194	FR846730	Unknown
Ppter_DR_283	FR846819	Dormancy/Auxin Associated Family Protein
Ppter_DR_303	FR846839	Acyl-Activating Enzyme 18
Ppter_DR_341	FR846877	Unknown
Ppter_DR_342	FR846878	Oxidoreductase, 2OG-Fe(II) Oxigenase Family Protein
Ppter_DR_349	FR846885	Unknown
No hits found		
Ppter_DR_63	FR846599	No Hits Found
Ppter_DR_107	FR846643	No Hits Found
Ppter_DR_112	FR846648	No Hits Found
Ppter_DR_200	FR846736	No Hits Found
Ppter_DR_223	FR846759	No Hits Found
Ppter_DR_284	FR846820	No Hits Found
Ppter_DR_307	FR846843	No Hits Found
Ppter_DR_327	FR846863	No Hits Found

with transcriptional upregulation during acclimation to stress, as indicated by the results of Watkinson et al., (2003) and Lorenz et al., (2005), our goal in this study was to identify reliable candidate genes induced by water deficit. For this purpose, we generated an SSH library using clonal material, which is a more exhaustive technique for the isolation of such candidate genes than has been used in previous works. We picked 4940 colonies that were transformed with the forward-subtractive PCR and amplified the inserts using nested primers. In all cases, the amplified inserts were more than 200 bp long. Differential screening by macroarray led us to the identification of 1718 clones that were presumably induced by water stress.

2.2. Annotation and functional classification of ESTs

We proceeded with the sequencing of the 1718 candidate clones. In total, 1099 good sequences were clustered into 386 unigenes. We used these nucleotide sequences and their translated amino acid sequences for annotation by performing homology searches with non-redundant databases from NCBI and the Gene Index (Pine and Spruce). Thirty-five unigenes showed a significant homology with plastid sequences. Of the 351 presumably nuclear sequences (included in the EST database of EMBL with accession numbers FR846537 to FR846887, as well as in EuroPineDB [14]; Supplementary Table S1), 124 have been reported in previous works that focused on water stress in *Pinus*, whereas other 40 have not been formerly described in *Pinaceae*. Based on the translated sequences, no significant homology was detected for up to 170 amino acid sequences (48% of the putative nuclear unigenes), which illustrates the current lack of genomic knowledge for gymnosperms. Putative functional categories were assigned to the other 181 amino acid sequences using FuncatDB [15] and Blast2GO [16] (Fig. 1, Supplementary Table S2, Supplementary Table S3).

The largest group corresponded to genes involved in metabolism (16%). Up to 42% of these were related to carbohydrate metabolism, whereas lipid, fatty acid and isoprenoid metabolism related sequences, which could be related with the synthesis and accumulation of hormones during water stress, comprised the 18% of the metabolism group. Another important group contained proteins involved in the transcription process. Some ESTs showed

