IDENTIFICATION OF DROUGHT RESPONSIVE GENES IN ALEPPO PINE (*Pinus halepensis*) AND LOBLOLLY PINE (*Pinus taeda*.L)

A Thesis

by

PRATHEESH SATHYAN

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2004

Major Subject: Molecular and Environmental Plant Sciences

IDENTIFICATION OF DROUGHT RESPONSIVE GENES IN ALEPPO PINE (*Pinus halepensis*) AND LOBLOLLY PINE (*Pinus taeda*.L)

A Thesis

by

PRATHEESH SATHYAN

Submitted to Texas A&M University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Approved as to style and content by:

Carol A. Loopstra
(Chair of Committee)

Marla L. Binzel
(Member)

Thomas D. Byram
(Member)

Charles T. Smith
(Head of Department)

Marla L. Binzel
(Chair of MEPS Faculty)

December 2004

Major Subject: Molecular and Environmental Plant Sciences

ABSTRACT

Identification of Drought Responsive

Genes in Aleppo Pine (*Pinus halepensis*) and Loblolly Pine (*Pinus taeda* L)

(December 2004)

Pratheesh Sathyan, B.S., Kerala Agricultural University

Chair of Advisory Committee: Dr. Carol A. Loopstra

Drought is a major constraint for attaining economic yield in tree crops. As an initial step to understand molecular response to water-deficit-stress in trees, gene expression in response to water stress was quantified using real-time RT-PCR. The specific objectives established for this to were I. to identify and characterize the genes induced by drought stress in Aleppo pine (*Pinus halepensis*) and II to identify and quantify the differentially expressed genes in different populations of Loblolly pine (*Pinus taeda*.L) due to water deficit (chapter III). Results of these studies may be used to identify candidate genes for future breeding programs against water-deficit-stress.

ACKNOWLEDGEMENTS

I sincerely thank my advisor, Dr. Carol Loopstra, for her help and guidance. She has been not only an advisor but also a great friend to me. I am also thankful to the other members of my graduate committee, Dr. Marla Binzel and Dr. Thomas Byram for their valuable help and advice. I also express my sincere gratitude to all my friends who helped me in all my efforts.

This study was supported by the Texas Higher Education Coordinating Board Advanced Technology Program (000517-0119-2001).

TABLE OF CONTENTS

CHAPTER Page

I INTRODUCTION: THE MOLECULAR RESPONSE OF PINES TO WATER-DEFICIT-STRESS 1

II GENES INDUCED BY WATER-DEFICIT-STRESS ARE DIFFERENTIALLY EXPRESSED IN TWO POPULATIONS OF ALEPPO PINE (*Pinus halepensis)* 5

Introduction 5
Materials and Methods 7
Plant Materials 7
Library Construction 9
Gene Sequencing 10
Extraction of RNA for Real-time RT-PCR 11
Real-time RT-PCR 11
Northern Blot Hybridizations 12
Results 12
Identification and Characterization Water-Stress-Responsive Genes 12
Quantification of Differential Expression Using Real-time RT-PCR 17
Effect of Provenance on Gene Expression 19
Discussions 20

III DIFFERENTIAL EXPRESSION OF WATER-DEFICIT-GENES IN TWO POPULATIONS OF LOBLOLLY PINE (*Pinus taeda* L.) AND CHARACTERIZATION OF LEA MULTIGENE FAMILY MEMBERS 27

Introduction 27
Materials And Methods 30
Extraction of RNA and Preparation of Samples for Real-time RT-PCR 31
Identification of Drought Responsive Gene Sequences 32
Real-time RT-PCR 33
Results 35
Differential Expression Due Water-Deficit-Stress 40
Effect of Provenances on Gene Expression 42
Discussions 51

IV CONCLUSIONS 58

LITERATURE CITED 60

VITA 67

LIST OF FIGURES

Page

Figure 1: Northern Blot showing expression of genes by water-deficit-stress in aleppo pine. 19

Figure 2: Comparison of LEA III protein from *P.taeda* at the amino acid level. (The bold capital letters indicate identical amino acids and the grey shading indicates the similar amino acids.).. 36

Figure 3: Comparison of PtLEAII protein with pfam members. (Identical amino acids are shaded black and similar amino acids are shaded grey.)............................. 37

Figure 4. Scatter Plots showing the induction of the genes due to water-deficit-stress treatment in Lost Pines (LP) and south Louisiana (SLA) 43

LIST OF TABLES

Page

Table 1: Plant proteins similar to aleppo pine proteins in the Genbank NR database as determined with tblastx. 15

Table 2: ΔCt (Cycle Threshold) values following control, moderate and high stress treatments of two populations of aleppo pine. 18

Table 3: Loblolly pine genes induced due to drought treatment 34

Table 4: Normalized ΔCt (Cycle Threshold) values following stress treatments of loblolly pine 41

CHAPTER I

INTRODUCTION: THE MOLECULAR RESPONSE OF PINES TO WATER-DEFICIT-STRESS

Plants are exposed to various kinds of abiotic stresses in their natural environment. Abiotic stresses include drought, salinity, freezing, nutrient deficiency and chilling. Drought is one of the most important abiotic stresses to which plants are exposed. Drought is a meteorological phenomenon that is described as a period of meager rainfall, which causes depletion of soil moisture (Kramer and Kozlowski, 1979). Plant responses to water stress are variable depending upon the age of the plant, species, genotype and other environmental conditions. Drought is a major constraint for growth of vegetation and has a differential effect on forest crops depending on age (Newton et al., 1991(a)). In young seedlings it causes seedling mortality and in mature crops it leads to reduced growth. Water deficit has been attributed to be one reason for a 57% loss of first year seedlings in pine forests (Wilinston, 1972). In case of severe drought periods, up to 65% of loblolly pine seedlings can be lost (Philips, 1998). In mature loblolly pine stands, 40% of variation in radial growth is due to water deficit (Chang and Aguilar, 1980).

This thesis follows the style of Plant Physiology.

Water deficit not only causes economical burdens but also leads to ecological degradation of the site. Initiation of plant response to water-deficit-stress begins with the recognition of stress, with loss of water from the cell perceived, followed by triggering of a cellular signal transduction pathway.

Plants respond to this initial signal by either dehydration tolerance or dehydration avoidance mechanism (Bray, 1997). Both these mechanisms involve changes in gene expression. In dehydration tolerance, the plants tolerate low water potentials while in dehydration avoidance plants delay the onset of water deficits in the tissues. Both these mechanisms are centered around the accumulation of certain differentially expressed gene products. Signaling pathways can be divided into ABA-dependent pathways and ABA- independent pathways. Gene products that are activated by these pathways can be classified as functional proteins and regulatory proteins (Shinozaki and Yamaguchi-Shinozaki, 1997). Functional proteins include water channel proteins, key enzymes for osmolyte biosynthesis, chaperones, LEA (late embryogenesis abundant) proteins, proteinases and detoxicating enzymes. Regulatory proteins include transcription factors, protein kinases, and phospholipases. Regulatory proteins are involved in the further regulation of signal transduction and gene expression in stress tolerance. Transgenic expression of genes that encode some of these proteins have been shown to increase drought tolerance in plants (Sivamani et al., 2000).

Water stress is one of the vital barriers that hampers yield in trees. Changes in the climatic regime due to the green house effect and unpredicted droughts aggravate the situation. These factors coupled with the vital importance of wood in today's world paved the way for research to understand the molecular responses of trees to abiotic stress. Our understanding of molecular response of tress to water-deficit-stress is still in its infancy, partly due to their long regeneration age and large size. In addition the genus Pinus is burdened with a huge genome size which ranges from 22 to 40 pg (Wakemiya et al., 1996) or haploid nuclear DNA which is 200-400 times the size of Arabidopsis (Somerville and Somerville, 2000). These differences as well as evolutionary divergence act as a hindrance for the use of model organisms such as *Arabidopsis thaliana.* Therefore understanding the molecular responses of pines is a big challenge for forest molecular biologists. New technologies such as microarraies and real-time RT-PCR allow studies of gene expression on minute scales. Real-time RT-PCR is a powerful and sensitive method to study gene expression (Bustin, 2000). Real-time RT-PCR can be used to quantify gene expression as well to differentiate between gene families, this attains special importance in pines which are reported to have presence of complex gene families (Kinlaw and Neale, 1997). In order to understand how gene expression varies in response to water-deficit-stress in two different species of pines from different populations the following specific objectives were established for this study: Objective I. Identification and characterization of genes induced by drought stress in aleppo pine (*Pinus halepensis*) (*chapterII);* Objective II. Identification and quantification of differentially expressed genes in different populations of loblolly pine (*Pinus taeda* L.)

due to water deficit (*chapter III*). Result of these studies may be useful to identify candidate genes for future breeding programs for resistance to water-deficit-stress.

CHAPTER II

GENES INDUCED BY WATER-DEFICIT-STRESS ARE DIFFERENTIALLY EXPRESSED IN TWO POPULATIONS OF ALEPPO PINE (*Pinus halepensis)*

Introduction

Forest species are subjected to numerous environmental stresses including water deficit. Drought can be an important factor affecting forests worldwide. In young seedlings, drought causes mortality (Wilinston, 1972). In mature trees, there is a drastic reduction in yield and occasionally tree death. Much of the variation in ring width in conifers can be attributed to water deficit (Zahner and Donnelly, 1967). Changes in the global environment due to various human activities have altered the global climatic regime to some extent and more dramatic changes to the environment, including increased drought, are expected in the future. Since drought has an immense influence on yield in many plant species, efforts to curtail this loss are an important area of research in plant molecular biology.

The molecular mechanisms by which trees respond to water-deficit-stress are not fully understood. Molecular and physiological aspects of drought tolerance in forest trees were reviewed by Newton et al. (1991(a)). The molecular mechanisms by which domesticated plants respond to water-deficit-stress have also been reviewed (Bray, 1997; Ingram and Bartel, 1996; Shinozaki and Yamaguchi-Shinozaki, 1997). Studies with model organisms such as *Arabidopsis* can be applied to trees but due to the very large genome sizes of some pines, potential relationships between genome size and water relations in tree species (Wakamiya et al., 1996) their evolutionary divergence and the

relatively long life spans of trees, the scope of this application is reduced to a basic level. Loblolly pine genes induced by water-deficit-stress were first reported by Chang et al. (1996(b)). More recently, a large number of loblolly pine ESTs representing genes induced by drought stress were sequenced by different groups. Many genes induced by exposure to polyethylene glycol (PEG) have been identified in hydroponically-grown *Pinus pinaster* (maritime pine) (Dubos and Plomion, 2003; Dubos et al., 2003). However, the results need to be verified with seedlings grown under field conditions.

In this paper we analyzed water-deficit-responsive genes in aleppo pine (*Pinus halepensis*) seedlings. Aleppo pine covers approximately 3.5 millions hectares in the Mediterranean area. It extends from 9° longitude west in Morocco to 36° longitude east in Jordan and from 45° latitude north in France to 31°30' latitude in Palestine and Israel (Schiller et al., 1986). Aleppo pines growing in countries like Israel and Jordan belong to the Mediterranean population of this species and have a remarkable capacity to survive water-deficit-stress. Due to its survival following drought, this species is extensively used for afforestation in these areas and there is considerable interest in finding the best populations for establishing plantations (Grunwald et al., 1986).

To gain a better understanding of the genes involved in responses to water-deficit-stress in aleppo pine, we constructed a cDNA library from stressed roots. We differentially screened the library to identify genes up- and down-regulated following water-deficit-stress and sequenced some of the highly induced genes. Real-time RT-PCR was used to quantify expression of selected genes in seedlings from two populations.

Materials and Methods

Plant Materials

Seed from two populations of aleppo pine within Israel were provided by Dr. Gabriel Schiller of the Volcani Center of the Agriculture Research Organization in Israel. Of the Israeli seed sources available to us, these two populations had the greatest difference in precipitation. Seed from plantations was rejected from consideration because these trees are often from unknown or multiple origins. Since a major goal was to identify drought tolerance genes, seedling progeny of a native stand of trees growing on a relatively dry site were used for library construction and screening. A native aleppo pine stand growing at the Yirka site in northern Israel was thought to be drought tolerant because the site has a mean annual rainfall of 685mm, which falls near the low end of the range (580 to 925mm) observed for natural stands within Israel. Progeny of trees from a site with greater rainfall were also included in this study. These trees would presumably be less tolerant than those native to the drier Yirka location. Seeds were chosen from the natural stand at Beit Jann, which receives the highest mean annual rainfall of any available source with 925mm. Yirka and Beit Jann are geographically closely located at approximately 32.5^{o}N latitude and 35°E longitude. The two stands grow at dissimilar elevations, with Yirka at 430m and Beit Jann at 850m on Mount Meron, the highest mountain in Israel.

Seeds of the two sources were sown in flats containing coarse vermiculite, and after germination they were transplanted into 15-cm diameter 1.8-liter pots containing two parts calcined clay for soil-like drying characteristics and one part coarse perlite for

aeration. Seedlings were grown in a glasshouse and irrigated every two to three days using purified water. A weekly application of complete nutrient solution was applied to encourage rapid, balanced growth until they reached eight months of age and typical stem lengths of 20-25 cm. Water-deficit-stress was imposed in the following fashion. At air temperatures ranging between 15°C minimum at night and 38°C maximum during the day, well-watered plants received enough purified water to reach full saturation every two days. Water was withheld from increasing numbers of individuals for three weeks. Thus, at the termination of the experiment, some seedlings had received no irrigation for 21 days, others for 19 days, others for 16 days, others for 14 days, et cetera, producing a wide range of water-deficit-stress levels. Using seedling side branches that were five to eight cm long, predawn water potentials were determined with a Scholander pressure chamber (Scholander et al., 1965). Roots were rinsed with water to remove the potting medium. Seedlings were quickly separated into roots, stems, and needles, frozen in liquid nitrogen, and stored at –80°C to await RNA isolation. For purposes of RNA pooling, well-watered seedlings (with water potentials ranging from -0.35 to -0.50 MPa) were included in the nonstressed pool. Seedlings whose water potentials ranged between –1.0 and –1.5 MPa were considered to be moderately stressed, and those between –2.0 MPa and –2.75 MPa were considered highly stressed. This plant material was used for library construction and screening and initial northern blot analyses. A new set of seedlings was raised and treated in a similar manner to provide materials for the real-time RT-PCR analyses with the exception that RNA from individual trees was used for the experiment rather than pooled RNA.

Library Construction

RNA extractions for the library construction were done using a guanidine thiocyanate extraction and cesium chloride pad centrifugation protocol modified from Sambrook et al. (1989) as previously described (Loopstra and Sederoff, 1995). The cDNA library was constructed using pooled RNA from roots of stressed seedlings whose water potentials ranged from –2.0 to –2.75 MPa. The cDNA library was constructed using a TimeSaver cDNA Synthesis kit with the directional cloning toolbox (Amersham Pharmacia, Piscataway, NJ). The cDNA inserts were ligated into the *EcoR* I and *Not* I sites of the pSport1 vector (Invitrogen, Carlsbad, CA). Approximately 21,500 white colonies were picked into 384-well plates containing LB with ampicillin and freezer buffer, grown overnight at 37^0 C and stored at -80^0 C.

