GENE EXPRESSION AND ASSOCIATION ANALYSES OF STRESS RESPONSES IN LOBLOLLY PINE (*PINUS TAEDA* L.)

A Dissertation

by

CANDACE MARIE SEEVE

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

DOCTOR OF PHILOSOPHY

December 2010

Major Subject: Molecular and Environmental Plant Sciences

Gene Expression and Association Analyses of Stress Responses

in Loblolly Pine (*Pinus taeda* L.)

GENE EXPRESSION AND ASSOCIATION ANALYSES OF STRESS RESPONSES IN LOBLOLLY PINE (*PINUS TAEDA* L.)

A Dissertation

by

CANDACE MARIE SEEVE

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

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Carol A. Loopstra
Committee Members,	Konstantin V. Krutovsky
	David N. Appel
	Dirk B. Hays
Chair of Interdisciplinary Faculty,	Dirk B. Hays

December 2010

Major Subject: Molecular and Environmental Plant Sciences

ABSTRACT

Gene Expression and Association Analyses of Stress Responses

in Loblolly Pine (*Pinus taeda L.*). (December 2010)

Candace Marie Seeve, B.S., University of Hawai'i at Hilo

Chair of Advisory Committee: Dr. Carol A. Loopstra

The molecular mechanisms underlying disease-resistance and drought-resistance in forest trees are not well understood. Linking variation in gene expression with genetic polymorphisms and with variations in disease- and drought-resistance phenotypes can provide information about these complex traits. We used real-time quantitative polymerase chain reaction (PCR) to detect variations in the expression of 88 disease- and drought-responsive genes within an association population of 354 loblolly pine trees (*Pinus taeda* L.). Using association genetics approaches, we then linked 3,938 single nucleotide polymorphisms (SNPs) in candidate genes with gene expression phenotypes to identify novel disease- and drought-responsive genes. To further examine differences in gene expression induced by drought, *Fusarium circinatum* (responsible for pitch canker disease), and drought + *F. circinatum*, the expression of 114 genes identified through comparative and association genetics approaches was analyzed on a subset of 24 loblolly pine trees possessing a range of pitch canker- and drought-resistance phenotypes. Significant differences in the uninduced expression of all 88 genes measured on the association population were observed among loblolly pine trees. Principal component analysis showed that some variation within the association population could be accounted for by population substructure of geographic origin. Hierarchical clustering of genes based on uninduced expression did not consistently group together functionally similar genes probably because

expression was collected on unstressed stem tissue. This was supported in the smaller expression study as correlations between expression values of genes in the same functional networks were usually stronger when induced by a treatment compared with correlations between the uninduced expression of genes in the control group. Gene expression frequently changed by up to 4-fold in response to one or more treatments, but PtMYB12 was the only gene that exhibited a statistically significant change in response to treatments. ANOVA analyses of gene expression controlling for pitch canker resistance and for water use efficiency phenotypes identified differentially expressed genes suggesting that they may be contributing to these phenotypes. Finally, association genetics approaches detected 101 significant associations between SNPs in 94 candidate genes potentially involved in stress responses and 27 gene expression phenotypes.

DEDICATION

This dissertation is dedicated to Ma and Pa Seeve, Ferf, Gram, Bunny, and all the rest.

Finally!!

ACKNOWLEDGEMENTS

I would like to thank my committee, Dr. Appel, Dr. Krutovsky, and Dr. Hays for their guidance throughout the course of my research. I would like to extend a special thank you to my committee chair, Dr. Loopstra, who has not only been a supportive advisor throughout my research, but has also been a role model for me both professionally and personally.

I would also like to thank my friends and the faculty at the University of Hawai'i at Hilo for encouraging me to pursue a graduate degree and for getting me started in a biology career. Thanks also go to my friends and colleagues and the MEPS department faculty and staff for making my time at Texas A&M University a great experience. I would like to extend my appreciation to Sreenath Palle and Jeff Puryear for their help and for making work that was sometimes very tedious, most enjoyable. I would also like to thank Ryan Hammons for showing me that there is always time to have a little fun and teaching me about all of (or at least most of) the Texas A&M traditions.

Finally (and most importantly), I would like to recognize my Mother, Father, and Brother for their encouragement *always*. Thanks to my Grandma Sandy for all of the letters and cartoon clips that she sent to me and always seemed to arrive just when I needed them. Thanks to my Aunt Bonnie for being a friend and a really cool aunt! You all have influenced my path more than you know.

TABLE OF CONTENTS

Page

ABSTRACT iii

DEDICATION v

ACKNOWLEDGEMENTS vi

TABLE OF CONTENTS vii

LIST OF FIGURES ix

LIST OF TABLES x

CHAPTER

I INTRODUCTION 1

II ASSOCIATION GENETICS OF STRESS-RESPONSIVE GENE EXPRESSION IN LOBLOLLY PINE (*Pinus taeda* L.) 7

- Introduction 7
- Materials and methods 10
  - Plant material 10
  - RNA extractions and cDNA synthesis 11
  - Gene selection and primer design 11
  - Quantitative real-time PCR 12
  - Amplification efficiency 12
  - Gene expression analysis 12
  - Principal component analysis 13
  - Clustering analysis 13
  - Genetic association methods 14
  - Model selection 14
- Results 15
- Discussion 34
- Conclusions 41

III THE EFFECTS OF CYCLIC DROUGHT AND PITCH CANKER EXPOSURE ON THE EXPRESSION OF ABIOTIC AND BIOTIC STRESS-RELATED GENES IN LOBLOLLY PINE (*Pinus taeda* L.) 43

- Introduction 43
- Materials and methods 47
  - Plant material 47
  - *Fusarium circinatum* inoculations 48

CHAPTER Page

Cyclic drought treatments 48
F. circinatum + drought treatments 49
Control treatments 49
Physiological measurements 49
RNA extractions and cDNA synthesis 50
Gene selection and primer design 50
Quantitative real-time PCR 51
Amplification efficiency 51
Gene expression analysis 52
Data analysis 52
Results 53
Discussion 64
Conclusions 68

IV CONCLUSIONS 70

REFERENCES 73

APPENDIX A 86

APPENDIX B 93

APPENDIX C 95

VITA 99

LIST OF FIGURES

FIGURE		Page
2.1	Range map for loblolly pine trees in the southern U.S.	10
2.2	Constitutive gene expression frequency distributions	17
2.3	Graphical representation of principal component analysis of gene expression results	19
2.4	Dendrogram produced by hierarchical clustering of gene expression	22
2.5	Gene expression grouped by associated SNP genotype	30
3.1	Diagram showing overlap in gene expression responses to disease and drought stresses	54
3.2	PtMYB12 expression among loblolly clones by treatment	55
3.3	Gene expression box plots by treatment	60
3.4	Correlations between levels of gene expression by treatment	63

LIST OF TABLES

TABLE		Page
2.1	Gene expression ranges by functional class	18
2.2	SNPs detected in association with expression phenotypes	25
2.3	Models predicting carbon isotope discrimination	32
2.4	Models predicting pitch canker lesion length	33
3.1	Gene grouped by putative function	56
3.2	Gene expression changes relative to the control group in response to treatments	57
3.3	Genes identified by ANOVA analyses controlling for water use efficiency or pitch canker resistance phenotypes	61

CHAPTER I

INTRODUCTION

Forests in the southern U.S. provide lumber and pulpwood, habitat for wildlife, carbon sequestration, and recreation among other immeasurable services (Hanson et al. 2010). The predominant pine species of southern forests is loblolly pine (*Pinus taeda* L.). Its range extends almost continuously from the Piney Woods region of Texas east to the Atlantic Coast and north to Virginia. Loblolly pine is a fast-growing evergreen that succeeds in a variety of site conditions making it an ideal species for both management purposes and timber production. Since the 1950's southern tree improvement cooperatives have made substantial gains in growth rate, wood properties, and disease resistance through selection and breeding (Schultz 1997). Today genetically improved loblolly pine is the most important and widely planted pine species in the southern U.S. (McKeand et al. 2003).

While traditional selection and breeding has certainly been successful, twenty or more years are required for a single generation of breeding and testing (Schultz 1997). The ADEPT2 ("Allele Discovery of Economic Pine Traits") project has been a collaborative effort by UC-Davis, UFL, NCSU, and Texas A&M to use population genomics approaches to identify genes controlling economically and ecologically important complex traits and to discover valuable allelic variants for marker assisted breeding to accelerate the selection of elite loblolly families. (ADEPT2 2008).

This dissertation follows the style of Tree Genetics & Genomes.

Most of the economically important traits in loblolly pine, such as wood properties, growth, pathogen resistance, and dehydration tolerance, are complex traits under the control of multiple genes. Quantitative trait loci (QTL) mapping strategies employed to dissect complex traits in other crop species is limited in loblolly pine and other forest trees (Strauss et al. 1992). High levels of genetic diversity maintained in loblolly pine reduce linkage between marker alleles and QTL alleles in populations (Strauss et al. 1992). Also, the long generation time of loblolly pine is not conducive for the establishment of a QTL mapping population by crossing and backcrossing (Strauss et al. 1992). In contrast with traditional linkage mapping, association genetics makes use of the high levels of genetic diversity present within loblolly pine populations to detect allelic affects that are responsible for variations in phenotypic traits. The use of a large population of unrelated individuals where recombination has accumulated over generations enables fine-scale mapping of allelic affects and also allows for the examination of multiple alleles present in the population (Neale and Savolainen 2004; González-Martínez et al. 2007).

Association genetics, or linkage disequilibrium mapping, identifies genetic determinants of complex traits by correlating genetic polymorphisms with phenotypic variations in a population of unrelated or distantly-related individuals (Neale and Savolainen 2004; Cordell and Clayton 2005). As its name implies, linkage disequilibrium, is central to association genetics. It dictates the mapping resolution and the marker density required to draw meaningful associations in a species (Neale and Savolainen 2004; Veyrieras et al. 2007). Linkage disequilibrium is the nonrandom association of alleles at different loci within a population. It is a population trait contingent on population structure, selection, drift, recombination, mutation and mating system (Flint-Garcia et al. 2003). Species such as loblolly pine that are distributed in large and continuous, outcrossing populations tend to possess linkage disequilibrium that decays more

rapidly (Flint-Garcia et al. 2003). The large number of markers required to cover the entire loblolly genome (>2.0 x 10^{10} bp) and absence of a full genome sequence makes genome-wide scans in loblolly unfeasible (Wakamiya et al. 1993; Neale and Savolainen 2004; Veyrieras et al. 2007). Instead, candidate gene-based association studies that target polymorphisms within candidate genes that may be influencing the trait of interest are well-suited for loblolly pine (Neale and Savolainen 2004).

Two of the most important traits that tree breeders seek to improve in loblolly pine are resistance to pests and pathogens and drought tolerance. Drought is a major cause of loblolly pine seedling mortalities and limits growth of mature trees (Schopmeyer 1939; Newton et al. 1991; Schultz 1997; Watkinson et al. 2003). Recent predictions that global climate change will intensify and prolong periods of drought has increased interests in the effects of drought and how interactions between drought and other disturbances will impact forest trees (Winnett 1998; Logan et al. 2003; Kliejunas et al. 2008).

At the physiological level, forest tree adaptations to avoid water deficits include accumulation of solutes for osmotic adjustment, early leaf abscission, limiting leaf area, cuticular wax accumulation, stomatal control, and increased rooting, among others (Newton et al. 1991). The molecular mechanisms underlying these physiological changes are not well understood, but are of great interest for improving drought tolerance. It is known that following perception of osmotic changes, plants induce signal transduction cascades (Shinozaki et al. 1997; Chaves et al. 2003; Shinozaki et al. 2007). There appears to be at least five signal transduction cascades involved response to dehydration. Three pathways are regulated by the accumulation of the hormone abscisic acid (ABA) and two pathways are ABA-independent (Chaves et al. 2003; Shinozaki et al. 2007). The genes induced in these pathways include genes coding for proteins that protect the cells from potential dehydration damage and transcription factors and regulatory

proteins that further control plant responses through secondary messengers, hormones, or phosphorylation (Shinozaki et al. 1997).

Similarly, plant resistances to pathogen infections are complex. Plants possess constitutive defenses that are physical barriers such as the plant cell wall and cuticle and induced defenses that are initiated through multiple signal transduction pathways. Induced defenses are elicited by pathogen proteins and are amplified through endogenous signaling molecules including reactive oxygen species, salicylic acid, ethylene, jasmonates, Ca^{2+}, H^{+} ions, lipids, nitric oxide (Yang et al. 1997; Kunkel and Brooks 2002). An additional layer of complexity is the specific interactions between specific plant and pathogen genotypes in gene-for gene resistances. In this type of resistance the plant must possess a dominant resistance (R) gene corresponding to the pathogen's dominant avirulence gene (Avr) to initiate defenses (Yang et al. 1997). The manifestations of defense responses include lignifications, callose depositions, or programmed cell death (hypersensitive response) around the infection site, production of antimicrobial compounds, and systemic acquired resistance to further infections (Yang et al. 1997). Rapid recognition, activation of the expression of defense-related genes, and downstream defense mechanisms appear to be key in resistance to infections (Yang et al. 1997).

There appears to be significant crosstalk among conserved signaling pathways that are induced by different biotic and abiotic stresses (Knight and Knight 2001). In a survey of the previous literature examining the effects of drought on plant-pathogen interactions, Desprez-Loustau et al. (2006) found a general trend for increase in disease incidence with increasing drought stress, however they emphasized that drought-disease interactions were dependent on multiple factors including severity and length of drought, disease type, genotypes of both the host and pathogen, and environment.

The necrotrophic fungal pathogen *Fusarium circinatum* is the causal agent of pitch canker infections of conifers almost worldwide (Wingfield et al. 2008). *F. circinatum* spores typically enter through wounds or are vectored by the eastern pine weevil (Barnard and Blakeslee 1987; Wingfield et al. 2008). The first symptoms of infection are reddening of the infected stem followed by the appearance of sunken, resin-soaked cankers and finally death of the stem as the fungus girdles the stem (Barnard and Blakeslee 1987). Main stem cankers can be lethal if the infection is severe enough to fully girdle the tree (Barnard and Blakeslee 1987). Pitch canker infections frequently occur in plantations and seed orchards where they may result in significant losses (Wingfield et al. 2008). Interactions between drought and pitch canker disease in loblolly pines are not known. Evidence for positive associations between drought conditions and pitch canker incidence include increases in infection rates during the late summer and autumn months and an outbreak of pitch canker infections in Florida during a period of severe drought (Schmidt et al. 1976; Barnard and Blakeslee 1987).

The complexity of disease- and drought-resistance responses can complicate the precise phenotyping of these traits (González-Martínez et al. 2008). Maximizing the power to detect gene-phenotype relationships sought in genetic studies requires both high-throughput genotyping capabilities and precise high throughput phenotypic measures to detect subtle phenotypic changes (Edmeades et al. 2004). Recent advances in high-throughput expression-profiling technologies have made the measurement of thousands of gene expression phenotypes in different genetic backgrounds or under different environmental conditions feasible. Similar to any complex trait, the genetic determinants that control the expression phenotype can then be identified through different genetics approaches and may provide detailed information about gene networks controlling important traits than broader phenotypic measures (Cheung and Spielman 2002; Edmeades et al. 2004).

Real-time quantitative PCR (RT-qPCR) has become a widely used method for analyzing the expression of a moderate number of genes (VanGuilder et al. 2008). Real-time RT-qPCR follows the amplification of PCR products in "real time" throughout the PCR reaction (VanGuilder et al. 2008). PCR products are detected by one of a variety of fluorescence-based technologies (e.g., fluorescent probes and intercalating dyes). Real-time quantitative PCR is highly sensitive and precise. Its specificity enables quantitation of the expression of individual gene family members (VanGuilder et al. 2008). In addition, detection of accumulating PCR products by fluorescent intercalating dyes such as SYBR® Green are flexible allowing expression to be collected in multiple genotypes making it suitable for molecular phenotyping of large numbers of individuals.

We have used RT-qPCR to detect variation in the expression of disease-responsive and drought-responsive genes among loblolly pine trees indicating that different disease and drought resistance phenotypes exist across the range of loblolly pine. Furthermore, using candidate gene association genetics approaches we have been able to identify single nucleotide polymorphisms (SNPs) associated with our expression phenotypes and, once verified, could be used as markers in future breeding programs selecting for disease resistance and drought resistance in loblolly pine trees.

CHAPTER II

ASSOCIATION GENETICS OF STRESS-RESPONSIVE GENE EXPRESSION IN LOBLOLLY PINE (*Pinus taeda* L.)

INTRODUCTION

Water availability and diseases are two of the largest hurdles for both managed stands and natural populations of forest trees today. Forest trees are not amenable to traditional breeding and selection strategies (Strauss et al. 1992) making progress toward dissecting and improving these traits a slow process. The development of genetic tools would significantly reduce costs for tree improvement programs to identify and select elite families and could even increase the sustainability and productivity of forest trees.

The economic and ecological importance of loblolly pine (*Pinus taeda* L.) in the southeastern U.S. has fostered the development of breeding programs and genetic resources that have positioned loblolly pine for genetic studies that could both facilitate the selection of elite loblolly families by marker assisted breeding and provide information relevant across the Pinaceae family (Krutovsky et al. 2004). Most of the important traits in forest trees such as wood quality and response to stress are quantitative traits under the control of multiple loci. The utility of traditional quantitative trait loci (QTL) mapping approaches in loblolly pines are limited though. Forest trees have long generation times that are not conducive to the establishment of a segregating population and are largely outcrossing which limits linkage between marker alleles and QTLs (Strauss et al. 1992). These traits, among others, while unsuitable for traditional QTL mapping, make loblolly pines good candidates for association genetics approaches (Neale and Savolainen 2004).

Association genetics detects statistically significant correlations between the occurrence of allelic polymorphisms and a phenotypic trait across a whole population. In contrast with QTL mapping studies that rely on known pedigrees and linkage between the marker alleles and QTLs, association genetics takes advantage of the recombination that has occurred in the whole population and broken up the linkage disequilibrium (LD) for fine mapping of QTLs in LD with the trait of interest (Flint-Garcia et al. 2003; Neale and Savolainen 2004). Linkage disequilibrium dictates the mapping resolution and the marker density required to achieve meaningful associations (Flint-Garcia et al. 2003). In loblolly, linkage disequilibrium measured in genic regions decayed rapidly—within 2 Kb—making the number of markers likely required to cover the whole genome unfeasible (Brown et al 2004; Neale and Savolainen 2004). Candidate gene approaches target allelic variation in previously identified genes thought to be functionally related to the trait of interest and is suitable for studies in loblolly pine (Neale and Savolainen 2004).

The main concern in designing association genetics studies is the presence of population substructure (Pritchard and Rosenberg 1999). Population substructure can result in spurious associations due to an over-representation of an allele in a segment of the population (Pritchard and Rosenberg 1999; Yu et al. 2005). If the extent of population substructure is known, software employing methods for correcting for population substructure are available (Yu et al. 2005; Bradbury et al. 2007). Loblolly pines are largely outcrossing species and are distributed almost continuously throughout the southeastern U.S. limiting the amount population substructure across the range (Neale and Savolainen 2004).

Gene expression studies have gained popularity for two reasons. First, recent studies of variation in gene expression among organisms have shown that it may be more significant in driving evolution than once thought (King and Wilson 1975; Fay et al. 2004; Stranger et al.

2007). Second, gene expression can serve as a precise phenotype in identifying genotypic variations that influence important traits. Maximizing the power to detect gene-phenotype relationships sought in genetic studies requires both high-throughput genotyping capabilities and precise high throughput phenotypic measures to detect subtle phenotypic changes (Edmeades et al. 2004). Recent advances in high-throughput expression-profiling technologies have made the measurement of thousands of gene expression phenotypes in different genetic backgrounds or under different environmental conditions feasible. Similarly to any complex trait, the genetic determinants that control the expression phenotype can then be identified through different genetics approaches. The use of gene expression phenotypes in genetic studies simplifies the interpretation of complex phenotypes and reveals very small phenotypic differences that may provide more detailed information about genes controlling important traits than broader phenotypic measures (Cheung and Spielman 2002; Edmeades et al. 2004).