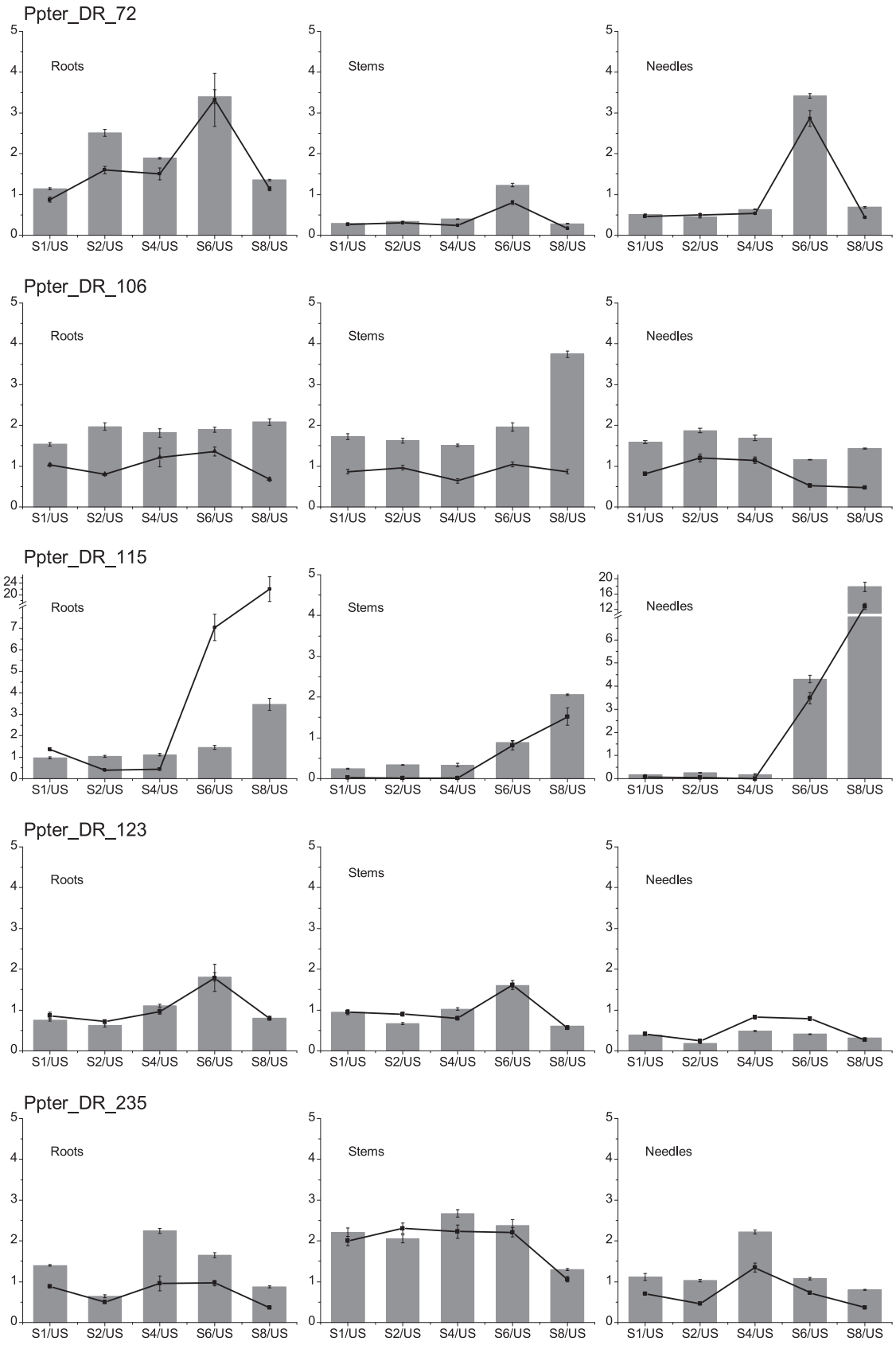

Fig. 2. Expression profiles obtained with microarray (histograms) and RT-PCR (lines) techniques for ten selected genes.

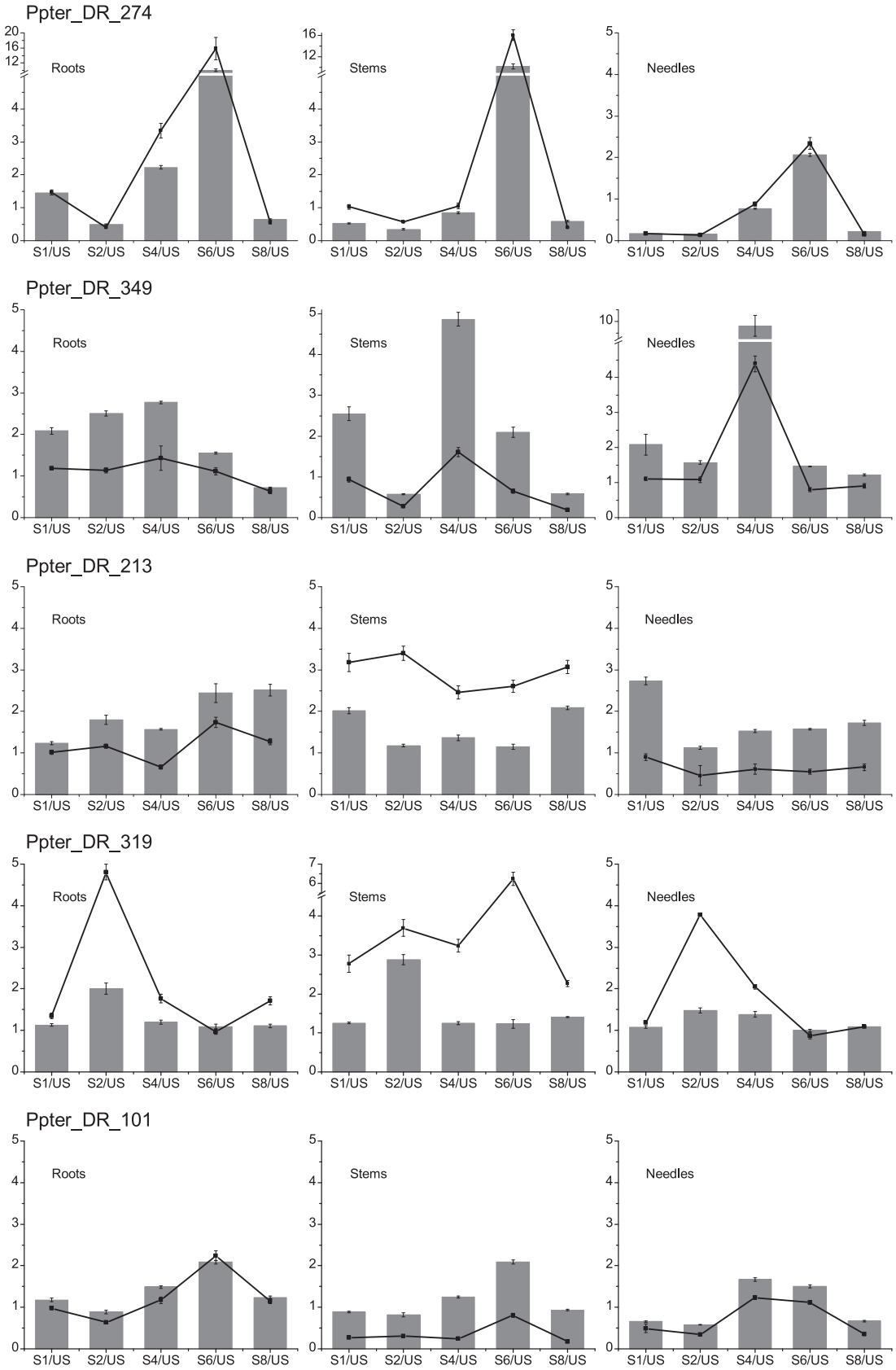

Fig. 2. (*continued*).

a high homology with transcription factors linked to environmental stresses, such as *DREB*, *bZIP* and NAC. The library also includes ESTs putatively involved in cell rescue and defence, such as *RD21* and heat shock proteins.

2.2.1. Metabolism related genes

One third of the ESTs with a putative function according to FunCat correspond to this group. Half of them (8) are included in the carbohydrate metabolism group. The accumulation of sugars has been correlated with the acquisition of desiccation tolerance in plants, probably because the sugars protect the structures from mechanical and metabolic stresses during dehydration [17].