A Biomek 2000 robotic workstation (Beckman, Fullerton, CA) was used to produce high-density filters containing DNA from 1536 colonies per filter in a duplicated pattern. The library was screened using two probes. One probe was produced using root mRNA from five nonstressed seedlings (-0.35 to -0.50 MPa). The other was produced using root mRNA from five stressed seedlings (–2.15 to –2.75 MPa). Cold first strand syntheses were done using 500 ng of each polyA RNA and oligdT as the primer using Superscript II (Invitrogen). Each of these cDNA pools was then extracted with phenol-chlorophorm:isoamyl alcohol (CIA) and CIA alone, and precipitated with 10 µg tRNA. The pellets were rinsed with 70% ethanol, and dried. The resulting cDNA pellets were resuspended in 20 µl water, boiled for 10 min and chilled on ice. Random primer DNA labeling was then performed using ^{32}P-labeled dCTP and dATP. Incorporation

was estimated at 60% for both nonstressed- and stressed-root probes following sephadex spin columns.

Two identically treated membrane sets were hybridized with the cDNA probes. Processed membranes were prewashed for 30 min at 37°C in 2x SSC with 1% SDS and prehybridized in Church buffer (1mM EDTA, 0.5M NaPO4 pH 7.2, 7% SDS, 1% BSA) with 200ug/ml herring sperm DNA overnight at 65° C. This buffer was drained and replaced with buffer containing boiled probes from either stressed or nonstressed roots. The probes were allowed to hybridize at 65 ^{o}C for 4 days in a rotary hybridization oven. The membranes were washed in 2x SSC with 1% SDS, followed by two high stringency washes in 0.2x SSC with 0.1% SDS at 65 ^{o}C. The autoradiography films were compared in order to identify clones of genes with altered levels of expression. (Dr. Loopstra kindly performed the library construction and screening for the project).

Gene Sequencing

Gene sequencing was done using a Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems (ABI), Foster city, CA) and an ABI 3100 Genetic Analyzer. M13 universal primers were used to obtain the first 500 bases in both the forward and reverse directions. Gene specific primers were designed using Primer Express software (version 2.0, ABI) so the complete sequence of both strands could be attained in both directions. The reaction conditions for the sequencing reactions were 96^{0}C for 10 sec, 50^{0}C for 5 sec, 60^{0}C for 4 min, 4^{0}C hold. For the gene specific primers the annealing temperature was 5^{0}C below the melting temperatures of the primers.

Extraction of RNA for Real-time RT-PCR

Extraction of good quality RNA from pine seedlings is difficult as phenolics are associated with stressed roots. In order to overcome this problem we used the protocol described by Chang et al. (1993) with the following modifications. The ratio of extraction buffer to tissue was increased and approximately 2g of the ground tissue were added to 20 ml of buffer. Additional CIA extractions were done until the interphase became thin. In order to remove the remaining traces of DNA, a DNA-free kit (Ambion Inc, Austin, Texas) was used according to the user's manual. Quantification of the RNA was done using a spectrophotometer, and samples with a 260/280 ratio of 1.8 or higher were used.

Real-time RT-PCR

The RNAs were reverse transcribed using a TAQMAN Reverse Transcriptase Kit (ABI) using the following conditions: 25^0 C for 10 min, 48^0 C for 30 min, and 95^0 C for 5 min. Real-time RT-PCR primers were designed using Primer Express (version 2.0, ABI) and regular custom primers ordered (Invitrogen). SYBR green was used to quantify the RT-PCR products as follows. Five microliters of SYBR green master mix (ABI), 3.5µl of depc H_20, 0.5µl of template, and 0.5µl (1µg) of each primer were dispensed into 384-well plates (ABI). An ABI 7900HT Sequence Detection System was used according to the manufacturer's recommendations. Melting Curve Analyses were done to ensure product specificity and helped differentiate between the true product and primer dimers.

We used Δ Ct (cycle threshold) values to compare the differential expression between the various treatments. In this method the Ct value of the normalizer is deducted from the Ct value of the sample to obtain the ΔCt. We used the pine ribosomal 18s gene as the normalizer.

$$\Delta Ct = Ct_{(Sample)} - Ct_{(normalizer)}$$

ΔCt values were calculated for each of the replicates, and the average of these values was calculated for the comparisons between treatments. A gene with a low ΔCt value signifies higher expression, and a high ΔCt signifies low expression. We conducted ANOVA with ΔCt values to verify the statistical significance of the different treatments.

Northern Blot Hybridizations

Northern blot hybridizations were done for a few highly expressed genes to confirm the real-time RT-PCR results using the protocol described by Loopstra and Sederoff (1995).

Results

Identification and Characterization Water-Stress-Responsive Genes

To identify genes differentially expressed due to water-deficit-stress, we screened a *P. halepensis* cDNA library of 21,500 clones. We identified 156 clones representing genes with increased expression levels in stressed roots and 56 with decreased expression in stressed roots. We partially sequenced sixty clones and found

some were repeated multiple times. Fourteen of these genes were completely sequenced in both directions and characterized using real-time RT-PCR. Searches of the NCBI (National Center for Biotechnology Information), Genbank NR (non-redundant) and dBEST databases revealed that several of the clones show similarity to previously identified genes with known functions, some of which are related to desiccation tolerance. The others were similar to hypothetical or unknown proteins in the NR or EST databases.

PhALDH shows significant similarity towards mitochondrial aldehyde dehydrogenases including those from *Lotus corniculatus* and *Arabidopsis thaliana.* Our predicted protein is 84.3% identical to the lotus protein over 51 amino acids. Analyses using PSORT (Nakai and Horton, 1999) and Mitoport II (Claros and Vincens, 1996) revealed that our predicted protein contains a mitochondrial targeting sequence.

PhLEA is most similar to two putative LEA (late embryogenesis abundant) proteins from *Picea glauca* (white spruce). LEAs are a large family of proteins expressed in plants in response to various stresses including water-deficit-stress. The closest angiosperm gene with a proposed function was a LEA gene identified from tobacco (e-value 0.006). In order to find how similar these clones are structurally, we compared the translated proteins of EMB32 (the most similar spruce LEA), PhLEA and the tobacco LEA. PhLEA and EMB32 are 88.4% identical over 95 amino acids. PhLEA and the tobacco LEA are only 33.7% identical over 95 amino acids.

PhCYC showed significant similarity with many cyclophilins. PhCYC is most similar to a *Picea abies* (Norway spruce) cyclophilin, CyPA, which is a cyctosolic form

expressed in both roots and leaf tissue (Luan et al., 1994(a)). PhCYC and the *Picea* protein are 91.9% identical at the amino acid level. The nearest angiosperm counterpart (*Catherunthus rosea)* is 83.1% identical over 172 amino acids. PSORT (Nakai and Horton, 1999) was used to look for a cytoplasmic signal sequence (Luan et al., 1994 (b)). However, we found that PhCYC has a signal sequence that directs it to the mitochondria or chloroplast.

PhCTL is very similar to several chitinases. This was also the gene we most repeatedly cloned from our differentially screened library, and therefore it can be taken as an indication of the abundance of this transcript in stressed roots. Chitinase genes are differentially expressed in response to various kinds of abiotic and biotic stress conditions. Following a BLAST search of the NR database, we found PhCTL to be most similar to a gene from *Persea americana* (avocado). PhCTL1 is 88.5% identical to the 26 amino acid N-terminal region characteristic to class I chitinases.

PhSUS is most similar to a *Pisum sativum* sucrose synthase with 78.8% identity over 151 amino acids. An A*rabidopsis thaliana* sucrose synthase (*SUS1*) was reported to be induced in response to both cold and drought stress (Dejardin et al., 1999).

PhPGRP is primarily composed of prolines (52.6%), glycines (31.6%) and histidines (12.3%). The amino acid sequence is: MHPGG**PPPP**GG**PPPP**GG **PPPP**GDHPHGHHH**PPPP**GPGG**PPPPP**GPGG**PPP**HCGPGGPGGRCC. It was also interesting to note the distribution of prolines and glycines in this protein. It contains four repeats of **PPPP**, one **PPPPP**, one **PPP**, seven repeats of GG and four repeats of GPG. PhPGRP is similar to various loblolly and maritime pine ESTs from normal,

Table 1: Plant proteins similar to aleppo pine proteins in the Genbank NR database as determined with tblastx.

Clone	Accession#	Similar Proteins	Accession # of the most similar protein	e-value
PhALDH	AY705795	Aldehyde dehydrogenase (Lycopersicon esculentum)	BT013599	3e-24
PhPGRP	AY705796	No hits		
PhMYB	AY705797	MYB Transcription factor (Arabidopsis thaliana)	NM_126168	3e-49
PhLEA	AY705798	LEA protein (Picea glauca)	L47743	e-116
Ph16	AY705799	Unknown protein (Oryza sativa)	AK121344	3e-44
PhCYC	AY705800	Cyclophilin (Picea abies)	AJ271126	e-116
Ph20	AY705801	No hits		
PhSUS	AY705802	Sucrose synthase (Pisum sativum)	AJ001071	3e-92
PhIP	AY705803	Inorganic pyrophosphates (Arabidopsis thaliana)	AY551439	e-149
PhCTL	AY705804	Endochitinase (Persea americana)	Z78202	e-118
Ph31	AY705805	No hits		
Ph4Cl	AY705806	Putative acyl-CoA synthetase/ 4-coumarate CoA ligase (Capsicum annuum)	AF354454	2e-56
Ph33	AY705807	No hits		
Ph38	AY705808	Hypothetical Norway spruce protein (Picea abies)	AJ132539	6e-87

compression, opposite, earlywood and latewood libraries but is not particularly similar to anything in the NR database except for other proline-glycine rich proteins that do not appear to be related.

PhMYB is similar to a myb transcription factor from *Arabidopsis thaliana*. Our predicted polypeptide and the *Arabidopsis thaliana* myb transfactor are 65.1% identical over 105 amino acids.

PhIP is similar to inorganic pyrophosphates and the predicted protein is 71.2% identical over 302 amino acids to an *Arabidopsis thaliana* inorganic pyrophosphate.

Ph4Cl has significant similarity to a putative acyl-CoA synthetase from *Capsicum annuum* (64% of 91 amino acids) and Arabidopsis putative 4-coumarate-CoA ligases (71% over 66 amino acids). However, we did not see significant similarity to a known 4Cl from loblolly pine xylem (Voo et al., 1995).

Several genes have no obvious similarity to genes encoding proteins with a known function. Within the NR database, Ph38 is most to similar to a gene encoding a hypothetical Norway spruce (*Picea abies*) protein (65.8% identical over 266 amino acids). Ph16 shows similarity to expressed proteins in arabidopsis and rice. Both Ph38 and Ph16 are very similar to ESTs from loblolly and maritime pines. Ph20, Ph31, and Ph33 are not particularly similar to any previously identified genes from angiosperms but are very similar to ESTs from loblolly and maritime pines. The maritime pine libraries were made from drought stressed roots (Dubos and Plomion, 2003) while the loblolly pine libraries were made from xylem (Whetten et al., 2001) drought stressed roots, well-watered roots and seedlings. Ph38 and Ph20 are also similar to genes from a

white spruce cambium and phloem library. The 14 clones, their accession numbers and the most similar proteins are given in Table 1.

Quantification of Differential Expression Using Real-time RT-PCR

The 14 selected genes were further characterized using real-time RT-PCR. All 14 genes had increased expression following the stress treatments. The average ΔCt value of the high stress treatments was deducted from the average ΔCt value of the control treatments and the results are given in Table 2. The differences in ΔCt values between the highly stressed and control treatments ranged from 1.99 to 8.00 cycles and represent the level of induction. The gene with the greatest difference in expression between the treatments was PhLEA (LEA gene) with a difference of 8.00 cycles in the Beit Jann population. An eight cycle difference reflects an approximate 256 fold difference in expression following the two treatments assuming the PCR is efficient and products double each cycle. This was followed by the sucrose synthase gene PhSUS (6.7 in Beit Jann), the chitinase gene PhCTL (6.5 in Beit Jann and the cyclophilin PhCYC (6.1 in Beit Jann).

Genes of unknown function had differences ranging from 1.99 to 4.64 cycles. The pattern of induction varied by gene (Table 2). For some genes, the moderate stress treatment resulted in expression levels almost equal to those caused by high stress (i.e. Ph33 in the Yirka population) while for others the expression was intermediate between the control and high stress treatments (i.e. PhLEA). In a few cases, moderate stress did not cause much induction over the control treatment (i.e. Ph4CL in the Yirka population).

Table 2: ΔCt (Cycle Threshold) values following control, moderate and high stress treatments of two populations of aleppo pine.

Clone Name	Yirka				Beit Jann			
	ΔCT control	ΔCT moderate	ΔCT high	ΔCTc-ΔCTh	ΔCT control	ΔCT moderate	ΔCT high	ΔCTc-ΔCTh
PhALDH	13.7	13.6	10.6	3.1	12.2	9.6	7.6	4.6
PhPGRP	19.0	17.5	15.9	3.1	16.5	16.5	11.4	5.1
PhMYB	17.7	15.2	12.5	5.1	15.3	12.7	11.4	3.9
PhLEA	13.8	11.2	7.9	5.9	15.3	10.6	7.3	8.0
Ph16	18.7	17.0	16.7	2.0	17.3	16.4	13.1	4.2
PhCYC	16.2	15.2	13.1	3.1	17.3	15.6	11.2	6.1
Ph20	19.9	18.3	17.5	2.3	20.5	17.4	18.1	2.4
PhSUS	20.8	16.6	14.7	6.0	18.3	15.3	11.6	6.7
PhIP	12.4	10.9	9.25	3.1	11.5	11.0	7.3	4.2
PhCTL	19.5	17.1	13.8	5.6	16.5	11.3	10.0	6.5
Ph31	12.6	9.4	7.9	4.6	7.5	5.9	3.0	4.5
Ph4Cl	20.9	20.6	16.7	4.2	18.3	16.2	14.0	4.3
Ph33	15.6	13.1	13.4	2.2	6.9	4.8	3.2	3.7
Ph38	11.0	9.0	9.0	1.9	9.2	8.0	4.9	4.3

Northern blot analyses were done for a few of the highly differentially expressed genes including PhLEA, PhCYC and PhSUS. In concurrence with the real-time RT-PCR data our northern blot analyses showed PhLEA to be highly differentially expressed. There was also a difference in the level of expression between the two treatments for the other genes (Fig 1).