This study aimed to: 1) identify natural variation in the expression of 89 disease- and drought-related genes in loblolly individuals that may be evolutionarily and/or economically important; to 2) use candidate gene association genetics approaches to relate gene expression phenotypes with single nucleotide polymorphisms (SNPs) in 3,938 candidate genes thought to be involved in drought and/or disease resistance in loblolly pine trees; to 3) assign putative functions to candidate genes; and to 4) examine SNP effects on expression phenotypes. To our knowledge this study, in cooperation with another association genetics study examining wood property traits (Palle et al. 2010), is the first that has sought to identify genetic determinants for gene expression phenotypes using association genetics approaches.

MATERIALS AND METHODS

Plant material: Four hundred seventy-three unrelated loblolly individuals (clones) with two biological replicates (ramets) each were provided as rooted cuttings from the ADEPT2 association population. This population included more than 500 unrelated loblolly individuals representing most of the natural range (Fig. 2.1) that were provided from seed lots by the three southern pine breeding cooperatives and was maintained at North Carolina State University. The rooted cuttings were potted and grown in green houses in a completely randomized design for four months. At the time of harvesting the trees, 449 clones with at least 2 biological replicates (ramets) appeared to be in healthy condition. The main stem, roots, and needles were collected separately in tubes, flash-frozen and stored at -80°C.

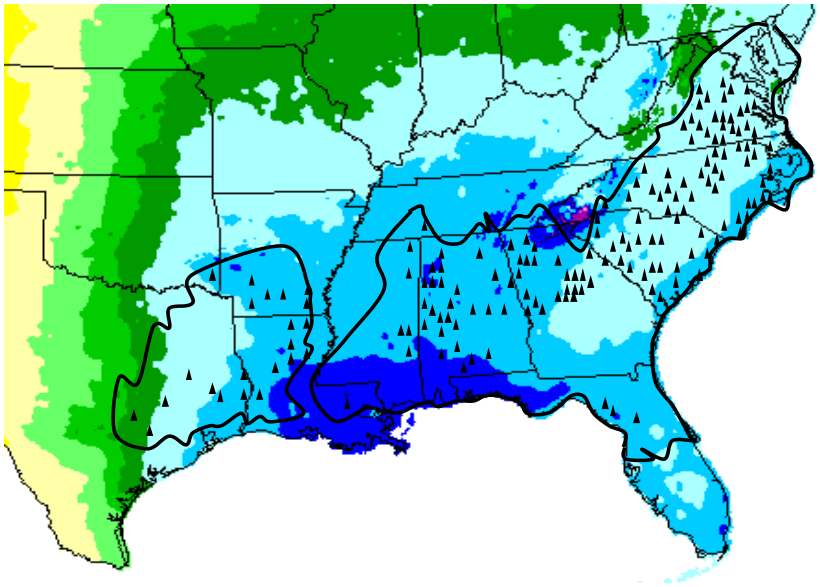

Fig. 2.1 Range map for loblolly pine trees in the southern U.S. The ADEPT2 association population represents most of the natural range of loblolly pine (denoted by the black border). Each marker on the map represents a county containing one or more individuals in the population.

RNA extractions and cDNA synthesis: Total RNA was extracted from finely ground pine stems according to Chang et al. (1993) with the addition of an extra chloroform extraction to fully eliminate protein contamination. The RNA was treated with DNA-free™ (Ambion, Austin, TX) to remove contaminating DNA. Total RNA was quantified using a Nanodrop 1000™. First strand cDNA was primed from 5µg RNA with random hexamer primers using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA) and amplified according to the manufacturer's instructions.

Gene selection and primer design: Genes for expression analysis were selected based on evidence of their involvement in disease and/or drought responses in the previous literature and from previous results generated in our laboratory. Many of the genes were identified in other species. Table A.1 (Appendix A) lists the putative function for each gene and refers to the study that indicated its involvement in stress responses. Homologous sequences in *P. taeda* were identified through blast searches of the loblolly pine ESTs within the NCBI Expressed Sequence Tag database (dbEST). The Unigene set linked with the ESTs that were most similar to the query sequence (based on the E-value and coverage of the query sequence) was selected (see Table A.1 for Unigene IDs). The software Sequencher 4.2 (Gene Codes Corp., Ann Arbor, MI) was used to align the Unigene EST's and generate contig sequences representing the putative orthologous genes in loblolly pine. The contig sequences were blasted again in NCBI using the BLASTN tool to identify other putative genes in *P. taeda* with high sequence similarity. In these cases, the gene was rejected for the association analysis since amplification of a single gene product appeared to be unworkable. Gene-specific primer pairs were designed using Primer Express 3.0 software (Applied Biosystems, Foster City, CA). Sequence parameters included: 1) primer T_m: 58-62°C; 2) primer %GC Content: 40-60%; 3) primer length: 18-22 nt; 4) amplicon length: ~75mer. Areas of the contig sequences that appeared to contain SNPs based on alignment of the

ESTs were avoided for primer design. Primers were supplied by Invitrogen (Carlsbad, CA). Primer sequences are listed in Appendix A, Table A.1.

Quantitative real-time PCR: Real-time quantitative PCR (RT-qPCR) reactions were performed using SYBR® green dye (Applied Biosystems, Foster City, CA) for quantification. With the exception of scaling down volumes to 8 µl, reactions were carried out as instructed in the manufacturer's protocol in 384-well plates and run on an ABI 7900 HT (Applied Biosystems, Foster City, CA). Expression was measured on two ramets by two technical replicates for every clone. No template and no reverse transcriptase reactions were used as negative controls to detect contamination of the PCR reagents and cDNA samples. Raw expression data was collected with the software SDS 2.3 (Applied Biosystems, Foster City, CA) and relative gene expression values ($\Delta\Delta C_T$) were calculated in RQ Manager 1.2 (Applied Biosystems, Foster City, CA) by subtracting the target gene expression from the expression of the endogenous control β-actin and then subtracting this ΔC_T value from the ΔC_T value of another clone selected as a calibrator sample. The selective amplification of a single gene was verified by the presence of a single, smooth peak in dissociation curve analyses.

Amplification efficiency: The primer binding sites for every gene-specific primer were sequenced for 2-3 of the highest and lowest expressing clones to ensure that SNPs in primer binding sites did not decrease the amplification efficiency (AE) of any gene. To further ensure that sample-to-sample variations in AE were not responsible for the expression differences, the amplification efficiencies for every gene were compared among the highest and lowest expressing clones using a one-way ANOVA included in the "Data Analysis for Real-Time PCR" (DART-PCR) Excel worksheet (Peirson et al. 2003).

Gene expression analysis: Delta-delta C_T values were manually examined for outliers that may have been due to experimental errors. Clones where expression between biological

replicates varied by more than 0.7 cycles were removed leaving 354 clones with at least two biological ramets and two technical replicates. Welch's ANOVA for unequal variances was used to test for significant variation in gene expression among the different clones. The mean expression value for each clone was calculated for every gene and this expression data was autoscaled by subtracting the mean $\Delta\Delta C_T$ for that gene from each data entry, and dividing by the gene's standard deviation to reposition the mean expression and standard deviation to 0 and 1, respectively (Stahlberg et al. 2008). These data were used for subsequent principal component, clustering, association analyses, and model selection.

Principal component analysis: Principal component analysis (PCA) was used to explore for the presence of population substructure that could explain some of the gene expression variation observed. PCA was performed based on the correlation matrix of gene expression values for each tree in PC-ORD v. 4.0 software (McCune and Mefford 1999). Missing expression values were replaced with the average expression for that gene. The significance of each PC was evaluated by the Broken Stick method (Jackson 1993) provided in PC-ORD. PCs with eigenvalues greater than the "broken stick" eigenvalue were considered significant for further interpretation. The loading scores for the significant components were saved for further interpretation.

Clustering analysis: Hierarchical clustering of genes based on expression profiles and subsequent bootstrap analysis to assess the strength of each cluster was performed using the pvclust package (Suzuki and Shimodaira 2006) in R software (R Development Core Team 2008) (method=ward; distance measure=Euclidean; nboot=10,000). Nodes with approximately unbiased boostrapping values (AU) higher than 90% were considered to be well supported and were highlighted using the pvrect function in pvclust.

Genetic association methods: SNP discovery was performed at UC-Davis by resequencing 7,508 candidate genes spanning the entire linkage map for loblolly pine in 18 unrelated loblolly pine haploid gametophytes (Eckert et al. 2010). Three thousand nine hundred thirty-eight SNPs spanning the entire linkage map of loblolly pine were selected for genotyping (UC Davis Genome Center) in the full ADEPT2 association population (Eckert et al. 2010).

Population substructure was inferred from two sets of molecular markers—3059 SNP markers and 23 SSRs (Eckert et al. 2010). Principal component analysis on the SNPs and use of the software STRUCTURE on both the SSR and SNP markers revealed minimal population structure (Eckert et al. 2010). The Q-matrix (k=5) was selected for inclusion as a covariate in a general linear model (GLM) to account for population substructure and control for false positive associations.

A general linear model was performed for every SNP-trait pair in the software TASSEL v. 2.0.1 (Bradbury et al. 2007). The 'qvalue' package in R (R Development Core Team 2008) was used to calculate q-values to control for false positives accumulated in multiple testing (Storey and Tibshirani 2003). A significance threshold value of q=0.05 was considered significant.

Model selection: The model selection package 'leaps' (Lumley and Miller 2004) in R (R Development Core Team 2008) was utilized to identify the best subset of gene expression phenotypes that predicted physiological measurements of water use efficiency or disease resistance in linear regression. Carbon isotope ratios (CID) were collected as a measure of water use efficiency. CID was measured on the entire ADEPT2 association population at North Carolina State University by Patrick Cumbie. CID values were provided as BLUP scores to reduce environmental noise (Cumbie, personal communication). Lesion length representing resistance to the fungal pathogen *Fusarium circinatum* was measured at the University of Florida

by Tania Quesada. Since the leaps function can only handle 31 variables, exhaustive searches were performed using the regsubsets function within 'leaps' to reduce the number of genes predicting CID or *F. circinatum* lesion size (nbest=10; nvmax=8; force.in=NULL; force.out=NULL, really.big=TRUE). These genes were included for further analysis with the leaps function (int=FALSE; method=c("Cp"), nbest=3). Leaps returned the three best subsets of each size for predicting either CID or lesion size in linear regression. The test statistic Mallow's C_p was selected to prevent overfitting of the model.

RESULTS

Significant variation in the expression of stress-related genes exists among 354 clones representing the natural range of loblolly pine.

The expression of 89 disease- and drought-related genes was measured using RT-qPCR. Amplification efficiency was examined in the highest and lowest expressing clones using a one-way ANOVA in the DART-PCR Excel file (Peirson et al. 2003) and by sequencing through the primer binding sites for each gene. In some cases SNPs were present in the primer binding sites of the low expressing clones, but ANOVA analysis showed no significant difference among amplification efficiencies ($p< 0.05$) and sequencing of the primer binding sites of the high expressing clones revealed SNPs also. Therefore, these SNPs were not believed to be affecting gene expression values. One gene, PtMLO1, had a SNP at the 3' end and significant differences in the amplification efficiency of the highest and lowest expression clones were detected with DART-PCR. Further sequencing of two high expressing clones and a mid expressing clone showed that the highest expressors did not have the SNP and the middle expressor was heterozygous for the SNP. This gene was removed for all further analyses.

Gene expression values were examined manually. In cases where expression between biological ramets differed by more than 1.6-fold, both ramets of that individual were removed, and when a single technical replicate differed by more than 1.6-fold (likely due to experimental error) it was removed leaving 354 clones with two ramets for expression and association analyses. For most genes, expression was normally distributed with the majority of clones varying from the mean by only about ±2-fold and fewer genotypes with expression values ±2.8-fold from the mean (Fig. 2.2). A few genes exhibited slightly skewed distributions and one gene, pyrroline-5-carboxylate dehydrogenase (P5CDH), had a bimodal distribution with more clones displaying extreme expression values (>2.8-fold from the mean) (Fig. 2.2). Welch's ANOVA analysis revealed significant differences in mean expression values among clones for every gene ($p=0.01$).

Genes were categorized by putative gene function (Table 2.1). Each category had a wide range of expression differences and there did not appear to be any differences between groups. The average range of expression differences between the highest and lowest expressing clones was 49.4-fold. The gene with the largest range between the highest and lowest expressing clones was PtEMB4, encoding a late embryogenesis abundant protein (LEA). Its expression varied by more than 9,000-fold. An outlying clone with very low expression levels would have extended this range to more than 500,000-fold. Even though the expression data for this individual seemed accurate as all replicates for both ramets differed by less than 1.6-fold the individual was not included in the expression or clustering analyses. However, the outlying individual was included in the association study to detect allelic variation in another gene regulating PtEMB4 that could provide a possible explanation for the very low expression levels. A drought-induced MYB transcription factor displayed the smallest range of expression—just over a 5-fold difference. A

full list of the average expression and expression range between the highest and lowest expressing clones for all genes is listed in Appendix B, Table B.1.

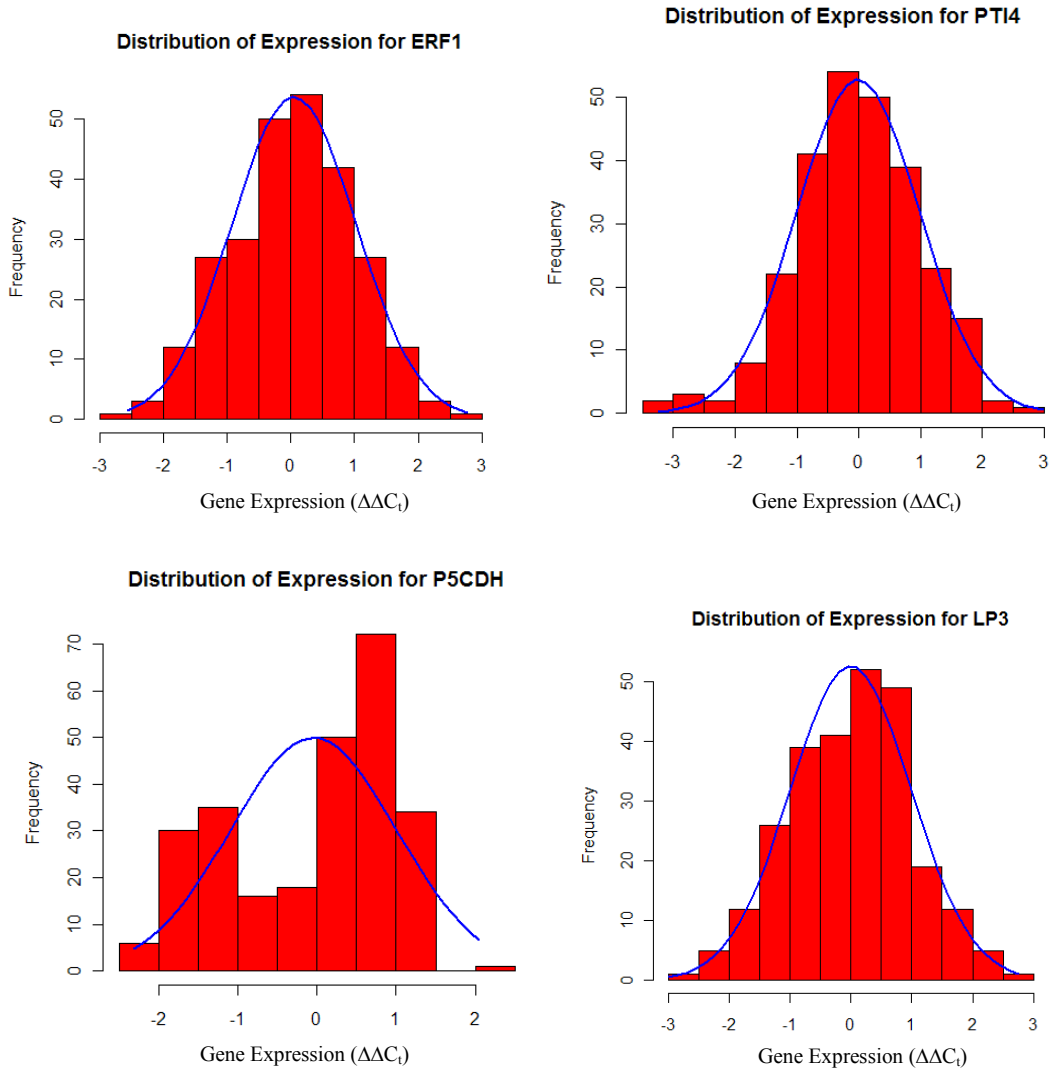

Fig. 2.2 Constitutive gene expression frequency distributions. Frequency distributions of gene expression values show that most genes were normally distributed with a standard deviation of about ±1 cycles (ΔΔCT) with one cycle being equivalent to a 2-fold difference in gene expression. More negative values (ΔΔCT) are equal to higher expression. Histograms were centered around zero for ease of reading. Gene expression was autoscaled for all further analyses to ensure that expression was normally distributed with a standard deviation of ±1 cycles and a mean of zero.

Table 2.1 Gene expression ranges by functional class. Genes have been grouped by putative function and are listed along with the extreme ranges of expression differences observed for genes in each class.

Putative gene functions	Gene names	Range of expression differences (fold-difference)
Cell wall related	BGLU1, CslA1	43.3-107
Detoxifying enzymes	GLX1, PtGSTF9, PtAldh, PtGSTU18,	6.43-83.4
Disease-responsive transcription factors	ERF1, PTI4, PtWRKY	27.1-33.5
Disease signaling	EB9D, ADR1, ARF, AOS, HPL, SamCMT, COI1, Avr-cf9	7.16-2060
Drought-responsive transcription factors	RAP2.1, RAP2.4, PtMYB6, bZIP, PtMYB9, NAC1, ATAF-1, HDZ32,	5.24-267
Drought responsive genes (previously identified in pine)	LP3, Pt31, Pt38, PtGRP, PtGTP, PtIP	13.3-995
Drought signaling	CPK3, RPK1, ERD3, ERECTA, ABI1, PtAN1, NCED	9.92-193
H_2O_2 signaling	PtCPN10, IMPA, HSP17.4-CIII	14.3-55.8
Late embryogenesis abundant proteins	PtLEA2, PtLEA3a, PtLEA3c, PtEMB1, PtEMB2, PtEMB3, PtEMB4	30.2-6100
Osmotic adjustment	PtGOLS1, P5CR, BALDH, P5CDH	5.95-627
Pathogenesis-related proteins	PR-2, PR-3, PR-5, PR-10	91.4-312
Peroxidases	PX-1, PX-2, PX-3, PX-4, PX-5, PtGPX1, PtGPX2, PtGPX3	10.8-381
Phenylpropanoid pathway	PtOMT, OXR2, CHI, PtPDIR1, ANR, CHS	12.5-141
ROS scavenging	PT1, OXR1,	7.97-42.1
Programmed cell death	PtBAG-1, PtBAG-4, BI-1, PtMLO1, PtMLO2	8.47-61.7
Terpenoid biosynthesis	CYPA, CYPB, CYPC, CYPD, STR_SYN, (+)pin-TPS, (-)pin-TPS, HMG-CoA, AOC, Sesqui-TPS, trp-TPS	15.1-123

Population substructure was analyzed by principal component analysis (PCA) on gene expression values for each clone. Principal component eigenvalues were determined to be

significant according to the “Broken stick eigenvalue” method. The results of the PCA showed that four significant principal components (PC) accounted for 29% of the cumulative variation in gene expression among loblolly individuals. Further inspection revealed a geographical organization of the first three principal components. The proportion of trees in each county that belonged to each of the first three components was calculated. Trees from the Atlantic Coast region loaded onto the first principal component most often, while trees from the Gulf Coast region and from west of the Mississippi tended to load onto the second and third components, respectively (Fig. 2.3).