Several ESTs in this group share homology with genes that have been previously reported to be related to drought stress tolerance, such as alkaline α-galactosidases (Ppter_DR_33 and Ppter_DR_120; [18]), malate synthases (Ppter_DR_276; [19]), glycosyltransferases (Ppter_DR_101; [20]) or chitinases (Ppter_DR_52; [21]). This group also includes a putative aldehyde dehydrogenase (Ppter_DR_4), which is presumably involved in the detoxification of aldehydes generated by alcohol metabolism. This result is consistent with the accumulation of ethanol in conifer seedlings during drought [22]. Similarly, the heterologous expression of maize *ZmALDH22A1* confers elevated stress tolerance in tobacco [23]. Ppter_DR_274, which was highly induced in the needles, stems and roots, is homologous to 12-oxo-phytodienoic acid reductase (OPR). This gene is induced by osmotic stress in *Zea mays*, and its heterologous expression in *Arabidopsis* improves the resistance to osmotic and salt-induced stress during seed germination [24].

2.2.2. Cell rescue and defence genes

This category includes putative late embryogenesis abundant (*LEA*) proteins (Ppter_DR_51, Ppter_DR_72 and Ppter_DR_95), which is a group of proteins whose involvement in desiccation tolerance is well known (f. i., Battaglia et al. (2008) [25] and references therein). Ppter_DR_136, Ppter_DR_152 and Ppter_DR_235 show homology with heat shock proteins of the HSP90 family, which are known as molecular chaperones and are related to different abiotic stresses (heat, cold, salt and heavy metals) [26]. Also in this group appeared an EST (Ppter_DR_106) putatively corresponding to a glutathione S-transferase, which are a group of genes induced by drought in *Arabidopsis* [27].

2.2.3. Transport

Transport processes play an important role in the mobilisation and accumulation of solutes and hormones and in cell detoxification pathways during adaptation to water stress. Among the selected ESTs, several putative transporters were also found. For instance, Ppter_DR_123 shows high homology with an inositol transporter, which could be involved in the transport of sugars through membranes to modify osmotic pressure under stress, a function that is consistent with the high capacity of maritime pine for osmotic adjustment as reported by López et al. [28]. Ppter_DR_162 shows homology with certain ABC transporters, which are involved in the response to different biotic and abiotic stresses [29].

2.2.4. Transcription-related genes

The expression of many of the genes mentioned above is likely controlled by different transcriptional regulatory pathways. Among the ESTs selected based on their expression patterns there were several putative transcription factors. For example, Ppter_DR_254 and Ppter_DR_319 show high homology with DREB2 factors, which are involved in an ABA-independent pathway induced by dehydration [30] and whose overexpression can increase water stress tolerance [31]. Ppter_DR_213 is homologous to NAC transcription factors, which are one of the largest families of plant specific factors that are also involved in drought tolerance [32]. Ppter_DR_296 has homology with a DEAD-box ATP-dependent RNA helicases, which have been reported to play a key role in stress responses in various organisms [33].

Nevertheless, almost one third (20) of the 67 selected ESTs (see below) were either homologous to sequences coding for unknown or unclassified proteins or lacked homologous sequences in the available databases. However, their high inducibility makes them interesting candidate genes. For instance, Ppter_DR_115 and Ppter_DR_163 are homologous to different members of the nodulin MtN3 family from *Arabidopsis*, which has not yet been assigned to any functional category. Nodulin-like proteins are plasmatic membrane aquoporins, and could be involved in the maintenance of water balance in plants. Both microarray and RT-PCR analyses showed a continuous accumulation of the Ppter_DR_115 transcript during the water stress experiment, with levels reaching 20-fold higher than in control samples. Another example is Ppter_DR_107, which was highly overexpressed in needles during long-term stress, and did not share homology with any angiosperm genes. Tentative annotation suggests it could represent a hydroxyproline-rich protein (PRP). Another proline-rich product was reported by Sathyan et al. (2005) to be induced by water stress in *P. halepensis*. Other upregulated ESTs include Ppter_DR_160, Ppter_DR334 and Ppter_DR_341, which were homologous to unknown proteins in *Picea*.

2.3. Microarray analysis and RT-PCR

Since its description by Diatchenko et al. [34] the SSH approach has proved to be a powerful tool for enriching libraries with differentially expressed genes. However, this approach generates quite a few false positives, and ESTs are subject to artefacts during construction of the cDNA library. For this reason, a careful analysis of the fragments obtained is needed. In this study, the selection of candidate genes from among the 351 nuclear unigenes was based on their overexpression during PEG-induced water stress. For this purpose we examined the transcription levels during the treatment in the needles, stems and roots of four genotypes independently using a customised microarray.

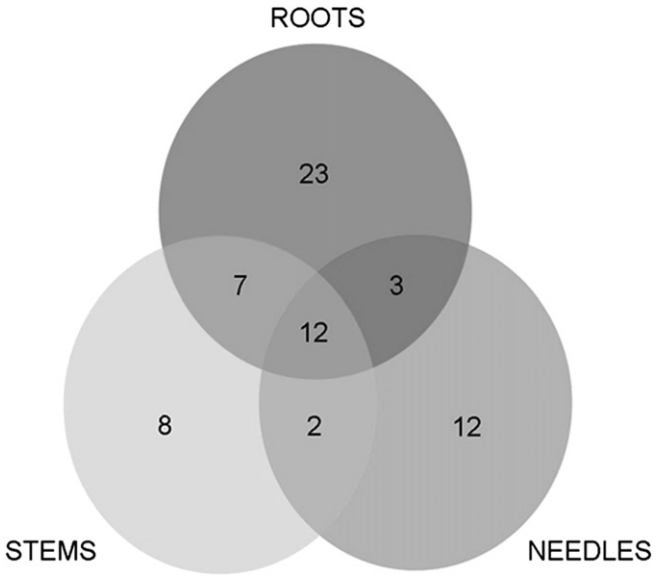

Fig. 3. Transcripts that are significantly upregulated by PEG stress in needles, stems and roots by PEG stress. In total, 67 unique ESTs were significantly upregulated in response to treatment.