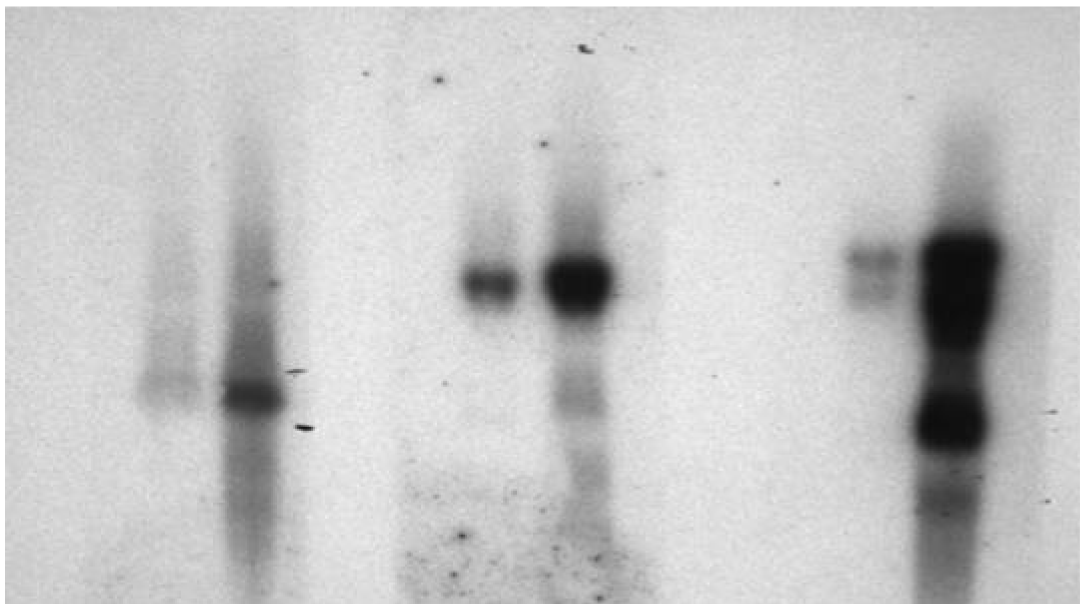


Figure 1: Northern Blot showing expression of genes by water-deficit-stress in aleppo pine.

Effect of Provenance on Gene Expression

Plants of the same species from different populations can differ in response to various environmental stresses. In this study, we utilized two aleppo pine populations.

These two populations are from locations with the greatest precipitation differences among the Israel aleppo pine sources available to us. The Beit Jann population is at an altitude of 850m and receives 925mm of annual precipitation. The Yirka population is at an elevation of 430m and receives 685mm. We calculated ΔCt values for each of the genes for each treatment within each population. These values were used to compare the differential expression of each gene in the two populations. For most genes, we see differences between the populations for control gene expression level, stress treatment expression level, and/ or level of induction. In most cases, the expression levels in nonstressed trees were higher in the Beit Jann population (ten genes). Two genes were expressed at higher levels in the Yirka trees (PhLEA, PhCYC). Following the high stress treatment,12 genes were more highly expressed in Beit Jann trees. The levels of induction (control expression minus high stress) were usually greatest in Beit Jann trees (eight genes). For two genes (Ph31, Ph33), Beit Jann trees had considerably higher levels of expression following all treatments (Table 2). The overall findings that expression of most genes is greater in Beit Jann trees agrees with physical observations we made while drought stressing the plants. We observed that on average, trees from the Yirka population wilted earlier than those from Beit Jann. However, the Yirka trees were usually larger.

Discussions

Several of the genes we characterized showed resemblance to previously identified water-deficit-stress-genes in angiosperms while some others are similar to unknown angiosperm genes. PhLEA was the most highly induced gene. This high

expression of a LEA was not unexpected as LEAs are commonly induced due to water-deficit-stress in angiosperms (Close et al., 1993), and there is universal distribution of this protein across different organisms ranging from bacteria to higher plants (Wise, 2003). PhLEA showed high similarity to putative gymnosperm LEAs but a low similarity to angiosperm LEA genes. Within the NR database, PhLEA is most similar to genes from white spruce that are expressed during somatic embryogenesis and are induced by ABA (Dong and Dunstan, 1997). Within angiosperms, it is most similar to genes labeled LEA or heat shock protein from solanaceous species but the e-values are quite low. The hydrophobicity plots of PhLEA, the spruce LEA and tobacco LEA proteins showed similar characteristics. Pfam classifies the PHLEA protein as a member of the LEAIII family (e-value 4.9e-36) (Bateman et al., 2004). Due to the high level of induction we observed following water-deficit-stress, the slight similarity to the angiosperm genes, the induction of the spruce genes by ABA and their expression during somatic embryogenesis, we concluded that the gymnosperm genes are LEAs but that the sequences have not been well conserved between angiosperms and gymnosperms.

None of the earlier studies of water-deficit-stress responsive genes in pines reported expression of LEA genes and in some cases, they reported the lack of induction (Costa et al., 1998). Some of the earlier studies focused on identification of water-deficit-stress-responsive genes from hydroponically grown maritime pine seedlings that were stressed using (PEG) polyethylene glycol (Dubos and Plomion, 2003; Dubos et al.., 2003). Chang et al. (1996(b)) did not find that LEAs were induced in water-deficit-stressed loblolly pines but their seedlings were not as severely stressed as the present

experiment. An earlier study reported instances were PEG-applied stress treatment failed to identify LEAs and related genes in other species (Leone et al., 1994(a)) and the same study indicated that there is a difference in the set of proteins which are expressed at various water-deficit-stress levels. In our study, pine seedlings were grown in fritted clay, which is a medium with similar drying characteristics as soil (Meier et al., 1992). We applied stress by withholding water, simulating natural water-deficit-stress conditions. RNAs were extracted from roots, which are one of the first organs to emanate water-deficit-stress signals in plants (Dubos and Plomion, 2003) and the library was made from highly stressed seedlings. All the above mentioned conditions helped us to emulate natural water-deficit- stress conditions to the maximum extent possible.

PhALDH shows similarity to aldehyde dehydrogenase genes from various angiosperms. Aldehydes are an ubiquitous class of highly reactive molecules involved in different physiological processes. Many of these are oxidized by aldehyde dehydrogenase. This group of enzymes has four functions: including detoxification, intermediary metabolism, osmotic protection and NADPH generation (Perozich et al., 1999). ABA induced ALDH genes have been isolated from *Arabidopsis thaliana* and *Craterostigma plantagin*eum in response to water-deficit-stress (Kirch et al., 2001). One suggested role of these enzymes during desiccation is protection of starch granules from oxidation by detoxifying aldehydes, which may interfere with sugar metabolism (Kirch et al., 2001). These results attain further importance based on the fact sugar metabolism is vital for dehydration tolerance (Ingram and Bartel, 1996).

PhCYC showed similarity to cyclophilins, which are conserved proteins across the plant and animal kingdoms. Their function as molecular chaperones during water-deficit-stress has been previously reported (Gething and Sambrook, 1992). Cyclophilins have been shown to interact with heat shock responsive proteins in fava bean (Luan et al., 1994(b)).

PhSUS is similar to angiosperm sucrose synthases. Sucrose synthase catalyses the conversion of sucrose uridine diphosphate into fructose and UDP-glucose. Although this reaction is reversible it is believed that this enzyme is mainly involved in breakdown of sucrose. Sucrose synthase accumulation is believed to satisfy the immediate glycolytic demand arising in plants due to water-deficit- stress (Dejardin et al., 1999). Sucrose synthase has been reported to be differentially expressed due to water-deficit-stress in a variety of species including the resurrection plant (Kleines et al., 1999). A sucrose synthase gene was shown to be more highly expressed in drought tolerant species of poplar compared to drought sensitive species confirming the role of this protein in the dehydration response (Pelah et al., 1997). Other observed roles of this enzyme are in cell wall biosynthesis (Delmer and Amor, 1995) and the respiratory pathway (Xu et al., 1989).

PhIP is extremely similar to angiosperm inorganic pyrophosphates. Over expression of inorganic pyrophosphate has been reported to increase the expression of sucrose synthase in potato tubers (Geingenberger et al., 1998). The role of this protein may be to act in conjunction with sucrose synthase.

PhMYB is similar to various MYB transcription factors in angiosperms. MYB proteins constitute a diverse family of plant transcription factors (Riechmann et al., 2000). MYB transcription factors are involved in a variety of functions including regulation of secondary metabolism, control of cellular morphogenesis, regulation of meristem formation, and regulation of gene expression upon dehydration (Jin and Martin 1999, Urao et al., 1993). A conserved DNA binding domain, the MYB domain, characterizes the MYB family. MYB factor binding proteins induced due to water deficit have been reported in a variety of plants including *Arabidopsis thaliana* (Urao et al., 1993). During dehydration MYB factor binding protein binds to the MYB recognization site in the promoter and activates the expression of dehydration resistance genes (Abe et al., 1997). Thus it seems that these proteins act as transcriptional activators in ABA signal transduction. Recent findings show that transgenic plants overexpressing a MYB factor protein have higher expression of various desiccation tolerance genes and thereby have higher drought resistance than wild type plants (Abe et al., 2003).

PhCTL is similar to various chitinase I genes from angiosperms. Chitinase genes are frequently expressed in plants due to pathogen attack. A chitinase was previously reported to be down-regulated in loblolly pines in response to water-deficit-stress (Chang et al., 1996 (b)), but that gene is not similar to the gene we identified. Chitinases are a diverse group of plant pathogenesis related genes which have the ability to hydrolyze chitin (Raikhel et.al., 1993). Differential expression of pathogenesis related genes has been reported in maritime pine (*Pinus pinaster*) in response to water-deficit-

stress (Dubos et al., 2003). The actual functional role of this protein during water-deficit-stress needs to be explored.

Ph4CL is a 4-coumarate CoA ligase (4CL). A 4CL catalyzes the activation of 4-coumaric acid, caffeic acid or ferulic acid and is considered a key enzyme in stress responses (Zhang and Chang, 1997). The role of this enzyme during stress is confirmed by higher expression during compression wood formation. Various lignin biosynthetic enzymes were previously reported to be differentially expressed in response to water-deficit-stress in a variety of species of plants including pines (Costa et al., 1998).

Results also include four genes (PhGRP, Ph20, Ph31 and Ph33) with no hits to angiosperm genes in the public databases. They are very similar to genes of unknown function from loblolly pine, maritime pine, and spruces. These genes may be specific to conifers or the sequences may have diverged so much that any similarities are not obvious. This is not unexpected due to the evolutionary distance between angiosperms and gymnosperms. PhLEA is an example of a gene that almost certainly encodes a LEA but has minimal similarity to the angiosperm genes.

Results also indicate observed differences in gene expression between the two populations of aleppo pine even though the two are quite close geographically. Earlier studies showed that the populations of aleppo pine in Israel are relicts of a single ancestor population which was split into various habitats (Schiller and Waisel, 1989). It was observed that in many cases seedlings from the Beit Jann population had higher expression than seedlings from Yirka. These results deviated from the originally thought that expression might be highest among Yirka trees since annual precipitation is lower.

However, our results agree with earlier seedling survival studies with these two populations indicating that Beit Jann trees have higher survival when subjected to water-deficit-stress (Schiller and Waisel, 1989). The Beit Jann population may be subjected to greater seasonal droughts and dehydration due to wind may be greater at the higher elevation.

It is believed that this is the first published attempt to examine the expression of water-stress-responsive genes between populations of pines. Earlier attempts to identify physiological parameters related to drought resistance that differ due to populations have been made in multiple species including maritime pine (Nguyen-Queyren and Bouchet-Lannat, 1989, Nguyen-Queyren and Bouchet-Lannat 2003) and loblolly pine (Van Buijtenen et al., 1976). Studies to examine gene expression differences between populations will provide insight into genes responsible for differences in drought tolerance and may be valuable to breeding programs or for genetic engineering. The development of real-time RT-PCR has facilitated such studies as it can be used to study more individuals than microarrays or northern blots, is more quantitative, and requires less RNA (Bustin, 2000).

CHAPTER III

DIFFERENTIAL EXPRESSION OF WATER-DEFICIT-GENES IN TWO POPULATIONS OF LOBLOLLY PINE (*Pinus taeda* L.) AND CHARACTERIZATION OF LEA MULTIGENE FAMILY MEMBERS

Introduction

Forest trees differ in their response to water deficit. The mechanism can be either drought avoidance or drought tolerance (Bray, 1997). Drought tolerance mechanisms involve a delayed initiation of water deficit in plant tissues while drought avoidance mechanisms cause plants to accumulate solutes so that the osmotic potential of the cell is kept high. Few of these responses are attributed to changes in gene expression. The genes, which are induced due to water deficit, can be classified based on their function as genes involved in avoidance, protection and repair (Bray, 1997). Avoidance function genes include transport proteins, proteins involved in ion channels and carriers. Protective function genes include LEA (Late Embryogenesis Abundant proteins). Repair function genes include the ubiquitins, chaperons and proteases.

Earlier experiments helped as to characterize genes induced by drought stress in aleppo pine (*Pinus halepensis,* Sathyan et al, unpublished) and were interested to see if the same genes were induced in loblolly pine. This project also generated an interest in the LEA multigene family. LEA proteins, which have a protective role, are among the most frequently induced proteins due water-deficit in angiosperms (Dure et al., 1989). LEA proteins are a large family of proteins expressed in plants in response to various stresses (Skriver and Mandy, 1990). LEAs were first identified from embryo tissues of

seeds during the desiccation phase (Dure et al., 1989) and are thereby named LEAs. During late stages of embryogenesis, when the seeds mature and begin to desiccate as part of the plant developmental cycle, LEA transcript levels increase. Due to preferential expression during this desiccation stage, it was predicted that these proteins have an adaptive role in plant stress. Several studies added more evidence on the adaptive role of these proteins during plant stress (Pelah et al., 1997). Overexpression of LEA genes in transgenic plants has lead to increased drought tolerance or water use efficiency compared to wild type plants (Sivamani et al., 2000). Based on the protein characteristics it was speculated that these LEAs protect proteins during stress by binding to them and protecting against dehydration stress (Dure et al., 1989) but the exact roles of these proteins are yet to be verified scientifically. With the increasing interest in genetically engineering crops against environmental stresses, many projects were conducted in the field of stress molecular biology. As result of these experiments, researchers have reported more 112 LEA-like genes from various species of plants (Wise, 2003). Other evidence for the universal role of LEA genes came from the reports of LEA gene expression in organisms other than plants including bacteria (Stacy and Aalen, 1998). Even though reports of expression of LEA genes in angiosperms are common, studies of gymnosperm LEA gene expression are limited. Close et al. (1993) used an antibody to a class II LEA consensus sequence to detect dehydrins in seeds of a large number of species including two gymnosperms: pinon pine and ginkgo. LEA genes induced during somatic embryogenesis were identified in Douglas-fir (*Pseudotsuga menziesii*) (Jarvis et al., 1997) and white spruce (*Picea glauca*) (Dong and Dunstan,

1997). A lack of reports of LEA expression in gymnosperms led to the speculation of an alternative pathway in response to water-deficit-stress (Dubos et al., 2003). However, in a study conducted in our laboratory with aleppo pine (*Pinus halepensis*) in which we reported preferential expression of LEA genes induced in roots following water-deficit-stress (Sathyan et al, unpublished). Additional studies are needed to confirm the roles of these proteins in gymnosperms.

