Photosynthetic Acclimation Is Reflected in Specific Patterns of Gene Expression in Drought-Stressed Loblolly Pine[1[w]]

Jonathan I. Watkinson, Allan A. Sioson, Cecilia Vasquez-Robinet, Maulik Shukla, Deept Kumar, Margaret Ellis, Lenwood S. Heath, Naren Ramakrishnan, Boris Chevone, Layne T. Watson, Leonel van Zyl, Ulrika Egertsdotter, Ronald R. Sederoff, and Ruth Grene*

Departments of Plant Pathology, Physiology, and Weed Science (J.I.W., C.V.-R., B.C., R.G.) and Computer Science (A.A.S., M.S., D.K., M.E., L.S.H., N.R., L.T.W.), Virginia Tech, Blacksburg, Virginia 24061; Forest Biotechnology Group, Department of Forestry, North Carolina State University, Raleigh, North Carolina 27695 (L.v.Z., R.R.S.); and Institute of Paper Science and Technology, 500 10th Street, N.W., Atlanta, Georgia 30318 (U.E.)

Because the product of a single gene can influence many aspects of plant growth and development, it is necessary to understand how gene products act in concert and upon each other to effect adaptive changes to stressful conditions. We conducted experiments to improve our understanding of the responses of loblolly pine (*Pinus taeda*) to drought stress. Water was withheld from rooted plantlets of to a measured water potential of −1 MPa for mild stress and −1.5 MPa for severe stress. Net photosynthesis was measured for each level of stress. RNA was isolated from needles and used in hybridizations against a microarray consisting of 2,173 cDNA clones from five pine expressed sequence tag libraries. Gene expression was estimated using a two-stage mixed linear model. Subsequently, data mining via inductive logic programming identified rules (relationships) among gene expression, treatments, and functional categories. Changes in RNA transcript profiles of loblolly pine due to drought stress were correlated with physiological data reflecting photosynthetic acclimation to mild stress or photosynthetic failure during severe stress. Analysis of transcript profiles indicated that there are distinct patterns of expression related to the two levels of stress. Genes encoding heat shock proteins, late embryogenic-abundant proteins, enzymes from the aromatic acid and flavonoid biosynthetic pathways, and from carbon metabolism showed distinctive responses associated with acclimation. Five genes shown to have different transcript levels in response to either mild or severe stress were chosen for further analysis using real-time polymerase chain reaction. The real-time polymerase chain reaction results were in good agreement with those obtained on microarrays.

Drought stress can limit plant growth, resulting in reduced crop yields. Many attempts to analyze plant responses to drought have focused on single genes. However, the product of a single gene can function in or act upon multiple response pathways, influencing many aspects of plant growth and development. Therefore, it is necessary to understand how gene products act in concert and upon each other to effect adaptive changes to stressful conditions. We have used microarrays and Expresso, a Next Generation Software computational system, to analyze the changes in gene transcript profiles in loblolly pine (*Pinus taeda*) in response to drought stress.

Microarrays have emerged as a prominent tool in the analysis of large scale gene expression. The spotting of cDNA onto glass slides allows for ascertaining the expression profile, as revealed by steady-state transcript levels, of thousands of genes at a single time. The technique promises to reveal networks of genes that contribute to the same biological response and provide new information on the functions of unknown genes (Somerville and Somerville, 1999). Microarrays have been used in plants to identify changes in transcript levels of genes associated with drought stress (Seki et al., 2001; Heath et al., 2002; Ozturk et al., 2002), cold stress (Seki et al., 2001; Fowler and Thomashow, 2002), salt stress (Kawasaki et al., 2001; Ozturk et al., 2002), pathogen interaction (Scheideler et al., 2001), and during development (Ruan et al., 1998; Girke et al., 2000). These studies cite large-scale changes in transcript profiles as being associated with stress, but none correlate these changes with adaptive responses. It is one of our goals to uncover the mechanisms underlying strategies to cope with and adapt to drought stress.

Microarrays present a challenge to researchers both in terms of their design and implementation (Kerr and Churchill, 2001) and in subsequent storage and analysis of data. Such challenges require automated computational assistance. With the voluminous amounts of data generated from each microarray experiment, effective and efficient access to this in-

[1] This work was supported in part by the National Science Foundation (grant nos. EIA–0103660 [Next Generation Software] and IBN–0219332 [Information Technology Research] to L.S.H, N.R., L.T.W., R.G., and B.C.).

[w] The online version of this article contains Web-only data.

* Corresponding author; e-mail grene@vt.edu; fax 540–231–5755.

http://www.plantphysiol.org/cgi/doi/10.1104/pp.103.026914.

formation becomes critical. Because new experiments bring additional insights into the genomic response of plants to specific conditions, it is advantageous to reanalyze previous experiments in light of these new insights. For example, Perez-Amador et al. (2001) used cluster analysis on 47 Arabidopsis gene expression profiles stored in the Stanford Microarray Database, an online database, and discovered novel expression patterns for several genes identified as having increased transcript levels in a mutant. Because the results of one microarray experiment can be used to plan subsequent experiments, access to data during design is essential.

Our Expresso system (Alscher et al., 2001; Heath et al., 2002; Sioson et al., 2003) addresses multiple phases of the microarray experiment lifecycle, including experimental design, microarray design and printing, data acquisition, image analysis, statistical analysis, and data mining. In contrast to many existing tools, Expresso is meant to systematize methodological issues in microarray experiments. Also, the evolution of Expresso is the consequence of close collaboration among biologists, computer scientists, and statisticians. Alscher et al. (2001) identified the three key design principles behind Expresso: (a) model-based design and management of experiments, (b) algorithms for "closing the experimental loop," and (c) a lightweight data management system that allows graceful changes to the underlying schema of the data. For instance, our several generations of loblolly pine experiments have had varying attributes and aspects, to which the Expresso framework has readily adapted. The data management system supports a variety of analysis algorithms; in this paper, we highlight the use of the relational data mining technique called inductive logic programming (ILP; Muggleton and Feng, 1990; Muggleton, 1999).

To date, a primary use of Expresso has been to study drought stress responses in loblolly pine, the predominant timber species in the southeastern U.S., covering about 13.4 million ha (Schultz, 1999). Loblolly pine is also an important timber crop in Africa, Asia, and South America, where tree growth can exceed that of trees grown in the United States. The faster growth of trees in these locales depends upon the use of suitable seed stock, and, even then, improvements can be made to optimize growth in each locale (Schultz, 1999). Tree improvement involves traits such as better wood quality and higher density, greater disease resistance, and improved growth under various environmental conditions (Schultz, 1999). Drought stress can limit tree growth and alter wood quality (Lev-Yadun and Sederoff, 2000). It is necessary to understand responses to drought stress to achieve the development of crops with increased resistance to drought. Loblolly pine constitutes an excellent system to model drought stress in softwood timber species and gymnosperms in general. Furthermore, the use of a gymnosperm will add considerable knowledge to comparative genomic studies, extending what is known about genome analysis from angiosperms to gymnosperms.

Drought stress has been correlated with expression changes in many plant genes. These include the heat shock proteins (HSPs; Ristic et al., 1998), late embryogenic-abundant proteins (LEAs; Iuchi et al., 1996), and aquaporins. Some of the more well-characterized responses include genes whose products regulate expression of drought responsive genes, the dehydration-responsive element-binding proteins (Liu et al., 1998), and abscisic acid-responsive element-binding proteins (Kang et al., 2002). Other genes involved in drought stress are those associated with lipid signaling such as the phospholipase gene families (Katagiri et al., 2001; Sang et al., 2001) and those associated with detoxification of reactive oxygen species (Alscher et al., 1997). Although these studies relate the expression of specific genes to drought imposition, no one has defined the role of stress-related genes in acclimation to the stress.

Results obtained using microarrays have broadened the analysis of gene expression in response to drought and confirmed the results of individual gene studies. For example, Seki et al. (2001) reported 44 drought-inducible genes in Arabidopsis using microarrays and identified putative targets of DREB1A. Ozturk et al. (2002) noted a change in expression using microarrays for several barley (*Hordeum vulgare*) genes identified as being drought responsive, including an LEA, a dehydrin, and a water channel protein. A water channel protein was also one of those identified by Seki et al. (2001) in Arabidopsis, suggesting at least some commonality between results from different species. We have previously used Expresso to analyze changes in transcript profiles of loblolly pine in response to drought stress. We designed and printed an array consisting of 384 cDNAs, chosen by us as being associated with stress responses. Using the 384-gene microarray, 72 genes showed increased transcript levels after four cycles of mild drought stress (Heath et al., 2002). This group of genes did not show increased transcript levels as a result of severe drought stress, making them candidate genes for drought tolerance mechanisms. The data mining technique of ILP incorporated in Expresso associated gene expression with membership in putative functional categories. From the ILP results, we identified functional categories of genes that responded to the stress, including "heat" (HSPs and LEAs) and "membrane transport proteins" (aquaporins). Other categories affected by the stress were "cell wall related" and "lignin biosynthesis." A group of genes categorized as isoflavone reductases were also identified as being stress regulated. Babiychuk et al. (1995) identified these genes as being involved in a glutathione-independent mechanism

implicated in protection of cells from oxidative stress, which has been correlated with drought stress imposition. These enzymes are also related to those involved in lignin biosynthesis and in the synthesis of biotic plant defense compounds (Gang et al., 1999), suggesting that they are involved in multiple stress-responsive pathways.

We have completed a second series of experiments to improve our understanding of loblolly pine response to drought stress. The arrays consist of 2,173 clones selected from five existing loblolly pine cDNA libraries (Stasolla et al., 2003). This paper focuses on the use of data mining algorithms in Expresso to analyze this newer set of data and to identify significant changes in gene expression.