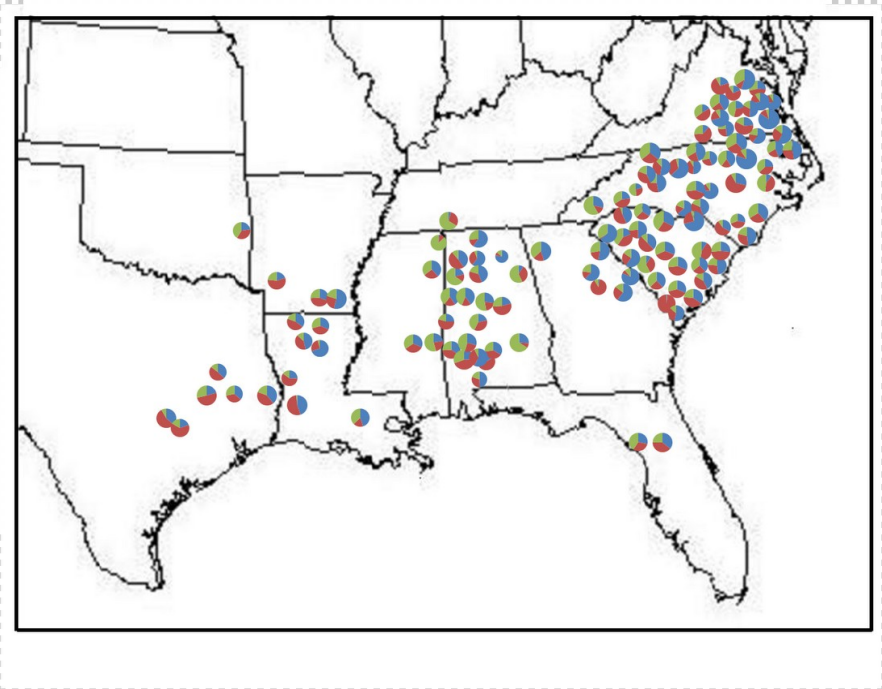

Fig. 2.3 Graphical representation of principal component analysis of gene expression results. Each pie chart on the map represents the proportion of trees in that county that belong to each of the first 3 principal components. Components appear to have a geographical basis dividing the range of loblolly pine into three regions—Atlantic Coast, Gulf Coast, and West of the Mississippi.

Hierarchical clustering of genes revealed weak clustering of gene family members and functionally similar genes.

To examine the hypothesis that functionally similar genes have similar expression profiles, genes were hierarchically clustered by centered correlation coefficients with bootstrapping for support. Nodes with approximately unbiased bootstrap values (AU) greater than or equal to 90 were deemed to be well-supported and are highlighted in red in Fig. 2.4. Two clusters, 8 and 10 did not have AU values greater than 90 (highlighted with green in Fig. 2.4), but were notable due to functional similarities identified in the previous literature and from previous experiments in our lab. Cluster 8 contained both MYB transcription factors and two more drought responsive genes. COI1 appears to be involved in MeJA signaling upstream of the MeJA-ABA signaling branch point where as ABI1 is a negative regulator of ABA signaling (Munemasa et al. 2007). Cluster 10 is composed mainly of dehydration-responsive genes formerly identified in loblolly pine. The genes in clusters 1, 2, and 3 are involved in stress responses through oxygen species. Cluster 4 contains BI-1 (BAX inhibitor 1) and two BAG (Bcl2-associated athanogene) domain genes.

All three of these genes are negative regulators of programmed cell death (Kawai-Yamada et al. 2001; Doukhanina et al. 2006; Kabbage and Dickman 2008). Cluster 5 contains a heat-shock protein and galactinol synthase 2. Nishizawa et al. (2006) found that over-expression of a stress-inducible heat-shock transcription factor resulted in strongly correlated increases in the expression of both of these genes suggesting that they may be target genes of this transcription factor. The genes in cluster 6 have been shown to be responsive to salt and dehydration stresses, except Avr-Cf9. Avr-cf9 is homologous to a disease resistance gene first isolated in Scots Pine (Li and Asiegbu 2004) and has not been shown to be responsive to abiotic stresses. Its inclusion into a cluster with other drought genes suggests that it could have multiple functions in loblolly pine. The two genes in cluster 7—a β-glucosidase and the ERECTA gene—are involved in lignification and cell wall synthesis, and cell patterning and growth. Cluster 12 contains genes involved in terpenoid synthesis—including cytochrome P450s and terpene synthases.

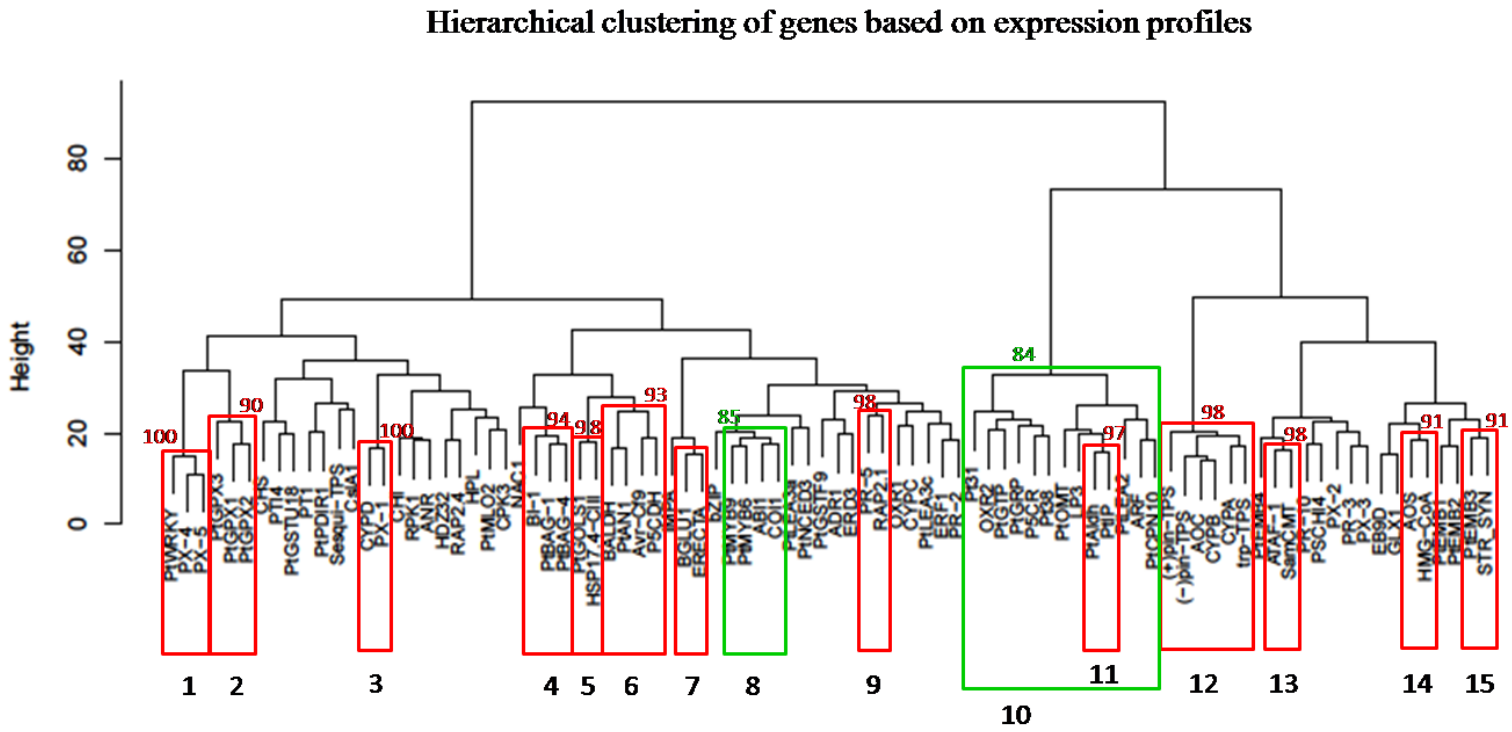

Fig. 2.4 Dendrogram produced by hierarchical clustering of gene expression. Genes clustered based on gene expression values reveal weak clustering of gene family members and functionally similar genes. Clusters outlined in red are supported by significant bootstrapping values (AU=90). Clusters outlined in green have high (but not significant) bootstrapping values. These clusters are supported by previous experimental evidence and using the web-based tool Aranet (Lee et al. 2010).

Cluster number	Genes	Putative function(s)
1	PtWRKY, PX-4, PX-5	Disease-response transcription factor; Stress response through oxygen species
2	PtGPX3, PtGPX1, PtGPX2	Oxygen radical scavenging
3	CYPD, PX-1	Cytochrome P450; Stress response through oxygen species
4	BI-1, PtBAG-1, PtBAG-4	Negative regulators of programmed cell death
5	PtGOLS1, HSP17.4-CIII	Stress-responsive target genes of heat shock transcription factor HsfA2 (*A. thaliana)*
6	BALDH, PtAN1, Avr-Cf9, P5CDH	Salt and dehydration responsive genes and transcription factor; osmoprotectants; disease resistance protein
7	BGLU1, ERECTA	Cell wall and lignin synthesis; cell growth and patterning
8	PtMYB9, PtMYB6, ABI1, COI1	Drought responsive transcription factors and genes
9	PR-5, RAP2.1	Thaumatin-like protein; DREB2 family transcription factor
10, 11	Pt31, OXR2, PtGTP, PtGRP, P5CR, Pt38, PtOMT, LP3, PtAldh, PtIP, PtLEA2, ARF, CPN10	Loblolly genes previously shown to be differentially expressed in response to drought;
12	(+) pin-TPS, , (-) pin-TPS, AOC, CYPB, CYPA, trp-TPS	Enzymes in terpenoid synthesis
13	ATAF-1, SamCMT	Jasmonate responsive genes
14	AOS, HMGCoA	MeJa synthesis; Jasmonate-responsive gene
15	PtEMB3, STR_SYN	LEA protein; terpene synthase

Fig. 2.4 continued.

Association genetics analyses revealed significant associations between gene expression and SNPs in candidate genes.

One hundred one significant associations ($q<0.05$) between SNPs and gene expression phenotypes were discovered using a general linear model. Of the 88 stress-related genes included for association analysis, 27 formed significant associations with SNPs in 94 putative candidate genes (Table 2.2). SNP amplicon sequences obtained from the Diversitree database (ADEPT2 2008) were blasted against loblolly pine ESTs in the NCBI EST database and contigs were assembled with Sequencher 4.2. These sequences were translated using the online Expasy Translate tool (Gasteiger et al. 2003) and the putative function of each amino acid sequence was ascertained through a blastp search of the NCBI protein databases (Table 2.2). Ten of the contigs were not translated due either to short contig length or poor homology with any ESTs in the NCBI loblolly pine EST database. Ten of the SNPs appeared to fall in the 5' UTRs of the candidate genes, 30 SNPs appeared to fall in the 3' UTRs, 5 appeared to be intronic SNPs, and 45 SNPs were in exons. Of the 45 exonic SNPs, 25 SNPs were nonsynonymous mutations and 8 SNPs were found to be in highly conserved regions using blastp searches.

Table 2.2 SNPs detected in association with expression phenotypes. Significant associations (q< 0.05) detected between SNPs and gene expression phenotypes are listed by the stress-related gene that expression data was collected on. SNPs are identifiable by their Diversitree database SNP I.D. (ADEPT2 2008). The putative functions of SNP-containing candidate genes were determined through blastp searches of the non-redundant protein sequences in the NCBI data base. Highlighted associations denote that evidence for interactions between the homologs of the candidate gene and the gene that expression was collected for were detected in Arabidopsis using Aranet (Lee et al. 2010).

Gene	Diversitree SNP locus	Amino acid substitution	Transcript region with SNP	Candidate gene putative function
ADR1	2-4252-02-301	nc*	3' UTR	Acyl carrier/ acyl-ACP thioesterase
	2-7961-01-49	L**	exon	Protein kinase family protein
	0-18897-02-515	nc	intron	Inositol phosphatase
AOC	0-8085-01-77	nc	3' UTR	Unknown transcribed locus
Avr-Cf9	UMN-4647-02-182	nc	5' UTR	Phosphoglyceride transfer protein
	0-13868-01-538	A	exon	Endomembrane system protein
	0-1828-01-184	nt***	nt	Unknown transcribed locus
PtBAG-4	0-9831-01-165	nc	5' UTR	RAB GTPase activator
	UMN-6523-01-130	nc	3' UTR	Unknown transcribed locus
	CL4232Contig1-04-352	D→N	exon	Aldo/keto reductase family protein
	0-9847-01-336	nc	3' UTR	Unknown transcribed locus
CHS	2-1768-01-128	nc	3' UTR	Stearoyl-CoA 9-desaturase protein
CPK3	2-7852-01-525	nc	3' UTR	ATP binding/ATPase protein
	0-15826-01-690	nt	nt	Unknown transcribed locus
	UMN-2473-01-75	nc	5' UTR	Auxin-responsive family protein
PtCPN10	0-8795-01-334	nc	3' UTR	Unknown transcribed locus
EB9D	CL1083Contig1-09-66	nc	5' UTR	Carbohydrate, sugar binding protein
	2-6183-01-544	nc	intron	Protein kinase
	0-18745-02-148	nc	3' UTR	MAP kinase
	0-18745-02-476	nc	3' UTR	MAP kinase
	UMN-582-02-373	nt	nt	Unknown transcribed locus
	0-17990-01-536	S→A	exon	Unknown transcribed locus

Table 2.2 continued.

EB9D	2-3989-02-265	Q→R	exon	WRKY transcription factor
	CL2663Contig1-05-172	M→L	exon	Cinnamoyl-CoA reductase protein
	UMN-4361-01-81	K	exon	GRAS transcription factor
	CL131Contig1-03-173	nc	3' UTR	S-adenosylmethionine synthetase
	0-10729-02-220	nc	3' UTR	C-myc-binding protein
	CL2446Contig1-03-157	nc	3' UTR	Unknown protein
	0-10262-01-558	Q→H	exon	Acyl-CoA-binding protein
ERD3	0-5629-01-304	L	exon	3-beta hydroxysteroid dehydrogenase
	0-4344-01-218	nc	3' UTR	ATP binding/ATPase protein
OXR1	UMN-2818-01-81	S→T	exon	Ralf-like signal transducer
	CL1714Contig1-04-215	L→I	exon	Unknown protein
	0-14120-03-165	nt	nt	Unknown transcribed locus
	0-11781-01-254	I→V	exon	UDP-d-glucuronate 4-epimerase
CslA1	0-7652-01-333	P→A	exon	NOL1/NOP2/sun family protein
	0-12219-01-579	nt	nt	Unknown transcribed locus
	0-13278-02-207	nc	5' UTR	MYND-type zinc finger protein
	0-1169-01-71	T→K	exon	ATP binding/protein kinase
	2-2270-01-79	nc	3' UTR	hesB domain protein
	0-2317-01-98	nc	3' UTR	Unknown transcribed locus
HPL	0-8089-01-393	nc	5' UTR	flavonoid 3'-monooxygenase/ oxygen binding protein
PtOMT	CL1241Contig1-01-118	nt	nt	Unknown transcribed locus
	2-3444-01-348	T	exon	Short-chain dehydrogenase/ reductase (SDR) protein
	2-7344-02-112	nc	3' UTR	Mov34 family protein
PtEMB1	2-4644-02-361	nc	3' UTR	Auxin-induced transcription factor
	2-7856-02-438	nc	intron	Auxin-induced transcription factor
	2-4749-01-281	K	exon	Heat shock transcription factor

Table 2.2 continued.

PtEMB1	2-2273-02-467	K→R	exon	Plastid developmental protein
	2-4207-01-230	S	exon	Amidophosphoribosyltransferase
	2-5073-01-179	nc	3' UTR	ATPase, transmembrane transporter
	UMN-582-02-373	nt	nt	Unknown transcribed locus
PtEMB3	UMN-3238-01-230	A→T	exon	C3HC4-type RING zinc finger protein
	UMN-6365-02-387	nc	nt	Unknown transcribed locus
	CL3727Contig1-03-213	E→A	exon	MYB transcription factor
PtEMB4	2-4644-02-361	nc	3' UTR	Auxin-induced transcription factor
	2-7856-02-438	nc	intron	Auxin-induced transcription factor
PtMLO2	0-7098-01-474	M→I	exon	Unknown transcribed locus
bZIP	0-14943-01-375	nc	intron	Unknown transcribed locus
	2-3296-02-82	nc	3' UTR	Electron carrier, copper ion binding
	CL3490Contig1-04-93	L→R	exon	pfkB-type carbohydrate kinase
	0-5297-02-38	nc	3' UTR	phosphodiesterase, transcription factor
	0-6605-01-171	nc	5' UTR	Unknown protein
	2-5483-02-355	G	exon	Unknown transcribed locus
P5CR	CL1027Contig1-04-410	nc	3' UTR	MYB transcription factor
PtPDIR1	CL863Contig1-03-162	P	exon	Phenylalanine ammonia-lyase
	2-1621-01-364	T→I	exon	Unknown transcribed locus
	CL4336Contig1-01-180	G	exon	RNA binding protein
	0-15023-01-102	R→C	exon	Unknown transcribed locus
	2-1784-02-439	A	exon	Brassinosteroid-mediated protein kinase
PR-5	UMN-3979-02-55	E→N	exon	Transmembrane kinase
Pt38	0-873-02-72	nt		Unknown transcribed locus
PtLEA3a	CL909Contig1-04-120	D→E	exon	RNA-binding protein
PX-5	UMN-5299-01-191	C	exon	Importin-α receptor/protein transporter
	CL383Contig1-01-157	L	exon	Ubiquitin-protein ligase

Table 2.2 continued.

SamCMT	UMN-2913-01-584	S	exon	Endomembrane system binding protein
	CL866Contig1-01-60	nc	3' UTR	Dihydrolipoamide S-acetyltransferase
	0-9265-01-46	nc	3' UTR	DNA binding protein
	0-9847-01-336	V	exon	Unknown transcribed locus
ANR	0-642-01-111	N→D	exon	Isocitrate dehydrogenase
	0-16070-01-66	nc	3' UTR	Unknown protein
	2-2420-01-355	P	exon	Unknown protein
	0-11087-01-123	nc	3' UTR	Homeobox-leucine zipper protein
	0-1191-01-405	nt	nt	Unknown transcribed locus
	0-7745-01-176	V	exon	Unknown protein
	CL2125Contig1-04-84	K	exon	Lipase/calmodulin-binding protein
	UMN-2818-01-81	S→T	exon	Ralf-like signal transducer
	CL2272Contig1-02-119	R	exon	Peptidase
	0-4645-01-65	nc	5' UTR	Unknown transcribed locus
	2-3279-01-58	N→K	exon	Leucine-rich repeat family
	UMN-4783-01-396	K	exon	DNA binding protein
	UMN-4111-01-150	nc	3' UTR	Ribonucleotide reductase
	0-3275-01-378	nc	3' UTR	Unknown protein
	0-16446-01-609	D→V	exon	Histone binding protein
	CL192Contig1-03-4	nt	nt	Unknown transcribed locus
	UMN-1209-02-122	nc	5' UTR	Unknown transcribed locus
	2-220-01-65	stop→S	exon	Unknown protein
	2-7524-02-36	nc	3' UTR	Unknown transcribed locus
	0-6817-01-156	nc	5' UTR	Endoplasmic reticulm protein of unknown function
	2-5264-02-453	N→K	exon	NAD+ isocitrate dehydrogenase
PtWRKY	2-3296-02-82	nc	3' UTR	Electron carrier, copper ion binding

* SNP in non-coding region.

** Amino acid substitutions with only one amino acid listed were synonymous mutations that did not change the amino acid.

*** Putative candidate gene sequence could not be translated.

We sought to utilize a comparative genetics approach to support the relationships between the SNP-containing candidate genes and our expression phenotypes detected in the association analysis. The most similar Arabidopsis homolog for all of the genes involved in positive associations were identified through blastn and blastp searches. The Arabidopsis homologs were entered into the online program AraNet (Lee et al. 2010). Significantly associated genes that were predicted by Aranet to be in the same functional gene networks in *Arabidopsis* were recognized for their probable interactions. Functional associations among the homologs of the SNP-containing candidate genes and the stress-related genes selected for expression analysis in *Arabidopsis* were identified in seven cases and are highlighted in Table 2.2.