The aim of this experiment was to identify various genes induced by drought stress in loblolly pine. Loblolly pine is one of the most important timber species in United States occupying an area of 13.4 million ha. Loblolly pine naturally occurs in the 15 southern and Mid-Atlantic States and has a wider geographic range than any other southern pine except shortleaf. Its range extends from latitude 39^0 21'N in Delaware and nearby coastal regions of New Jersey and eastern Maryland south to latitude 28 ^{0} N in central Florida and west to eastern Texas, the southwestern tip of Oklahoma, and southern Arkansas. Small outlying populations occur in southwestern North Carolina, east central Arkansas, and northeastern and southeastern Texas. The later population in Texas is called the "Lost Pines" and is at the western boundary of the species range (Shultz, 1997). The Lost Pines are seen in the counties of Bastrop, Fayette and Caldwell counties of central Texas. This disjoint population is 100 miles away from the contiguous range of loblolly pines. The population receives 10 to 20 inches less rainfall annually and four-six inches less in July-August, which happen to be the driest months, than trees in the pine belt of Texas (Bilan, 1961). Several studies were undertaken to understand the superior adaptations of this population to water stress. Based on these

studies, conclusions were made that these populations adopt various drought avoidance mechanisms to evade the harsh climatic conditions existing in the locality (Van Buijtenen et al., 1976) but nothing is known about the molecular mechanisms by which these populations respond to water deficit.

To better understand the roles of various LEA gene family members and other drought responsive genes in two different populations of loblolly pine, real-time RT-PCR was conducted using gene-specific primers. Seedlings from Lost Pines and south Louisiana populations were analyzed. The different genes we studied showed preferential expression due to treatments and population. Identification of drought-induced genes in different populations of loblolly pine may assist with breeding programs and will contribute to our understanding of the molecular basis of drought resistance.

Materials and Methods

Seedlings from two populations, The Lost Pines (LP, Bastrop County, TX) and south Louisiana (SLA, Livingston and Washington Parishes) were included in the study. The Lost Pines are in an area with the lowest level of precipitation within the loblolly range (865 – 965 mm) while the Louisiana population is from the wettest part of the range (1500 – 1700 mm). Three families were planted per population. The Bastrop seeds came from open pollinated families while the Louisiana population was pollinated with an Atlantic coastal polymix. The seeds were collected and stored in a refrigerator. Seeds were soaked in 20% bleach for 20 minutes, rinsed in water twice, and soaked in 1% H_2O_2 for 24 hours before a 45 day cold stratification at 4° C. The seeds were germinated

in plastic trays containing a mixture of perlite, sand and peat. Seedlings were transplanted to pots containing fretted clay and were transferred to a glasshouse. The seedlings were irrigated every two to three days using purified water, with a weekly application of complete nutrient solution to encourage rapid, balanced growth until they reached eight months of age. Drought stress was imposed in the following fashion: At air temperatures ranging between 15° C minimum night and 38° C maximum day, well-watered plants received enough purified water to reach full saturation every two days. The water was withheld from increasing numbers of individuals for three weeks. Thus, at the termination of the experiment, some seedlings had received no irrigation for 16 days, others for 14 days, others for 12 days, et cetera, producing a wide range of drought stresses. Predawn water potentials using seedling side branches four to eight cm long were determined with a Scholander pressure chamber (Scholander et al., 1965). Roots were rinsed with water to remove the potting medium. Seedlings were quickly separated into roots, stems, and needles, then frozen in liquid nitrogen and stored at -80° C to await RNA isolation. Based on the water potentials, plants were divided into three classes control (0.3-0.7 MPa), moderate (1.0-1.5 MPa) and highly stressed (2.0 - 2.9 MPa).

Extraction of RNA and Preparation of Samples for Real-time RT-PCR

Extraction of good quality RNA from pine seedlings can be difficult due to the presence of phenolics. We used the method developed by Chang et al. (1993) for extraction of RNA.Modifications were made in the procedure by performing the day one precipitation at -20^{0}C rather than at 4^{0}C. In order to remove trace amounts of DNA from

the samples we used DNA-free (Ambion Inc, Austin) according to the user's manual. The RNA was later quantified using a spectrophotometer and samples having 260/280 values greater than 1.8 were used for the experiment. To check for integrity, 1µg of each sample was run on a 1% agarose gel stained with ethidium bromide.

Identification of Drought Responsive Gene Sequences

Genes were chosen for characterization based on two previous projects in our laboratory.The loblolly pine EST databases were used to design primers for real-time RT-PCR. Earlier research helped to characterize genes induced by drought stress in aleppo pine (*Pinus halepensis*, Sathyan et al. unpublished) and in the present part of the project the interest was to see if the same genes were induced in loblolly pine. The aleppo pine project also generated an interest in the LEA genes. A microarray project to identify genes differentially expressed in earlywood and latewood in two populations of loblolly pine identified additional candidate genes for this project (Yang and Loopstra, unpublished). To identify the loblolly pine sequences, aleppo pine sequences were used to perform blast searches at the Genomics of Wood Formation in Loblolly Pine (http://pinetree.ccgb.umn.edu) and Transcriptome Responses to Environmental Conditions in Loblolly Pine (http://fungen.org/Projects/Pine/Pine.htm) databases. To identify other LEA gene family members, blast search with conserved sequences from the Pfam database were searched against the above mentioned loblolly pine EST databases.

Real- time RT-PCR

The RNAs were reverse transcribed using a TAQMAN Reverse Transcriptase Kit (Applied Biosystems, ABI), Foster City CA) using the following conditions: 25^0 C for 10 minutes, 48^0 C for 30 minutes, and 95^0 C for 5 minutes. Real-time RT-PCR primers were designed using Primer Express (version 2.0, ABI) and custom primers ordered (Invitrogen, Carlsbad, CA). SYBR green was used to quantify the RT-PCR products as follows. Five microliters of SYBR green master mix (ABI), 3.5µl of depc H_20, 0.5µl (10µg/µl) of template, and 0.5µl (1µg) of each primer, were dispensed into 384-well plates (ABI). An ABI 7900HT Sequence Detection System was used according to the manufacturers recommendations. In our experiment we used Ct (Cycle threshold) values to compare the differential expression between various treatments. With this method, the Ct value of the normalizer is deducted from the Ct value of the sample to obtain the ΔCt. We used the pine ribosomal 18s gene as the normalizer.

$$\Delta Ct = Ct_{(Sample)} - Ct_{(Normalizer)}$$

ΔCt values were calculated for each of the replicates and the averages were calculated for the comparisons between treatments. Thus a gene with a low ΔCt value signifies higher expression and a high ΔCt signifies low expression. We conducted ANOVA (Nested factorial) to verify the effect due to treatments and population.

Table 3: Loblolly pine genes induced due to drought treatment

Clone	Contig/Accession number	Putative Function
PtPGRP	CF470651	Glycine-Proline rich
PtHS	Contig 7670	Heat Shock protein
PtLEA (c)	Contig 7780	LEA Protein
PtLEA (a)	Contig 6215	LEA Protein
PtLEA (b)	Contig 7031	LEA Protein
PtLEAII	CF471443	LEA Protein
PtCYC	CF472800	Cyclophilin
PtSUS	Contig 7744	Sucrose Synthase
PtIP	CF392675	Inorganic Pyrophosphates
PtCTL	CF395720	Chitinase
Pt31	RTFEPL17H05.b1A029	Unknown Protein
PtMYB	HEAT11A04.g1A029	Myb Factor Binding Protein
PtGTP	Contig 6650	GTP Binding Protein
PtALDH	CF671734	Aldehyde Dehydrogenase

Results

Transcript levels of 13 out of the 14 genes analyzed showed preferential expression due to water-deficit-stress. Ten out of the 14 genes characterized in this experiment have sequence similarity to previously identified drought-responsive genes in angiosperms. The putative function of each of these genes are shown in Table 3. In Table 3, the Contig number indicates whether the sequence came from a contig in the Genomics of Wood Formation in Loblolly Pine (http://pinetree.ccgb.umn.edu). Accession number indicates that the sequences came from EST sequencing project Transcriptome Responses to Environmental Conditions in Loblolly Pine (http://fungen.org/Projects/Pine/Pine.htm) databases for those sequences which are not yet submitted databases we have given the clone id in the site.

We identified three members of the LEA III gene family. Each of these showed similarity to previously identified white spruce LEAs (*Picea glauca*) and a lower level of similarity to angiosperm LEAs. PtLEA III (c) (http://pinetree.ccgb.umn.edu/) showed greatest similarity to the spruce gene EMB28 (e-value 7e-27). These two proteins are identical at 65/83 amino acids. PtLEA III (b) has greatest similarity to spruce gene EMB32 (e-value 4e-26). These two proteins are identical at 94/142 amino acids. PtLEA III (a) is similar to the spruce gene EMB11 (8e-28). These two proteins are identical at 45/50 amino acids. The most similar angiosperm LEA for all three pine LEAs is a group 5 LEA from tobacco (*Nicotiana tabacum*) but the levels of similarity were quite low with highest e-value of 0.026.

```
                     *        20         *        40         *        60
LEAIII(a) : MAGRLLSAYRLSSLLTDIKISHARQYTAAASDAMRSSGAA-----DAMRSSAGGGNDKRR : 55
LEAIII(b) : MAKRLLSVHRLSSLLSDIKNFHSRQYTAAAAEAMRSSGVVAREFPEGTKAGGGGGNTKRT : 60
LEAIII(c) : MARRLLSAHALSSLLSDIRIFHARQYTAAAAEAMRSSGAAGREFPEVSKAGRGGGNNKG- : 59
            MA RLLS   LSSLL3DI4  H RQYTAAA  AMRSSG          4   GGGN K
```

Figure 2: Comparison of LEA III protein from *P.taeda* at the amino acid level. (The bold capital letters indicate identical amino acids and the grey shading indicates the similar amino acids.)

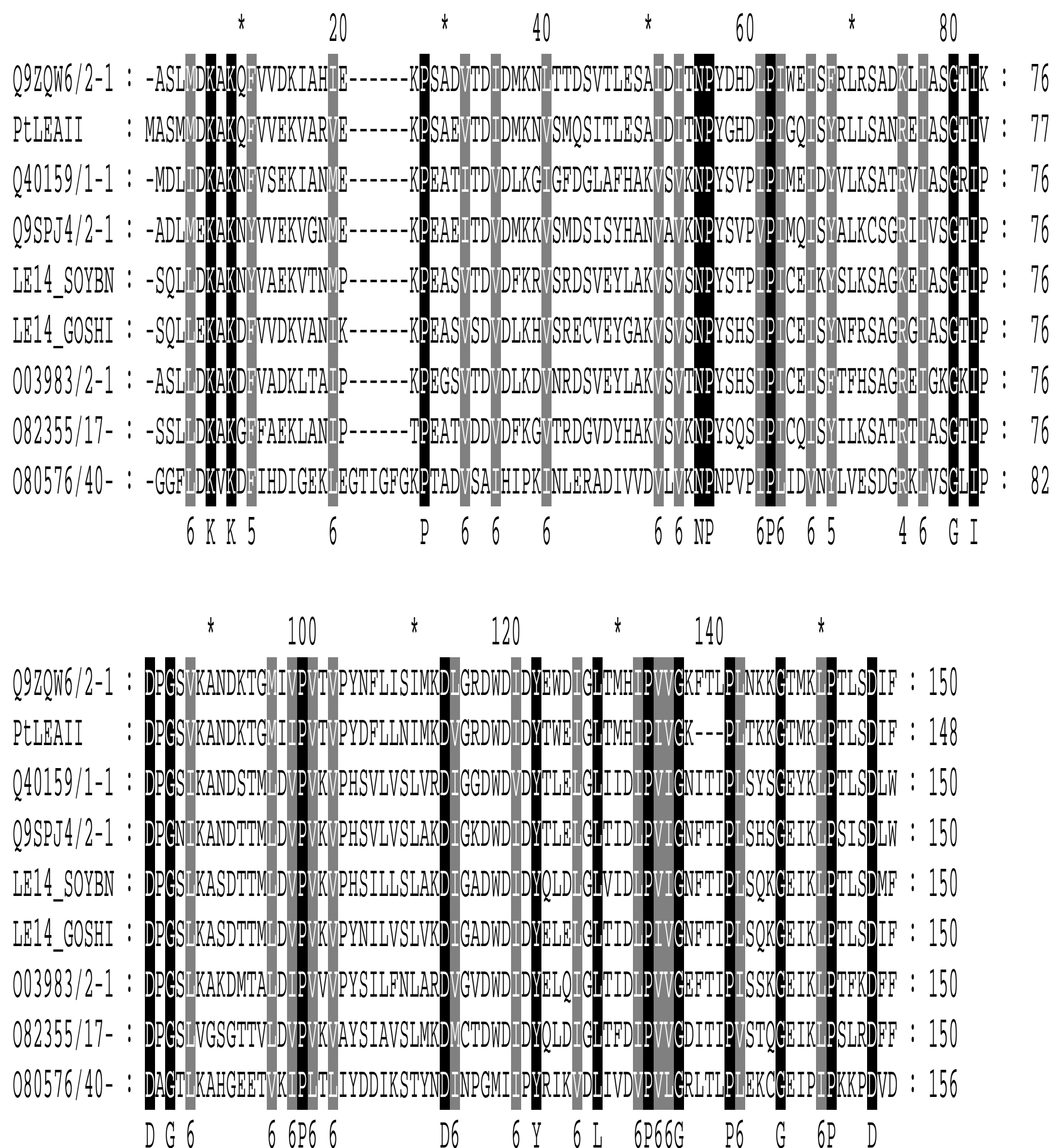


Figure 3: Comparison of PtLEAII protein with pfam members. (Identical amino acids are shaded black and similar amino acids are shaded grey)

PtLEAII is most similar to a Douglas-fir (*Pseudotsuga menziesii)* LEA gene (e-value 1e-81). These two proteins are identical at 126/163 amino acids. Unlike the LEA III members, this gene had high similarity to previously identified angiosperm LEAs. Besides the Douglas-fir gene, all other genes in the NR data base with a similarity greater than an e-value of 6e-30 are angiosperm genes. The most similar is LEA -14 from rice (*Oryza sativa* e-value 4e-55). These two are identical at 75 /147 amino acids. Other angiosperm genes which have high similarity include LEAs from *Gossypium hirsutum* (e-value 3e-54) and *Arabidopsis thaliana* (e-value 2e-50). The amount of similarity between angiosperm and gymnosperm LEA II members is evident in Fig 3. The amount of similarity between different members of LEA III family in loblolly pines are evident from Fig 2.

PtCYC is similar to various gymnosperm and angiosperm cyclosporin genes. The amount of conservation across the genera was quite high for this gene. PtCYC is very similar to the spruce cyclosporin-A where the identity is 121/152 amino acids. The angiosperm protein from *Ricinus communis* is identical at 91/116 amino acids.

PtSUS is similar to various sucrose synthase genes from angiosperms. The most similar is from *Pisum sativum* which is identical at 190/238 amino acids.

PtCTLis similar to endochitinase genes from various angiosperms. The highest similarity was with an *Oryza sativa* endochitinase with an identity of 91/142 amino acids.

PtHS is similar to a Douglas-fir (*Pseudotsuga menziesii)* heat shock protein. The two gymnosperm proteins are very similar with an identity of 103/129 amino acids. The angiosperm protein with the highest similarity is a carrot heat shock protein with an identity of 45/58 amino acids.