RESULTS

Drought Stress and Physiological Measurements

Daily water potential measurements indicated that trees reached the desired water potential for each cycle of stress at 3 to 4 d after withholding water for mild stress and at 6 to 7 d after withholding water for severe stress. Trees were subjected to three cycles of stress (Fig. 1). Control trees were watered normally throughout the experiment and were maintained at a water potential of −0.4 MPa or above. Trees grown under mild stress showed little alteration in growth and, as did control trees, continued to produce new flushes of growth. Trees grown under severe stress had fewer new flushes of growth compared with controls, indicating that severe stress limited tree growth. Photosynthetic measurements were taken at the time of peak drought stress to identify the effect of the stress on photosynthesis. Both mild and severe stress levels led to reduced photosynthesis during the first cycle of stress, with a much greater photosynthetic reduction in trees grown under severe stress (Table I). In subsequent cycles of stress, trees under mild stress showed recovery of photosynthetic rate to levels at or approaching those of controls, a result we considered to be photosynthetic acclimation to the stress. Trees grown under severe stress remained much below control levels, a result we noted as photosynthetic failure.

Gene Transcript Profiles

The two-stage, linear mixed model (Wolfinger et al., 2001) used by Expresso is intended to remove global effects from the computed intensity values. It also allows the estimation of significant changes in gene transcript levels below the standard 2.0- or 2.5-fold increase. All reported results are significant at $\alpha = 0.05$. Between mild cycles 1 and 2, there was an increase in the number of genes showing regulation (from 133–213 positively affected clones and from 94–159 negatively affected clones; Fig. 2), which cor-

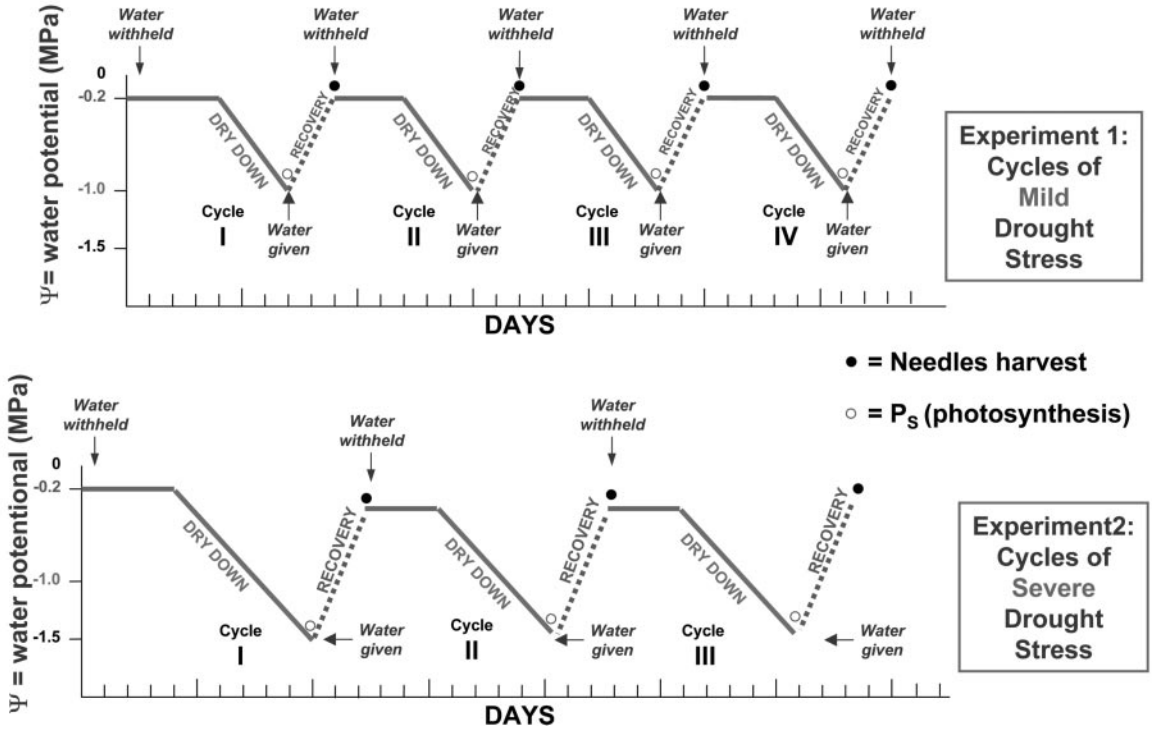

Figure 1. Graphical representation of experimental design for mild and severe drought stress treatments (does not reflect actual time period). Water was withheld from treated trees until a desired water potential was reached (−1 MPa for mild stress, and −1.5 MPa for severe stress) as determined by predawn, water potential measurements. At maximum drought stress, photosynthesis was measured and plants were rewatered. Needles and stems were harvested 24 h after rewatering.

Table I. *Effect of mild or severe drought stress on net photosynthesis in 1-year-old loblolly pine rooted cuttings*

Condition	Net Photosynthesis[a]		
	Cycle	Control	Stressed
		$\mu mol\ CO_2\ m^{-2}\ s^{-1}$	
Mild	1	4.28 ± 0.42	2.48 ± 0.40
	2	3.54 ± 0.48	3.82 ± 0.61
	3	4.75 ± 0.44	3.28 ± 0.54
Severe	1	3.67 ± 0.32	0.88 ± 0.40
	2	3.00 ± 0.49	0.19 ± 0.41
	3	2.90 ± 0.54	0.77 ± 0.94

[a]Measurements were made on the first fully mature fascicle in each case, using the LI-COR 6400 (LI-COR, Lincoln, NE). Three or four repeated measurements were made in each case. SEs are given.

relates with photosynthetic acclimation. Trees grown under severe stress did not show such an increase and the number of genes showing a positive response remained consistent across all three cycles of severe stress. Figure 2 shows the genes that are regulated in both mild and severe stress. Generally, less than 20% of genes showing a response are shared between mild and severe stress cycles, which suggests that different response pathways are initiated for the different levels of stress. Many of the genes identified as being positively regulated were noted in our previous experiment (Heath et al., 2002) and include Rubisco-binding protein, aquaporin, and isoflavone reductase (for a complete list of genes that responded, see Supplemental Data, available in the online version of this article at http://www.plantyphysiol.org).

To aid in data mining by ILP, clones were assigned to the functional categories described by Heath et al. (2002). Although the clones were selected based on the Munich Information Center for Protein Sequences (MIPS) functional categories, we used the gene annotations and information from the literature to recategorize the genes in a hierarchy more reflective of stress response characteristics, e.g. the categories were based on plant responses to stress and the processes protected by stress defense mechanisms (Fig. 3). Notice that the categories have a hierarchical structure: "Trafficking" is a component of the parent category "cells," which is in turn a component of the parent category "development" (Fig. 3). In addition, a gene may be placed in multiple categories.

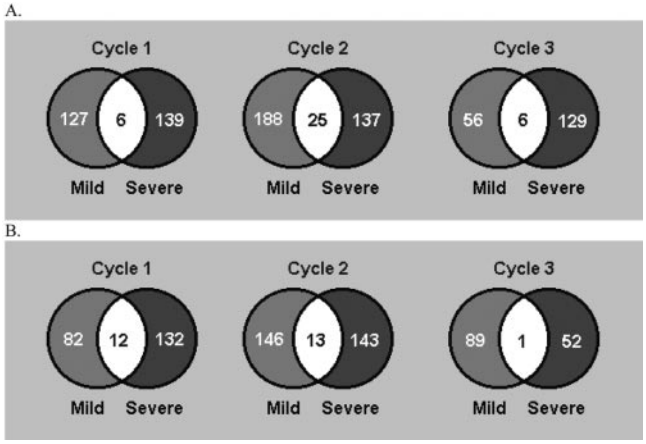

Figure 2. Venn diagrams representing number of unique clones responding either positively (A) or negatively (B) to mild or severe stress. Genes showing significant changes in transcript level were compared between mild and severe stress. The results are presented in the form of Venn diagrams here.

The primary input to ILP is the set of genes that show a statistically significant change in gene expression level for the various treatments. Membership of genes in functional categories is also used by ILP to obtain relationships among gene expression, treatments, and functional categories. ILP was applied to the statistically significant changes in gene expression levels for the various treatments together with the functional categories to obtain relationships between gene category and treatment. Signature patterns of changes in transcript levels relating to the functional categories were found using ILP (Figs. 4 and 5). An example of ILP rules demonstrating the transcript levels of genes whose products participate in trafficking between the endoplasmic reticulum and the Golgi were elevated during acclimation (mild cycles 1 and 3) is:

$$expression(X,mc1,\text{“ + ”}) : - category(X,trafficking)$$

$$expression(X,mc3,\text{“ + ”}) : - category(X,trafficking)$$

These rules should be read from right to left; for instance, the first rule above states that if a gene (X) is in category "trafficking," then it is up-expressed in experimental condition "mc1" (mild stress cycle 1). Genes falling into the categories of "nitrogen and sulfur metabolism," "respiratory electron transport," and "cell membranes" were also identified by ILP as showing significant changes in transcript levels in response to drought stress. We further focused on three groups of genes that were identified using ILP. One group, which can be found in the intersection of the categories "Gene Expression" and "ROS and Stress," includes "protection and repair" genes. These responded positively during acclimation. The second group of genes is included in the subcategories of "carbon metabolism" and showed a positive response during all three cycles of mild stress with highest expression during acclimation. The third group falls under the category of "phenylpropanoid metabolism" and showed a positive response during the first and third cycles of mild stress.