Table 2
Fold change in the expression of the 67 selected candidate genes upregulated in different organs during treatment. The table shows data with a significant difference (FDR < 0.05 and fold change ≥ 1.6) in relative level.

ID_sequence	Roots					Stems					Needles				
	S1/US	S2/US	S4/US	S6/US	S8/US	S1/US	S2/US	S4/US	S6/US	S8/US	S1/US	S2/US	S4/US	S6/US	S8/US
Ppter_DR_101				1,96					1,95					2,54	
Ppter_DR_115					2,82					6,24			2,27	10,34	45,22
Ppter_DR_120			3,01	3,94					2,75				2,92	7,84	10,16
Ppter_DR_160				2,46					3,35					6,07	3,94
Ppter_DR_274			4,46	10,28					7,97					8,10	3,19
Ppter_DR_276				3,49	3,99				2,36			1,95		4,03	4,92
Ppter_DR_305				3,42	3,09			1,97	4,16					2,97	4,04
Ppter_DR_33				2,80					2,63					5,50	5,08
Ppter_DR_349			2,87					3,55					2,52		
Ppter_DR_6			2,46	3,75	3,22				2,61					2,32	
Ppter_DR_61				2,71					3,36					4,27	2,61
Ppter_DR_72			2,51	3,40					3,62					8,67	
Ppter_DR_112								2,45							
Ppter_DR_122							2,41		−2,92	−4,65					
Ppter_DR_136							2,18								
Ppter_DR_152							2,15								
Ppter_DR_278							2,90								
Ppter_DR_41							2,25								
Ppter_DR_77						−1,92			1,61						
Ppter_DR_90							2,15								
Ppter_DR_107												1,74			6,38
Ppter_DR_239														2,61	
Ppter_DR_244														2,68	
Ppter_DR_284													2,35	2,39	
Ppter_DR_303															2,27
Ppter_DR_315															2,35
Ppter_DR_342															2,97
Ppter_DR_4															3,20
Ppter_DR_48															2,31
Ppter_DR_51											3,24		5,41		
Ppter_DR_80															2,38
Ppter_DR_95														3,46	2,49
Ppter_DR_123				2,27					3,33						
Ppter_DR_200					3,18		1,95	2,52	2,71						
Ppter_DR_235			2,14					2,32							
Ppter_DR_242				2,99					2,43						
Ppter_DR_334		2,43	4,33	6,11	4,75				2,14						
Ppter_DR_37				1,88	5,22		2,36								
Ppter_DR_63				2,00	6,75		2,06								
Ppter_DR_106			2,52	1,95											
Ppter_DR_13				2,09											
Ppter_DR_137					2,97										
Ppter_DR_162			2,30												
Ppter_DR_163				2,16											
Ppter_DR_180				2,00											
Ppter_DR_187				2,12											
Ppter_DR_194			2,69												
Ppter_DR_198			3,33												
Ppter_DR_213					2,84										
Ppter_DR_22				2,08											
Ppter_DR_223				1,86											
Ppter_DR_227				2,25											
Ppter_DR_254			2,04												
Ppter_DR_262					3,07										
Ppter_DR_287					2,74										
Ppter_DR_296			2,16												
Ppter_DR_307				2,29											
Ppter_DR_319		2,38													
Ppter_DR_327			2,20	2,76											
Ppter_DR_329				1,76											
Ppter_DR_341					2,86										
Ppter_DR_99			2,10												
Ppter_DR_275		−2,63		2,01										3,58	2,66
Ppter_DR_283					3,16										3,20
Ppter_DR_44				2,64	3,43									2,90	4,56
Ppter_DR_159									2,74					3,72	
Ppter_DR_52								2,95	5,44					3,67	4,10

We selected 67 positive candidate genes whose average expression for the four genotypes, in at least one sampling point and in at least one organ, was more than 1.6 fold higher than in the control (Table 1). This represents 19% of the presumed 351 nuDNA ESTs identified, which is a proportion similar to that obtained in previous studies with forest trees, as reported by Bae et al. in *Populus* [35]. The largest functional groups among the 67 selected genes coincided with those for the complete 351 EST set.

To validate the expression patterns detected by microarray, 10 candidate genes covering the main functional groups and expressional clusters (see below) were subjected to RT-PCR. This time, only a single genotype was used in the analysis. RT-PCR analysis confirmed the microarray expression patterns (Fig. 2). For most genes, both techniques had a Pearson correlation value higher than 90%, although lower values were found for a few genes. This kind of discrepancy has been previously reported, mainly for genes with low induction levels, and has been attributed to the different dynamic range of the two techniques, with a higher accuracy attributed to RT-PCR [36].

2.4. Expression pattern of candidate genes

Of the 67 unigenes that were significantly upregulated by PEG-induced water stress, 29 were identified in needles, 29 in stems and 45 (two thirds of the total) in roots (half of them, i. e., 23, exclusively in this organ), which is consistent with the key role played by roots in water stress (Fig. 3). Significant overexpression levels are reflected in Table 2.

Hierarchical clustering of the expression patterns detected with microarrays (from the four genotypes) established six groups in the stems, and five groups in the roots and needles (Supplementary data Fig. 1). Differences among the genotypes regarding these profiles were also detected. Nevertheless, the expression profiles were fairly consistent between the microarray and qRT-PCR techniques within the same genotype.