PtALDH is similar to *Lotus corniculatus* and Arabidopsis aldehyde dehydrogenase genes. It is identical at 94 out of 116 amino acids with the lotus gene. It also has considerable identity with mammalian aldehyde dehydrogenase genes with the most similar being an aldehyde dehydrogenase from rat (*Rattus norvegicus*, e-value 6e-37).

PtGTP is similar to a GTP binding protein from sunflower (*Helianthus annuus)* with an identity of 81/86 amino acids. Other angiosperm GTP binding proteins are also quite similar.

PtGRP is not particularly similar to previously identified proteins in the NR database. There are many copies of the gene in the EST database from various root and xylem libraries from loblolly pine and maritime pine. Other proline-glycine rich genes from various species are somewhat similar but are lacking the histidine residues found in PtGRP. The structure of PtGRP is interesting as it mostly consists of proline (52.6%), glycine (31.6%) and histidine (12.3%) residues and the following amino acid sequence M H P G G P P P P G G P P P P G G P P P P G D H P H G H H H P P P P G P G G P P P P P G P G G P P P H C G P G G P G G R C C.

PtIP is similar to inorganic pyrophosphatases from many diverse organisms including plants, animals, and yeast. The most similar one is from *Arabidopsis thaliana* with an identity of 136/157 amino acids.

Pt38 is similar to a spruce (*Picea abies*) hypothetical protein with a similarity of 92/125 amino acids. Other proteins which showed similarity included a hypothetical protein from *Saccharum officinarum* with an identity of 29/56 amino acids. Pt31 is similar to pine ESTs including some from a stressed root library but there were no hits in the NR database leading to possibility that this may be a gymnosperm specific protein.

Differential Expression Due Water-Deficit-Stress

To test the statistical significance due to the treatments, we calculated the ΔCt values for each treatment and gene. ANOVA was conducted on normalized transcript levels. Thirteen out of 14 genes had statistically significant differences ($p<0.01$) in expression between treatments. The one which did not show statistically significant difference between the treatments was PtLEA III (a). The ΔCt values of the different genes for the various treatments are given in table 4. The differences in ΔCt values between the control and high stress treatments varied from 0.6 (PtLEA III (c) in SLA) to 6.2 (PtLEAIII(c) in LP) cycles.

Table 4: Normalized ΔCt (Cycle Threshold) values following stress treatments of loblolly pine

Clone Name	Lost Pines				South Louisiana			
	ΔCT control	ΔCT moderate	ΔCT high	ΔCTc- ΔCT h	ΔCT control	ΔCT moderate	ΔCT high	ΔCTc- ΔCTh
PtALDH	24.08	22.4	18.4	5.6	24.3	22.4	20.5	3.7
PtPGRP	No exp	23.5	20.5		22.5	20.7	19.6	2.8
PtLEA III(c)	15	13.7	8.8	6.2	12.1	11.3	11.5	0.6
PtLEAIII (b)	17.7	16.2	12.2	5.5	17.5	15.2	14.8	2.7
Pt LEAII	13.9	12.3	10.5	3.4	15.7	14.5	12.7	3.0
PtCY	17.0	15.2	14.1	2.8	16.2	14.8	14.9	1.4
PtSUS	24.6	24.1	19.9	4.7	24.1	22.7	20.0	4.1
PtIP	22.1	20.1	18.1	4.0	22.9	21.7	21.1	1.8
PtCTL	13.3	10.8	9.4	3.9	15.0	11.7	9.3	5.7
Pt31	19.0	17.3	14.3	4.7	18.3	16.7	15.5	2.8
Pt HS	12.5	11.2	9.4	3.1	11.8	11.6	9.8	2.0
PtGTP	12.4	11.2	8.9	3.5	14.1	12.9	11.5	2.6
PtMYB	16.6	13.8	12.1	4.5	17.8	17.1	14.4	3.4

If PCR is 100 percent efficient, a six cycle difference is equivalent to an approximately 64 fold difference in expression. Other genes which were highly induced following water stress include Pt LEA III (b) (5.5 cycles, LP), PtSUS (4.7 cycles, LP), and Pt33 (4.7cycles, LP).

The pattern of induction varied with the different genes. For some genes, the moderate stress treatment resulted in expression levels almost equal to those caused by high stress (i.e. PtCYC in SLA) while for others the expression was intermediate between the control and high stress treatments (i.e. PtIP in LP). In a few cases, moderate stress did not cause much induction over the control treatment (i.e. Pt HS in SLA).

Effect of Provenances on Gene Expression

Differences in expression levels or amount of induction in populations adapted to different environmental conditions may provide insight into the importance of these genes for stress resistance. In this study, the two populations we used came from the extremes of climatic regimes in the loblolly pine natural range. The south Louisiana population receives high rainfall (1500 - 1700 mm) while The Lost Pines (865 – 965 mm) are a disjoint population with superior drought tolerance (Van Buijtenen et al., 1976). Previous research demonstrated that this population is highly adapted for drought stress and plants make alterations in physiological parameters to cope up with the harsh climatic conditions on the site (Bilan, 1961).

A

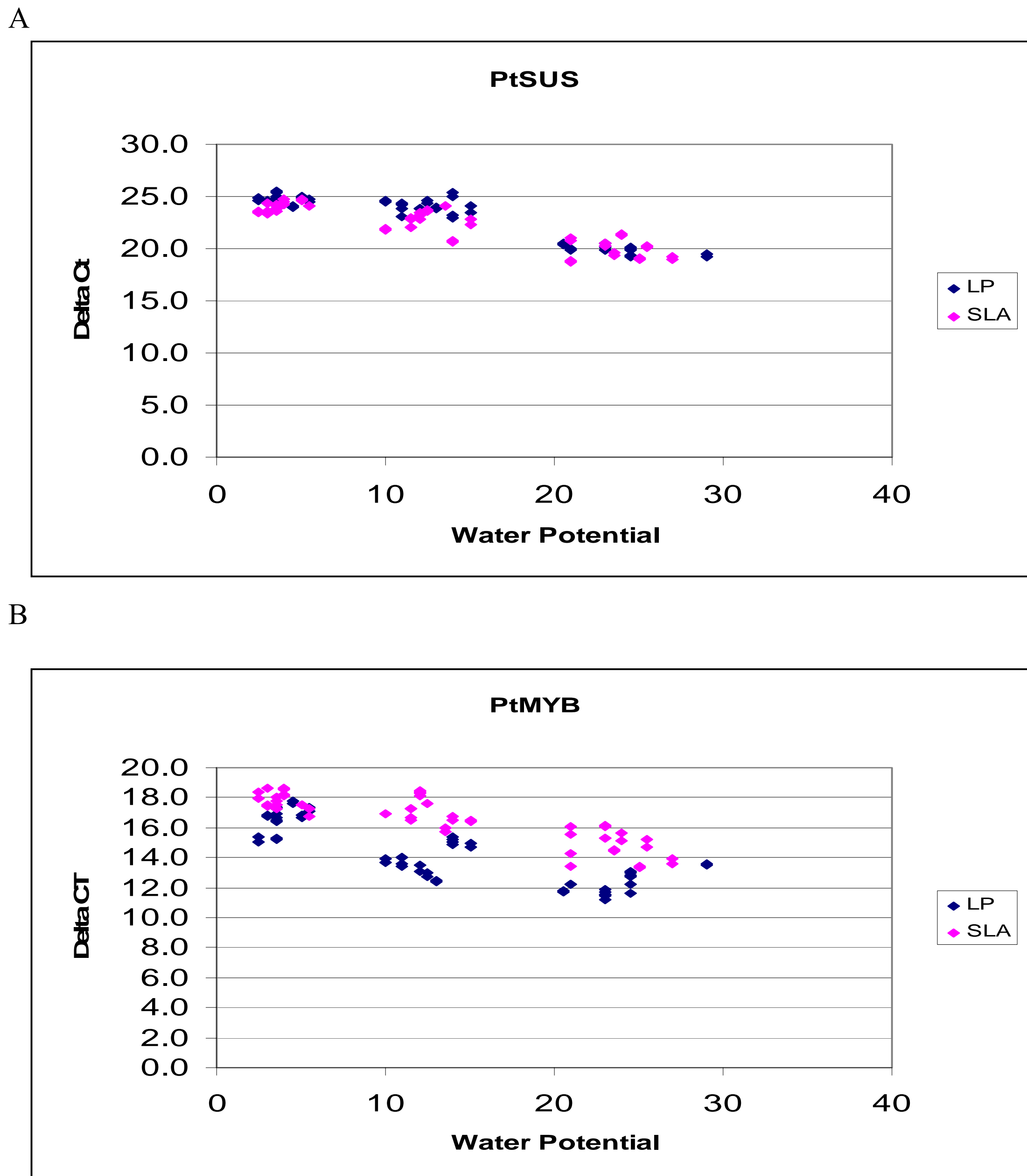
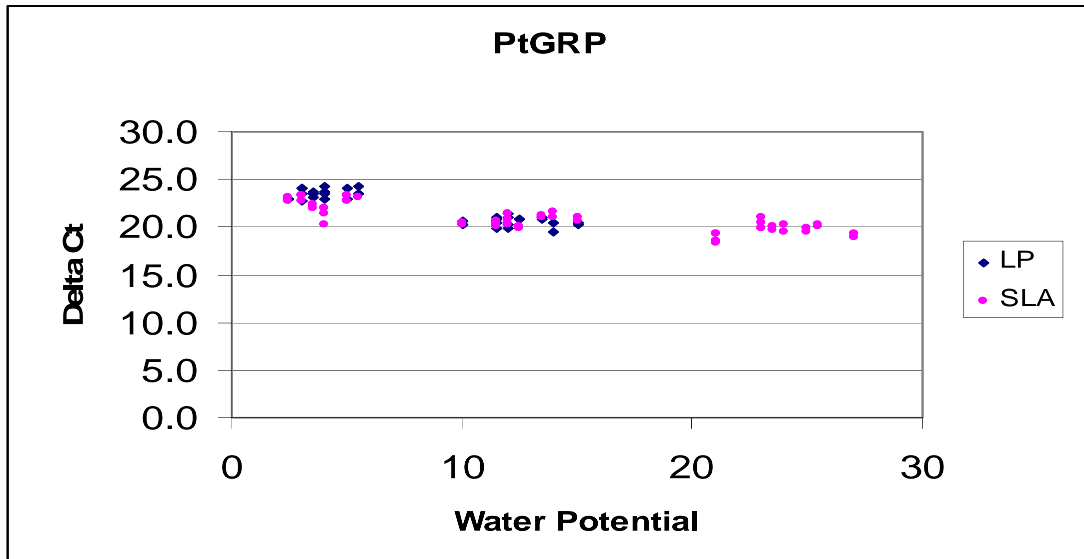


Figure 4. Scatter Plots showing the induction of the genes due to water-deficit-stress treatment in Lost Pines (LP) and south Louisiana (SLA)

C



D

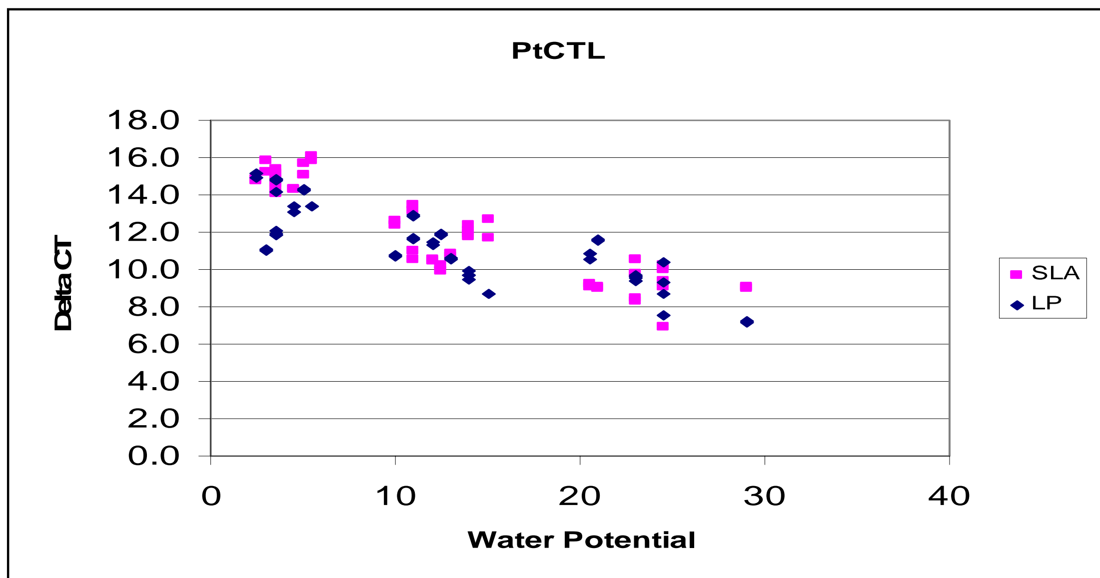
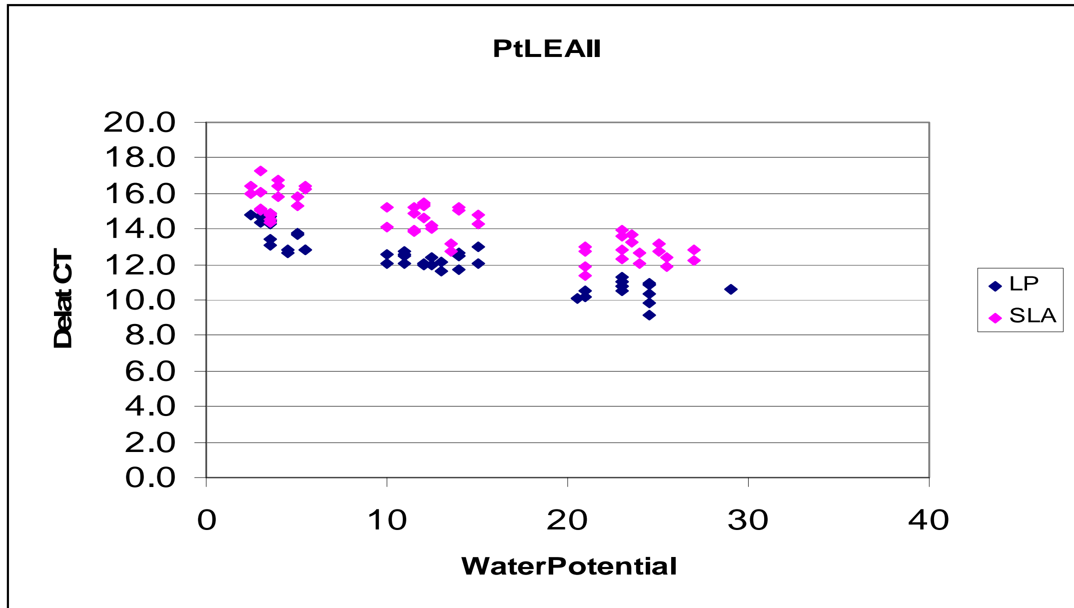