Protection and Repair Genes

Genes encoding the HSPs and LEAs can be found in the subcategories of "protection and repair," "drought," "heat," and "cold" within "ROS and Stress" (Fig. 3). In trees grown under mild stress, more genes within "heat" and "protection and re-

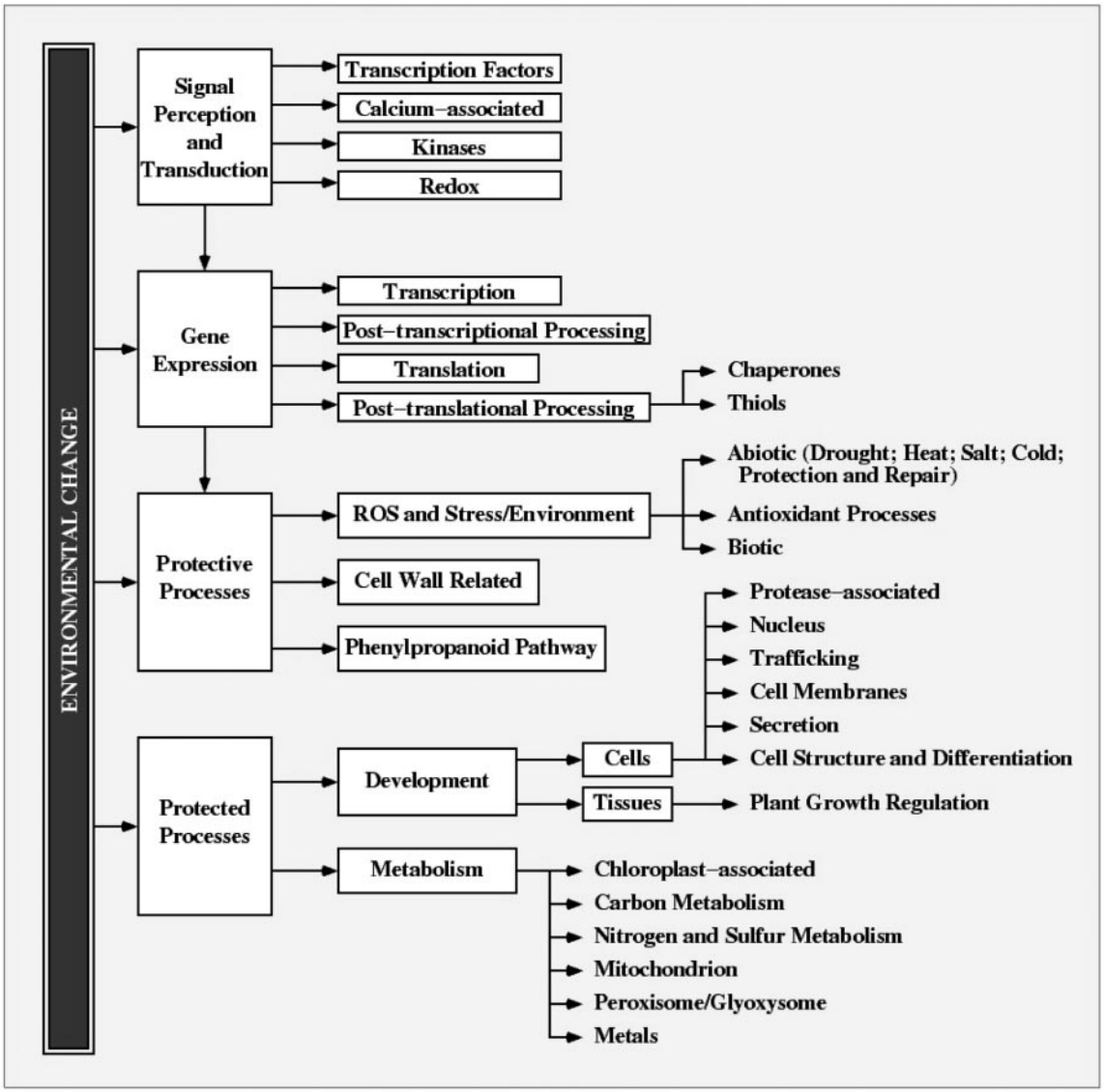

Figure 3. Schematic diagram of hierarchy of gene functional categories used in Expresso. Clones were assigned to functional categories based on the gene annotation and information from the literature. The categories are based on known plant stress responses and the process that are thought to be protected by those stresses. By virtue of being placed in a category, a clone automatically becomes a member of all parent categories. In addition, a clone can be placed in multiple functional categories that are not direct ancestors of each other.

pair" had increased transcript levels compared with severe stress conditions (Figs. 4 and 5). However, "protection and repair" genes correlated with cold stress were higher in trees grown under severe stress. A more detailed analysis of some of the HSP and LEA genes identified as having altered transcript levels revealed that LEA group 2 homologs were associated more with mild stress conditions, whereas LEA group 3 homologs were associated more with imposition of severe stress (Fig. 6). Figure 6 also shows that specific clones of HSP 70 and HSP 90 respond differently to mild and severe stress. For example, an HSP 70, DNA k-type homolog (clone NXSI_117_C08) and an HSP 90 homolog (NXSI_116_B04) show increased transcript levels during mild stress and no response during severe stress. Conversely, a mitochondrial HSP 70 (NXCI_022_G01) and a different HSP 90 (NXNV_149_E10) show increased transcript levels during severe stress and no response during mild stress.

Carbon Metabolism

Genes in the category of "Carbon Metabolism" were identified by ILP as having increased transcript levels in trees grown under mild stress. We focused on the genes associated with core carbon metabolism (Fig. 7). Genes that showed positive regulation in mild stress include pyruvate kinase, which leads to the formation of pyruvate from phospho*enol*pyru-

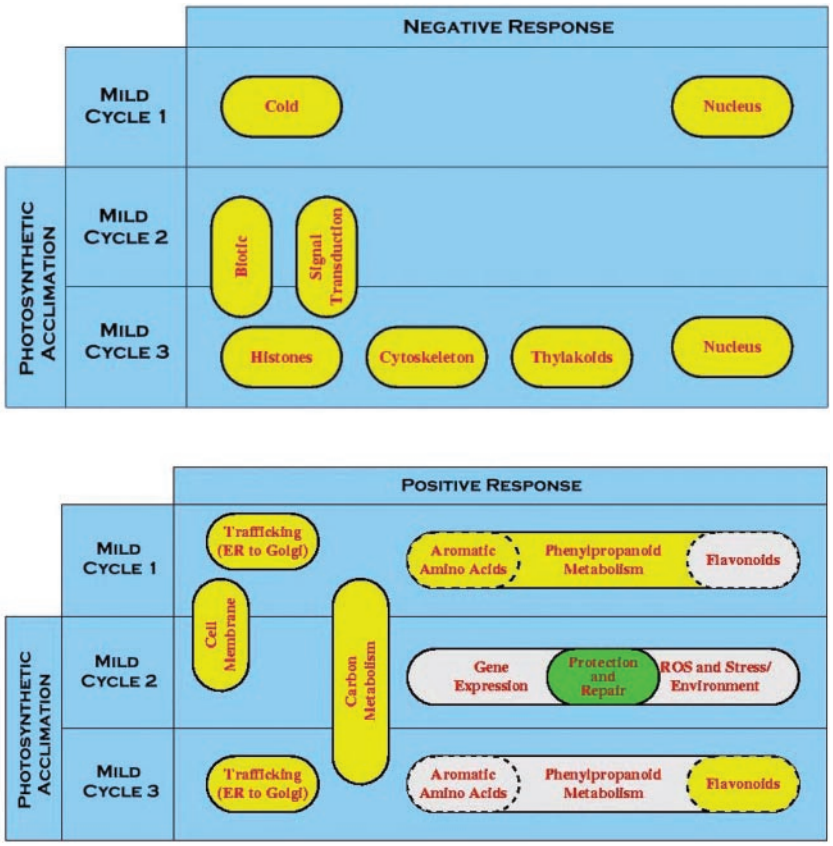

Figure 4. Signature patterns of changes in gene expression profiles for trees grown under mild drought stress. The data generated through microarray analysis were mined using ILP for rules relating expression patterns and functional categories. The rules generated by ILP are represented as colored ovals in the diagram. Subcategories are represented as ovals within larger ovals. The green oval represents a subcategory that responded and is shared between two main categories.

vate, and pyruvate dehydrogenase, which feeds carbon from glycolysis into the tricarboxylic acid (TCA) cycle. Other genes that show increased transcript levels in mild stress include those in the oxidative pentose phosphate pathway such as Rib-5-phosphate epimerase, transaldolase, and transketolase. Genes associated with the reductive pentose phosphate pathway such as Rubisco and two plastidic isoforms of glyceraldehyde-3-phosphate show increased transcript levels in severe stress.

Phenylpropanoid Pathway and Aromatic Amino Acids

ILP produced rules concerning expression of genes involved in aromatic amino acid biosynthesis and in the phenylpropanoid pathway (Fig. 8). Because there is a link between one branch of the aromatic amino acid pathway and the phenylpropanoid pathway, these groups were examined together. Genes associated with the entry of carbon into these pathways, DHAP-dihydroxyacetone phosphate synthase and 3-dehydroquinate synthase, had higher transcript levels in trees grown under mild stress compared with severe stress. Phe ammonia lyase did not show a response in mild stress, whereas cinnamate-4-hydroxlase had positive expression in cycles 1 and 3 of mild stress. Nearly all the genes involved with flavonoid biosynthesis that were represented on the array showed an increase under mild stress. Genes associated with lignin generally showed no response to mild or severe drought stress. Of the 12 laccase clones present on the array, only four showed a response with one isoform of laccase showing a positive response in mild stress, whereas another showed a positive response under severe stress (Fig. 8).