To examine the relationship of the SNP-containing candidate gene with the associated expression phenotype, the average gene expression was calculated based on the SNP genotype (Fig. 2.5). In most cases the rare homozygote had lower expression values for the associated gene. Also, in most cases the heterozygote had expression values similar to the common homozygote suggesting that a single wild type allele is dominant and compensates for reduced activity of the second allele. Allelic insufficiencies were observed in a few cases where a single allele did not compensate for the weaker allele and the heterozygote had an intermediate phenotype (Fig. 2.5).

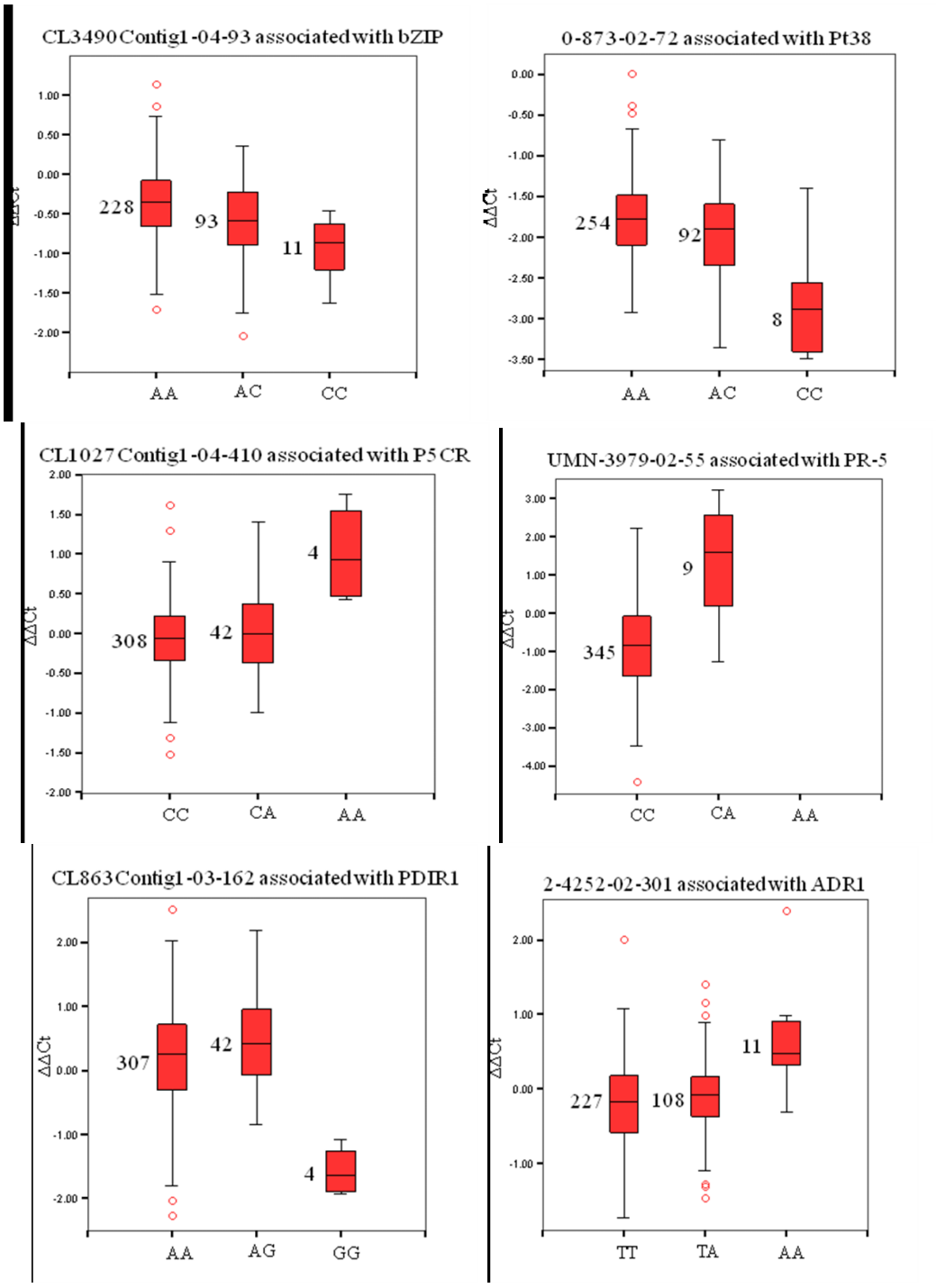

Fig. 2.5 Gene expression grouped by associated SNP genotype (identified by Diversitree SNP ID). The midline of each box plot represents the median expression for that SNP genotype and outlying expression values are represented by circles outside the whiskers. The number of individuals possessing each SNP genotype is also listed next to each box.

Model selection identified two subsets of genes that best predicted physiological measurements representing water use efficiency and resistance to pitch canker disease.

Subsets of genes predicting either carbon isotope discrimination (CID) (as a measure of water use efficiency) or pitch canker lesion length (as a measure of resistance to *F. circinatum* infection) in linear regression were identified using the package 'leaps' in R software. The 3 best models predicting each trait were identified based on the test statistic Mallow's C_p and were further assessed for fit by examining the literature. Genes were identified in two steps: twenty-seven genes were initially identified for inclusion in subset selection for predicting CID with the regsubsets function in 'leaps' (Table 2.3) and the leaps function was then used to identify the best model for predicting CID. According to the C_p statistic in the leaps output the model best predicting CID in linear regression included 7 genes (Table 2.3). All of the seven genes are known to be drought-responsive genes in either loblolly pine or other plant species. The model including 10 genes also seemed to be a good fit. Models including more than 10 genes had less significant C_p values and included genes not mentioned in the literature to be responsive to drought. The selected models seem to maximize the number of genes that have previously been shown to be involved in dehydration responses and have the most significant Mallow's C_p values.

The procedure was repeated for subset selection of genes predicting pitch canker lesion length. Regsubsets identified 31 genes for inclusion in subset selection and the three best models generated by leaps include 9, 10, and 11 genes, respectively (Table 2.4). Based on the C_p values the model containing 10 genes appears to be the best model. Genes in the three best models included both putative drought- and disease-responsive genes. PtCPN10, ERD3, PtGPX2, RAP2.4, and PtEMB1are drought-related genes that were originally identified in other organisms, and PtAldh and PtGRP were identified in drought-stress studies in pine. Expression

of drought-related genes may be correlated with pitch canker resistance by cross-talk between abiotic and biotic stress response pathways.

Table 2.3 Models predicting carbon isotope discrimination. The three best models for predicting carbon isotope discrimination data based on gene expression were identified using the package 'leaps' in R software. According to test statistic Mallow's C_p the model containing 7 genes best predicted CID. Models including 6 and 10 genes have similar C_p values and contain genes relevant to dehydration responses and are listed also.

	Genes predicting CID: 3 best models based on Mallow's C_p (C_p=)		
Genes identified with regsubsets	6 (C_p= -0.211)	**7 (C_p= -0.322)**	10 (C_p= -0.224)
ARF			
PtCPN10			
ERD3			X
ERECTA			
IMPA			X
PtOMT			
LP3			
P5CR			X
PtGRP		**X**	
PtGTP	X	**X**	X
PtLEA3a			
ABI1			
Avr-Cf9			
BALDH			
PtBAG-1			
PtBAG-4			
BI-1			
PtMYB9			
PtAN1			
PR-3	X	**X**	X
PtMYB6			
PX-1	X	**X**	X
PX-3			
PX-4	X	**X**	X
PX-5			
ATAF-1	X	**X**	X
CPK3	X		X

Table 2.4 Models predicting pitch canker lesion length. The three best models for predicting lesion length due to infection with *Fusarium circinatum* based on gene expression were identified using the package 'leaps' in R software. According to test statistic Mallow's C_p the model containing 10 genes best predicted lesion length. Models including 9 and 11 genes have similar C_p values and are listed also.

	Genes predicting pitch canker lesion length: 3 best models based on Mallow's C_p (C_p=)		
Genes identified with regsubsets	9 (C_p= -0.599)	**10 (C_p= -0.934)**	11 (C_p= -0.405)
PtCPN10	X	**X**	X
ERD3	X	**X**	X
PtGOLS1			
IMPA			
PtALDH	X	**X**	X
PtGRP	X	**X**	X
Av9-Cf9			
PtBAG-1			
BI-1	X	**X**	X
CHI	X	**X**	X
ERF1			
PR-2	X	**X**	X
PR-3			
PR-5			
CHS			
CYPA			
CYPD			
PtGPX1			
PtGPX2	X	**X**	X
PtMYB6			
PtPDIR1			
PTI4			
RAP2.4			X
trp-TPS			
PtWRKY			
AOS			
ATAF-1			
EB9D			
HPL		**X**	X
PtEMB3			
PtEMB1	X	**X**	X

DISCUSSION

Drought resistance and disease resistance are two important traits in loblolly pine. Several studies have shown the genetic and adaptive basis of these traits (Morgenstern 1996; González-Martínez et al. 2007; Baltunis et al. 2008). Few studies have examined the extent of variation in the expression of disease- and drought-related genes or its adaptive significance in loblolly pine trees. Several studies in other organisms have demonstrated that gene expression differences may serve as a substrate for natural selection driving evolution (King and Wilson 1975; Fay et al. 2004; Stranger et al. 2007). Population-wide gene expression studies could provide valuable insights into the molecular mechanisms of adaptations for disease and drought resistances and opportunities for discovery of novel alleles for improving these two economically important traits.

In this study, statistically significant variation in the expression of all 88 disease- and drought-responsive genes was detected among 354 loblolly pine trees representing the species' natural range in the southeastern U.S. The average range of expression differences for a gene observed in this study was larger than the gene expression differences in xylem development genes observed by Palle et al. (2010). Examination of gene expression in *S. cerevisiae* also found the expression of stress-related genes to be more variable than that of growth genes (López-Maury et al. 2008). Additional variation in expression may be a function of specific regulatory elements frequently associated with stress-related genes and could act as an adaptive strategy for surviving a wide variety of environmental disturbances (López-Maury et al. 2008). Most of the gene expression profiles were normally distributed with a majority of the individuals having expression values laying 2-3 fold from the mean. On average, 20% of individuals had extreme expression values that varied by more than 3-fold from the mean expression of a gene. Similar gene expression differences among individuals have been observed in other species (Oleksiak et

al. 2002) and in those cases variation was contributed mainly to genetic drift or random changes that occur in finite populations. The largest range in expression phenotypes for a single gene was remarkable—more than 9,000 fold between the highest and lowest expressing clones. Such extreme expression phenotypes suggest that a mutation in a cis- or trans- regulatory site may be significantly affecting the expression of that gene.

Staubach et al. (2009) found strong correlations between gene expression variation and population divergence suggesting that population substructure could account for some variation in gene expression. To identify the presence of substructure within our population based on gene expression variation, principal component analysis was performed. Principal component analysis detected three geographically-based principal components that accounted for 24% of the cumulative variation in gene expression among loblolly individuals (Fig. 2.2). These results correspond to the results detected by Eckert et al. (2010) using PCA analysis of SNP markers. Eckert et al. (2010) found, using both SNP and SSR markers, that population substructure across the range of loblolly pine was largely explained by the Mississippi river discontinuity with a weaker division east of the Mississippi grouping trees into an "Atlantic Coast cluster" and "Gulf cluster" (Eckert et al. 2010). While there is still a considerable amount of variation in gene expression not explained in these components, the clear agreement between our results and with those arrived at by Eckert et al. (2010) follow the correlation between population differentiation and expression polymorphisms reported by Staubach et al. (2009).

Using cluster analyses to detect groups of genes that are similarly expressed between loblolly individuals can provide information about their biological functions and their regulation (Eisen et al. 1998; Slonim 2002; Wang et al. 2004). Hierarchical clustering of 88 disease- and drought-related genes assembled thirty-five of the genes into 13 groups with significant AU boostrapping values (red boxes in Fig. 2.3). Two more groups of genes with nearly significant

AU values were also noted (green boxes in Fig. 2.3) in view of the fact that the genes in these clusters were previously shown to be involved in the same functional pathways and/or have similar expression patterns. The results of the clustering analysis revealed that only small clusters of genes were supported by significant AU values, and that members of gene families or functionally similar genes did not cluster together as consistently as expected. This could be explained by a variety of reasons: 1) Clustering of expression profiles assumes that genes are regulated at the transcriptional level. Genes regulated by post-transcriptional modifications may disrupt clustering of functionally similar genes (Adryan and Schuh 2004). 2) Genes analyzed for expression were almost entirely stress-responsive genes and expression was collected in unstressed tissues. The constitutive expression of these genes is likely more variable than the induced expression. Expression of gene family members and functionally similar genes transcriptionally regulated by a common signal in response to stress will likely be more apparent in stressed tissues. 3) Many of the genes analyzed were homologs of genes identified in other organisms and their involvement in stress responses in loblolly pine have not been confirmed. It may be that these genes have alternate roles in loblolly pine. Although, the first two explanations seem more likely as putative gene sequences were usually compiled from loblolly pine EST libraries that had been subjected to stress and the association analysis showed that the expression of several of the genes were frequently correlated with SNPs in other putative stress-related genes.

QTL mapping is a favored approach for studying complex traits. However, obstacles for using linkage mapping in natural populations exist including: 1) limited genetic tools, 2) absence of a mapping population, 3) limited linkage between marker alleles and quantitative trait loci (QTL) (Strauss et al. 1992). Association genetics is a powerful tool for linking natural phenotypic variation with QTL in natural populations where QTL mapping is inappropriate.

Loblolly pine has a short history of domestication preserving much of its genetic diversity and has the genetic resources required for association genetics studies (Neale and Savolainen 2004). In this study we have utilized a candidate gene association genetics approach to link allelic variation of candidate genes with gene expression phenotypes.

One hundred one significant associations between gene expression phenotypes and SNPs in candidate genes were detected. SNPs in candidate genes can alter gene expression phenotypes in a variety of ways. While associations between nonsynonymous mutations that result in amino acid changes in the exons of candidate genes and expression phenotypes are more easily explained, these were not the majority of SNPs detected. The majority of the SNPs were in noncoding regions of the transcripts. Such SNPs may affect elements regulating transcription or translation, RNA stability, or by influencing splicing. A surprisingly large number of SNPs were found in the 3' UTR in comparison with the other non-coding regions. Translational regulatory elements are abundant in eukaryotic 5' and 3' UTRs (Wilkie et al. 2003). SNPs in regulatory elements can affect or even abolish translation (Wilkie et al. 2003). A family of 3'UTR-binding miRNAs has also recently been found in the plant kingdom (Arteaga-Vazquez et al. 2006). MiRNAs play a central role in regulating physiology from growth to stress responses across kingdoms (Arteaga-Vazquez et al. 2006). In loblolly pine, repressed expression of miRNAs and their disease-related targets has been discovered in association with infections by the rust pathogen *Cronatrium quercuum* (Lu S et al. 2007). Several of the genes predicted to be targets of loblolly pine miRNAs by Lu S et al. (2007) had the same putative functions as the candidate genes associated with gene expression identified in this study. For example, a mov34 protein, several MYB and auxin-induced transcription factors, a flavonoid 3'-monooxygenase/ oxygen binding protein, and a leucine-rich repeat family protein were common in both studies (Lu S et al. 2007). SNPs were commonly located in UTR regions in these genes. SNPs in introns may

similarly alter intronic regulatory elements and or may alter intronic sequences important for splice site recognition. Inaccurate splicing can affect mRNA stability, transport, or size of the mature mRNA (Brown and Simpson 1998). In any of these cases reduced translation of transcriptional regulatory proteins, such as protein kinases, phosphatases, transcription factors, heat shock proteins, can affect downstream gene expression.

The gene that formed the most associations was an anthocyanidin reductase (ANR) similar to the Arabidopsis BANYULS gene, a negative regulator of the branch of the flavonoid pathway leading to catechin biosynthesis (Xie et al., 2003). ANR formed associations with SNPs in 21 candidate genes (Table 2.2). The large number of associations for this particular gene is not surprising as the flavonoid biosynthetic pathway is highly complex and is involved in a variety of physiological processes in plants. The gene with the largest range in expression differences was a late embryogenesis abundant protein (LEA), PtEMB4, which formed associations with SNPs in two auxin-induced (IAA) transcription factors (Table 2.2). The same individual was homozygous for the rare allele in both IAA transcription factors. Expression of PtEMB4 in the rare homozygote was 9,000-fold less than in the common homozygotes or heterozygotes. Heterozygotes had expression values similar to the common homozygous genotype suggesting a dominant relationship of the common allele to the rare allele. Unfortunately, because only one rare homozygous individual was present in our population for expression analysis interpretation of these results are problematic. Expression of PtEMB4 needs to be confirmed in additional individuals. If confirmed, this genotype could be valuable for dissecting biochemical pathways and physiological traits that the IAA transcription factors contribute to in loblolly pine. Many genetic approaches that are common in many model organisms (such as knock-outs or transgenic lines) are difficult to manage in loblolly pine. This genotype could be economically important as

auxin is a key signaling hormone in traits such as plant growth and development, and stress responses.

The web-based tool Aranet (Lee et al. 2010) was used in a comparative genetics approach to identify the putative functions of candidate genes and to identify genes connected in functional networks that would support the positive associations that were detected. Aranet is a genome-wide functional gene network for *Arabidopsis* based on an extensive collection of experimental data and statistically associated candidate genes (Lee et al. 2010). Lee et al. (2010) showed that Aranet was capable of incorporating data from other organisms and correctly predicting the processes they were involved in. This was especially true for abiotic and biotic stress responses. Aranet supported functional connections between associated genes in only 7 cases. This low success rate was not unexpected due to 1) the large evolutionary distance between *Arabidopsis* and *Pinus*, and 2) the occurrence of associations between gene expression and SNPs where the SNP was linked with the causal SNP in a nearby gene.

Aranet did provide useful insights for developing a model for interactions among associated genes in at least one example though: Among the candidate genes associated with PtEMB1 were two Aux/IAA transcription factors. One of the loblolly IAA transcription factors, IAA2, has previously been characterized by Goldfarb et al. (2003) and it is homologous to *Arabidopsis* IAA7, an auxin-responsive transcription factor involved in plant growth and development and in stress response (Song et al., 2009). It appears that Aux/IAA transcription factors regulate auxin responses by dimerizing with auxin-responsive transcription factors (ARFs) and preventing them from activating transcription of downstream auxin-responsive genes (Gray et al. 2001; Tiwari et al. 2001; Tiwari et al. 2004). As auxin increases in the plant Aux/IAA-ARF heterodimers dissociate and Aux/IAAs are targeted by SCF^{TIR1} (Skp1-Cullin-F-box E3 ligase) ubiquitination and 26s degradation freeing ARFs to initiate transcription (Tiwari

et al. 2001; Tiwari et al. 2004). Further experimentation revealed a corepressor required for full Aux/IAA repression—a leucine-rich EAR (ethylene response factor-associated amphiphilic repression) motif that interacts with TPL, a transcriptional repressor (Tiwari et al. 2004; Szemenyei et al. 2008; Long et al. 2006).

PtEMB1 was also associated with a heat shock transcription factor, an amidophosphoribosyltransferase, and an ATPase/transmembrane transferase protein that were found to be connected in functional networks using Aranet. Further review of the literature showed that the heat shock transcription factor, AtHSFB4, also contained an EAR repressor domain. There was no report in the literature or evidence from Aranet that Aux/IAA interacted with any of the other candidate genes associated with PtEMB1. A reasonable hypothesis explaining the connection among these four genes may be that TPL represses both IAA2 and the heat shock transcription factor (and subsequently its downstream genes) detected in the association analysis. If this were the case, a SNP in TPL in linkage disequilibrium with these candidate genes, or a SNP in a linked gene at a nearby locus could result in the indirect association of SNPs detected in these genes.