Figure 4 continued

E



F

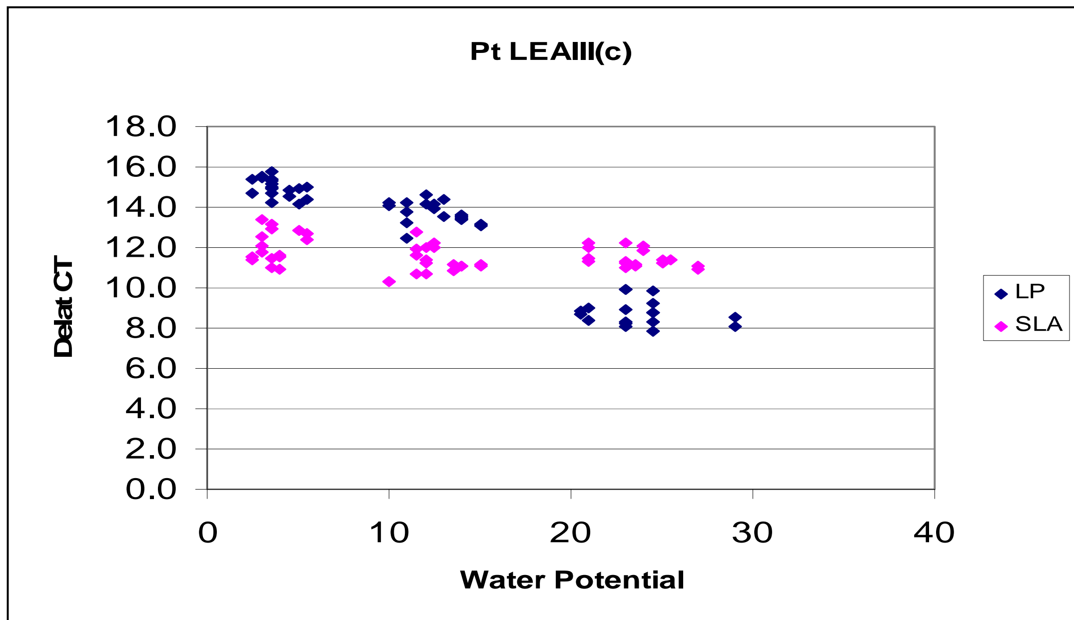
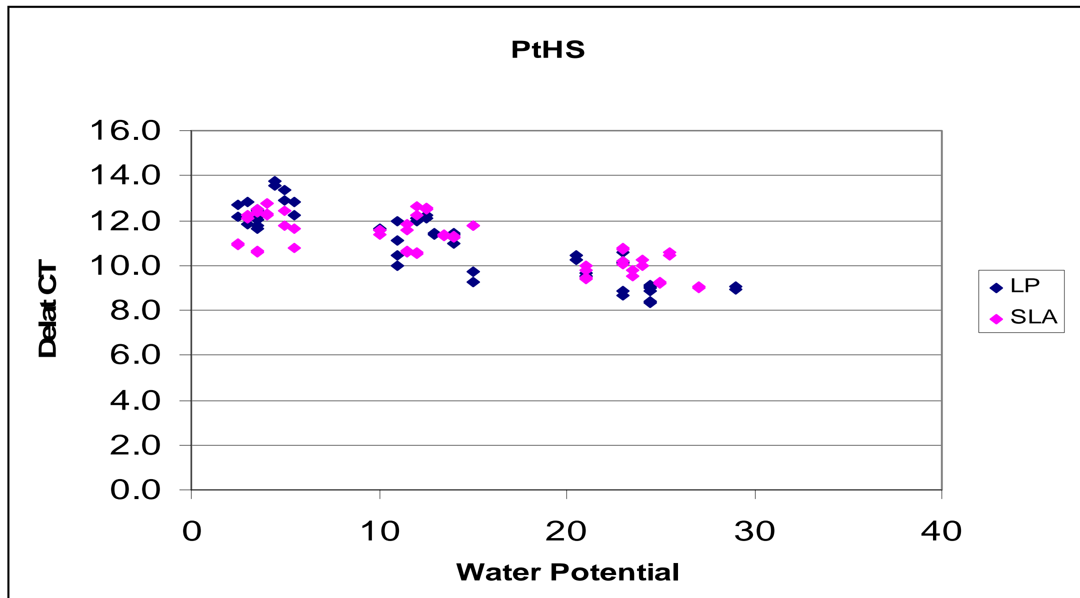


Figure 4 continued

G



H

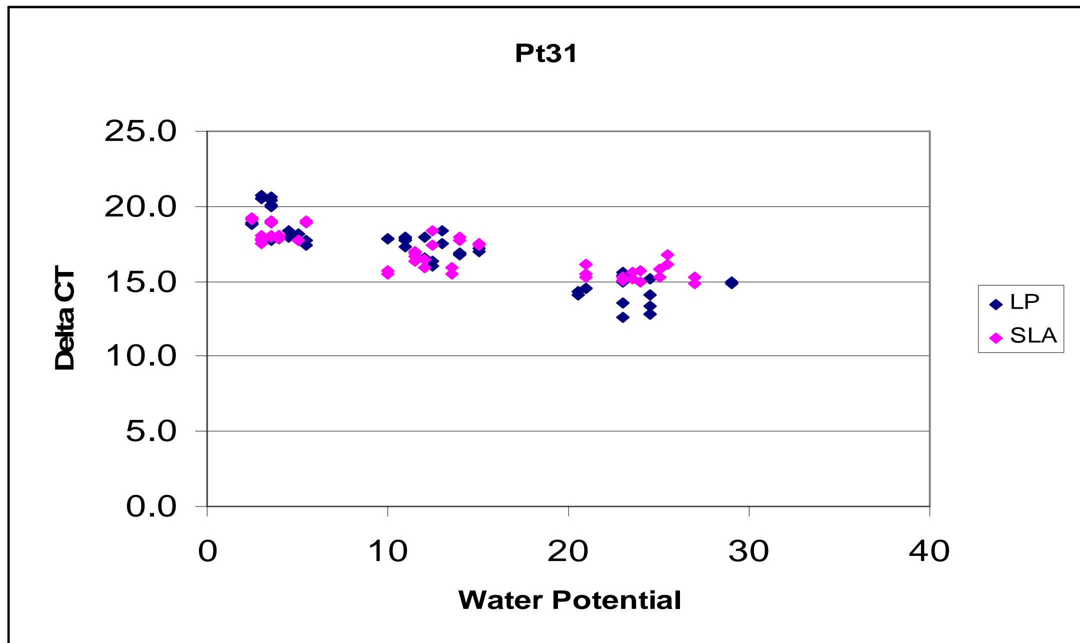
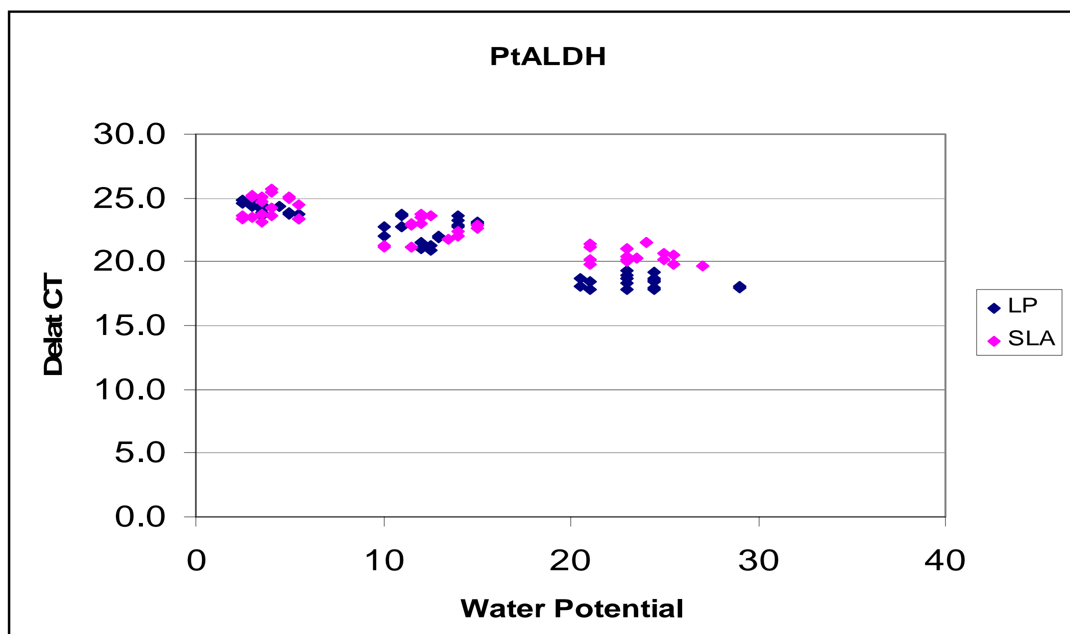


Figure 4 continued

I



J

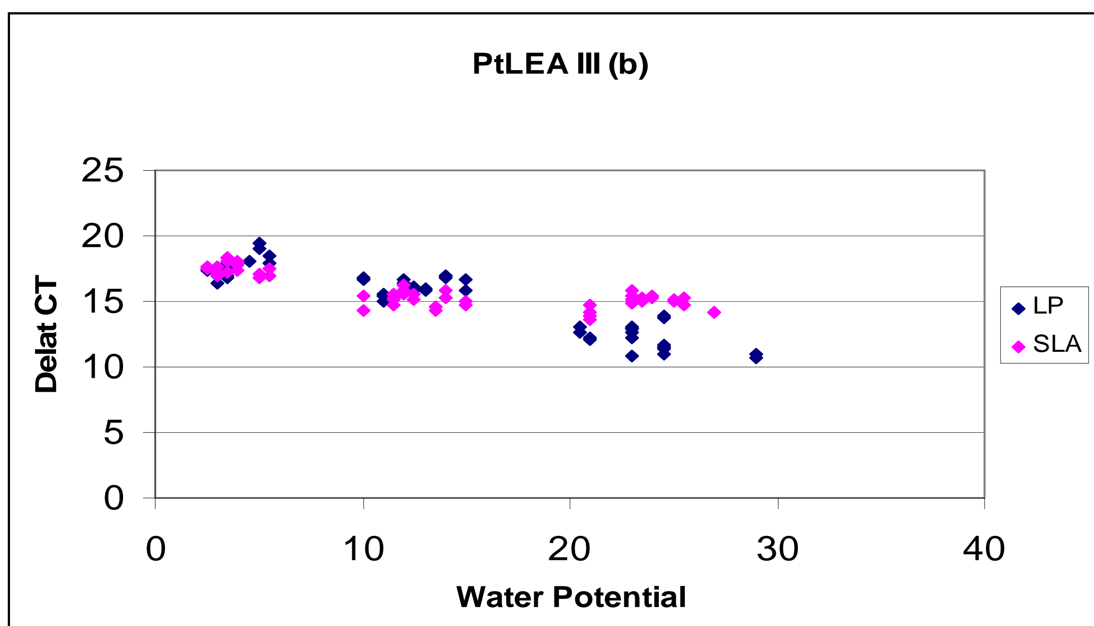
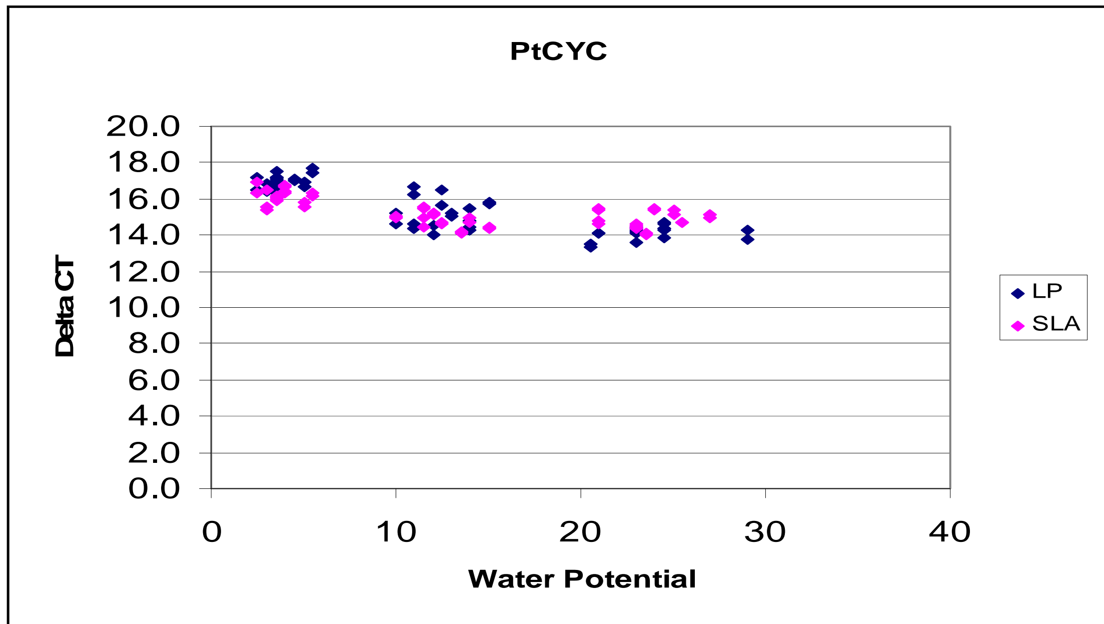