Confirmation of Gene Transcript Profiles

Five genes shown to have different transcript levels in response to either mild or severe stress were chosen for further analysis using real-time (RT)-PCR. Two genes encoding LEA proteins were found to have similar expression profiles using RT-PCR as compared with the microarrays. Pine LEA g3 (P14G04) was shown to be up-expressed in all three cycles of severe stress using microarray analysis (Fig. 6). A similar result was seen with the RT-PCR where there was a 2.5-fold increase in transcript level during the second cycle of severe stress and a 1.7-fold increase in transcript level during the third cycle of severe stress (Table II). The results generated for pLEAg2 by RT-PCR were also similar to those generated by microarray analysis (Table II; Fig. 6) with an increase in transcript levels during the third cycle of mild stress but no change during severe stress. The flavonoid genes chalcone isomerase and naringenin-2-oxo dioxygenase (3-flavonone hydroxylase) were also found to have similar transcript profiles during

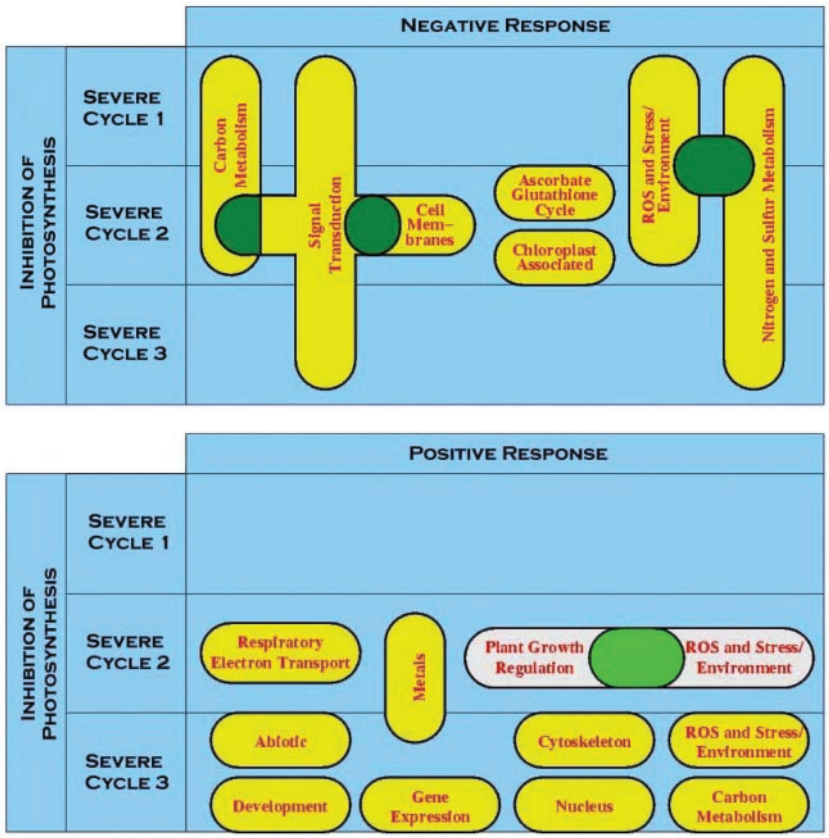

Figure 5. Signature patterns of changes in gene expression profiles for trees grown under severe drought stress. Diagrams were developed as in Figure 4.

drought stress when analyzed using RT-PCR compared with the results obtained from microarrays (Table II; Fig. 8).

DISCUSSION

The analysis of drought response in loblolly pine using microarrays has implications for many aspects of the timber industry, including crop improvement. Results obtained with pine serve as a model to expand the knowledge of angiosperm gene analysis into gymnosperms. We have demonstrated previously that 72 of 384 stress-associated genes show increased transcript abundance in loblolly pine grown under mild stress, whereas 69 of those genes showed no response in severe stress (Heath et al., 2002). These expression patterns were correlated with acclimation to the stress. To further analyze the response of loblolly pine to drought, we used a 2,173-element array, comprised of genes selected to reflect all 15 MIPS functional categories. Daily water potential measurements were taken to determine the level of drought stress. We also took photosynthetic measurements at the time of tissue harvest. Pines grown

Clone ID	Annotation	Class	Mild			Severe		
			1	2	3	1	2	3
14D07	Putative Dehydrin	LEA Group 2	0	+	0	0	0	0
NXCI_002_C10	Putative Dehydrin	LEA Group 2	0	0	+	0	0	0
NXCI_006_H04	Embyonic abunndant protein (white spruce)	LEA Group ?	0	0	+	0	0	0
PC23D04	LEA76 homolog	LEA Group 3	0	0	0	0	+	0
ST01E01	Putative LEA	LEA Group 3	0	0	0	0	+	+
PC14C08	LEA76 homolog	LEA Group 3	0	0	0	0	0	+
PC14G04	LEA76 homolog	LEA Group 3	-	0	-	+	+	+
PC05A11	LEA76 homolog	LEA Group 3	0	0	-	+	+	0
PC08E04	LEA76 homolog	LEA Group 3	0	0	-	+	0	+
39A03	DNA J-like protein	DNA J	0	+	0	-	0	0
40F04	Low molecular weight heat shock protein	sHSP	+	0	-	0	0	0
NXNV_132_E06	DNA J homolog	DNA J	+	0	0	0	0	0
NXSI_116_B04	Heat Shock Protein 82	HSP 90	0	+	0	0	0	0
NXSI_117_C08	DNA K type molecular chaperone hsc 70	HSP 70	0	+	0	-	0	0
NXNV_149_E10	Putative Heat Shock Protein 90	HSP 90	0	0	0	0	+	0
NXCI_022_G01	Heat Shock 70 Kda protein, mitochonrial	HSP 70	0	0	0	+	+	0
NXNV	Heat Shock Protein, 82K, precursor	HSP 90	0	0	0	+	0	0
13C12	Putative peptidyl-prolyl cis-trans isomerase	Ppiase	0	0	0	+	0	0
38H03	Peptidyl proline isomerase	Ppiase	0	+	0	0	0	0
39H08	Peptidyl prolyl cis-trans isomerase	Ppiase	0	+	0	0	0	0

Figure 6. Response of protection and repair genes to mild and severe drought stress as discovered through microarrays and ILP. Protection and repair genes are grouped according to the class in which they fall. Changes in transcript profiles are shown as positive (+), negative (−), or unchanged (0).

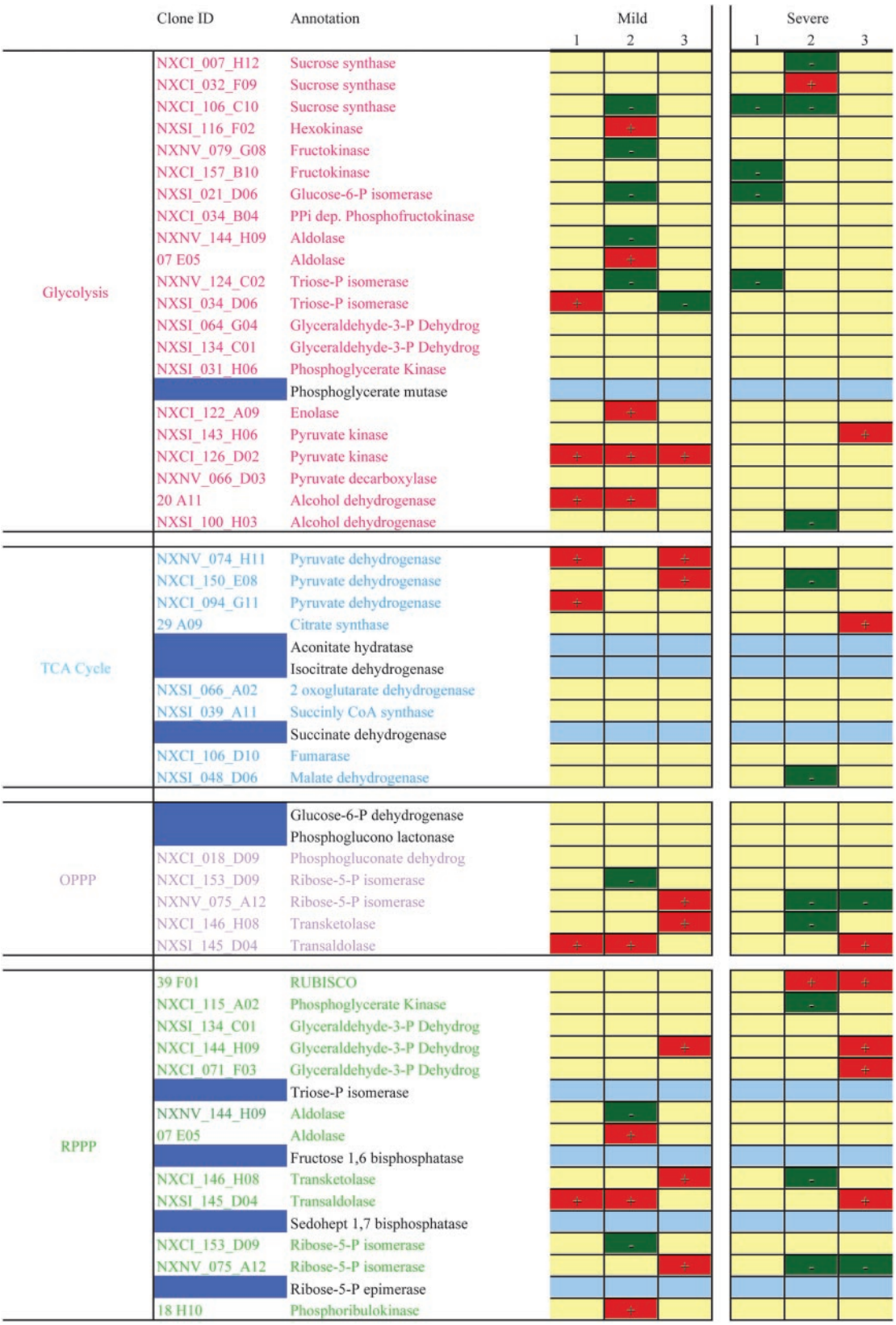

Figure 7. Transcript profiles of genes in the category of carbon metabolism during three cycles of mild or severe drought stress. Genes within the categories of "carbon metabolism" were identified by ILP as having increased transcript levels in trees grown under mild stress. The genes within these categories that are present on the array are represented here. Expression is indicated as positive (+), negative (−), or unchanged (0). Blue squares indicate genes that were not represented on the array.