The following general considerations can be drawn from an analysis of the expression clusters:

- Most PEG-induced genes reached a maximum transcription level at S6 (48 h from the beginning of the experiment, and after 36 h at −1.6 MPa). In this group, the most noticeable induction was seen for Ppter_DR_274 in the roots and stems.
- Some of these genes were induced faster in the roots, with the maximum induction seen at S4 (1 h after reaching −1.6 MPa). These include genes putatively involved in cell rescue and defence based on their homology with heat shock proteins (Ppter_DR_235) or glutathione S-transferases (Ppter_DR_106), and the putative transcription factor Ppter_DR_254. This pattern is consistent with the role of roots in detecting and triggering the response to water stress.
- Several genes showed a fast response, reaching their maximum induction at S2 (1 h after the second change of hydroponic solution, at −0.8 MPa). Among them, a putative transcription factor was detected in the roots (Ppter_DR_319) and several sequences had high homology with the HSP90 heat shock protein family (Ppter_DR_136, 152).
- The induction of other genes continued to increase until the end of the experiment. This group mainly includes genes putatively related to carbohydrate metabolism, as well as some putative transcription factors in the roots (Ppter_DR_296, Ppter_DR_213).
- The prolonged and even increasing induction of most of these genes during the water stress treatment is consistent with the conditions faced by Mediterranean trees during a real drought in nature.

3. Conclusion

Here, we have identified a set of 67 reliable expressional candidate genes that may be useful for further studies on water stress response in conifers, which is a plant division for which reliable molecular tools are still scarce and for which transfer from angiosperms is difficult and very often impossible. Of 67 candidate genes, 37 correspond to new genes that were not reported in previous works on water stress in this genus. The expression patterns of these genes revealed by microarray analysis were confirmed by RT-PCR for a subset of genes, including the most abundant expression clusters and functional groups.

As a follow-up to this study, the expression of these candidate genes under drought stress induced by soil water depletion will be examined, allowing us to discriminate between water stress effects and the effects derived specifically from the treatment with PEG.

4. Materials and methods

4.1. Plant materials and treatment conditions

We have used clonal material from the Oria provenance (Almería, southeastern Spain). This provenance has been previously shown to have a good inducible response to water stress because it has adapted to low and irregular annual precipitations and frequent droughts (f. i., Sánchez-Gómez et al.,2010). Taking into account the individual variability found within populations, which is common for conifers (see, f. i., [37] and references therein), ten different genotypes were included in this study. The plants were grown in a greenhouse in hydroponic culture with an aerated nutrient solution under controlled conditions (24 °C day/22 °C night, 12 h photoperiod, relative humidity: 60% by day and 80% by night). The nutrient solution (30 l/45 plants; NPK 90:41:72; pH ~ 6.5) was changed twice a week.

Water stress was induced by adding polyethylene glycol (PEG, MW 8000) to the culture solution. For this purpose the hydroponic solution was changed every 4 h, with progressively increasing concentrations of PEG added and a 0.4 MPa decrease in the water potential of the solution each time until a final water potential of −1.6 MPa was reached. One plant per genotype was collected separately 1 h after every change of the hydroponic solution (sampling points S1-S4), and 24 h (S5), 48 h (S6), 10 days (S7) and 21 days (S8) after the beginning of the treatment (for a total of 8 sampling points). The needles, stem and roots from each plant were collected separately, immediately frozen in liquid nitrogen and stored at −80 °C. Plants of each genotype kept without PEG in the hydroponic solution were harvested as controls.

4.2. Isolation of RNA and the subtraction technique

The total RNA for each sampling point was extracted separately from the roots, stem and needles of each sampled plant following the CTAB–LiCl precipitation method [38]. Equal amounts of the total RNA from the roots, stems and needles of all of the plants were mixed to form an RNA pool. To identify sequences putatively regulated by water stress more efficiently, we constructed a subtractive cDNA library using a PCR-Select™ cDNA Subtraction kit (Clontech, CA, USA) and following the manufacturer's protocols. A mix of cDNA from plants treated with PEG for 1 h to 21 days was used as tester and the cDNA from control plants was used as driver. Subtracted PCR products were ligated into the pGEM® T-easy vector (Promega, WI, USA) and transformed into *Escherichia coli* DH5α cells. We then picked 4940 clones. The presence and size of the inserts were determined by direct amplification from crude bacterial lysates using the nested PCR primers 1 and 2R, which were provided in the PCR-selected cDNA subtraction kit.

4.3. Differential screening of the subtracted library by cDNA macroarray

As a first selection of positive clones, differential screening was performed using a PCR-Select™ differential screening kit (Clontech, CA, USA). Amplified inserts from each clone were spotted onto Hybond-N^+ nylon membranes (GE Healthcare BioSciences, WI, USA). Membranes were screened with four probes labelled with dioxigenin-dUTP (Roche, Basel, Switzerland), i.e., two subtracted cDNA probes (forward and reverse subtractions) and two unsubtracted probes (PEG-stressed and unstressed). Detection was performed using the DIG DNA labelling and detection kit (Roche, Basel, Switzerland). The intensities of the hybridisation signals were visualised using a Molecular Imager®ChemiDoc™XRS System (BioRad, CA, USA). Clones with a mean expression ratio (forward subtracted probe intensity divided by reverse probe intensity) of more than 1.5 were selected as upregulated genes. All clones of differentially expressed genes were selected for sequencing.