Figure 4 continued

K



L

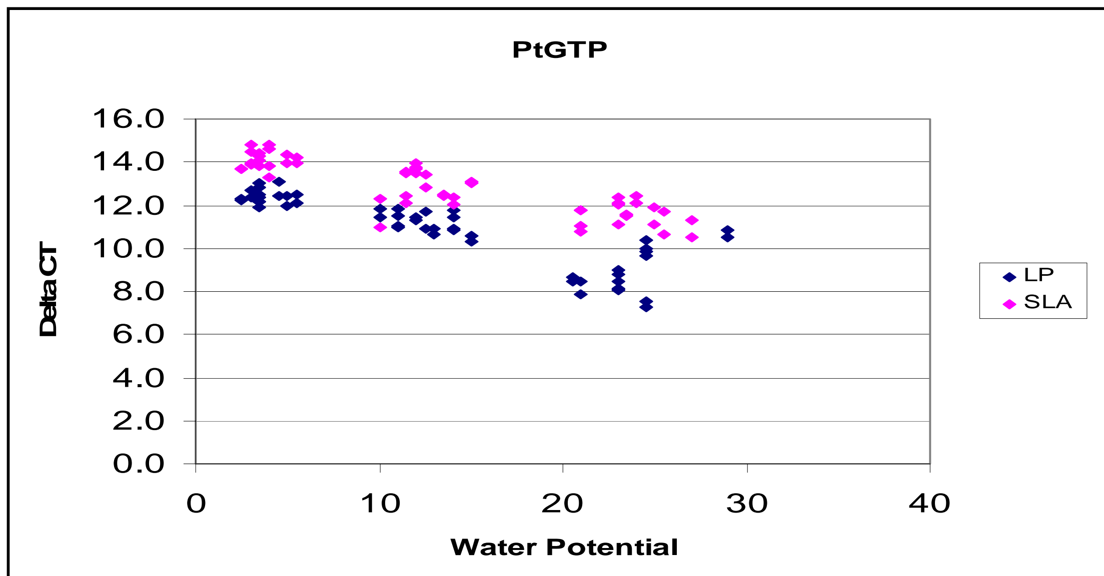


Figure 4 continued

M

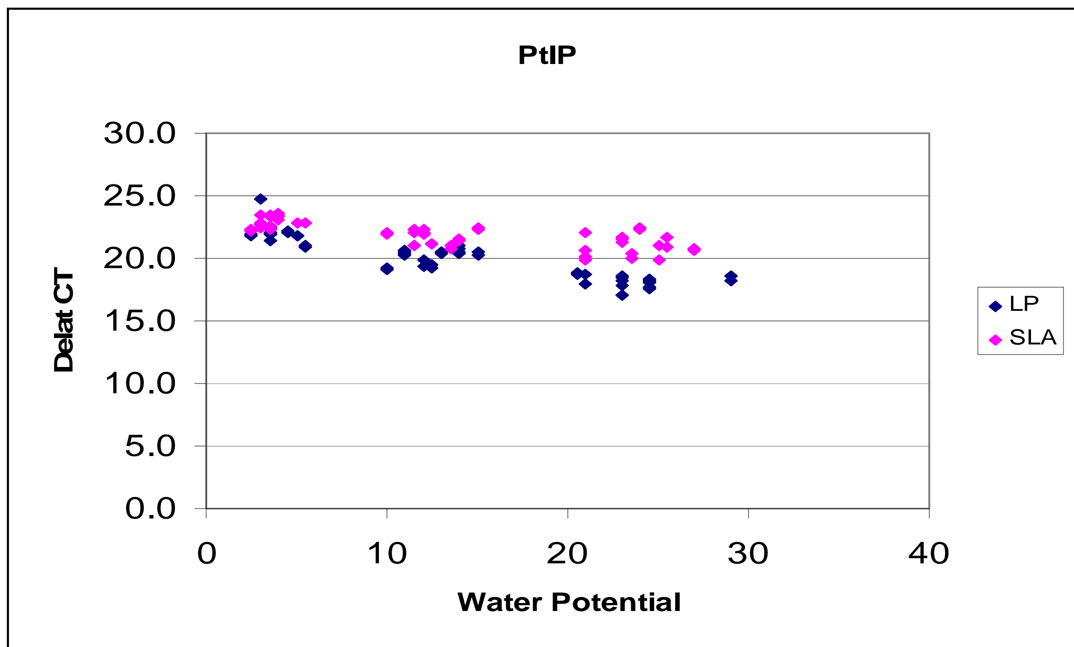


Figure 4 continued

To identify changes at the transcript level, we compared normalized ΔCt values for each gene within the two populations. We used ANOVA to statistically test the significance of the variation between the two populations using a conservative p value of 0.01. Within well-watered plants, we did not usually see large differences in expression between the two populations. For seven genes, there was less than a one cycle difference. Two genes were expressed at higher levels in SLA trees and four genes were expressed at higher levels in LP trees (Table 4). The expression patterns of various genes from different populations are compared using a scatter plots fig 4. The PtLEA III(c)

SLA ΔCt was 2.9 cycles less than that for LP indicating expression in SLA trees is approximately eight times higher. The greatest difference was for PtPGRP, which was not expressed in well-watered LP trees at a level high enough for detection following 40 PCR cycles. Nine plants (three individuals from three families) were each tested at least twice and expression was never detected (a ΔCt greater than 40 cycles). Following the high stress, five genes were expressed at similar levels in trees from both populations (less than a one cycle difference) and nine genes were expressed at higher levels in LP trees. The majority of the genes (nine) were induced more in LP than SLA trees (ΔCt control - ΔCt high). The highest difference in the level of induction was PtLEA III (c). Within the Lost Pines, we observed a 6.2 cycle difference between highly stressed and control plants. Within the south Louisiana trees, we observed a 0.6 cycle difference. If this difference is converted to numerical terms making the assumption that the PCR was 100 percent efficient, we estimate that LP trees are induced 64 fold while SLA trees are induced only marginally (less than two fold). The next greatest difference in induction was for PtLEAIII (b). We observed a 5.5 cycle difference between highly stressed and control LP trees and a 2.7 cycle difference between SLA trees. It is interesting to note that both the above mentioned genes are members of the LEA III gene family. The only gene which showed greater induction in south Louisiana was PtCTL with a 5.7 cycle difference vs. 3.9 cycles in the Lost Pines. The overall results indicate greater induction of genes within Lost Pines individuals compared to those from the south Louisiana population, coinciding with the fact that The Lost Pines population is more drought tolerant than south Louisiana population.

Discussions

An earlier experiment identified and characterized drought response genes in aleppo pine. In this experiment we studied the expression of some of the same genes using two populations of loblolly pine. In addition, this experiment also characterized two genes identified in an earlier microarray experiment to examine expression in earlywood and latewood from two loblolly pine populations and four LEA genes. Most of the genes studied showed similarity to previously identified water-responsive genes in angiosperms although one gene appears to be conifer specific.

Several members of the LEA multigene family were characterized in loblolly pine. LEAs are the most frequently induced proteins in response to water deficit stress (Close et al. 1989, Ingram and Bartel, 1996). Although LEA genes are commonly reported in angiosperms in response to water-deficit-stress (Close et al. 1989) they have rarely been reported in gymnosperms. This is the first report of expression of multiple gymnosperms LEA gene family members induced in response to water deficit. In this experiment three members of the LEAIII gene family and one member of the LEAII family based on Pfam (Bateman et al. 2004) classifications. Multigene families which are common in eukaryotes, were reported by Kinlaw and Neale, (1997) in pines. To differentiate how transcript levels vary within the members of the gene family, real-time RT-PCR was conducted using gene specific primers. To design the primers, the DNA and protein sequences of each of the members of the gene family were analyzed with care taken to separate allelic differences from multigene members. Allelic variation in pines is very high (Whetten et al., 2001). Two out of the three LEA III genes showed

increased expression in response to water deficit. Both genes were more highly induced in Lost Pines seedlings than in south Louisiana seedlings. One LEA III gene was not induced following water-deficit stress. Increased expression of LEA gene members in response to water-deficit was previously reported in rice (Shinozaki et al., 1989). Preferential expression of multigene family members can be attributed as one way to adapt to the environment (McAdam and Arkin 1999; Meagher et al., 1999). We also identified one member of the LEA II family that also seems to be induced in stressed roots. Unlike the LEA III gene family, this gene is very similar to both angiosperm and gymnosperm LEA II genes.

The LEA proteins are highly complex and universal with reports of expression in plants, fungi, nematodes and bacteria (Wise, 2003). Due to its complex nature and the large number of proteins identified to date, these proteins still need to be classified functionally. Its presence following stress and the nature of the proteins suggests various probable functions. These proteins have a complex structure with hydrophilic and hydrophobic regions with coiled motifs (Ingram and Bartel, 1996). Coiled motifs are used by these proteins to bind to water thereby reducing its loss during desiccation. Another conceivable function of these proteins is to act in solvation of cytosolic structures (Baker and Langdon, 1990). The structural and functional complexity of these proteins is evident from the fact that this protein has been divided into different families and sub families based on similarities and differences between the various members (Wise, 2003). Modern classification of this protein is based on small structurally conserved amino acid sequences (Wise, 2003).

Earlier studies with gymnosperms in response to water deficit were either conducted with PEG (polyethylene glycol) stressed seedlings (Dubos and Plomion, 2003; Dubos et al., 2003) or at a lower level of water stress than in this study (Chang et al., 1996(b)). Leone et al. (1994(a)) reported that a PEG applied stress treatment failed to identify LEAs and related genes and the same study indicated that there is difference in the set of proteins which are expressed at various water stress levels. In our experiments we emulated natural conditions to the maximum extent possible. Seedlings were grown in medium having similar characteristics as soil and stress was applied by withholding water. RNA was isolated from roots which are shown to be one of the first organs to radiate water stress signals (Dubos and Plomion, 2003).

The characterized genes belonged to various classes including sugar metabolism (PtSUS, PtIP, PtALDH), chaperons and heat shock proteins (PtCYC, PtHS), LEA proteins (three PtLEA IIIs and PtLEA II), secondary pathway (PtCTL), transcription factors (PtMYB), GTP binding proteins (PtGTP), and genes without any identified function (Pt 31). Most of the genes were similar to previously-identified water-deficit genes in angiosperms. Sugar metabolism proteins are one of the largest groups of proteins that have been reported to be differentially expressed in response to drought stress (Ingram and Bartel, 1996). Sucrose synthase is involved conversion of sucrose uridine diphosphate into fructose and UDP-glucose and this is believed to satisfy the immediate glycolytic demand arising in plants due to water stress (Dejardin et al., 1999). Other observed roles of this enzyme are in cell wall biosynthesis (Delmer and Amor, 1995) and the respiratory pathway (Xu et al., 1989). Over expression of inorganic

pyrophosphate has been reported to increase the expression of sucrose synthase in potato tubers (Geigenberger et al., 1998). The role of this protein may be to act in conjunction with sucrose syntheses. Aldehydes are a ubiquitous class of highly reactive molecules involved in different physiological processes. Many of these are oxidized by aldehyde dehydrogenase. This group of enzymes has four function, detoxification, intermediate metabolism, osmotic protection and NADPH generation (Perozich et al., 1999). One suggested role of these enzymes during desiccation is protection of starch granules from oxidation by detoxifying aldehydes, which may interfere with sugar metabolism (Kirch et al., 2001). Accumulation of soluble sugars in response to desiccation seems to be a common mechanism adopted by a variety of organisms including plants, animals and microorganisms. Accumulated sugars serve as compatible solutes permitting osmotic adjustment. This mechanism has been seen to work in drought sensitive and drought tolerant species. Trehalose has been found to be the most effective sugar but the amount of this sugar in plants is very low so sucrose seems to be a viable alternative in plants (Leprince et al., 1993). One plausible mechanism by which sugars work in desiccation tolerance is by glass formation that prevents the breakdown of cellular machinery in the event of desiccation (Koster, 1991).

The cyclophilins and chaperons identified were PtCYC and PtHS. Cyclophilins act as molecular chaperones by accelerating the rate limiting steps in protein folding (Gething and Sambrook, 1992). Chaperons are molecules that help the proper folding of disturbed protein structures during environmental disturbances. Heat shock responsive cyclophilins have been previously reported in favabean (Luan et al., 1994(b)). Heat

shock proteins are predominately produced in plants in response to heat shock and in response to various conditions including water deficit. Heat shock proteins that are produced in cells as a response to environmental disturbance have been reported to interact with cyclophilins in proper folding of disturbed protein structures (Bose et al., 1996).

PtMYB appears to be a MYB transcription factor. MYB proteins constitute a diverse family of plant transcription factors and control diverse pathways (Riechmann et al., 2000). MYB factor binding proteins induced due to water deficit have been reported in variety of plants including *Arabidopsis thaliana* (Urao et al., 1993). During dehydration, MYB factor binding proteins bind to the MYB recognization site in the promoter and activate the expression of dehydration resistance genes (Abe et al., 1997). Thus it seems that these proteins act as transcriptional activators in ABA signal transduction. MYB genes have been identified to control the expression of various genes including aldehyde dehydrogenase (Abe et al., 2003). These factors may be very important from an evolutionary point of view because important changes in evolution are attributed to changes due to transcription factors (Whetten et al., 2001).

Secondary pathway genes identified include an endochitinase. Endochitinases are predominately expressed in response to plant-pathogen interactions. Chitinases are a diverse group of plant pathogenesis related proteins with the ability to hydrolyze chitin (Raikhel *et al.* 1993). Reports of expression of same genes in multiple stress pathways have been established in angiosperms (Shinozaki et al., 1997). Differentially expressed

pathogenesis related genes have been reported in maritime pine (*Pinus pinaster*) in response to water-deficit-stress (Dubos et al., 2003).

Drought is an important hindrance for loblolly pine in attaining its maximum economic yield. Therefore, identifying molecular responses to drought may result in increases in yield following marker assisted breeding or genetic engineering. The functional roles of the genes need to be identified for this purpose. Ascertaining the functional roles of genes in pines is a difficult task due to the existence of multigene families (Kinlaw and Neale, 1997) the enormous size of the pine genomes (Whetten et al., 2001) and difficult transformation systems. The functional roles of genes with similarity to angiosperm genes can be identified making use of model organisms like Arabidopsis (Whetten et al. 2001). Another plausible way is to compare gene expression in highly divergent populations to understand the adaptive role of the genes (Kahn, 1993). In this study, we analyzed populations from the extremes of the climatic regimes within the loblolly pine natural range. The Lost Pines come from an area of very low rainfall. Average yearly rainfall for most of the loblolly pine range is between 1020 and 1520 mm. The Lost Pines receive less precipitation than the lower part of the range (average of 930 mm per year). South Louisiana receives a more rain (1500 – 1700 mm). Therefore these two populations were well suited for transcript level comparisons within the loblolly pine range. Several previous studies compared the morphological and physiological characters of the Lost Pines population. Trees from The Lost Pines have been reported to have higher survival capacity than trees from other provenances (Wells and Wakeley, 1966; Zobel and Goddard, 1955) but nothing is known about how these

changes are reflected at the molecular level. Therefore, we compared the expression patterns in the two populations to better understand the adaptive roles of the genes of interest and how the higher drought survival capacity of The Lost Pines is reflected at the molecular level. Our results indicate that at the highest level of stress, expression is similar within the two populations or is higher in the drought hardy population than in the drought sensitive population. These results coincide with earlier results from poplars were it was observed that there is preferential expression of various angiosperm water-stress-responsive genes in drought hardy species than in drought sensitive species (Pelah et al., 1997). Changes in gene expression within in species to adapt to the environment have been earlier reported in several species including spruce (Clapham et al., 1994). These changes are mostly attributed to changes in cis-regulatory elements which accounts for the largest heritable changes from an evolutionary point of view (Whetten, 2001).

The molecular response to water-deficit-stress in plants is a complex trait which involves genes unique to the pathway as well as genes common to other stress pathways. We have identified adaptive role of various genes involved in water-deficit-stress in loblolly pine by comparing gene expression in two populations. Results of this experiment can be used for selecting candidate genes as well as populations for future breeding programs.

CHAPTER IV

CONCLUSIONS

Molecular response to water-deficit-stress is a complex phenomenon involving genes in various pathways. To understand these complex responses we studied the gene expression in two species of pines using real-time RT-PCR.

In Chapter II, drought responsive genes in *Pinus halepensis.*were identified and the expression levels quantified. Various genes, including cyclophilins, sucrose synthase, LEA and chitinase genes which were previously reported in angiosperms in response to water-deficit-stress were identified. Genes unique to gymnosperms were also found. These results showed that water-deficit-stress pathway in gymnosperms have both similarities and differences to those found in angiosperms. To understand how gene expression varies within a species we compared gene expression in two populations: Yirka and Beit Jann. We observed difference in levels of expression, with the Beit Jann population showing higher levels of induction compared to Yirka in most cases. In chapter III, most genes identified in the earlier project were to be differentially expressed in drought stressed loblolly pines. Gene specific primers were designed to quantify the gene expression of LEA gene family members. In total, four LEA gene members were identified. Gene expression levels from two geographically divergent populations were compared. Lost pines which are disjoint population with more meager rainfall than found in the loblolly pine natural range showed higher level of expression of these genes than the South Louisiana source which comes from a region of high precipitation. Results from the two experiments indicate variation in gene expressions both due to

populations as well as due to treatments. Changes in gene expression due to populations may be an adaptation to the climatic conditions prevalent in the area of origin.