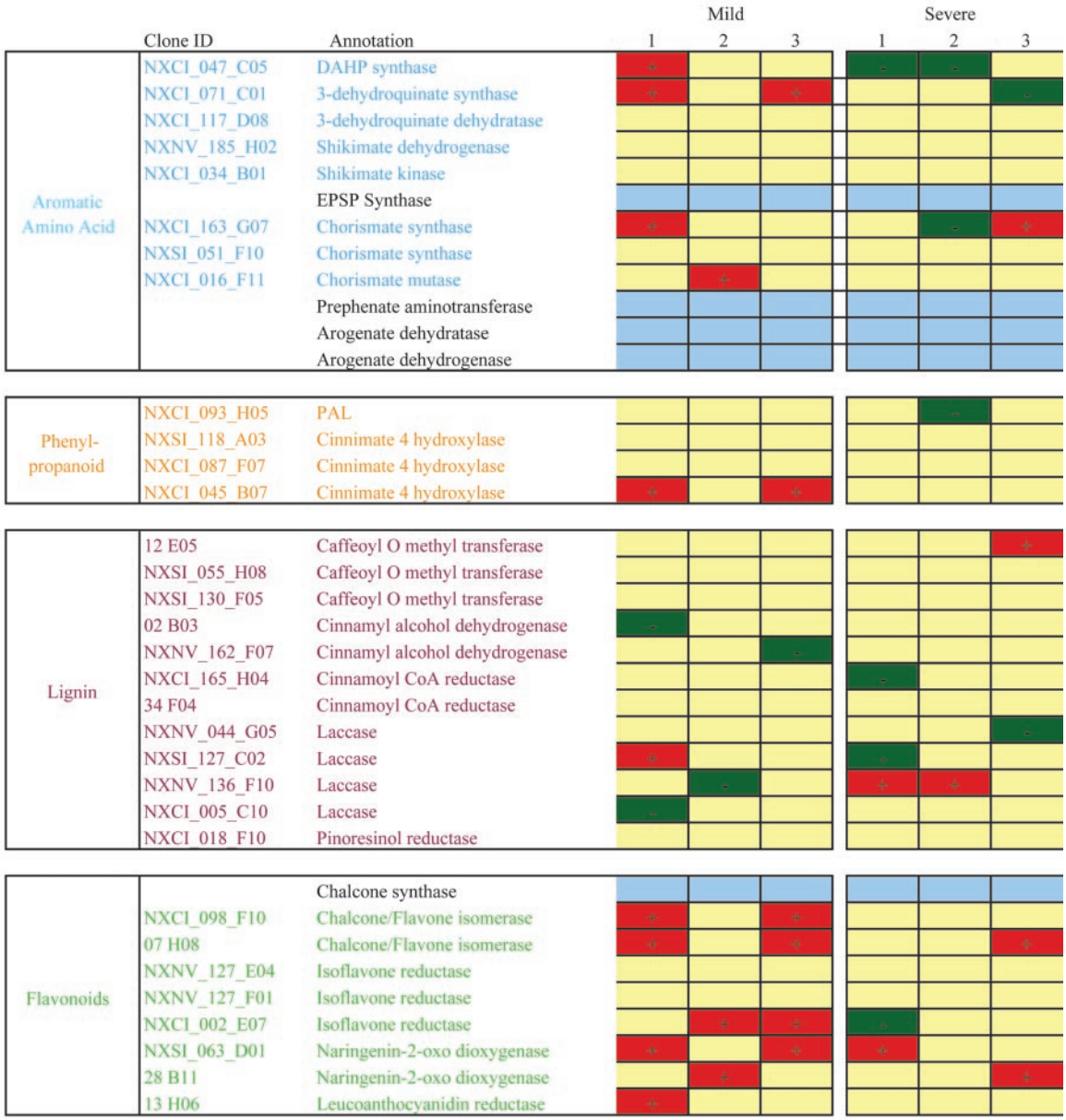

Figure 8. Transcript profiles of genes in the categories of "phenylpropanoid metabolism" and "aromatic amino acids" during three cycles of either mild or severe drought stress. Genes within the categories of "phenylpropanoid metabolism" and "aromatic amino acids" were identified by ILP as having increased transcript levels in trees grown under mild stress. The genes within these categories (identified by annotation) that are present in a pathway are represented here. Expression is indicated as positive (+), negative (−), or unchanged (0). Blue squares indicate genes that were not represented on the array.

under mild stress (−1 MPa) showed an initial reduction in photosynthetic rate, with recovery to near control levels in subsequent drought cycles. The recovery in photosynthetic rate is an expression of photosynthetic acclimation to the drought stress.

Quantitative Changes in Transcript Profiles Reflect Unique Acclimatory Responses under Mild Stress

The change in RNA transcript profiles of loblolly pine due to drought stress can be correlated with physiological data reflecting photosynthetic acclimation to mild stress or photosynthetic failure during imposition of severe stress. The number of clones showing either positive or negative changes in transcript levels increases from 5.8% to 8.6% between mild cycles 1 and 2, suggesting that the behavior of these genes is correlated with acclimation. The near doubling of the number of genes showing negative changes (3.8%–6.7%) suggests that negative gene expression shares importance with positive gene expression in acclimation to the imposed stress. There

Table II. *Real-time PCR results of selected genes shown by microarray analysis to have altered transcript levels during either mild or severe drought stress*

Clone Identification	Mild Drought Stress		Severe Drought Stress	
	Cycle 2	Cycle 3	Cycle 2	Cycle 3
07 H08	0.899[a] ± 0.326	3.66 ± 0.469	0.993 ± 0.249	2.06 ± 0.371
NXSI_063_D01	0.407 ± 0.182	2.1 ± 1.73	1.28 ± 1.5	1.59 ± 0.369
PC14G04	0.845 ± 0.836	0.62 ± 0.332	2.54 ± 1.01	1.74 ± 0.421
NXCI_002_G10	0.79 ± 0.122	2.13 ± 0.562	1.45 ± 0.327	1.75 ± 0.486
ST40F04	0.51 ± 0.506	2.93 ± 0.909	1.00 ± 1.41	1.29 ± 1.53

[a] Indicates fold change (where 1 is equal to no change, nos. greater than 1 represent a positive change in gene expression, and nos. less than 1 represent negative changes in gene expression) SD ± 1 in RNA level of the target gene in treated samples as compared with control samples.

was no comparable change in trees grown under severe water deficit, where changes in genes showing altered transcript profiles remained at 6%. However, in both mild and severe stress, the genes with altered transcript profiles dropped in the third cycle of stress. Again, it is shown for the first time, to our knowledge, that the greater change in gene transcript levels under mild stress is associated with physiological acclimation.

The large number (approximately 80% of genes showing changes in transcript profiles) of genes unique to each stress level suggests that the plant is able to sense the degree of stress and activate different response pathways to cope with that stress. These different response pathways contain related genes; however, clones with the same annotation appear to be differentially expressed members of a multigene family. For example, one clone of laccase is positively regulated during mild stress and negatively regulated during severe stress, whereas another laccase gene is negatively regulated during mild stress and positively regulated in severe stress. Other examples include genes that encode annexin, a Ca^{2+}-binding protein involved in diverse cellular functions (Gerke and Moss, 2002); ADP-ribosylation factor, a GTPase involved in vesicle budding and membrane trafficking (Takeuchi et al., 2002); and myo inositol-1-phosphate synthase (see Supplemental Data). The differential response of members of multigene families to stress as revealed by microarray analysis has not been reported previously, but these results suggest that the microarray technique can yield results allowing differentiation among some members of a multigene family. RT-PCR data of selected genes corroborate the microarray data and give credence to our findings (Table II).

Group 2 LEAs Are Specifically Up-Regulated during Photosynthetic Acclimation under Mild Drought Stress

The expression of genes associated with protection and repair mechanisms have been correlated with responses to stress imposition (Xu et al., 1996; Lee and Vierling, 2000; Zhang et al., 2000). We previously identified protection and repair genes as being positively regulated in response to drought and suggested that they aided in adaptation to drought stress by protecting processes within the cell (Heath et al., 2002). Lee and Vierling (2000) suggested that one protection afforded by HSPs, one group of protection and repair genes, is manifested through the interaction of small HSPs with denatured proteins, preventing aggregation and promoting refolding of denatured proteins. Through ILP, we have found that different groups of protection and repair genes were expressed in either mild or severe stress (Fig. 6). We generated a phylogenetic tree to determine how the loblolly pine LEAs are related to each other and to LEAs from other species (Fig. 9), and pine clones associate with group 2 LEAs (dehydrins) from cotton (*Gossypium hirsutum*), Arabidopsis, and white spruce (*Picea abies*). Studies of dehydrin (LEA group 2) in cowpea (*Vigna unguiculata*) point to a possible function in membrane protection (Ismail et al., 1999). Furthermore, Iuchi et al. (1996) found that group 2 LEAs were highly expressed under drought stress in a drought-tolerant cowpea. Although this group of LEAs has been correlated with drought stress, it was shown that the group 2 LEAs from pine had higher transcript levels during mild stress, suggesting that they are specifically associated with acclimation. Several pine clones show close homology to group 3 LEAs from Arabidopsis and cotton (Fig. 9). A group 3 LEA from barley (HVA1) confers tolerance to water and salt stress when expressed in rice (*Oryza sativa*; Xu et al., 1996). Again, correlative evidence suggests that the group 3 LEAs are associated with stress, but it is shown for the first time, to our knowledge, that the pine group 3 LEAs are specifically associated with severe stress. We previously identified HSPs as being responsive to drought stress, and our results suggested that different homologs responded to the different levels of stress (Heath et al., 2002). The results of this study corroborate our previous experiment showing that the changes in expression of LEAs and HSPs are important to drought stress responses in general. However, some, such as the group 2 LEAs, are more important during acclimation, whereas the group 3 LEAs may be more important to protection during severe stress.