Carbon isotope discrimination (CID) is a favored measure of water use efficiency in loblolly pines because it is heritable, replicable, and can be evaluated in immature trees (González-Martínez et al. 2008). The regsubsets function provided 27 genes to include in a further exhaustive search of all possible models including these genes by the leaps function. The putative orthologs of all seven genes in linear regression with the variations in CID have been reported in the literature to be involved in dehydration responses. PR-3 is a chitinase typically considered a pathogen response gene but is also responsive to abiotic stresses (Pihakaski-Maunsbach et al. 2001; Loopstra and Sathyan 2004; Seo et al. 2008). The model containing 8 genes also seemed to be a good fit. It includes 3 additional genes—an early response to

dehydration gene (ERD3), a proline biosynthesis gene (P5CR), and, an importin α gene (IMPA). IMPA is part of a heterodimer that shuttles proteins between the nucleus and cytoplasm and accumulates in the nucleus in response to a variety of stresses (Miyamoto et al. 2004).

Pitch canker lesion length is an easy quantitative measure of *F. circinatum* resistance that has been used successfully in previous genetic studies (Quesada et al. 2010). The leaps package regsubsets function returned 31 genes that were included for model selection with the leaps function to select the three best models predicting pitch canker lesion length. Based on C_p values, the model including 10 genes fit variations in pitch canker lesion length best. The three best models all contained a combination of drought- and disease-responsive genes. This may indicate that cross talk between abiotic and biotic stress signaling pathways is plays a role in disease resistance. Significant cross talk among pathways has been shown (Knight and Knight 2001; Desprez-Loustau et al. 2006). In some cases it appears that environmental disturbances can "prime" plants for biotic stresses (Desprez-Loustau et al. 2006). Differences in the constitutive expression of drought-response genes may manifest in differences in disease resistance through cross talk among signaling pathways.

CONCLUSIONS

Expression studies revealed a substantial amount of variation in the expression of disease and drought responsive genes across the native range of loblolly pine. It seems that some of the gene expression variation can be accounted for by substructure within the population similar to that detected by Eckert et al. (2010) using genotypic markers. The idea that patterns of gene expression correlate with genotypic variation and may imply evolutionary patterns is not fully supported. Our results suggest that there is at least a weak correlation. Functionally similar

genes were only weakly correlated. To confirm that gene function is conserved in loblolly pine future studies should concentrate on characterizing expression patterns in response to stresses.

Association genetics approaches linking SNPs with gene expression phenotypes identified 101 SNPs in candidate genes. Future studies verifying these associations are required to confirm associations and to examine the relationship between genetic polymorphisms, gene expression variation, and differences in disease- and drought-resistant physiological traits.

Regarding future association genetics studies in loblolly pine, a larger population would increase power. In several cases, the rare homozygote represented less than 10% of the population, and in a few cases only one rare homozygote individual was in the population. When examining stress-responsive pathways, expression analysis of stressed tissues would clarify results and may produce more significant associations. Finally, a larger coverage of the transcriptome and the availability of a complete genome sequence for increasing coverage of the genome with SNP markers would yield many more associations and a more holistic model building approach including interactions between genes to dissect interactions of genes and whole biochemical pathways that control important traits.

CHAPTER III

THE EFFECTS OF CYCLIC DROUGHT AND PITCH CANKER EXPOSURE ON THE EXPRESSION OF ABIOTIC AND BIOTIC STRESS-RELATED GENES IN LOBLOLLY PINE (*Pinus taeda* L.)

INTRODUCTION

The forests of the southern United States are important both ecologically and economically. The south has become the largest lumber- and pulpwood-producing region and has been largely influenced by plantings of genetically improved seedlings (Howard 2005). One of the predominant species in plantations and managed stands of southern forests is loblolly pine (*Pinus taeda* L.). Loblolly pine makes up a large percentage of the softwood timber and pulpwood sales (Hanson et al. 2010). Loblolly pine forest communities dominate one-quarter (54.8 million acres) of southern forests providing immeasurable services such as habitat for wildlife, carbon sequestration, and recreational areas (Hanson et al. 2010). Demand for timber products are predicted to remain steady or increase (Howard 2005) while suburban and agricultural encroachment, climate change, and pathogens threaten to decrease the range and health of southern forests (Hanson et al. 2010). Ideally, managed stands of genetically improved trees would provide enough gains to keep up with increases in the consumption of wood products and leave areas of natural forests for conservation and other purposes (Hanson et al. 2010). Two of the most important factors affecting southern forests are diseases and water availability (Hanson et al. 2010). Tree improvement cooperatives have already made some gains in these traits through breeding (Schultz 1997). However, very little is known about the underlying genetic mechanisms controlling these traits in loblolly pine.

Plants possess both constitutive and induced defenses to cope with environmental disturbances. Constitutive defenses are physical barriers including the plant cuticle and cell wall. Pines also constitutively produce an array of terpenes and other phenolic compounds that act as chemical barriers to pathogens (Ralph et al. 2006b). Induced defenses are initiated by signal transduction cascades that activate gene expression to further elicit physiological changes to avoid, tolerate, or resist the stress conditions (Shinozaki et al. 1997; Chaves et al. 2003; López-Maury et al. 2008). Gene expression changes are the first regulatory step in modifying phenotypic traits. Transcriptome studies have shown that increases in the expression of stress-responsive genes are balanced by decreased expression of growth genes and it appears that cross-talk among signal transduction pathways moderate responses when faced with multiple disturbances to maintain homeostatic conditions (Watkinson et al. 2003; López-Maury et al. 2008). Although expression is regulated at multiple levels, several studies have found significant variation in gene expression among individuals within and among populations (Oleksiak et al. 2002; Storey et al. 2002; López-Maury et al. 2008). It has been suggested that variations in gene expression could provide substrate for natural selection and may have a significant role in driving adaptations to environmental conditions (King and Wilson 1975; López-Maury et al. 2008). Studying disease- and drought-responsive gene expression profiles could shed light on the molecular mechanisms of adaptations to these two environmental disturbances, and on the trade-offs associated with selecting for stress-resistant traits.

The signal transduction pathways induced in response to disease and drought are very complex. Drought responsive genes appear to be under the control of at least five major pathways: two ABA-independent pathways, one of which appears to interact negatively with ethylene signaling (Sharp and LeNoble 2002; Chaves et al. 2003), two ABA-dependent pathways, and a third ABA-dependent pathway that appears to interact with jasmonate signaling,

a hormone classically identified as a pathogen- or wound-responsive hormone (Shinozaki et al. 2007). Additional pathways appear to be initiated by other secondary signaling molecules such as reactive oxygen species (ROS), sugars, and nitric oxide (Chaves et al. 2003).

Responses to pathogens may be regarded as even more complex involving multiple signaling pathways and specific gene-for-gene interactions dependent on the host and pathogen genotypes (Yang et al. 1997). Disease responsive genes are induced by hormones salicylic acid, ethylene, and jasmonates with evidence for cross talk among all three pathways (Yang et al. 1997; Kunkel and Brooks 2002). As mentioned previously, ABA signaling also interacts with jasmonate and ethylene signaling pathways by repressing the expression of defense-responsive genes in these pathways (Fan et al. 2009). Similarly to drought-responses, other secondary signaling molecules (ROS, lipids, sugars, Ca^{2+}) induce expression in response to pathogen infections (Yang et al. 1997; Shah 2005). Most of the studies examining disease resistance in forest trees have focused on terpenoid accumulation and/or terpenoid volatile emissions leaving much to be learned by examining other pathogen-induced signaling pathways (Ralph et al. 2006b).

Crosstalk among conserved signaling pathways induced by different biotic and abiotic stresses may be inversely regulated or they may be jointly regulated with one disturbance "priming" the plant for tolerance to others (Knight and Knight 2001; Pastori and Foyer 2002; Desprez-Loustau et al. 2006). Key convergence points likely consist of regulatory molecules such as mitogen-activated protein kinases (MAPKs), cyclin-dependent kinases (CDPKs), and WRKY, NAC, and ERF transcription factors (Fujita et al. 2009) and secondary signaling molecules such as Ca^{2+}, ROS, NO that are common to both abiotic and biotic signaling (Fujita et al. 2009). In a survey of previous literature, Desprez-Loustau et al. (2006) reported a general trend for increased disease incidence with increasing water stress, but found that the type, length,

and severity of the infections influenced whether drought was either positively and negatively associated with disease incidence.

The fungal pathogen *Fusarium circinatum* is responsible for causing pitch canker disease on loblolly pine trees almost worldwide. The first recorded observation in the southeastern United States was made in 1946 (Hepting and Roth 1946). *F. circinatum* infections in loblolly pine are identifiable by reddening of the infected stem and the appearance of resin-soaked cankers. Girdling of lateral stems by the pathogen resulting in death of the tissue from the infection to the branch tip is common. Infections of the main stem can result in top-kill as well, depending on the severity of the infection (Barnard and Blakeslee 1987; Wingfield et al. 2008). Outbreaks of pitch canker infections of loblolly pine in managed stands and seed orchards can result in severe economic losses (Barnard and Blakeslee 1987; Wingfield et al. 2008). Hammerbacher et al. (2009) showed that abiotic stresses significantly reduce resistance of pine seedlings to pitch canker infections. This coincides with previous observations in the field that pitch canker outbreaks increase in occurrence in association with periods of drought, during the late summer and fall when water availability is limited, and in association with early frosts (Schmidt et al. 1976; Dwinell et al. 1985; Barnard and Blakeslee 1987).

Variation in drought tolerance and genetic resistance to *Fusarium circinatum* infections have been demonstrated among loblolly pine individuals (Bilan et al. 1977; Kayihan et al. 2005; González-Martínez et al. 2008; Quesada et al. 2010). Responses to both drought stress and pitch canker infections in loblolly pine are controlled by multiple genes (Morse et al. 2004; Kayihan et al. 2005; Quesada et al. 2010). Genes responsive to drought or pitch canker infections have been identified in loblolly pine (Chang et al. 1996; Watkinson et al. 2003; Morse et al. 2004; Lorenz et al. 2006; González-Martínez et al. 2008; Quesada et al. 2010). However, disease and drought resistance traits have primarily been examined in model organisms and important crop species

(Ingram and Bartels 1996). Using comparative and association genetics approaches, 114 putative drought- and disease-related genes have been identified in loblolly pine. Expression of these genes was measured in 24 loblolly pine trees representing a range of disease- and drought-resistant phenotypes exposed to cyclic drought stress and/or infection with *F. circinatum* inoculations. Expression was examined to confirm their responsiveness to abiotic and/or biotic stresses, to identify significantly different gene expression responses among the different phenotypes, and to examine the effects of multiple stresses on gene expression induction. The results of this study may help to identify differentially expressed genes in loblolly pine that contribute to disease resistance, drought resistance, or cross-tolerance to both stresses.

MATERIALS AND METHODS

Plant material: Twenty-four unrelated loblolly individuals (clones) belonging to the NCSU association population were selected from the tails of carbon isotope ratios (as a measure of water use efficiency with the most water use efficient individuals having the lowest carbon isotope values, and vice versa) collected at North Carolina State University (NCSU; Patrick Cumbie, unpublished) and from clones shown to be either highly susceptible or resistant to the pathogen *Fusarium circinatum* in experiments at the University of Florida (UF; Quesada et al. 2010). At least 16 biological replicates (ramets) of each clone were provided as rooted cuttings from hedges maintained at North Carolina State University. The rooted cuttings were planted in fritted clay and grown in green houses in a completely randomized design for almost five months. During this period all trees were uniformly well-watered and fertilized. Each ramet was assigned to one of four treatment groups—cyclic drought, *F. circinatum* inoculation, drought + *F. circinatum*, or control—so that 4 ramets per clone were subjected to each treatment. At the time of harvest, all physiological measurements and tissue collections were performed before

dawn. The top 6 inches of the main tree stem was collected from each plant for gene expression analyses and fresh, green needles were collected for extractions of epicuticular waxes. Tissues were stored in 15-ml tubes, flash frozen in liquid N_2, and placed in -80°C for long-term storage.

Fusarium circinatum *inoculations:* *F. circinatum* isolate S45 (provided by Dr. John Davis's lab, UF) was maintained on clarified V8 agar media (50 ml clarified V8 broth containing 1.63 g. dissolved $CaCO_3$, 15 g. agar per 1 L deionized water). *F. circinatum* isolates were transferred to liquid clarified V8 media and grown to approximately 450 spores/µl. Four ramets per clone were inoculated with the *F. circinatum* using a variation of the methods described in Davis et al. (2002). The main shoot tip from each tree was clipped and 2 µl of the inoculum was applied to the wound using a pipettor. Additionally, a small subset of clones was inoculated in the same matter with water to survey gene expression changes that may be occurring in response to wounding during the inoculation procedure rather than the actual pathogen infection. Each inoculated shoot tip was covered with a cotton ball and wrapped with parafilm to maintain a humid environment around the wound. Twenty-four hours after inoculations, the cotton balls and parafilm were removed. Trees remained in the greenhouse for approximately 4 weeks before harvesting.

Cyclic drought treatments: Four ramets per clone were subjected to cyclic drought treatments by withholding water until the predawn water potentials of four randomly selected trees in this treatment group all reach at least -1.5 MPa where upon the trees were well-watered for the next two days. This process was repeated for 4 weeks. At the end of the treatment period the main stem was collected from each tree having a water potential of at least -1.2 MPa. Trees that had not reached this criterion were replaced in the greenhouse and harvested later that week and the following week.

F. circinatum + *drought treatments*: Four ramets per clone were exposed to drought stress treatments and pitch canker inoculations simultaneously. These trees were first inoculated with *F. circinatum* as previously described and watered for two days post-inoculation. At this time, drought stress cycles were imposed as described above.

Control treatments: A control group of four ramets per clone remained well-watered and fertilized throughout the treatment period.

Physiological measurements: As was briefly mentioned, carbon isotope data and pitch canker lesion length data was previously measured at NCSU (Patrick Cumbie, unpublished) and UFL (Tania Quesada, Quesada et al. 2010) on biological replicates of the same clones included for expression analysis. Predawn water potentials were monitored by randomly selecting four individuals within the drought and drought + *F. circinatum* treatments for measurements throughout the experiment. At the time of harvest, predawn water potentials were recorded for all individuals as well as the tree circumference at pot height. Pitch canker symptoms such as reddening of the stem or the presence of cankers were noted and measured on the inoculated trees.

Epicuticular wax was extracted by methods adapted from Lütz et al. (1990). Approximately 2 grams of frozen needles were weighed and transferred to a pre-weighed glass culture tube. Fifty-five milliliters of pre-heated chloroform was poured into each tube and waxes were extracted for 15 minutes in a 60˚C water bath. The chloroform with dissolved waxes was decanted into a pre-weighed beaker and evaporated in a fume hood until dry. The beakers were subsequently weighed to determine the net weight of waxes extracted. Needles were also dried for 24 hours at 70˚C and weighed in order to calculate wax per unit dry needle weight. This unit was used for all statistical analyses.

RNA extractions and cDNA synthesis: The main tree stem collected from each plant was ground to a fine powder under liquid N_2 in preparation for RNA extractions. Total RNA was extracted according to Chang et al. (1993) with the addition of an extra chloroform extraction to fully eliminate protein contamination. RNA was treated with DNA-free™ (Ambion, Austin, TX) to remove contaminating DNA. The Nanodrop 1000™ was used to quantify and to estimate the purity of each RNA sample. Any sample suspected of containing contaminating DNA and protein were subjected to a final DNAse or chloroform treatment, respectively. All samples had OD_{260}/OD_{280} ratios between 1.8 and 2.1. First strand cDNA was primed with random hexamer primers from 5µg RNA using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, CA) and amplified according to the kit's protocol.

Gene selection and primer design: Eighty-eight of the genes included in this expression analysis were selected from the literature and their expression was previously measured in unstressed xylem tissue in the association genetics study of disease- and drought-responsive genes described in Chapter II. Ten genes containing SNPs that were significantly associated with the expression of stress-related genes in Chapter II were also included for expression analysis. The remaining 16 genes were selected based on evidence of their involvement in disease and/or drought responses in the literature. Ten genes involved in cuticular wax biosynthesis responsive to water deficiency were also selected from Kosma et al. (2009) and their expression was measured only in the control and drought treatment groups. Table A.1 in Appendix A lists the putative function, *P. taeda* contig ID, primer sequence, and study that the gene was originally identified in.

To design primers for the genes identified in the previous association genetics study and for genes later selected from the literature, contig sequences were assembled from loblolly ESTs identified by blast searches of the NCBI Expressed Sequence Tag database (dbEST) using

Sequencher 4.2 (Gene Codes Corporation, Ann Arbor, MI). Contig sequences were blasted again against loblolly pine sequences available in NCBI to ensure that these genes were not other genes with high sequence similarity to prevent potential amplification of multiple gene products. Gene-specific primers were designed with Primer Express 3.0 software (Applied Biosystems, Foster City, CA). Parameters included: 1) primer T_m: 58-62°C; 2) primer %GC Content: 40-60%; 3) primer length: 18-22 nt; 4) amplicon length: ~75mer. Areas of the contig sequences that appeared to contain SNPs based on alignment of the ESTs were avoided for primer design. Primers were supplied by Invitrogen (Carlsbad, CA).

Quantitative real-time PCR: Real-time quantitative PCR (RT-qPCR) reactions using SYBR® green dye (Applied Biosystems, Foster City, CA) for quantification were set up in 384-well plates and run on an ABI 7900 HT (Applied Biosystems, Foster City, CA). Reaction volumes were decreased to 8 µl but were otherwise carried out according to the manufacturer's protocol. In all four treatment groups, expression was collected on three ramets by two technical replicates for all twenty-four clones. No template and no reverse transcriptase reactions were used as negative controls to identify contamination between PCR reagents and cDNA samples. Raw expression data collected with SDS 2.3 software (Applied Biosystems, Foster City, CA) were loaded into RQ Manager 1.2 (Applied Biosystems, Foster City, CA) where relative gene expression ($\Delta\Delta C_T$) was calculated by subtracting the ΔC_T (Endogenous control (β-actin) gene expression -Target gene expression) from the ΔC_T value of a normalizer sample. Amplification of a single gene was confirmed by dissociation curve analysis.

Amplification efficiency: The primer binding sites of the 88 genes analyzed in the association genetics study were previously sequenced to detect SNPs that could decrease the amplification efficiency (AE). For the remaining genes, the amplification efficiencies for every gene were compared among the highest and lowest expressing clones using a one-way ANOVA

included in the "Data Analysis for Real-Time PCR" (DART-PCR) Excel worksheet (Pierson et al. 2003). Genes that appeared to exhibit extreme expression differences among individuals or between treatments were also selected for sequencing of their primer binding sites to confirm that sample-to-sample variations in AE was not increasing expression differences.

Gene expression analysis: Delta-delta C_T values were manually examined for cases where the expression values varied by more than 0.7 cycles between technical replicates. In cases where the expression of one of the replicates was clearly an outlier in comparison with the other two ramets, it was removed from the analysis. If the expression among ramets was not similar enough to identify an outlying technical replicate value, the ramet was removed and further analyses on the expression values for that clone were based on just two ramets. To reposition the mean expression value and standard deviation of each gene within each treatment to 0 and 1, respectively, the mean $\Delta\Delta C_T$ value was calculated for every gene. This mean gene expression value was subtracted from the expression of each individual and then divided by the gene's standard deviation (Stahlberg et al. 2008). This data was used for all subsequent statistical analyses.

Data analysis: ANOVA and Welch's ANOVA for genes exhibiting unequal variances (as determined by Levene's test) were performed in SPSS to test for 1) significant variations in gene expression among clones within a treatment, for 2) significant variations in gene expression between treatments, and for 3) significant variations in stem circumference and water potential measurements taken at harvest among treatments. ANOVA and Welch's ANOVA were also used to test for 1) significant differences in gene expression between individuals classified as either "resistant" or "susceptible" to pitch canker infections based on pitch canker lesion lengths and between "water use efficient" and "water use inefficient" individuals based on measure of carbon isotope discrimination, and for 2) significant differences in circumference or water

potentials measured at harvest among "resistant" or "susceptible" disease and drought phenotypes. Fisher's least significant difference (LSD) or Dunnett's T3 post hoc tests (for genes with unequal variances) were performed to identify which group was significantly different in the ANOVA tests. Partial correlation analysis was performed to identify significant relationships between expressed genes while controlling for different treatments. Spearman rank correlations were performed to identify correlations between gene expression profiles collected on trees in the control group and carbon isotope values (Patrick Cumbie, unpublished), pitch canker lesion length (Tania Quesada, unpublished), and wax content. Partial correlations were also performed to correlate gene expression with circumference and water potential measurements while controlling for treatments.