LITERATURE CITED

Abe H, Urao T, Ito T, Seki M, Shinozaki K, Yamaguchi-Shinozaki K (2003) *Arabidopsis* AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activator in abscisic acid signaling. Plant Cell **15**: 63-78

Abe H, Yamaguchi-Shinozaki K, Urao T, Iwasaki T, Hosokawa D, Shinozaki K (1997) Role of *Arabidopsis* MYC and MYB homologs in drought and abscisic acid regulated gene expression. Plant Cell **9**: 859-1868

Baker JB, Langdon OG (1990) Pinus taeda L. Loblolly pine. In Silvics of NorthAmerica, Vol. 1. Conifers. Agriculture Handbook 654. Forest Service, United States, Department of Agriculture, Washington, DC, 497-512

Bateman A, Coin L, Durbin R, Finn RD, Holich V, Griffiths-Jones S, Khanna A, Marshall M, Moxon S, Sonnhammer ELL, Studholme DJ, Yeats C , Sean R (2004) The Pfam protein families database. Nucl Acid Res **32**: D138-D141

Bilan MV (1961) Effect of planting date on regeneration and development of roots of loblolly pine seedlings. Proceedings XIII congress of International Union of Forest Research Organizations Vienna Austria Part 2 Volume 1 Sec 22-15

Bose S, Weikl T, Bugl H and Buchner J (1996) Chaperone function of Hsp90-associated proteins. Science 274: 1715-1717

Bray EA (1997) Plant response to water deficit. Trends Plant Sci **2**: 48-53

Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription PCR assays. J Mol Endo **25**: 169-193

Chang M, Aguilar GJR (1980) Effect of climate and soil on the radial growth of loblolly pine (Pinus taeda L.) in humid environment of southeastern U.S.A. Forest Ecology and Management **3:** 141-150

Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Rep **11**: 114-117

Chang S, Puryear J, Dias MADL, Funkhouser EA, Newton RJ, Cairney J (1996(a)) Cloning of DNA for chitinase homologue which lacks chitin binding sites and is down regulated by water stress and wounding. Plant Mol Biol **31**: 693-699

Chang S, Puryear J, Dias MADL, Funkhouser EA, Newton RJ, Cairney J (1996(b)) Gene expression under water deficit in loblolly pine (Pinus taeda): isolation and characterization of cDNA clones. Physiol Plant. **97**: 139-148

Clapham DH, Arnold SV, Dormling I, Ekberg I, Eriksson G, Larsson CT, Norell L and Qamaruddin M (1994) Variation in total and specific RNA during inwintering of two contrasting populations of *Picea abies* . Physiol Plant **90:** 504-512

Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported proteins and their targeting sequences. Eur J Biochem **241**: 770-786

Close TJ, Fenton RD, Moonan F (1993) A view of plant dehydrins using antibodies specific to the carboxy terminal peptide. Plant Mol Biol **23:** 279-286

Close TJ, Kortt AA, Chandler PM (1989) A cDNA based comparison of hydration-induced proteins in barley and corn. Plant Mol Biol **13:** 1033-1034

Costa P, Bahrman N, Frigerio JM, Kremer A, Plomion C (1998) Water-deficit-responsive proteins in maritime pine. Plant Mol Biol **38:** 587-596

Dejardin A, Sokolov LN, Kleczkowki LA (1999) Sugar/osmoticum level modulate abscisic acid–independent expression of two stress responsive sucrose synthase genes in Arabidopsis. Biochem Journal **344:** 503-509

Delmer DP, Amor Y (1995) Cellulose biosynthesis. Plant Cell **7:**987-1000

Dong JZ, Dunstan DI (1997) Characterization of cDNAs representing five abscisic acid-responsive genes associated with somatic embryogenesis in *Picea glauca* and their responses to abscisic acid stereostructure. Planta **203:** 448-453

Dubos C, Plomion C (2003) Identification of water deficit responsive genes in maritime pine (*Pinus pinaster* Ait.) roots. Plant Mol Biol **51:** 249-262

Dubos C, Provost GL, Pot D, Salin F, Lalane C, Madu D, Frigerio JM, Plomion C (2003) Identification and characterization of water stress responsive genes in hydroponically grown maritime pine (*Pinus pinaster*) seedlings. Tree Physiol **23:**169-179

Dure III L, Crouch M, Harada J, Ho THD, Mondy J, Quantrano R, Thomas T, Sung ZR (1989) Common amino acids sequences domains among the LEA proteins of higher plants. Plant Mol Biol **12:** 475-486

Geigenberger P, Hajirezaei M, Geiger M, Dieting V, Sunnewald V, Stitt M (1998) Overexpression of pyrophosphate leads to increased sucrose degradation and starch

synthesis, increased activities of enzyme for sucrose-starch interconversions, and increased levels of nucleotides in growing potato tubers. Planta **205:** 428-437

Gething MJ, Sambrook J (1992) Protein folding in the cell. Nature **355:** 33-35

Grunwald C, Schiller G, Conkle MT (1986) Isozyme variation among nature stands and plantations of Aleppo pine in Israel. J Bot **35:** 161-174

Ingram J, Bartel D (1996) The molecular basis of dehydration tolerance in plants. Annu. Rev. Plant. Physio. Plant Mol Biol **47:** 377-403

Jarvis SB, Taylor MA, Bianco J, Corbineau F, Davies HV (1997) Dormancy – breakage in seeds of Douglas fir (Pseudotsuga menziesii) (Mirb.) Franco.) Support for the hypothesis that LEA gene expression is essential for this process. J Plant Physiol **151:** 457-464

Jin H, Martin C (1999) Multifunctionality and diversity within the plant MYB-gene family. Plant Mol Biol **41:** 577-585

Kahn TL (1993) Characterization of expression of drought and abscisic acid-regulated tomato genes in drought-resistant species *Lycopericon pennellii*. Plant Physiol **103:** 597-605

Kinlaw C, Neale D (1997) Complex gene families in pine genomes. Trends Plant Sci **2**:356-359

Kirch HH, Nair A, Bartels D (2001) Novel ABA- and dehydration inducible aldehyde dehydrogenase genes isolated from the resurrection plant *Craterostigma plantagineum* and *Arabidopsis thaliana*. Plant J **28(5):** 555-567

Kleines M, Elster RC, Rodrigo MJ, Blerracq AS, Sulamini F, Bartels D (1999) Isolation and expression analysis of two stress-responsive sucrose-synthase genes from resurrection plant *craterostigma plantagineum* (Hochst). Planta **209:** 13-24

Koster KL (1991) Glass formation and desiccation tolerance in seeds. Plant Physiol **96:** 302-324

Kramer PJ, Kozlowski TT (1979) Physiology of Woody Plants. Academic Press, New York, 811p

Leone A, Costa A, Tucci M, Grillo S (1994) Comparative analysis of short and long-term changes in gene expression caused by low water potential in potato (*Solanum tuberosum*) cell suspension cultures. Plant Physiol **106:**703-712

Leprince O, Hendry GAF, McKersie BD (1993) The mechanisms of desiccation tolerance in developing seeds. Seed Sci Res **3:** 231-246

Loopstra CA, Sederoff RR (1995) Xylem-specific gene expression in loblolly pine. Plant Mol Biol **27:**277-291

Luan S, Albers MW, Schreiber SL (1994(a)) Light regulated tissue-specific immunophilins in a higher plant. Proc Natl Acad Sci USA **91:** 984-988

Luan S, Lane WS, Schreiber SL (1994(b)) PcypB: A chloroplast–localised heat shock–responsive cyclophilin from Fava bean. Plant Cell. **6:**885-892

McAdam HH, Arkin A (1999) Its noisy business Genetic regulation at the nanomolar scale. Trends Genet **15:** 65-69.

Meagher RB, Mckinney EC, Kandasamy MK (1999) Isovariant dynamics expand and buffer the responses of complex systems: the diverse plant actin gene family. Plant cell **11:** 995-1005

Meier CE, Newton RJ, Puryear JD, Sen S (1992) Physiological responses of loblolly pine (*Pinus taeda* L.) seedlings to drought stress, osmotic adjustment and tissue elasticity. J Plant Physiol **140:** 754-760

Nakai K, Horton P (1999) PSORT: a program for detecting the sorting signals of proteins and predicting their subcellular localization. Trends Biochem Sci **24(1):** 34-35

Newton RJ, Funkhouser EA, Fong F, Tauer CG (1991(a)) Molecular and physiological genetics of drought tolerance in forest species. For Ecol Mang **43:** 225-250

Newton R J, Funkouser EA, Tauer CG (1991(b)) Physiology of woody plants. Academic Press, New York, 811P

Nguyen-Queyren A, Bouchet-Lannat F (1989) Variation in growth and osmotic regulation of roots of water–stressed maritime pine (*Pinus pinaster*.Ait) provenaces. Tree Physiol **5:** 123-133

Nguyen-Queyren A, Bouchet-Lannat F (2003) Osmotic adjustment in three-year-old seedlings of five provenances of maritime pine (*Pinus pinaster*) in response to drought. Tree Physiol **23(6):** 397-404

Pelah D, Wang W, Altman A, Shoseyov O, Bartels D (1997) Differential accumulation of water stress-related proteins, sucrose synthase and soluble sugars in populus species that differ in their water stress response. Physiol Plant **99:** 153-159

Perozich, J, Nicholas H, Lindahl R, Hempel J (1999) The big book of aldehyde dehydrogenase sequences: An overview of the extended family In: Advances in experimental Medicine and Biology volume 463 (Weiner, H et.al, eds).: Kluwer Academic New York USA Pp 1-7

Philips K 1998 Drought chops forests by $342 million say AgNews, Texas A&M University agricultural program September, 1998 http://agnews.tamu.edu/stories/DRGHT

Raikhel NV, Lee H-I, Broekaert WF (1993) Structure and function of chitin binding protein. Annu. Rev. Plant. Physiol. Plant Mol Biol **44:** 591-615

Riechmann JL, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe OJ, Samaha RR, Creelman R, Pilgrim M, Broun P, Zhang JZ Ghandehari D, Sherman BK and Yu G (2000). Arabidopsis transcription factors: Genome wide comparative analysis among eukaryotes. Science **290:** 2105-2110

Sambrook J, Frilsch EF, Maniatus T (1989) Molecular Cloning: A laboratory manual Cold Spring Harbor Laboratory Press. Cold Spring Harbor NY

Sathyan P, Puryear JD, NewtonRJ, Loopstra CA (2004) Genes induced by water-deficit-stress are differentially expressed in two populations of aleppo pine (*Pinus halepensis)*(Unpublished)

Schiller G, Conkle MT, Grunwald C (1986) Local differentiation among Mediterranean populations of aleppo pine in their isoenzymes. Silvae Genet **35:** 1-8

Schiller G, Waisel Y (1989) Among–provenance variation in *Pinus halepensis* in Israel. For Ecol Manag **28:** 141-151

Schultz RP (1997) Loblolly Pine: The ecology and culture of loblolly pine (Pinus taeda L.) USDA Agricutural Handbook 713

Scholander PF, Hemmingsen E, Garey W (1965) Sap pressure in vascular plants. Science **148:** 339-346

Shinozaki KY, John M, Chua NH (1989) Four tightly linked rab genes are differentially expressed in rice .Plant Molecular Biology 14, 29-39

Shinozaki K, Yamaguchi-Shinozaki K (1997) Gene expression and signal transduction in water-stress response. Plant Physiol **115:** 327-334

Sivamani E, Bahielden A, Wraith JM, Al-Niemi T, Dyer W E, Ho T H D, Qu R (2000) Improved biomass productivity and water use efficiency under water deficit conditions in transgenic wheat constitutively expressing the barley HVAI gene. Plant Science **155:** 1-9

Skriver K, Mandy J (1990) Gene expression in response to abscisic acid and osmotic stress. Plant Cell **2:** 2555-2556

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

Stacy RAP, Aalen RB (1998) Identification of sequence homology between the internal hydrophilic repeated motifs in group 1 late-embryogenesis-abundant proteins in plants and hydrophilic repeats of the general stress protein GsiB of *Bacillus subtilis.* Planta **206:** 476-478

Urao T, Yamaguchi-Shinozaki K, Urao S, Shinozaki K (1993) An Arabidopsis myb homolog is induced by dehydration stress and its gene product binds to the conserved myb recognisation sequence. Plant Cell **5:** 1529-1539

Van Buijtenen JP, Bilan MV, Zimmerman RH (1976) Morphological characteristics related to drought resistance in *Pinus taeda*. In M.G.R. Cannel and F.T. Last (editors). Tree Physiology and Yield Improvement. Academic Press New York 349-359

Voo KS, Whetten RW, O'Malleg DM, Sederoff RR (1995) 4-Coumarate:coenyme A ligase from loblolly pine xylem: isolation, characterization and complementary DNA cloning. Plant Physiol **108:** 85-97

Wakamiya IH, Price J, Messina MG, Newton RJ (1996) Pine genome size, diversity and water relations. Physiol Plant **96:** 13-20

Wells OO, Wakeley PC (1966) Geographic variation in survival, growth and fusion-rust infection of planted loblolly pine. For Sci Monogr 11 40p

Whetten R, Sun YH, Zhang Y, Sederoff RR (2001) Functional genomics and cell wall biosynthesis in loblolly pine. Plant Mol Biol **47:** 275-291

Wilinston HL (1972) The question of adequate stocking. Tree Planters Notes **23(1):** 2PP

Wise MJ (2003) LEAping to conclusions: a computational reanalysis of late embroyogensis abundant protein and their possible roles. BMC Bioinformatics **4:** 52.0

Xu DP, Sung SJS, Loboda T, Kormanik PP, Black CC (1989) Characterization of sucrolysis via the uridine diphospahate and pyrophosphate-dependent sucrose synthase pathway. Plant Physiol **90:** 635-642

Yang SH ,Loopstra CA Unpublished report

Zahner R, Donnelly JR (1967) Refining correlation of rainfall and radial growth in young red pine. Ecology **48:** 525–530

Zhang XH, Chang VL (1997) Molecular cloning of 4-coumarate: coenzyme A ligase in loblolly pine and the roles of this enzyme in the biosynthesis of lignin in compression wood. Plant Physiol **113:** 65-74

Zobel BJ, Goddard RE (1955) Preliminary results on tests of drought-hardy strains of loblolly pine (*Pinus taeda* L.). Texas For Serv Res Note No.14

VITA

In September 2000, Pratheesh Sathyan received his B.S. (Forestry) degree from The Kerala Agricultural University, Kerala, India. In 2002, he started his graduate studies at Texas A&M University with specialization in molecular biology of plant species. He is expecting to graduate with his M.S. (Molecular and Environmental Plant Sciences) degree in December 2004 from Texas A&M University, College Station, TX. He can be contacted at:

Nandhanam House
Kp10/98
NCC Road, Trivandrum
Kerala, India-695005
E-mail: pratheeshsathyan@neo.tamu.edu