The LEAs that responded under severe stress have also been associated with cold responses (Dong et al.,

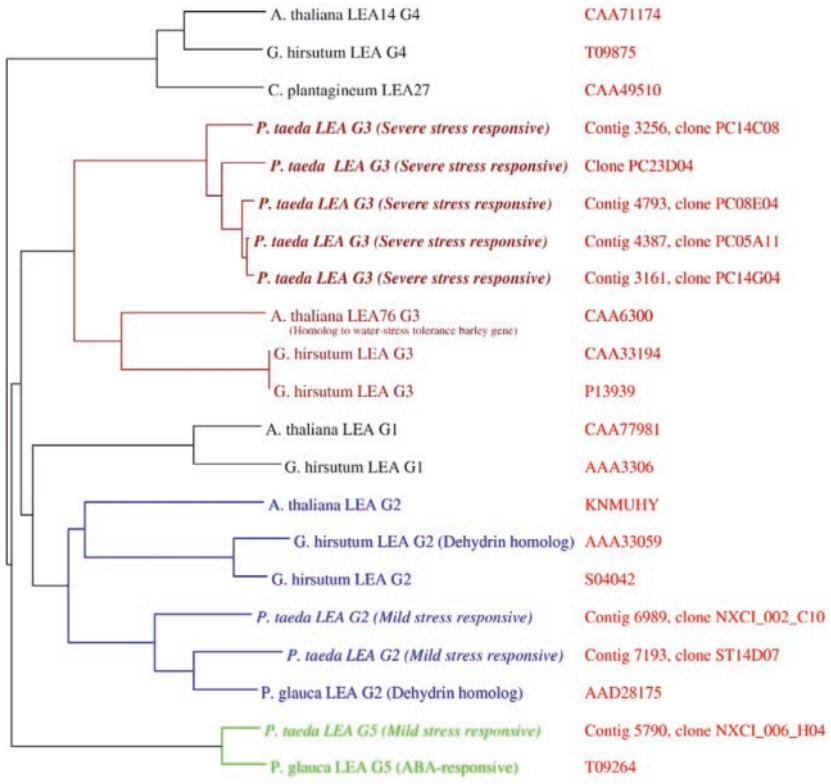

Figure 9. Phylogenetic tree of different groups of LEA proteins from loblolly pine compared with those from other species. The tree was generated from a multiple sequence alignment using pine contigs (that contain the expressed sequence tags [ESTs] with differential expression) and homologous sequences in GenBank (obtained through BLASTX). Alignment of translated pine sequences (translated using the SIXFRAME tool in WorkBench; http://workbench.sdsc.edu), and homologous sequences was performed using the ClustalW tool of WorkBench. The parameters used were the default parameters from WorkBench. The njplot tool was used to visualize the resulting tree and to generate a postscript file. Color and numbering was added to the postscript file. Pine clones are identified by contig number and clone identification, and other plant proteins are identified by GenBank numbers. Group 4 (G4) LEAs make up the first main group in black. Group 2 (G2) LEAs are colored in blue, Group 3 (G3) LEAs are colored in burgundy, and Group 5 LEAs are colored in green. The Group 1 (G1) LEAs are colored in black and lie between Groups 2 and 3.

2002), and ILP indicated that other cold-associated genes showed higher transcript levels in severe stress than in mild. Although the cold response has been linked with the drought response (Seki et al., 2001), we show for the first time, to our knowledge, that there is a clear distinction in the response to mild drought stress as compared with severe drought stress and cold stress. This suggests that severe drought stress activates different response pathways and that the plant can sense the degree of stress and respond accordingly.

Changes in Transcript Levels of Genes Encoding Enzymes of Core Carbon Metabolism Are Reflective of Acclimation under Mild Drought Stress

ILP found several rules indicating significant changes in carbon metabolism in response to mild drought stress. Our data suggest that carbon metabolism has a role in acclimation to drought stress. Using cluster analysis, Scheideler et al. (2001) found that genes categorized in carbon metabolism tended to have higher transcript levels in Arabidopsis treated with *Pseudomonas syringae* pv. *tomato*. They suggested that transcriptional regulation of genes encoding enzymes involved in carbon metabolism could direct the flow of carbon into specific pathways necessary to the response. That appears to be the case here, where at least two genes encoding enzymes involved in carbon metabolism, which could provide carbon skeletons to other metabolic processes, showed increased transcript levels during mild stress. One of these genes encodes pyruvate kinase, which may be involved in providing carbon skeletons for amino acid biosynthesis. Because it is also an ATP-producing enzyme, its activity may be necessary to augment the lower ATP levels caused by reduced photosynthetic capability. Another gene encodes pyruvate dehydrogenase, which is important for shuttling carbon from glycolysis into the TCA cycle. None of the other enzymes of the TCA cycle show increased transcript levels, which suggests that there is no net increase in flow through the TCA cycle. It is important to note, however, that regulation of these enzymes could be posttranscriptional. Because the TCA cycle provides carbon skeletons for many biosynthetic reactions, the increased transcript level of pyruvate dehydrogenase could be providing carbon for diverse uses within the plant.

The maintenance and possible increase of general carbon metabolic pathways could be reflective of the acclimation process taking place during mild drought stress. Trees grown under severe stress had reduced transcript levels of genes associated with the reductive pentose phosphate pathway, suggesting that little photosynthate was being generated, which is in agreement with the photosynthetic data. Under severe stress, there is a general increase in expression of genes associated with mitochondrial electron transport (see Supplemental Data), suggesting that the plant is compensating for drought-induced lack of reducing equivalents from photosynthetic electron

transport. Our previous array had no genes associated with core carbon metabolism; therefore, the results of the current array have given us a glimpse into the possible regulation of carbon flux during stress. With greater representation of genes encoding enzymes of carbon metabolism on future arrays, it will be possible to develop a detailed picture of gene expression related to carbon fluxes within the plant in response to drought.

Transcripts of Genes Encoding Flavonoid Enzymes Are Specifically Up-Regulated during Photosynthetic Acclimation under Mild Drought Stress

Other rules found by ILP suggest that nitrogen and sulfur metabolism and phenylpropanoid metabolism are important to acclimation to drought stress. Within nitrogen and sulfur metabolism, an increase in transcripts encoding enzymes of aromatic amino acid biosynthesis was found. An increase in transcript levels of several genes in the oxidative pentose phosphate pathway under mild stress (Fig. 7) suggests that carbon is being shuttled to aromatic amino acid biosynthesis. A similar increase is not seen in severe stress cycles. Because there is no overall positive trend in the genes of lignin biosynthesis, we suggest that carbon from aromatic amino acids is being channeled through the phenylpropanoid pathway into flavonoids. A rise in transcript levels of some genes important to flavonoid biosynthesis is seen under mild stress. Flavonoids have been implicated in stress responses (Winkel-Shirley, 2002), and we previously noted that isoflavone reductase was positively regulated in trees grown under mild stress but not trees grown under severe stress (Heath et al., 2002). An increase in dihydroflavonol-4-reductase transcripts was noted during osmotic, ionic, and heat stress in yeast (*Saccharomyces cerevisiae*; Garay-Arroyo and Covarrubias, 1999). Dihydroflavonol-4-reductase also showed an increase in transcript levels during dehydration in a drought-resistant cowpea (Iuchi et al., 1996), suggesting that dihydroflavonols are important to resistance to the stress. Flavonone-3-hydroxylase (naringenin-2-oxo dioxygenase) is the first enzyme in the production of dihydroflavonols and showed increased transcript levels in all three cycles of mild stress. Because photosynthetic acclimation only occurred in mild stress, it is possible that the flow of carbon to the flavonoids and specifically the dihydroflavonols is important to photosynthetic acclimation to mild drought stress. Although other research has implicated individual flavonoid genes in stress responses, we specifically show that a coordinate increase in transcript levels of several genes of the flavonoid pathway is associated with physiological acclimation of loblolly pine to mild drought stress.

Drought Regulation of Genes Encoding Polyamine Biosynthetic Enzymes Is Reflective of Senescence and Dormancy in Response to Severe Stress

Another gene associated with nitrogen and sulfur metabolism whose expression was affected by mild stress is Arg decarboxylase. This gene product is involved in polyamine biosynthesis and was noted as being drought responsive by Ozturk et al. (2002). Arg decarboxylase had higher transcript levels in the third cycle of mild stress when the plants had acclimated. Arg decarboxylase and *S*-adenosyl-Met (SAM) decarboxylase had lower transcript levels in trees grown under severe stress and were identified by ILP as being negatively regulated during severe cycles 1 and 2 (see Supplemental Data). Other researchers have correlated SAM decarboxylase with drought stress, suggesting that polyamine biosynthesis is important in plant responses to drought stress (Li and Chen, 2000a, 2000b). Spermine and spermidine, two polyamines in the SAM decarboxylase pathway, were found to induce elongation growth, increase photosynthetic capacity, and reduce membrane leakage in drought-stressed jack pine (*Pinus banksiana*; Rajasekaran and Blake, 1999), pointing to a direct role of polyamines in cellular responses to drought stress. The negative regulation of Arg decarboxylase and SAM decarboxylase during severe drought stress may suggest that synthesis of the polyamines is reduced and that the polyamines are not functional in the protection of cellular processes during severe stress.

Though polyamines have been associated with drought stress, we show that genes involved in their production are positively affected during photosynthetic acclimation under mild drought stress. The lower transcript levels of SAM decarboxylase and Arg decarboxylase under severe drought stress would indicate that polyamines are not as important to protection under severe drought stress. However, SAM decarboxylase is important to ethylene synthesis because it diverts SAM away from the ethylene biosynthetic pathway. Therefore, the down-regulation of SAM decarboxylase may reflect an increase in ethylene production in trees grown under severe stress. Ethylene has been associated with many stress responses in plants (e.g. Chen et al., 2002) and with abscission and senescence (Taylor and Whitelaw, 2001). Because ethylene is associated with senescence, the trees grown under severe stress may be switching to a senescence program. Two senescence-related genes were found to have higher transcript levels in trees at the third cycle of severe stress (see Supplemental Data). It is interesting to note that loblolly pine will drop its oldest needles at about the time it enters winter dormancy. The lack of growth of trees under severe stress and initiation of a senescence program reflects the entry of severely stressed trees into a state of dormancy. Again, we note that there is a distinction between mild and

severe drought stress in terms of gene expression as associated with photosynthetic acclimation.

CONCLUSIONS

The experimental results presented here reflect changes in transcript profiles of loblolly pine at two levels of drought stress, over several cycles. It can be seen that the changes in transcript profiles of loblolly pine as discovered through the use of microarrays are both quantitative and qualitative. A large number of genes showed changes in transcript levels due to mild stress, and there was an increase in the genes showing changes in expression reflective of unique acclimatory responses under mild stress. Many of the genes identified have been correlated previously with plant stress responses including LEAs and HSPs. However, we show for the first time, to our knowledge, that different members of individual gene families respond differently to the two levels of drought stress. We also show that group 2 LEAs are specifically up-regulated during photosynthetic acclimation under mild drought stress. Again, this indicates that the plant is able to sense the degree of stress and may activate different response pathways accordingly. It also delineates among the LEAs in terms of those that are expressed to allow acclimation to occur. Changes in transcript levels of genes encoding enzymes of core carbon metabolism are also reflective of acclimation during mild drought stress. This suggests that alteration of core metabolism can play a significant role in the stress response, perhaps in directing carbon or energy into pathways necessary for acclimation. We also show for the first time, to our knowledge, that genes of the flavonoid pathway are specifically up-regulated during mild drought stress and, thus, are associated with photosynthetic acclimation. Furthermore, regulation of genes encoding polyamine biosynthesis is indicative of senescence and dormancy in response to severe drought stress. The increase in transcripts profiled on the 2,173-element array allowed us to define roles for specific groups of genes in photosynthetic acclimation and develop a snapshot of how certain pathways are involved in the drought stress response. The addition of cDNAs encoding more members of these pathways will allow us to develop a greater understanding of how they function in the acclimatory response to mild drought stress.