4.4. Sequence analysis

The selected positive clones were sequenced using a 3730 XL DNA analyser (Applied Biosystems; Life Technologies, CA, USA) at Macrogen (Seoul, Korea). All unique ESTs were annotated on the basis of the existing annotation of non-redundant databases at the NCBI using BLASTN and BLASTX. ESTs without significant protein homology were then compared with the pine and spruce databases included in the Gene Index Project (http://compbio.dfci.harvard.edu/cgi-bin/tgi/Blast/index.cgi) and EuroPineDB (http://www.scbi.uma.es/pindb/). Homologies with e-values lower than 1e-05 for more than 100 nucleotides were considered significant. Functional classification of the ESTs was performed according to the functional categories of *A. thaliana* proteins (http://mips.helmholtz-muenchen.de/proj/funcatDB/) [15]. Blast2GO was used to identify gene ontology (GO) terms associated with identificated genes [16].

4.5. Microarray design and hybridation

In total, 351 unigenes selected from the SSH library were included in the microarray design (Agilent 8 × 15 K, Agilent Technologies, CA, USA). For each unigene, one to four 60-bp-long probes were designed and spotted at least three times on the slide. Probes designed for other pine, spruce and human ESTs available in public databases were included as negative controls. To select reliable candidate genes that would be widely useful for future studies, four unrelated genotypes from the Oria provenance were used as experimental biological replicates. RNA from sampling points S1, S2, S4, S6 and S8 and control plants was hybridised to the microarrays. RNA amplification and labelling were performed as described by Adie et al. [39]. RNA was purified by using the Qiagen RNeasy kit (QIAGEN, CA, USA). "The manual two-colour microarray based gene expression analysis" protocol (Agilent Technologies, CA, USA) was followed for hybridisations. Images from Cy3 and Hyper5 channels were equilibrated and captured with a GenePix 4000B (Axon, CA, USA), and spots were quantified using the GenPix software (Axon, CA, USA).

4.6. Data analysis

Background correction and normalisation of expression data were performed using LIMMA (Linear Models for Microarray Data) [40,41]. LIMMA is part of "Bioconductor, an R language project" (www.bioconductor.org). For local background correction and normalisation, the methods "normexp" and loess in LIMMA were used, respectively. To achieve a similar distribution across arrays and consistency among arrays, log-ratio values were scaled using the median-absolute-value as scale estimator. Differentially expressed genes were evaluated by the non-parametric algorithm 'Rank Products' available as the "RankProd" package at "Bioconductor, an R language project" [42,43]. This method detects genes that are consistently high ranked in a number of replicated experiments independently of their numerical intensities. The results are provided in the form of *p-values* defined as the probability that a given gene is ranked in the observed position by chance. The expected false discovery rate was controlled to be less than 5%.

Changes in the expression of a gene relative to control plants were estimated using the average signal intensity across the four data sets (four genotypes). Based on the statistical analysis, a gene was considered to be significantly upregulated if it met all three of the following criteria: (1) FDR RankProd < 0.05; (2) the fold change was ≥ 1.6 at any sampling point and in any organ and (3) the trend was consistent for all data. Hierarchical clustering of upregulated genes in the different organs was performed using the log-ratio data and the Euclidean distance (complete linkage and threshold 2.5) options of the MeV 4.4 software [44].

4.7. Real-time quantitative PCR

The expression pattern of several genes was confirmed by RT-PCR using RNA from one of the genotypes used as a biological replicate in the microarrays. For this purpose, the RNA was treated with DNAse Turbo (Ambion; Applied Biosystems, Life Technologies, CA, USA). First-strand cDNA was synthesised from 2 μg total RNA from each sample using PowerScriptIII reverse transcriptase (Invitrogen) according to the supplier's manual. 18S rRNA was used as a control, after verifying that the signal intensity remained unchanged across all treatments. The primers for experimental genes were designed using Primer Express version 3.0.0 (Applied Biosystems Life Technologies, CA, USA) and are shown in Supplementary Table S4. Polymerase chain reactions were performed in an optical 96-well plate with a CFX 96 Detection system (BIO-RAD), using EvaGreen to monitor dsDNA synthesis. Reactions containing 2× SsoFast EvaGreen Supermix reagent (BIO-RAD, CA, USA), 12.5 ng cDNA and 500 nM of primers in a final volume of 10 μl were subjected to the following standard thermal profile: 95 °C for 3 min, 40 cycles of 95 °C for 10 s and 60 °C for 10 s. Three technical replicates were performed for each PCR run. To compare the data from different PCR runs or cDNA samples, the mean of the CT values of the three technical replicates was normalised to the mean CT value of Ri18S. The expression ratios were then obtained using the ΔΔCT method corrected for the PCR efficiency for each gene [45].

Acknowledgements

The authors would like to thank Dr. Luis Gil, for technical and scientific support, Dr. Juan Majada and Dr. Tania Velasco, from SERIDA for providing the plant material and helping us with the hydroponic treatments, and Dr. Carmen Díaz-Sala (UAH) for help with RT-PCR analysis. This work has been funded through the projects AGL2006-03242/FOR (Spanish Ministry of Education and Science), CCG07-UPM/AMB-1932 and CCG10-UPM/AMB-5038 (Madrid Regional Government – UPM). PP has a pre-doctoral fellowship from the Spanish Ministry of Education and Science.

Appendix. Supplementary data

Supplementary data related to this article can be found online at doi:10.1016/j.plaphy.2011.09.022.