RESULTS

Real-time quantitative PCR was used to examine the expression of 114 putative disease- and drought-responsive genes in loblolly pine trees challenged with four treatments—cyclic drought, *F. circinatum*, cyclic drought + *F. circinatum*, and no treatment. Ten cuticular wax biosynthetic genes were additionally analyzed only in the drought and control treatments. The expression of all genes changed in response to drought, *F. circinatum*, and cyclic drought + *F. circinatum* treatments in comparison with their expression in the control group. Eighty-eight of the genes responded similarly to both drought and *F. circinatum* treatments (Fig. 3.1). Sixty-one genes were upregulated by both *F. circinatum* and drought treatments, and 27 genes were downregulated by both treatments, leaving just 16 genes that were inversely regulated by *F. circinatum* and drought treatments (Fig. 3.1). Analysis of variance of gene expression among clones within a treatment showed that 58 genes exhibited significantly different expression ($p<0.05$) among clones in every treatment group. To examine whether any differences in the

variability of expression existed based on gene function, genes were grouped into 16 functional classes. The average gene expression for each class was calculated and the smallest and largest range of expression differences for any gene in each class was identified (Table 3.1) for all treatments. No trends in the variability of genes were apparent. The pathogenesis-related genes did exhibit much higher average expression induced by the three different treatments than the other classes of genes (Table 3.1).

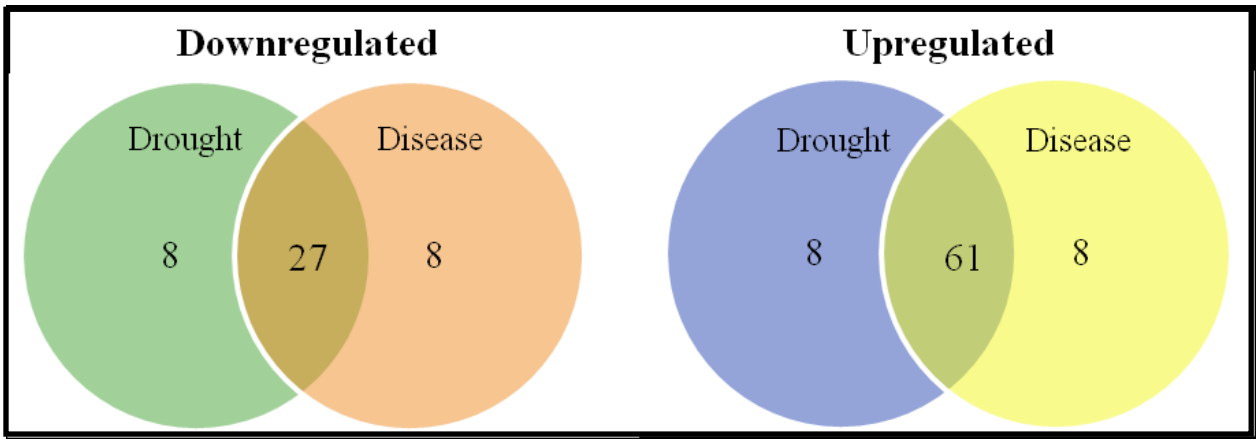

Fig. 3.1 Diagram showing overlap in gene expression responses to disease and drought stresses. Twenty-seven genes were downregulated by both cyclic drought and *F. circinatum* inoculations and 61 genes were upregulated by both treatments. The remaining 16 genes were inversely regulated by the two treatments.

Analysis of variance of gene expression among treatments revealed limited variation. The only gene showing significant differences ($p<0.01$) in expression between treatments was PtMYB12 (Fig. 3.2). Further post hoc testing showed that expression of PtMYB12 in the drought + *F. circinatum* was approximately 5-fold higher than its expression in the other treatments.

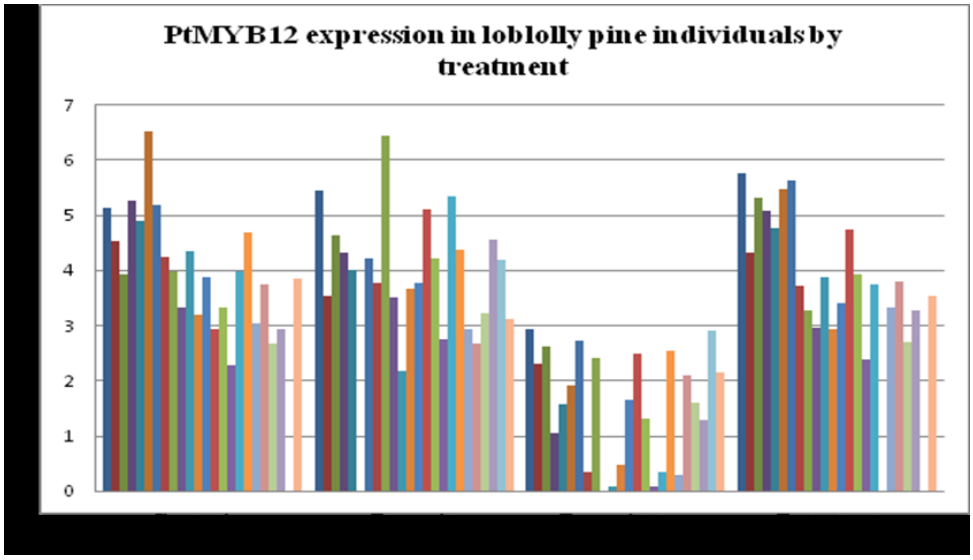

Fig. 3.2 PtMYB12 expression among loblolly clones by treatment. Expression values have been adjusted to improve readability. Lower $\Delta\Delta C_T$ values indicate higher gene expression.

Table 3.1 Genes grouped by putative function. The average expression and range of expression differences for genes in each class was calculated for all treatments.

Functional class	Control	Drought	*F. circinatum*	Drought + *F. circinatum*
Cell wall related				
Average expression	-0.610	0.200	-0.377	0.341
Range of expression differences (fold)	8.23-8.79	8.96-64.9	5.02-6.08	9.11-56.1
Cuticle biosynthesis				
Average expression	-0.008	0.441	na	na
Range of expression differences (fold)	3.06-34.7	3.05-56.9	na	na
Detoxifying enzymes				
Average expression	-0.112	-0.022	-0.687	-0.546
Range of expression differences (fold)	4.83-29.2	3.86-13.5	3.50-12.3	10.1-23.6
Disease response transcription factors				
Average expression	0.121	-0.201	-0.396	-0.516
Range of expression differences (fold)	5.22-53.4	4.08-1157	2.69-247.6	6.25-75.4
Disease signaling				
Average expression	0.063	-0.368	-0.592	-0.533
Range of expression differences (fold)	4.51-49.8	2.43-369.1	2.45-111.5	2.49-158.2
Drought response transcription factors				
Average expression	-0.116	-0.789	-0.478	-0.671
Range of expression differences (fold)	3.36-33.9	3.83-171.4	4.32-256	4.51-40.9
Drought signaling				
Average expression	0.004	-0.496	-0.029	-0.517
Range of expression differences (fold)	3.96-45.2	2.76-50.1	3.48-19.9	3.77-59.2
H_2O_2 signaling				
Average expression	-1.275	-1.388	-1.234	-1.359
Range of expression differences (fold)	5.94-141.7	11.3-16.3	4.24-10.1	4.96-21.5
Late embryogenesis abundant proteins				
Average expression	-0.757	-1.777	-1.418	-2.262
Range of expression differences (fold)	4.68-112.4	6.19-131.5	5.76-743	10.0-4050
Osmotic adjustment				
Average expression	-0.206	-0.317	-0.708	-0.345
Range of expression differences (fold)	3.07-29.9	3.48-52.4	2.46-9.23	3.75-137.4
Pathogenesis-related proteins				
Average expression	-0.709	-2.519	-3.934	-5.003
Range of expression differences (fold)	28.2-217.5	12.6-432.8	19.0-305	27.4-380.7
Peroxidases				
Average expression	0.038	-0.259	-0.608	-0.275
Range of expression differences (fold)	5.52-34.6	6.17-84.4	5.78-46.5	5.79-255.8
Phenylpropanoid pathway				
Average expression	0.041	0.090	-0.464	-0.161
Range of expression differences (fold)	3.75-52.9	7.41-41.5	1.41-30.9	5.26-539.9
Pine drought responsive genes				
Average expression	-1.094	-1.160	-0.968	-1.13
Range of expression differences (fold)	4.94-71.5	4.43-42.7	3.12-113.3	3.89-34.1
Programmed cell death				
Average expression	-0.352	-0.353	-0.321	-0.373
Range of expression differences (fold)	6.09-16.0	6.95-9.56	4.90-8.76	7.90-18.7
Terpenoid biosynthesis				
Average expression	0.038	1.361	-0.258	1.090
Range of expression differences (fold)	4.62-264	9.73-1353	3.60-237.9	16.8-4725

Although this was the only gene where differences in expression between treatments was supported by ANOVA analysis, the average expression of many genes changed by at least 4-fold in response to one or more treatments (Table 3.2), but may not have been supported statistically because of the range of expression differences between clones within a treatment (Fig. 3.3). Changes in gene expression in a subset of 6 genotypes were surveyed in response to water-inoculations. Ten genes displayed expression changes in response to wounding and inoculations with water that were very similar to the changes in expression observed in response to wounding and *F. circinatum* inoculations. These genes are noted in Table 3.2.

Table 3.2 Gene expression changes relative to the control group in response to treatments. Green represents a decrease in expression. The darkest green indicates that the average expression of that gene decreased by at least 4-fold. Medium green shows a 1.4- to 4-fold decrease and the lightest green is a less than 1.4-fold decrease. Red represents an increase in average expression with the darkest red being greater than a 4-fold increase, medium red is a 1.4- to 4-fold increase, and the lightest red shows an increase less than 1.4 fold.

Gene name	Putative function	Drought	*F. circinatum*	Drought + *F. circinatum*
(-)pin-TPS	Monoterpene synthase	-	-	+
(+)pin-TPS	Monoterpene synthase	-	-	+
ABA1	Zeaxanthin epoxidase	+	+	+
ABI1	Serine/threonine phosphatase 2C	+	+	+
ADR1	Activated disease resistance protein	+	-	-
ALDH7B4	Aldehyde dehydrogenase/oxidoreductase	+	+	+
ANR	Anthocyanidin reductase	+	-	+
AOC	Diterpene synthase	-	-	+
AOS	Allene oxide synthase	+	+	+
ARF	ADP-ribosylation factor	-	+	+
ATAF-1	NAC transcription factor	+	+*	+
PtNAP	Non-intrinsic ABC protein	-	-	+
Att1	CYP86A subfamily cytochrome p450	-		
Avr-Cf9	Avr9/Cf-9 LRR domain protein	+	-*	-
PtBAG-1	Class I Bcl-2 associated athanogene	-	-	-
PtBAG-4	Class I Bcl-2 associated athanogene	-	+	-
BALDH	Betaine aldehyde dehydrogenase	+	+	+
BGLU1	Glycosyl hydrolase family 1 protein	-	+	+
BI-1	Bcl-2 associated (Bax)-inhibitor protein	+	+*	+
bZIP	β-Zip domain transcription factor	+	+	+
CAB	Ubiquitin protein ligase	+	+	+
CER1	*Eceriferum* wax biosynthetic gene	-		
CER4	Fatty acyl-CoA reductase	-		

Table 3.2 continued.

CHI	Chalcone isomerase	-	+	+
CHS	Naringenin-chalcone synthase	-	-	+
CMyc	C-Myc binding protein	-	-	-
COI1	Ubiquitin protein ligase	+	+*	+
CPK3	Calcium-dependent protein kinase	+	+	+
CslA1	Cellulose synthase-like A1	-	-	-
CYPA	Cytochrome p450 monooxygenase	-	-	+
CYPB	Cytochrome p450 monooxygenase	-	-	-
CYPC	Cytochrome p450 monooxygenase	+	+*	+
CYPD	Cytochrome p450 monooxygenase	-	-	-
EB9D	14-3-3 protein	-	-	+
ERD3	Early-response to dehydration protein	+	+	+
ERECTA	Transmembrane receptor protein kinase	-	-	-
ERF1	Ethylene-responsive transcription factor	+	+	+
GLX1	Glyoxylase 1	+	+	+
GPAT6	Glycerol-3-phosphate acyltransferase	-		
GRAS	GRAS domain transcription factor	-	-	-
HDZ32	Homeodomain/leucine-zipper protein	+	+	+
HMG-CoA	Hydroxymethylglutaryl-CoA-synthase	+	+*	+
HPL	Hydroperoxide lyase	-	-	-
HSP17.4CIII	Class III heat shock protein	+	-	-
IAA2	Auxin-induced transcription factor	-	-	-
IAA3	Auxin-induced transcription factor	-	-	-
IMPA	Importin, alpha isoform	+	+	+
IPK	Inositol polyphosphate kinase	-	+	+
KCS4	3-ketoacyl-CoA synthase	-		
LACS6	Long-chain acyl-CoA synthetase	+		
LACS9	Long-chain acyl-CoA synthetase	+		
LP3	ABA/water deficit inducible protein	+	+	+
MYND	Zinc finger (MYND type) protein	-	+	-
NAC1	NAC domain transcription factor	+	+	-
NCED	9-*cis*-epoxycarotenoid dioxygenase	+	+	-
OXR1	2OG-Fe(II) oxygenase family protein	+	+*	+
OXR2	2OG-Fe(II) oxygenase family protein	-	-	-
P5CDH	1-pyrroline-5-carboxylate dehydrogenase	-	-	-
P5CR	Pyrolline-5-carboxylate reductase	+	+	+
PR-1	Pathogenesis-related anti-fungal protein	+	+	+
PR-2	β-1,3-glucanase	+	+*	+
PR-3	Basic chitinase	+	+	+
PR-5	Thaumatin-like protein	+	+	+
PR-10	Pathogenesis-related family 10 protein	+	+*	+
PSCHI4	Extracellular chitinase	+	+	+
PT1	Phenylcoumaran benzylic ether reductase	+	+	+
Pt31	Class I small heat shock protein	+	-	-
Pt38	Hypothetical protein	+	+	+
PtAldh	3-chloroallyl aldehyde dehydrogenase	-	+	+
PtAN1	AN1-like zinc finger protein	+	+	+
PtCPN10	Mitochondrial chaperonin	+	+	+
PtEMB1	Late embryogenesis abundant protein	+	+	+

Table 3.2 continued.

PtEMB2	Late embryogenesis abundant protein	+	+	-
PtEMB3	Late embryogenesis abundant protein	+	+	+
PtEMB4	Late embryogenesis abundant protein	+	+	+
PtGOLS1	Galactinol synthase	+	+	+
PtGPX1	Glutathione peroxidase	+	+	+
PtGPX2	Glutathione peroxidase	+	+	+
PtGPX3	Glutathione peroxidase	+	+	+
PtGRP	Glycine-rich protein	+	+	-
PtGSTF9	Class phi glutathione *S*-transferase	-	+	+
PtGSTU18	Class tau glutathione *S*-transferase	+	+	+
PtGTP	GTP-binding protein	-	-	-
PTI4	Ethylene-response element binding factor	-	-	-
PtIP	Inorganic pyrophosphatase	+	-	+
PtLEA2	Late embryogenesis abundant protein	+	+	+
PtLEA3a	Late embryogenesis abundant protein	+	+	-
PtLEA3c	Late embryogenesis abundant protein	+	+	+
PtMLO2	Mlo-family protein	+	+	+
PtMYB6	MYB domain transcription factor	+	+	+
PtMYB9	MYB domain transcription factor	+	+	+
PtMYB12	MYB domain transcription factor	-	-	-
PtMYB13	MYB domain transcription factor	+	+	+
PtOMT	*o*-methyltransferase	+	-	-
PtPAL	Phenylalanine ammonia-lyase	+	+	+
PtPDIR1	Dirigent-like protein	-	-	+
PtWRKY	WRKY domain transcription factor	+	+	+
PtWRKY2	WRKY domain transcription factor	+	+	+
PX-1	Class III secretory peroxidase	-	-	-
PX-2	Class III secretory peroxidase	+	+*	+
PX-3	Class III secretory peroxidase	+	+	+
PX-4	Class III secretory peroxidase	+	-	+
PX-5	Class III secretory peroxidase	+	-	+
RAP2.1	AP2 domain transcription factor	+	+	+
RAP2.4	AP2 domain transcription factor	+	+	+
RPK1	Protein kinase family protein	-	-	-
SamCMT	Jasmonic acid carboxyl methyltransferase	+	+	+
Sesqui-TPS	Sesquiterpene synthase	-	-	+
STR_SYN	Jasmonate-responsive alkaloid enzyme	+	+	-
trp-TPS	Monoterpene synthase	-	-	+
TTA7	Flavonoid 3'-monooxygenase	+	+	+
WBC11	ATPase/fatty acid transporter	+		
WBC12	ABC transporter	-		
WIN1	Ethylene-responsive transcription factor	-		

* Gene expression changes resulting from water-inoculations were similar to the expression changes observed following *F. circinatum* inoculations in the same clones.

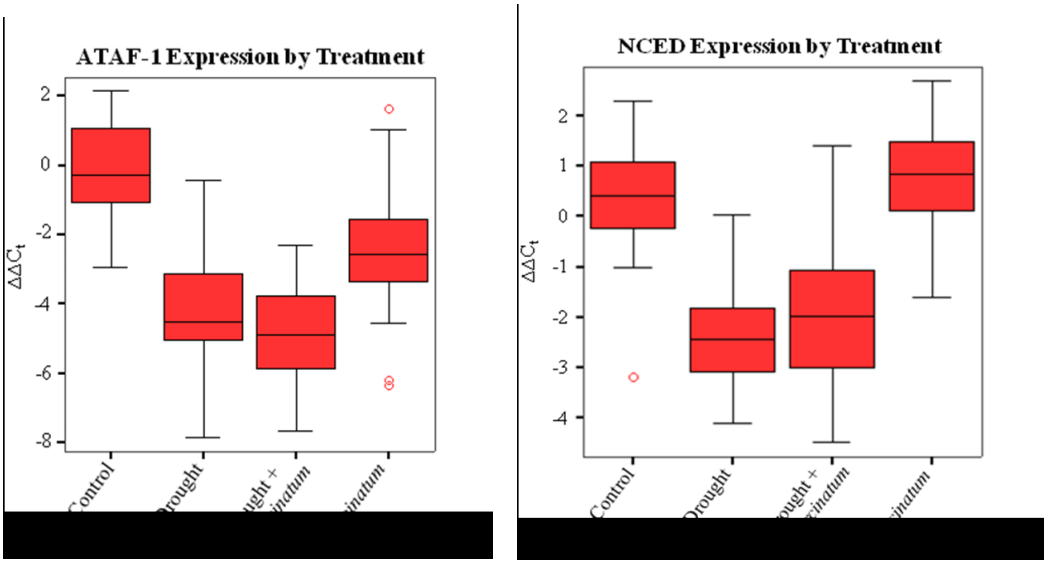

Fig. 3.3 Gene expression box plots by treatment. The midlines in these figures are the median expression values and outliers are denoted with circles. Many genes exhibited changes in expression in response to treatments, but were not statistically significant. Two examples are ATAF-1, a transcription factor induced by both biotic and abiotic stresses, and NCED, a key gene regulating ABA biosynthesis.

Since loblolly clones were selected based on previous evidence that they were either resistant or susceptible to pitch canker infections or had extreme water use efficiency phenotypes. ANOVA analyses were performed to identify genes that were differentially expressed among loblolly individuals classified as “resistant” or “susceptible” to pitch canker and “resistant” or “susceptible” to dehydration. Several genes were differentially expressed between both resistant and susceptible pitch canker phenotypes and resistant and susceptible to dehydration phenotypes (Table 3.3). Resistant and susceptible phenotypes (for both pitch canker and water use efficiency) exhibited differences in expression in multiple genes when exposed to cyclic drought (Table 3.3).