One of the major advantages of data mining, as employed in Expresso, is the support for incorporating partial information into the analysis. Such partial information can take many forms: results from previous years' experiments, prior data about which genes are implicated in which physiological processes, and information from the literature. Data mining techniques should have the ability to flexibly incorporate such partial information and build from them, rather than function as isolated algorithms. The ILP technique described in this paper is only one of the many approaches that can harness partial information. We are now supporting the construction and mining of multimodal networks to help piece together parts of biological pathways. This approach uses algorithms from graph theory and probabilistic modeling to support the mining of such networks. Together, these approaches help address some of the major challenges in bioinformatics data mining.

MATERIALS AND METHODS

Plant Material

Rooted plantlets of loblolly pine (*Pinus taeda*) from the Atlantic Coastal Plain were produced clonally by Dr. Barry Goldfarb (North Carolina State University, Raleigh). These were transported to Virginia Tech (Blacksburg) and grown in a mixture of peat moss:perlite:sand (1:1:1 [v/v]) under natural daylight in a greenhouse. Supplemental lighting (mercury vapor and high-pressure sodium lamps) was used for 6 h in the evening to maintain 16-h daylength from July to the termination of the experiment in September. Temperature was maintained at 26°C ± 4.5°C during the day and 18°C ± 3.0°C at night. Plants were watered as needed and fertilized once a week with one-half-strength Hoagland solution. For drought stress experiments, water was withheld until a desired water potential had been reached: −1 MPa for mild stress and −1.5 MPa for severe stress. Trees were dried down to the desired water potential and then rewatered. Needles and stems were collected 24 h after rewatering. The stress was then repeated by again withholding water. Imposition of each level of drought stress was repeated three times with harvesting of needles and stems 24 h after rewatering (Fig. 1). Water potential of needles was measured in the morning (predawn) using a Plant Water Status Console (model 3005, Soilmoisture Equipment Corp., Santa Barbara, CA). Net photosynthesis was measured at the time of peak drought stress for each level of stress at light saturation on a LI-COR 6400.

Microarrays

A set of 2,173 clones was selected from five pine cDNA libraries to include representative genes from all 15 functional categories assigned by MIPS to Arabidopsis (Stasolla et al., 2003). The cDNA libraries were generated from xylem (normal wood vertical, compression wood inclined, side-wood inclined), pollen cones, and shoot tips (http://web.ahc.umn.edu/biodata/nsfpine/contig_dir6). The compression wood library is a library composed of ESTs derived from mechanically stressed tissue. Clone amplification, cleaning and spotting was carried out at North Carolina State University (Stasolla et al., 2003). Each clone was replicated four times on a slide. The clones selected to be on the array were placed within a hierarchy of functional categories designed by us to reflect processes that are affected by water deficit and the mechanisms that are thought to be employed to protect these processes.

RNA Extraction and Hybridization

RNA was extracted from needles according to the method of Chang et al. (1993) as modified in our laboratory. In brief, after addition of the extraction buffer to the tissue sample, the tissue was homogenized for 1 h with a polytron tissue homogenizer at full speed. After centrifugation (9,000*g* for 10 min), the supernatant was transferred to a new tube and one-fifth volume of 5% (w/v) cetyl-trimethyl-ammonium bromide was added (5% [w/v] cetyl-trimethyl-ammonium bromide and 0.7 M NaCl). This was heated to 65°C for 20 min. An equal volume of chloroform:isoamyl alcohol (24:1 [v/v]) was added, and the sample was centrifuged. The supernatant was re-extracted with an equal volume of chloroform:isoamyl alcohol (24:1 [v/v]). The sample was centrifuged, and the supernatant was transferred to a new tube. The RNA was precipitated overnight with one-half volume of 10 M LiCl. Each pair of RNAs to be compared (treated versus control for each time point) were reverse transcribed and labeled with Cy3 and Cy5 dyes (Stasolla et al., 2003). Reciprocal labeling of each comparison was implemented to control

for variation due to the dye and resulted in eight replicates per clone per comparison. Hybridizations were carried out according to Stasolla et al. (2003). A total of 24 slides were used for the microarray analysis. A modified loop design (Kerr and Churchill, 2001) was used to formulate comparisons between treated and control samples. For each degree of stress, 12 slides were used that compared treated samples with two control samples using reciprocal labeling. For example, mild treated 1 was hybridized to mild control 1 and mild control 4, and mild treated 2 was compared with mild control 2 and mild control 1. This resulted in a total of 16 replicate spots per treatment.

RT-PCR

RT-PCR was used to confirm the data obtained from the microarray experiments. The relative abundance of five ESTs (chalcone isomerase [ST07H08], naringenin-2-oxo dioxygenase [flavonone-3-hydroxylase, NXSI_063_D01], pine LEA group 2 [NXCI_002_C10], pine LEA group 3 [PC14G04], and HSP18 [ST40F04]) shown to be differentially expressed in the second and/or third cycles of either mild or severe stress in the microarray experiments was tested. The same total RNA samples used for the microarrays were used here. The concentration of total RNA was measured using the RiboGreen RNA Quantitation Reagent and Kit (Molecular Probes, Eugene, OR). First strand cDNA was reverse transcribed from 300 ng of total RNA using Taq-Man Reverse Transcription Reagents (Applied Biosystems, Foster City, CA) according to the manufacturer's instructions. Gene-specific primers were designed by the Primer Express 1.0 program (Applied Biosystems). The relative transcript abundance was monitored on an Applied Biosystems 7700 Sequencer using SYBR Green PCR Master Mix (Applied Biosystems). The adenosine kinase amplicon was used as an internal control for normalization.

Analysis

Hybridized slides were scanned using a ScanArray 5000 (Perkin Elmer, Boston). The resulting slide images were analyzed using QUANTARRAY software (Perkin Elmer). For each spot, the non-normalized intensity value was collected, with the exception of a few spots that QUANTARRAY flagged as unacceptable.

The sensitivity of individual genes to the experimental treatments—three cycles each of mild or severe stress—was estimated using the two-stage analysis of Wolfinger et al. (2001). The first stage addresses global effects, whereas the second stage estimates the interaction between individual genes and experimental treatments. The technique is highly sensitive to small differential changes in relative expression values (Jin et al., 2001) and was utilized by Jin et al. (2001) and Stasolla et al. (2003) in the analysis of their data.

In the first stage, the values of the major global factors are estimated by fitting the intensity data to this ANOVA normalization model:

$$y_{gijkmn} = \mu + T_i + A_j + D_k + P_m + (A \times P)_{jm} + \epsilon_{gijkmn}$$

The y_{gijkmn} values are $\log_2$-transformed intensity signals with indices for gene (g), treatment (i), array (j), dye (k), printing pin (m), and spot (n). The value μ is the mean of all the y_{gijkmn} values. The remaining variables are: treatment (T_i; two levels, treated and control), array (A_j; six arrays), dye (D_k; two dyes, Cy3 and Cy5), pin (P_m; four printing pins), the interaction [$(A \times P)_{jm}$] of array (j) and pin (m), and the model residual (ϵ_{gijkmn}).

The second stage of the analysis uses the residual values (ϵ_{gijkmn}) computed in the first stage to estimate the interaction between individual genes and treatments (stress level and control) at $\alpha = 0.05$. The ANOVA model for this stage is:

In these equations, the variables are: the mean (G_g) of the residual values for gene (g), the interaction [$(G \times T)_{gi}$] of gene (g) and treatment (i); the interaction [$(G \times A)_{gj}$] of gene (g) and array (j), the interaction [$(G \times D)_{gk}$] of gene (g) and dye (k), and the interaction [$(G \times S)_{gn}$] of gene (g) and spot (n). The γ_{gijkmn} are the residual stochastic errors. Both models were implemented using SAS (SAS/STAT Software version 8.2, SAS Institute Inc., Cary, NC). The least square means of $(G \times T)_{gi}$ and the differences in least square means between the stress level and control were calculated for each gene. Each difference constitutes the estimated $\log_2$-fold change for that gene.

With three cycles, each with two stress levels, a total of six experiments were analyzed. Nominally, each microarray contained 2,180 clones, each replicated four times. Of those, seven were controls and 2,173 were pine clones. Of the pine clones, the intensity of 74 signals reached the maximum 16-bit intensity value of 65,535. Because the intensity signals for these genes were not accurately recorded, they were eliminated from further analysis. For each of the remaining 2,099 pine clones, a hypothesis test between control and treated was constructed. To assure an experiment-wide false positive rate of 0.05, the P value cutoff was set at the Bonferroni value of $0.05/2{,}099 \approx 2.4 \times 10^{-5}$, as suggested by Wolfinger et al. (2001). A gene with a positive (respectively, negative) estimated fold change after Bonferroni adjustment is marked as positively (respectively, negatively) expressed. The remaining genes are marked unchanged.

Data Mining by ILP

After the analysis of variance, the data was mined via ILP, a data mining technique for finding relationships among data (Dzeroski and Lavrac, 2001). Computational tools for returning relationships in the form of rules are readily available. Because the space of potential rules is actually infinite, tuning the search parameters to obtain an efficient data mining procedure for biologist-friendly rules is a challenge that was specifically addressed by computer scientists within the Expresso project for this data set. Signature patterns of gene expression changes across experiments were also derived from ILP.