References

[1] FAO, Situacion de los Bosques del Mundo 2007 (2007) Roma.
[2] A.M. Morse, D.G. Peterson, M.N. Islam-Faridi, K.E. Smith, Z. Magbanua, S.A. Garcia, T.L. Kubisiak, H.V. Amerson, J.E. Carlson, C.D. Nelson, J.M. Davis, Evolution of genome size and complexity in *Pinus*, PLoS ONE 4 (2009) e4332.
[3] IPCC, Climate change 2007. in: S. Salomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, K.L. Mille (Eds.), The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, UK and New York, NY, USA, 2007, p. 996.
[4] I. Aranda, R. Alia, U. Ortega, A.K. Dantas, J. Majada, Intra-specific variability in biomass partitioning and carbon isotopic discrimination under moderate drought stress in seedlings from four *Pinus pinaster* populations, Tree Genetics and Genomes 6 (2010) 169–178.
[5] D. Sánchez-Gómez, T. Velasco-Conde, F.J. Cano-Martín, M. Ángeles Guevara, M. Teresa Cervera, I. Aranda, Inter-clonal variation in functional traits in response to drought for a genetically homogeneous Mediterranean conifer, Environmental and Experimental Botany 70 (2010) 104–109.
[6] L.S. Heath, N. Ramakrishnan, R.R. Sederoff, R.W. Whetten, B.I. Chevone, C.A. Struble, V.Y. Jouenne, D.W. Chen, L. van Zyl, R. Grene, Studying the functional genomics of stress responses in loblolly pine with the expresso microarray experiment management system, Comparative and Functional Genomics 3 (2002) 226–243.
[7] J.I. Watkinson, A.A. Sioson, C. Vasquez-Robinet, M. Shukla, D. Kumar, M. Ellis, L.S. Heath, N. Ramakrishnan, B. Chevone, L.T. Watson, L. van Zyl, U. Egertsdotter, R.R. Sederoff, R. Grene, Photosynthetic acclimation is reflected in specific patterns of gene expression in drought-stressed loblolly pine, Plant Physiology 133 (2003) 1702–1716.
[8] W.W. Lorenz, F. Sun, C. Liang, D. Kolychev, H.M. Wang, X. Zhao, M.M. Cordonnier-Pratt, L.H. Pratt, J.F.D. Dean, Water stress-responsive genes in loblolly pine (*Pinus taeda*) roots identified by analyses of expressed sequence tag libraries, Tree Physiology 26 (2006) 1–16.
[9] W.W. Lorenz, R. Alba, Y.-S. Yu, J. Bordeaux, M. Simoes, J. Dean, Microarray analysis and scale-free gene networks identify candidate regulators in drought-stressed roots of loblolly pine (*P. taeda* L.), BMC Genomics 12 (2011) 264.
[10] P. Sathyan, R.J. Newton, C.A. Loopstra, Genes induced by WDS are differentially expressed in two populations of aleppo pine (*Pinus halepensis*), Tree Genetics and Genomes 1 (2005) 166–173.
[11] C. Dubos, G. Le Provost, D. Pot, F. Salin, C. Lalane, D. Madur, J.-M. Frigerio, C. Plomion, Identification and characterization of water-stress-responsive genes in hydroponically grown maritime pine (*Pinus pinaster*) seedlings, Tree Physiology 23 (2003) 169–179.
[12] C. Dubos, C. Plomion, Identification of water-deficit responsive genes in maritime pine (*Pinus pinaster* Ait.) roots, Plant Molecular Biology 51 (2003) 249–262.
[13] E. Eveno, C. Collada, M.A. Guevara, V. Leger, A. Soto, L. Diaz, P. Leger, S.C. Gonzalez-Martinez, M.T. Cervera, C. Plomion, P.H. Garnier-Gere, Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses, Molecular Biology and Evolution 25 (2008) 417–437.
[14] N. Fernández-Pozo, J. Canales, D. Guerrero-Fernández, D.P. Villalobos, S. Díaz-Moreno, R. Bautista, A. Flores-Monterroso, M.A. Guevara, P. Perdiguero, C. Collada, M.T. Cervera, A. Soto, R. Ordás, F.R. Cantón, C. Avila, F.M. Cánovas, M.G. Claros, EuroPineDB: a high-coverage web database for maritime pine transcriptome, BMC Genomics 12 (2011) 366.
[15] A. Ruepp, A. Zollner, D. Maier, K. Albermann, J. Hani, M. Mokrejs, I. Tetko, U. Guldener, G. Mannhaupt, M. Munsterkotter, H.W. Mewes, The FunCat, a functional annotation scheme for systematic classification of proteins from whole genomes, Nucleic Acids Research 32 (2004) 5539–5545.
[16] A. Conesa, S. Götz, J.M. García-Gómez, J. Terol, M. Talón, M. Robles, Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research, Bioinformatics 21 (2005) 3674–3676.
[17] M.J. Oliver, J.C. Cushman, K.L. Koster, Dehydration tolerance in plants. in: R. Sunkar (Ed.), Plant Stress Tolerance, vol. 639. Humana Press, 2010, pp. 3–24.
[18] M. Hara, K. Tokunaga, T. Kuboi, Isolation of a drought-responsive alkaline alpha galactosidase gene from New Zealand spinach, Plant Biotechnology 25 (2008) 497–501.
[19] S. Micheletto, L. Rodriguez-Uribe, R. Hernandez, R.D. Richins, J. Curry, M.A. O'Connell, Comparative transcript profiling in roots of *Phaseolus acutifolius* and *P. vulgaris* under water deficit stress, Plant Science 173 (2007) 510–520.
[20] J. Wang, B. Hou, Glycosyltransferases: key players involved in the modification of plant secondary metabolites, Frontiers of Biology in China 4 (2009) 39–46.