Table 3.3 Genes identified by ANOVA analyses controlling for water use efficiency or pitch canker resistance phenotypes. Loblolly pine individuals were classified as resistant or susceptible to *F. circinatum* and resistant or susceptible to dehydration based on physiological measurements. The table lists genes that were expressed differently ($p<0.05$) between the different phenotypic classes and the average fold-difference in expression between the resistant (R) and susceptible (S) individuals. There were no differences in gene expression among the resistant and susceptible phenotypes in the *F. circinatum* treatment group.

ANOVA controlling for water use efficiency							
Control		**Drought**		***F. circinatum***		**Drought + F. circinatum**	
Gene	**R-S (fold)**	**Gene**	**R-S (fold)**	**Gene**	**R-S (fold)**	**Gene**	**R-S (fold)**
Pt31	2.03	ERF1	2.43	na	na	na	na
CslA1	3.23	TTA7	2.31				
		PtBAG-4	2.24				
		NCED	2.48				
		PR-2	3.32				
		PtWRKY	2.98				
ANOVA controlling for pitch canker resistance							
Control		**Drought**		***F. circinatum***		**Drought + F. circinatum**	
Gene	**R-S (fold)**	**Gene**	**R-S (fold)**	**Gene**	**R-S (fold)**	**Gene**	**R-S (fold)**
(-)pin-TPS	-2.23	(-)pin-TPS	-2.89	na	na	na	na
(+)pin-TPS	-1.78	(+)pin-TPS	-2.82				
HDZ32	-2.21	trp-TPS	-3.02				
trp-TPS	-3.01	PtPDIR1	-2.27				
PR-10	-1.98	PtMYB12	-2.90				
PR-3	-2.19	PX-1	-2.12				
		CYPA	-3.65				
		CYPD	-2.74				

This suggests that differences in the induction of stress-related gene expression may be important in producing resistant or susceptible phenotypes. Also, loblolly pine trees varying in resistance to pitch canker exhibited expression differences when exposed to drought. Different responses to drought by the various pitch canker resistance phenotypes may help to dissect the molecular mechanisms underlying the observations of increased susceptibility to diseases during environmental stresses and assist in the selection of individuals resistant to pitch canker disease

under environmental stresses for breeding purposes. No differences in gene expression in response to *F. circinatum* or *F. circinatum* + drought were observed between susceptible and resistant phenotypes. Disease-responsive gene expression may have acclimatized during the 4-week period between inoculations and tissue collection making differences in gene expression between different phenotypes statistically insignificant.

Partial correlation analysis of gene expression controlling for treatments showed strong linear correlations between several genes (Appendix C, Table C.1). Correlations between genes varied among treatments. Further examination showed that in the majority of cases, correlations increased compared with the control in response to one or more treatments. Genes with highly correlated expression in response to one or more treatments may be induced by the same signal or may be involved in the same pathways. Since correlations between genes operating in the same functional pathways was a common trend observed throughout the association analyses, it was predicted that candidate genes containing SNPs associated with gene expression profiles in the association genetics analysis in Chapter II would be correlated. Only two genes identified by association genetics approaches in Chapter II were strongly correlated with the gene that they were originally associated with. PtNAP was found to be associated with the expression PtEMB1 and a GRAS transcription factor was associated with EB9D expression. The partial correlation coefficients between these two pairs of genes were PR=0.63 and PR=0.54, respectively. Correlations between PtNAP and PtEMB1 were exceptionally high (R^2=0.808) in the drought + *F. circinatum* treatments (Fig. 3.4). Correlations between all of the other candidate genes and the stress-related genes that they were originally associated with were very low. In these cases where expression of genes discovered through association genetics studies was not correlated with the expression of the previously associated gene, raise the question of whether these associations may have been due to another SNP at a linked locus.

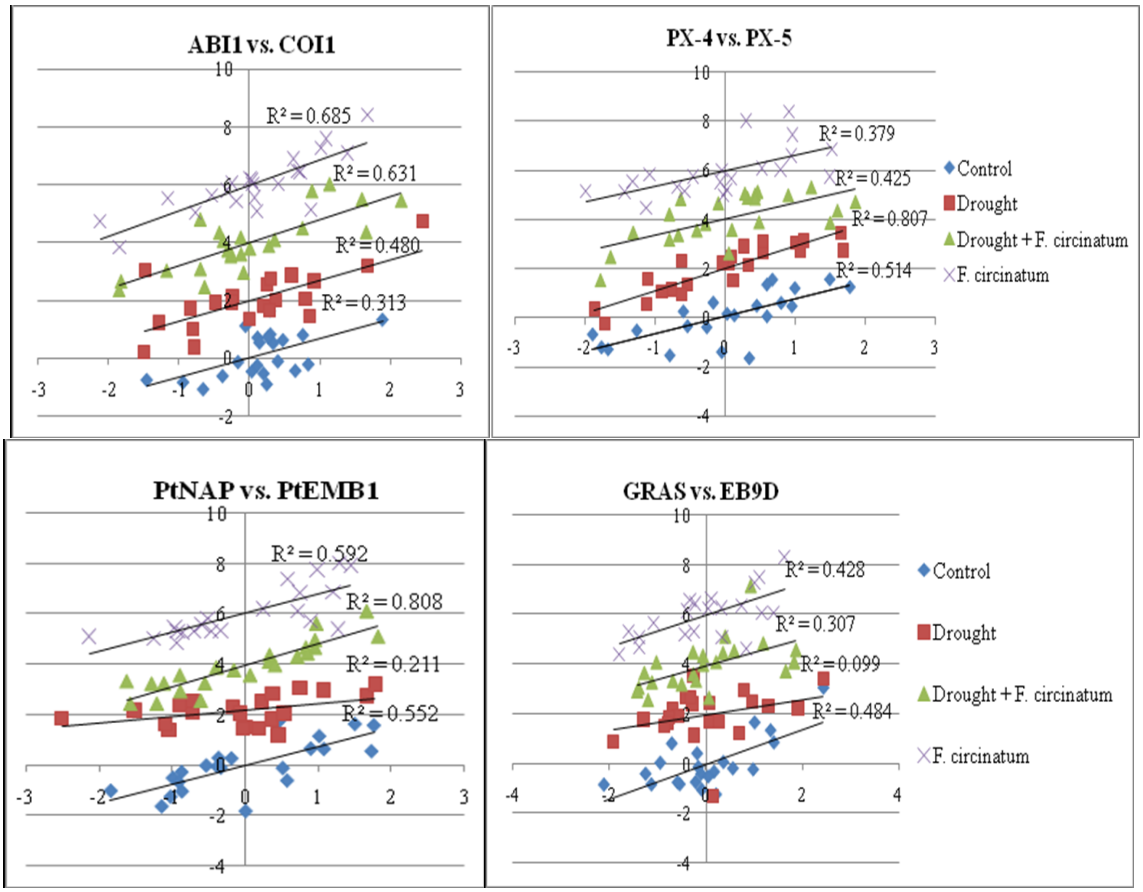

Fig. 3.4 Correlations between levels of gene expression by treatment. R^2 values close to 1 indicate stronger correlations between gene expression levels.

At the time of harvest, the tree circumferences and water potentials were measured on all trees. Welch's ANOVA analysis followed by Dunnett's T3 post hoc testing showed that there were no significant differences in either measure within any treatment group, but that trunk circumferences of trees in the control group were significantly larger ($p<0.01$) than the trunk circumferences of trees in the drought and drought + *F. circinatum* groups. It is well known that plants reallocate resources during periods of stress. Plant growth slows and above-ground biomass decreases during dehydration stress (Teskey et al. 1987; Schultz 1997).

Epicuticular wax content extracted from needles was quite variable among ramets and no significant differences in epicuticular wax among clones were observed. A transcription

factor, WIN1, was consistently lower by about 4-fold in response to drought. All other wax biosynthetic genes exhibited smaller changes in expression. No relationship between the expression of any specific wax biosynthetic gene and epicuticular wax content of needles was found. These results are not surprising since gene expression was measured in xylem tissue rather than in needles.

DISCUSSION

Drought stress is often the most important environmental factor limiting forest distribution and productivity (Newton et al. 1991; Winnett 1998; Kliejunas et al. 2008). Drought conditions result in seedling mortality, slowed growth of mature trees, and increased susceptibility to insects and diseases (Newton et al. 1991; Winnett 1998; Coder 1999). With the risk for climate change and the frequent introduction of exotic pathogens, significant losses by forest pathogens and pests are projected to increase also (Howard 2005). In an effort to identify and characterize genes responsive to drought stress, infection by the pathogen *Fusarium circinatum*, or a combination of these two stresses in loblolly pine, gene expression values were collected on loblolly pine trees possessing a range of disease- and drought-resistance phenotypes. Gene expression analysis is a useful tool for examining the molecular mechanisms underlying complex traits especially when a precise and reproducible phenotypic measurement is difficult to obtain (Stranger et al. 2005; Sackton et al. 2010).

Significant variations in gene expression were observed among loblolly individuals within the control group and within all three treatment groups. The variability in expression that was observed is in agreement with previous studies in other organisms that have described highly variable expression of stress-related genes (López-Maury et al. 2008) and with the results generated in the association genetics study in Chapter II. Expression of stress-related genes

appeared to be more variable in the control group with fewer genes exhibiting significant differences in expression among individual in the other treatment groups. The initiation and execution of signal transduction in response to biotic and abiotic stresses in a timely manner appears to be important for mounting effective responses (Sackton et al. 2010; Walley and Dehesh 2010). Thus, induced expression is likely tightly controlled and ineffective activation of defenses would be selected against in nature.

The results of the correlation analysis also suggested that regulation of induced gene expression was less variable than regulation of uninduced expression. Correlations between genes tended to be stronger in the treatment groups than in the control group. Partial correlation analysis controlling for treatments revealed genes with strongly correlated (R^2>0.5) profiles. Functionally related genes are predicted to have correlated expression patterns (Wang et al. 2004). In order to identify correlated genes that were functionally related, the genes with homologs in *A. thaliana* were entered into the functional gene network tool Aranet (Lee et al. 2010). In almost all of the cases, genes correlated with R^2>0.5 were predicted to be in the same functional networks in Aranet.

Ten of the genes analyzed were originally identified as SNP-containing genes that were correlated with the expression of putative disease-resistance and drought-resistance genes. Their expression in response to disease and dehydration stress was examined to confirm their involvement in stress responses. It was expected that expression of a SNP-containing gene would be correlated with the expression of the gene it was originally associated with in at least one of the treatments. PtNAP and PtEMB1, and GRAS and EB9d were the only gene pairs identified through association genetics studies that were significantly correlated. The possibility that SNPs in other genes linked with the SNP-containing candidate genes are actually contributing to the different gene expression phenotypes needs to be examined for the other candidate genes (cMyc,

IAA2, IAA3, MYND, PtPAL, and TTA7) that were identified in the association genetics study in Chapter II. It is possible that these genes are not stress-resistance genes, but that changes in their expression in these studies are secondary effects resulting from cross-talk among different pathways or reallocation of resources in response to stress (Watkinson et al. 2003). Only 23 genes exhibited average expression changes greater than 4-fold in response to at least one treatment. Most genes expression changes were moderate (1.4-4-fold), and only one gene, PtMYB12, showed significant differences in expression among treatments. Even small variations in signal transduction genes can be amplified by altering the expression of downstream genes to produce large phenotypic effects (Sackton et al. 2010). The majority of genes were either upregulated or downregulated by both cyclic drought stress and *F. circinatum* infections suggesting the use of common signaling elements by both abiotic and biotic stress pathways (Pastori and Foyer 2002). Responses to drought + *F. circinatum* were not cumulative responses of the two different stresses. As the data shows, abiotic and biotic signaling pathways may be inversely or jointly regulated (Desprez-Loustau et al. 2006) and result in either higher or lower expression values in comparison to expression resulting from treatment with a single stress. The most likely circumstance is that plants employ different responses to different combinations of environmental stresses and pathogens.

Loblolly individuals in this study were selected based on carbon isotope discrimination values and pitch canker lesion length. Carbon isotope discrimination has been shown to be an accurate phenotypic measure of water use efficiency in pine (González-Martínez et al. 2008) and pitch canker lesion length has been shown to be a good measure of *F. circinatum* resistance (Quesada et al. 2010). ANOVA analyses were performed to identify genes that were expressed differently between resistant or susceptible to pitch canker phenotypes and resistant or susceptible to dehydration phenotypes.

In the ANOVA analysis of gene expression among dehydration resistant and susceptible phenotypes, NCED (9-cis-epoxycarotenoid dioxygenase) was expressed 2-fold higher in resistant individuals than susceptible individuals. NCED is the key regulatory step in ABA biosynthesis (Iuchi et al. 2001; Hao et al. 2009), a key hormone in initiating stress responses in several dehydration-responsive pathways (Shinozaki et al. 2007; Hao et al. 2009). NCED-overexpressing lines of *A. thaliana* that accumulated higher levels of ABA also exhibited decreased transpiration rates and higher expression levels of drought-responsive genes (Iuchi et al. 2001). Less drought tolerant pine individuals with lower levels of ABA may take longer to initiate ABA-mediated drought-responsive pathways or may exhibit lower levels of ABA-induced gene expression resulting in less effective physiological responses.

In the ANOVA analyzing differences in gene expression among different pitch canker resistant and susceptible phenotypes, 5 of the genes expressed differently among pitch canker resistant phenotypes are involved in the production of terpenoids including (-)pin-TPS, (+)pine-TPS, trp-TPS, CYPA, and CYPD. Rapid accumulation of terpenoids in response to insect and pathogens is a key defensive element in multiple species of the Pinaceae family (McKay et al. 2003; Keeling and Bohlmann 2006; McKay et al; 2006; Wallis et al. 2008). Drought stress reduces resin flow and alters the concentrations of monoterpenes making trees more susceptible to pests and diseases (Schultz 1997). Expression of the terpene synthetic genes was approximately 3-fold lower in pitch canker resistant individuals than susceptible individuals subjected to cyclic drought. Also, uninduced expression of three terpene synthases was lower in resistant individuals than in susceptible individuals. These ANOVA analyses do not necessarily signify causal relationships between gene expression and a physiological trait. Lower expression levels of terpene synthases in pitch canker resistant individuals experiencing drought stress could be a secondary effect due to negative cross talk between ABA-responsive drought-induced

pathways and jasmonate signaling pathways that induce terpene synthesis (Shinozaki et al. 2007; Huber et al. 2005). Or, trees under drought stress may reallocate resources shared among multiple physiological processes. Either of these strategies may preserve the fitness of the tree making them more able to tolerate additional stresses.

Another notable gene expressed differently among resistant and susceptible pitch canker phenotypes exposed to drought was PtMYB12. PtMYB12 expression was nearly 3-fold lower in pitch canker resistant individuals. PtMYB12 was interesting because it was the only gene that showed significant differences in expression induced by drought + *F. circinatum* treatments. PtMYB12 was originally identified through association genetics analyses in Chapter II. PtMYB12 is homologous to AtMYB6, a MYB transcription factor in *Arabidopsis* that is responsive to both jasmonic acid and salicylic acid suggesting that this MYB could be involved primarily in disease responses. The large changes in expression observed in response to *F. circinatum* infections during dehydration stress suggest that this gene could be important for responses to multiple stresses. Further characterization of its expression and its effects on plant phenotype is needed.

CONCLUSIONS

In this study we were able to confirm significant variation in both uninduced and induced expression of disease- and drought-responsive genes among loblolly pines possessing varying degrees of resistance to *Fusarium circinatum* infections and drought stress. Through our gene expression analyses we were able to identify a MYB transcription factor, PtMYB12, that is significantly upregulated in response to a combination of cyclic drought treatments and *F. circinatum* inoculations. Through correlation analyses we were able to identify genes that were likely functioning in the same, or linked pathways. We were able to confirm some of these

connections by identifying the homologous genes in *Arabidopsis* and using the Aranet tool to determine if the homologs were acting in functional pathways together. Previous analyses of disease- and drought-resistance traits of the loblolly pines included in this study allowed us to associate variations in gene expression with the various phenotypes. Terpene synthases appeared to be significant in determining disease resistance, particularly during drought stress, and the expression of the ABA biosynthetic gene, NCED, appeared to be influential in drought tolerance. The genes highlighted in the results of this study may be helpful in improving disease- and drought-resistance traits in loblolly pine.

CHAPTER IV

CONCLUSIONS

Loblolly pine is an economically and ecologically important tree in the southern U.S. Genetic improvement of breeding stocks by marker-assisted selection is desirable as it drastically shortens the rotation time that would typically be required for evaluating important traits. Association genetics approaches take advantage of fine linkage disequilibrium within a population to link genetic polymorphisms with phenotypic variations. Gene expression phenotyping is advantageous as it serves as a precise, easy to evaluate phenotype for dissecting highly complex traits such as disease and drought responses.

In Chapter II, abundant variation in uninduced gene expression of stress-responsive genes was detected in loblolly pines. The variation in gene expression detected in loblolly pines from across the natural range appeared to have a geographical component. Differences in expression among loblolly individuals from the Atlantic Coast region, Gulf Coast region, and from West of the Mississippi River accounted for some of the variation that was observed. Genes involved in the same, or connected, signaling pathways were predicted to be correlated and cluster together. Clustering of uninduced gene expression was not strong and revealed only moderate clustering of gene families and functionally similar genes probably due to the fact that expression was collected on unstressed tissues. Correlations of gene expression values induced by drought treatments, *F. circinatum* inoculations, or drought treatments + *F. circinatum* treatments in Chapter III were much stronger, supporting the hypothesis that uninduced expression of stress-related genes is more variable. Association genetics studies between 3937 SNPs and 88 stress-related gene expression phenotypes in 354 individuals revealed 101 significant associations between SNPs in 94 candidate genes and 27 gene expression phenotypes.

The majority of SNPs were in introns and untranslated regions highlighting that SNPs in regulatory elements that are abundant in these regions most likely influence important traits.

In Chapter III, the expression of 104 putative stress-related genes was examined in control conditions and in response to drought stress, *Fusarium circinatum*, and drought + *F. circinatum*. Ten wax biosynthesis genes were additionally examined under uninduced conditions and in response to drought treatments. Expression was collected on a subset of 24 loblolly individuals that were selected based on evidence that they possessed a range of disease- and drought-resistant phenotypes.

Some of the genes identified through the association genetics study were questionable as to whether they are really involved in stress responses. Their expression was not strongly correlated with the expression of the gene that they were originally associated with. SNPs at nearby linked genetic loci may actually be causing phenotypic differences. Fewer genes exhibited significant differences among individuals when induced by stress treatments. Only 58 of the 114 genes were significantly different among individuals in all treatment groups. Most of the genes responded to one or more stress treatment with at least a 1.4-fold change in expression, but only one gene showed statistically significant different gene expression. PtMYB12 expression in response to drought + *F. circinatum* was different from any of the other treatments. This gene was also differentially expressed between pitch canker resistant and pitch canker susceptible individuals subjected to drought stress. Other genes that appeared to be contributing to pitch canker phenotypes determined by ANOVA analysis were terpene synthases. ANOVA analysis was also performed to examine genes expressed differently between drought resistance phenotypes scored by carbon isotope discrimination. The most interesting gene from this analysis was NCED, a key regulatory step in ABA biosynthesis.

These studies have provided an essential first step toward dissecting disease- and drought-resistance phenotypes in loblolly pine. The results of these studies have confirmed that most of the putative stress-related genes identified through comparative genetics and some of the genes identified with association genetics approaches exhibit substantial changes in expression in response to biotic and abiotic stresses. All of the SNPs in candidate genes identified through the association genetics studies need to be verified. Once verified, linking these allelic polymorphisms and gene expression variation with whole plant physiology can provide a more intimate understanding of the molecular mechanisms underlying disease- and drought-resistance traits and SNPs can be implemented in marker-assisted selection to reduce time and cost of breeding superior pine trees.