Data mining functionality in Expresso is based on a collection of relational database tables (15 tables for the current data set), implemented using the Postgres database management system. These tables summarize information such as experimental conditions (the different stress conditions), cDNA details (accession no., annotation, and putative functional categories), water potential and photosynthesis measurements, and expression levels. The Expresso query interface allows the biologist to perform a series of structured query language queries on these tables. Because the number of plausible structured query language queries is huge, the biologist more directly benefits from use of data mining techniques incorporated in Expresso. The ILP data mining technique (Muggleton, 1999) searches through the space of possible relationships among fields in the database to automatically identify the most promising rules (Heath et al., 2002). Rules are logical implications that summarize patterns found in the experimental and other data.

Expresso allows the biologist to specify the criteria for mining these rules. In the experiments reported here, we require that every implication have a strength of at least 80% and a support of at least 2. (In the rule A:−B, the strength is the fraction of genes satisfying B that also satisfy A, whereas the support is the number of genes that satisfy both A and B.) The results obtained are summarized in Figures 5 and 6 as cartoons. It is important that the criteria of 80% not be interpreted as a measure of statistical significance. Because we use the rules in a descriptive rather than a predictive fashion, the rules summarize the data and suggesting new and testable biological hypotheses.

Received May 17, 2003; returned for revision July 7, 2003; accepted August 28, 2003.

LITERATURE CITED

Alscher R, Chevone B, Heath L, Ramakrishnan N (2001) Expresso: a problem solving environment for bioinformatics: finding answers with microarray technology. *In* A Tentner, ed, Proceedings of the High Performance Computing Symposium, Advanced Simulation Technologies Conference. Society for Computer Simulation International, Seattle, pp 63–69

Alscher RG, Donahue JL, Cramer CL (1997) Reactive oxygen species and antioxidants: relationships in green cells. Physiol Plant **100:** 224–233

Babiychuk E, Kushnir S, Belles-Boix E, Van Montagu M, Inze D (1995) Arabidopsis thaliana NADPH oxidoreductase homologs confer tolerance of yeasts towards the thiol-oxidizing drug diamide. J Biol Chem **270:** 26224–26231

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

Chen WQ, Provart NJ, Glazebrook J, Katagiri F, Chang HS, Eulgem T, Mauch F, Luan S, Zou GZ, Whitham SA et al. (2002) Expression profile

matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. Plant Cell **14:** 559–574

Dong CN, Danyluk J, Wilson KE, Pecock T, Hunner NPA, Sarhan F (2002) Cold-regulated cereal chloroplast late embryogenesis abundant-like proteins: molecular characterization and functional analysis. Plant Physiol **129:** 1368–1381

Dzeroski S, Lavrac N, editors (2001) Relational Data Mining. Springer, Berlin

Fowler S, Thomashow MF (2002) Arabidopsis transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. Plant Cell **14:** 1675–1690

Gang DR, Kasahara H, Vander Mijnsbrugge K, Bauw G, Boerjan W, Van Montagu M, Davin LB, Lewis NG (1999) Evolution of plant defense mechanisms: relationships of phenylcoumaran benzylic ether reductases to pinoresinol-lariciresinol and isoflavone reductases. J Biol Chem **274:** 7516–7527

Garay-Arroyo A, Covarrubias AA (1999) Three genes whose expression is induced by stress in *Saccharomyces cerevisiae*. Yeast **15:** 879–892

Gerke V, Moss SE (2002) Annexins: from structure to function. Physiol Rev **82:** 331–371

Girke T, Todd J, Ruuska S, White J, Benning C, Ohlrogge J (2000) Microarray analysis of developing Arabidopsis seeds. Plant Physiol **124:** 1570–1581

Heath LS, Ramakrishnan N, Sederoff RR, Whetten RW, Chevone BI, Struble CI, Jouenne VY, Chen D, van Zyl L, Grene R (2002) Studying the functional genomics of stress responses in loblolly pine with the Expresso microarray experiment management system. Comput Funct Genomics **3:** 226–243

Ismail AM, Hall AE, Close TJ (1999) Purification and partial characterization of a dehydrin involved in chilling tolerance during seedling emergence of cowpea. Plant Phsyiol **120:** 237–244

Iuchi S, Yamaguchi-Shinozaki K, Urao T, Terao T, Shinosaki K (1996) Novel drought inducible genes in the highly drought-tolerant cowpea: cloning of cDNAs and analysis of the expression of the corresponding genes. Plant Cell Physiol **37:** 1073–1082

Jin W, Riley MR, Wolfinger RD, White KP, Passador-Gurgel G, Gibson G (2001) The contributions of sex, genotype and age to transcriptional variance in *Drosophila melanogaster*. Nat Genet **29:** 389–395

Kang J, Choi H, Im M, Kim SY (2002) Arabidopsis basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling. Plant Cell **14:** 343–357

Katagiri T, Takahashi S, Shinozaki K (2001) Involvement of a novel Arabidopsis phospholipase D AtPLD, in dehydration-inducible accumulation of phosphatidic acid in stress signaling. Plant J **26:** 595–605

Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ (2001) Gene expression profiles during the initial phase of salt stress in rice. Plant Cell **13:** 889–906

Kerr MK, Churchill GA (2001) Statistical design and the analysis of gene expression microarray data. Genet Res **77:** 123–128

Lee GJ, Vierling E (2000) A small heat shock protein cooperates with heat shock protein 70 systems to reactivate a heat-denatured protein. Plant Physiol **122:** 189–198

Lev-Yadun S, Sederoff R (2000) Pines as model gymnosperms to study evolution, wood formation and perennial growth. J Plant Growth Regul **19:** 290–305

Li ZY, Chen SY (2000a) Isolation and characterization of a salt- and drought-inducible gene for *S*-adenosylmethionine decarboxylase from wheat (*Triticum aestivum* L.). J Plant Physiol **156:** 386–393

Li ZY, Chen SY (2000b) Differential accumulation of the *S*-adenosylmethionine decarboxylase transcript in rice seedlings in response to salt and drought stresses. Theor Appl Genet **100:** 782–788

Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, in Arabidopsis. Plant Cell **10:** 1391–1406

Muggleton S (1999) Scientific knowledge discovery using inductive logic programming. Commun ACM **42:** 42–46

Muggleton S, Feng C (1990) Efficient induction of logic programs. *In* S Arikawa, S Goto, S Ohsuga, T Yokomori, eds, Proceedings of the First International Conference on Algorithmic Learning Theory. Japanese Society for Artificial Intelligence, Tokyo, pp 368–381

Ozturk ZN, Talame V, Deyholos M, Michalowski CB, Galbraith DW, Gozukirmizi N, Tuberosa R, Bohnert HJ (2002) Monitoring large-scale changes in transcript abundance in drought- and salt-stressed barley. Plant Mol Biol **48:** 551–573

Perez-Amador MA, Lidder P, Johnson MA, Landgraf F, Wisman E, Green PJ (2001) New molecular phenotypes in the dst mutants of Arabidopsis revealed by DNA microarray analysis. Plant Cell **13:** 2703–2717

Rajasekaran LR, Blake TJ (1999) New plant growth regulators protect photosynthesis and enhance growth under drought of jack pine seedlings. J Plant Growth Regul **18:** 175–181

Ristic Z, Yang GP, Martin B, Fullerton S (1998) Evidence of association between specific heat shock proteins and the drought and heat tolerance phenotype in maize. J Plant Physiol **153:** 497–505

Ruan Y, Gilmore J, Conner T (1998) Towards Arabidopsis genome analysis: monitoring expression profiles of 1400 genes using cDNA microarrays. Plant J **15:** 821–833

Sang Y, Zheng S, Li W, Huang B, Wang X (2001) Regulation of plant water loss by manipulating the expression of phospholipase D. Plant J **28:** 135–144

Scheideler M, Schlaich NL, Fellenberg K, Beissbarth T, Hauser NC, Vingron M, Slusarenko AJ, Hoheisel JD (2001) Monitoring the switch from housekeeping to defense metabolism in *Arabidopsis thaliana* using cDNA microarrays. J Biol Chem **277:** 10555–10561

Schultz RP (1999) Loblolly: the pine for the twenty-first century. New Forests **17:** 71–88

Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K (2001) Monitoring the expression patterns of 1300 Arabidopsis genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell **13:** 61–72

Sioson A, Watkinson JI, Vasquez-Robinet C, Ellis M, Shukla M, Kumar D, Ramakrishnan N, Heath LS, Grene R, Chevone BI et al. (2003) Expresso and Chips: creating a next generation microarray experiment management system. *In* Proceedings of the Next Generation Software Workshop, 17th International Parallel and Distributed Processing Symposium, Nice, France. IEEE Computer Society, Los Angeles, CA

Somerville C, Somerville S (1999) Plant functional genomics. Science **285:** 380–383

Stasolla C, van Zyl L, Egertsdotter U, Craig D, Liu W, Sederoff RR (2003) The effects of polyethylene glycol on gene expression of developing white spruce somatic embryos. Plant Physiol **131:** 49–60

Takeuchi M, Ueda T, Yahara N, Nakano A (2002) Arf1GTPase plays a role in the protein traffick between the endoplasmic reticulum and Golgi apparatus in tobacco and Arabidopsis cultured cells. Plant J **31:** 499–515

Taylor TE, Whitelaw GA (2001) Signals in abscission. New Phytol **151:** 323–339

Winkel-Shirley B (2002) Biosynthesis of flavonoids and effects of stress. Curr Opin Plant Biol **5:** 218–223

Wolfinger RD, Gibson G, Wolfinger ED, Bennett L, Hamadeh H, Bushel P, Afshari C, Paules RS (2001) Assessing gene significance from cDNA microarray expression data via mixed models. Comput Biol **8:** 625–637

Xu D, Duan X, Wang B, Hong B, Ho THD, Wu R (1996) Expression of a late embryogenesis abundant protein gene, HVA1, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiol **110:** 149–156

Zhang L, Ohta A, Takagi M, Imai R (2000) Expression of plant group 2 and group 3 lea genes in *Saccharomyces cerevisiae* revealed functional divergence among LEA proteins. J Biochem **127:** 611–616