[21] A. Rabello, C. Guimaraes, P. Rangel, F. da Silva, D. Seixas, E. de Souza, A. Brasileiro, C. Spehar, M. Ferreira, A. Mehta, Identification of drought-responsive genes in roots of upland rice (*Oryza sativa* L), BMC Genomics 9 (2008) 485.
[22] D.K. Manter, R.G. Kelsey, Ethanol accumulation in drought-stressed conifer seedlings, International Journal of Plant Sciences 169 (2008) 361–369.
[23] D. Huang, W. Wu, S.R. Abrams, A.J. Cutler, The relationship of drought-related gene expression in *Arabidopsis thaliana* to hormonal and environmental factors, Journal of Experimental Botany 59 (2008) 2991–3007.
[24] D. Gu, X. Liu, M. Wang, J. Zheng, W. Hou, G. Wang, J. Wang, Overexpression of ZmOPR1 in *Arabidopsis* enhanced the tolerance to osmotic and salt stress during seed germination, Plant Science 174 (2008) 124–130.
[25] M. Battaglia, Y. Olvera-Carrillo, A. Garciarrubio, F. Campos, A.A. Covarrubias, The enigmatic lea proteins and Other hydrophilins, Plant Physiology 148 (2008) 6–24.
[26] W.X. Wang, B. Vinocur, O. Shoseyov, A. Altman, Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response, Trends in Plant Science 9 (2004) 244–252.
[27] B. Kumar, S.L. Singla-Pareek, S.K. Sopory, Glutathione homeostasis: crucial for abiotic stress tolerance in plants. in: A. Pareek (Ed.), Abiotic Stress Adaptation in Plants. Springer, Netherlands, 2010, pp. 263–282.
[28] R. Lopez, I. Aranda, L. Gil, Osmotic adjustment is a significant mechanism of drought resistance in *Pinus pinaster* and *Pinus canariensis*, Investigacion Agraria-Sistemas Y Recursos Forestales 18 (2009) 159–166.
[29] D. Wanke, H.U. Kolukisaoglu, An update on the ABCC transporter family in plants: many genes, many proteins, but how many functions? Plant Biology 12 (2010) 15–25.
[30] P. Agarwal, P. Agarwal, M. Reddy, S. Sopory, Role of DREB transcription factors in abiotic and biotic stress tolerance in plants, Plant Cell Reports 25 (2006) 1263–1274.
[31] M. Chen, Q.-Y. Wang, X.-G. Cheng, Z.-S. Xu, L.-C. Li, X.-G. Ye, L.-Q. Xia, Y.-Z. Ma, GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants, Biochemical and Biophysical Research Communications 353 (2007) 299–305.
[32] L.-S.P. Tran, K. Nakashima, Y. Sakuma, S.D. Simpson, Y. Fujita, K. Maruyama, M. Fujita, M. Seki, K. Shinozaki, K. Yamaguchi-Shinozaki, Isolation and functional analysis of *Arabidopsis* stress-inducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter, Plant Cell 16 (2004) 2481–2498.
[33] S. Rocak, P. Linder, DEAD-box proteins: the driving forces behind RNA metabolism, Nature Reviews. Molecular Cell Biology 5 (2004) 232–241.
[34] L. Diatchenko, Y. Lau, P. Siebert, Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries, Proceedings of the National Academic Sciences of the United States of America 93 (1996) 6025–6030.
[35] E.-K. Bae, H. Lee, J.-S. Lee, E.-W. Noh, Isolation and characterization of osmotic stress-induced genes in poplar cells by suppression subtractive hybridization and cDNA microarray analysis, Plant Physiology and Biochemistry 48 (2009) 136–141.
[36] P. Dallas, N. Gottardo, M. Firth, A. Beesley, K. Hoffmann, P. Terry, J. Freitas, J. Boag, A. Cummings, U. Kees, Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR – how well do they correlate? BMC Genomics 6 (2005) 59.
[37] E.T. Hamanishi, M.M. Campbell, Genome-wide responses to drought in forest trees, Forestry 84 (2011) 273–283.
[38] S. Chang, J. Puryear, J. Cairney, A simple and efficient method for isolating RNA from pine trees, Plant Molecular Biology Reporter 11 (1993) 113–116.
[39] B.A.T. Adie, J. Perez-Perez, M.M. Perez-Perez, M. Godoy, J.J. Sanchez-Serrano, E.A. Schmelz, R. Solano, ABA is an essential signal for plant resistance to pathogens affecting JA biosynthesis and the activation of defenses in *Arabidopsis*, Plant Cell 19 (2007) 1665–1681.
[40] G.K. Smyth, T. Speed, Normalization of cDNA microarray data, Methods 31 (2003) 265–273.
[41] G.K. Smyth, Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. The Berkeley Electronic Press, 2004.
[42] R. Breitling, P. Armengaud, A. Amtmann, P. Herzyk, Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments, FEBS Letters 573 (2004) 83–92.
[43] F. Hong, R. Breitling, C.W. McEntee, B.S. Wittner, J.L. Nemhauser, J. Chory, RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis, Bioinformatics 22 (2006) 2825–2827.
[44] A.I. Saeed, N.K. Bhagabati, J.C. Braisted, W. Liang, V. Sharov, E.A. Howe, J. Li, M. Thiagarajan, J.A. White, J. Quackenbush, K. Alan, O. Brian, TM4 Microarray Software Suite, Methods in Enzymology, vol. 411, Academic Press, 2006, pp. 134–193.
[45] M.W. Pfaffl, A new mathematical model for relative quantification in real-time RT-PCR, Nucleic Acids Research 29 (2001) e45.

Further readings

[46] http://compbio.dfci.harvard.edu/cgi-bin/tgi/Blast/index.cgi
[47] http://www.scbi.uma.es/pindb/
[48] http://mips.helmholtz-muenchen.de/proj/funcatDB/