REFERENCES

ADEPT2 (Allele Discovery of Economic Pine Traits) (2008) ADEPT2 overview. Available via http://dendrome.ucdavis.edu/adept2/index.html. Accessed 9 Nov 2009.

Adryan B, Schuh R (2004) Gene-ontology-based clustering of expression data. Bioionforma Appl Note 20(16): 2851-2852.

Arteaga-Vázquez M, Caballero-Perez J, Vielle JP (2006) A family of microRNAs present in plants and animals. Plant Cell 18: 3355-3369.

Barnard EL, Blakeslee GM (1987) Pitch canker of southern pines. Plant Pathol Circular 302, 4 pp.

Baltunis BS, Martin TA, Huber DA, Davis JM (2008) Inheritance of foliar stable carbon isotope discrimination and third-year height in Pinus taeda clones on contrasting sites in Florida and Georgia. Tree Genet & Genomes 4: 797-807.

Bilan MV, Hogan CT, Carter HB (1977) Stomatal opening, transpiration, and needle moisture in loblolly pine seedlings from two Texas seed sources. For Sci 23: 457-462.

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, et al. (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinforma 23: 2633-2635.

Brown GR, Gill GP, Kuntz RJ, Langley CH, Neale DB (2004) Nucleotide diversity and linkage disequilibrium in loblolly pine. PNAS 101: 15244-15260.

Brown JWS, Simpson CG (1998) Splice site selection in plant pre-mRNA splicing. Annu Rev of Plant Physiol and Plant Mol Biol 49: 77-95.

Buschges R, Hollricher K, Panstruga R, Simons G, Wolter M, et al. (1997) The barley Mlo gene: A novel control element of plant pathogen resistance. Cell 99: 695-705.

Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. Plant Mol Biol Report 11: 113-116.

Chang S, Puryear JD, Dias DL, Funkhouser EA, Newton RJ, et al. (1996) Gene expression under water deficit in loblolly pine (*Pinus taeda*): isolation and characterization of cDNA clones. Physiol Planta 97: 139-148.

Chaves MM, Maroco JP, Periera JS (2003) Understanding plant responses to drought—from genes to the whole plant. Funct Plant Biol 30: 239-264.

Cheung VG, Spielman RS (2002) The genetics of variation in gene expression. Nat Genet 32: 522-525.

Chini A, Grant JJ, Seki M, Shinozaki K, Loake GJ (2004) Drought tolerance established by enhanced expression of the CC-NBS-LRR gene, ADR1, requires salicylic acid, EDS1 & ABI1. Plant J 38: 810-822.

Coder KD (1999) Drought damage to trees. University of Georgia, Daniel B. Warnell School of Forest Resources Extention publication, 5 pp.

Cordell HJ, Clayton DG (2005) Genetic association studies. The Lancet 366: 1121-1131.

Davis JM, Wu H, Cooke JEK, Reed JM, Luce KS, Michler CH (2002) Pathogen challenge, salicylic acid, jasmonic acid regulates expression of chitinase gene homologs in pine. Mol Plant-Microbe Interact 15: 380-387.

Desprez-Loustau ML, Marcais B, Nageleisen LM, Piou D, Vannini A (2006) Interactive effects of drought and pathogens in forest trees. Ann For Sci 63: 597-612.

Deuschle K, Funck D, Forlani G, Stransky H, Biehl A, et al. (2004) The role of [delta]1-pyrroline-5-carboxylate dehydrogenase in proline degradation. Plant Cell 16: 3413-3425.

Dixon RA, Paiva NL (1995) Stress-induced phenylpropanoid metabolism. Plant Cell 7: 1085-1097.

Dixon DP, Lapthorn A, Edwards R (2002) Plant glutathione transferases. Genome Biol 3: 1-10.

Dong JZ, Dunstan DI (1999) Cloning and characterization of six embryogenesis-associated cDNAs from somatic embryos of *Picea glauca* and their comparative expression during zygotic embryogenesis. Plant Mol Biol 39: 859-864.

Doukhanina EV, Chen S, van der Zalm E, Godzik A, Reed J, et al. (2006) Identification and functional characterization of the BAG protein family in Arabidopsis thaliana. J Biolog Chem 281: 18793-18801.

Du H, Zhang L, Liu L, Tang XF, Yang WJ, et al. (2009) Biochemical and molecular characterization of plant MYB transcription factor family. Biochem (Moscow) 74: 1-11.

Dwinell LD, Barrows-Broaddus J, Kuhlman EG (1985) Pitch canker: a disease complex of southern pines. Plant Dis 69: 270-276.

Eckert AJ, Van Heerwarden J, Wegrzyn JL, Nelson CD, Ross-Ibarra J, et al. (2010) Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). Genet 185: 969-982.

Edmeades GO, McMaster GS, White JW, Campos H (2004) Genomics and the physiologist: bridging the gap between genes and crop response. Field Crop Res 90: 5-18.

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. PNAS 95: 14863-14868.

Fan J, Hill L, Crooks C, Doerner P, Lamb C (2009) Abscisic acid has a key role in modulating diverse plant-pathogen interactions. Plant Physiol 150: 1750-1761.

Fay JC, McCullough HL, Sniegowski PD, Eisen MB (2004) Population genetic variation in gene expression is associated with phenotypic variation in *Saccharomyces cerevisiae*. Genome Biol 5: R26.

Feussner I, Wasternack C (2002) The lipoxygenase pathway. Annu Rev of Plant Biol 53: 275-297.

Flint-Garcia SA, Thornsberry JM, Buckler IV ES (2003) Structure of linkage disequilibrium in plants. Annu Rev Pl Biol 54: 357-74.

Fujimoto SY, Ohta M, Usui A, Shinshi H, Ohme-Takagi M (2000) Arabidopsis ethylene-responsive element binding factors act as a transcriptional activators or repressors of GCC box-mediated gene expression. Plant Cell 12: 393-404.

Fujita M, Fujita Y, Takahashi F, Yamaguchi-Shinozaki K, Shinozaki K (2009) In: Hirt H (ed) Plant stress biology: from genomics to systems biology. Wiley-VCH Verlag CmbH & Co, Weinheim, pp 67-79.

Gang DR, Kasahara H, Xia ZQ, Mijnsbrugge KV, Bauw G, et al. (1999) Evolution of plant defense mechanisms. Relationships of phenylcoumaran benzylic ether reductases to pinoresinol-lariciresinol and isoflavone reductases. J Biol Chem 274: 7516-7527.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, et al. (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis Nucleic Acids Res. 31:3784-3788.

Gillingham AK, Munro S (2007) The small G proteins of the Arf family and their regulators. Annu Rev Cell Dev Biol 23: 579-611.

Goldfarb B, Lanz-Garcia C, Lian Z, Whetten R (2003) Aux/IAA gene family is conserved in the gymnosperm, loblolly pine (*Pinus taeda*). Tree Physiol 23: 1181-1192.

González-Martínez SC, Wheeler NC, Ersoz, E, Nelson, CD, Neale DB (2007) Association Genetics in *Pinus taeda* L. I. Wood Property Traits. Genet 175: 399-409.

González-Martínez SC, Huber D, Ersoz E, Davis JM, and Neale DB (2008) Association genetics in *Pinus taeda* L. II. Carbon isotope discrimination. Heredity 101: 19-26.

Goujon T, Sibout R, Pollet B, Maba B, Nussaume L, et al. (2003) A new *Arabidopsis thaliana* mutant deficient in the expression of O-methyltransferase impacts lignins and sinapoyl esters. Plant Mol Biol 51: 973-989.

Gray WM, Kepinski S, Rouse D, Leyser O, Estelle M (2001). Auxin regulates SCF^{TIR1}-dependent degradation of AUX/IAA proteins. Nat 414: 271-276.

Hanson C, Yonavjak L, Clarke C, Minemeyer S, Boisrobert L, et al. (2010) Southern forests for the future. World Resources Institute. Available via http://www.seesouthernforests.org/files/sff/SouthernForestsfortheFuture.pdf. Accessed 27 July 2010.

Hammerbacher A, Wright LP, Wingfield BD, Wingfield MJ, Coutinho TA (2009) Factors affecting pine pitch canker modeled on Michaelis-Menten kinetics. Botany 87: 36-42.

Hao GP, Zhang XH, Wang YQ, Zhong-Yi W, Huang CL (2009) Nucleotide variation in the NCED3 region of *Arabidopsis thaliana* and its association study with abscisic acid content under drought stress. J Integr Plant Biol 51: 175-183.

Hematy K, Cherk C, Somerville S (2009) Host-pathogen warfare at the plant cell wall. Current Opin Plant Biol 12: 406-413.

Hepting GH, Roth ER (1946) Pitch canker, a new disease of some southern pines. J For 44: 742-744.

Howard JL (2005) U.S. timber production, trade, consumption, and price statistics 1965 to 2005. U.S. Dept of Agriculture Forest Service research paper FLP-RP-637.

Hu H, Dai M, Yao J, Xiao B, Xianghua L, Zhang Q, et al. (2006) Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. PNAS 103: 12987-12992.

Huber DPW, Phillippe RN, Madilao LL, Sturrock RN, Bohlmann J (2005) Changes in anatomy and terpene chemistry in roots of Douglas-fir seedling following treatment with methyl jasmonate. Tree Physiol 25: 1075-1083.

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

Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, et al. (2001) Regulation of drought tolerance by gene manipulation of 9-*cis*-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. Plant J 27: 325-333.

Jackson DA (1993) Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches. Ecol 74: 2204-2214.

Kabbage M, Dickman MB (2008) The BAG proteins: a ubiquitous family of chaperone regulators. CMLS 65: 1390-1402.

Kanyuka K, Praekelt U, Franklin KA, Billingham OE, Hoole R, et al. (2003) Mutations in the huge Arabidopsis gene BIG affect a range of hormone and light responses. Plant J 35: 57-70.

Kawai-Yamada M, Jin L, Yoshinaga K, Hirata A, Uchimiya H (2001) Mammalian bax-induced plant cell death can be down-regulated by overexpression of Arabidopsis bax inhibitor-1 (AtBI-1). PNAS 98: 12295-12300.

Kayihan GC, Huber DA, Morse AM, White TL, and Davis JM (2005) Genetic dissection of fusiform rust and pitch canker disease traits in loblolly pine. Theoretical Appl Genet 110: 948-958.

Keeling CI, Bohlmann J (2006) Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defense of conifers against insects and pathogens. New Phytol 170: 657-675.

Kesarwani M, Yoo J, Dong X (2007) Genetic interactions of TGA transcription factors in the regulation of pathogenesis-related genes and disease resistance in Arabidopsis. Plant Physiol 144: 336-346.

King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. Sci 188: 107-116.

Kliejunas JT, Geils BW, Glaeser JM, Goheen EM, Hennon P, et al. (2008) Climate and forest diseases of western North America: a literature review. U.S. Dept of Agriculture Forest Service, pp 1-44.

Knight H, Knight MR (2001) Abiotic stress signaling pathways: specificity and cross-talk. Trends Plant Sci 6: 262-267.

Kosma DK, Bourdenx B, Bernard A, Parsons EP, Lü S, et al. (2009) The impact of water deficiency on leaf cuticle lipids of Arabidopsis. Plant Physiol 151: 1918-1929.

Kotchoni SO, Bartels D (2003) Water stress induces the up-regulation of a specific set of genes in plants: aldehyde dehydrogenases as an example. Bulg J Plant Physiol, special issue: 37-51.

Krutovsky KV, Troggio M, Brown GR, Jermstad KD, Neale DB (2004) Comparative mapping in the Pinaceae. Genet 168: 447-461.

Kunkel BN, Brooks DM (2002) Cross talk between signaling pathways in pathogen defense. Curr Opin Plant Biol 5: 325-331.

Lapointe G, Luckevich MD, Seguin A (2001) Investigation of the induction of 14-3-3 in white spruce. Plant Cell Rep 20:79-84.

Lawrence SD, Dervinis C, Novak N, Davis JM (2006) Wound and insect herbivory responsive genes in poplar. Biotechnol Letters 28: 1493-1501.

Lee B, Henderson DA, Zhu JK (2005) The Arabidopsis cold-responsive transcriptome and its regulation by ICE1. Plant Cell 17: 3155-3175.

Lee J, Nam J, Park HC, Na G, Miura K, et al. (2007) Salicylic acid-mediated innate immunity in Arabidopsis is regulated by SIZ1 SUMO E3 ligase. Plant J 49: 79-90.

Lee I, Ambaru B, Thakkar P, Marcotte EM, Rhee SY (2010) Rational association of genes with traits using a genome-scale gene network for Arabidopsis thaliana. Nat Biotechnol 28: 149-156.

Li G, Asiegbu FO (2004) Induction of *Pinus sylvestris* PsACRE, a homology of Avr9/Cf-9 rapidly elicited defense-related gene following infection with root rot fungus *Heterobasidion annosum*. Plant Sci 167: 535-540.

Liepman AH, Nairn CJ, Willats WG, Sorenson I, Roberts AW, et al. (2007) Functional genomic analysis supports conservation of function among cellulose synthase-like A gene family members and suggests diverse roles of mannans in plants. Plant Physiol 143: 1881-1893.

Liu JJ, Ekramoddoullah AKM (2004) Characterization, expression and evolution of two novel subfamilies of *Pinus monticola* cDNAs encoding pathogenesis-related (PR)-10 proteins. Tree Physiol 24: 1377-1385.

Logan JA, Régnière J, Powell JA (2003) Assessing the impacts of global warming on forest pest dynamics. Front Ecol Environ 1: 130-137.

Long JA, Ohno C, Smith ZR, Meyerowitz EM (2006) TOPLESS regulates apical embryonic fate in Arabidopsis. Sci 312: 1520-1523.

Loopstra CA, Sathyan P (2004) Genes induced by water-deficit-stress are differentially expressed in two populations of Aleppo pine (*Pinus halepenensis*). Tree Genet & Genomes 1:166-173.

López-Maury L, Marguerat S, Bahler J (2008) Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. Nat Rev Genet 9: 583-593.

Lorenz WW, Sun F, Liang C, Kolychev D, Wang H (2006) Water stress-responsive genes in loblolly pine (*Pinus taeda*) roots identified by analyses of expressed sequence tag libraries. Tree Physiol 26: 1-16.

Lu PL, Chen NZ, An R, Su Z, Qi BS, Ren F, et al. (2007) A novel drought-inducible gene, ATAF1, encodes a NAC family protein that negatively regulates the expression of stress-responsive genes in Arabidopsis. Plant Mol Biol 63: 289-305.

Lu S, Sun YH, Amerson H, Chiang VL (2007) MicroRNAs in loblolly pine (*Pinus taeda* L.) and their association with fusiform rust gall development. Plant J 51: 1077-1098.

Lumley T, Miller A (2004) Leaps: regression subset selection. R package version 2.7. Available via http://cran.r-project.org/web/packages/leaps/index.html. Accessed 3 Sep 2010.

Lütz C, Heinzmann U, Gülz PG (1990) Surface structures and epicuticular wax composition of spruce needles after long-term treatment with ozone and acid mist. Environ Pollut 64: 312-322.

Marjamaa K, Hilden K, Kukkola E, Lehtonen M, Holkeri H, et al. (2006) Cloning, characterization and localization of three novel class III peroxidases in lignifying xylem of Norway spruce (*Picea abies*). Plant Mol Biol 61: 719-732.

Martin DM, Faldt J, Bohlmann J (2004) Functional characterization of nine Norway spruce *TPS* genes and evolution of gymnosperm terpene synthases of the *TPS-d* subfamily. Plant Physiol 135: 1908-1927.

Masle J, Gilmore SR, Farquhar GD (2005) The ERECTA gene regulates plant transpiration efficiency in Arabidopsis. Nat 436: 866-870.

Matos AR, Gigon A, Laffray D, Petres S, Zuily-Fodil Y, Pham-Thi AT (2008) Effects of progressive drought stress on the expression of patatin-like lipid acyl hydrolase genes in Arabidopsis leaves. Physiol Planet 134: 110-120.

McCune B, Mefford MJ (1999) Multivariate analysis of ecological data, version 4.26. MjM Software Design, Gleneden Beach, Oregon.

McKay SAB, Hunter WL, Godard KA, Wang SX, Martin DM, et al. (2003) Insect attack and wounding induced traumatic resin duct development and gene expression of (-)-pinene syntase in Sitka spruce. Plant Physiol 133: 368-378.

McKay SAB, Godard KA, Toudefallah M, Martin DM, Alfaro R, et al. (2006) Wound-induced terpene synthase gene expression in Sitka spruce that exhibit resistance or susceptibility to attach by the white pine weevil. Plant Physiol 140: 1009-1021.

McKeand S, Mullin T, Byram T, White T (2003) Deployment of genetically improved loblolly pine and slash pines in the south. J For 101: 32-37.

Merlot S, Gosti F, Guerrier D, Vavasseur A, Giraudat J (2001) The ABI1 and ABI2 protein phosphatases 2C act in a negative feedback regulatory loop of the abscisic acid signaling pathway. Plant J 25: 295-303.

Milla R, Maurer A, Rodriguez-Huete A, Gustafson JP (2003) Glutathione peroxidase genes in *Arabidopsis* are ubiquitous and regulated by abiotic stresses through diverse signaling pathways. Plant J 36: 602-615.

Miyamoto Y, Saiwaki T, Yamashita J, Yasuda Y, Kotera I, et al. (2004) Cellular stresses induce the nuclear accumulation of importin α and cause a conventional nuclear import block. JCB 165: 617-623.

Morgenstern EK (1996) Geographic variation in forest trees. UBC Press, Vancouver, BC.

Mori IC, Murata Y, Yang Y, Munemasa S, Wang YF, et al. (2006) CDPKs CPK6 and CPK3 function in ABA regulation of guard cell s-type anion- and Ca^{2+} permeable channels and stomatal closure. PLoS Biol 4: 1749-1762.

Morse AM, Nelson CD, Covert SF, Holiday AG, Smith KE, et al. (2004) Pine genes regulated by the necrotrophic pathogen *Fusarium circinatum*. Theor Appl Genet 109: 922-923.

Mukhopadhyay A, Vij S, Tyagi AK (2004) Overexpression of a zinc-finger protein gene from rice confers tolerance to cold, dehydration, and salt stress in transgenic tobacco. PNAS. 101: 6309-6314.

Munemasa S, Oda K, Watanabe-Sugimoto M, Nakamura Y, Shimoishi Y, et al. (2007) The coronatine-insensitive 1 mutation reveals the hormonal signaling interaction between abscisic acid and methyl jasmonate in Arabidopsis guard cells. Specific impairment of ion channel activation and second messenger production. Plant Physiol 143: 1398-1407.

Neale DB, Savolainen O (2004) Association genetics of complex traits in conifers. Trends Plant Sci 9: 325-330.

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

Nishizawa A, Yabuta Y, Yoshida E, Maruta T, Yoshimura K, et al. (2006) Arabidopsis heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. Plant J 48: 535-547.

Ohashi-Ito K, Fukuda H (2003) HD-Zip III homeobox genes that include a novel member, *ZeHB-13 (Zinnia)/ATHB-15 Arabidopsis)*, are involved in procambium and xylem cell differentiation. Plant Cell Physiol 44: 1350-1358.

Oleksiak MF, Churchill GA, Crawford DL (2002) Variation in gene expression within and among natural populations. Nat Genet 32: 261-266.

Palle S, Seeve C, Eckert A, Cumbie W, Goldfarb B, Loopstra C (2010) Natural variation in expression of genes involved in xylem development in loblolly pine (*Pinus taeda* L.). Tree Genet & Genomes 6: 1-14.

Pastori GM, Foyer CH (2002) Common components, networks, and pathways of cross-tolerance to stress. The central role of "redox" and abscisic acid-mediated controls. Plant Physiol 129: 460-468.

Pauw B, Memlink J (2004) Jasmonate-responsive gene expression. J Plant Growth Regul 23: 200-210.

Peirson SN, Butler JN, Foster RG (2003) Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. Nucleic Acids Res 31: e73.

Pihakaski-Maunsbauch K, Moffatt B, Testillano P, Risueno M, Yeh S, Griffith M, Maunsbach AB (2001) Genes encoding chitinase-antifreeze proteins are regulated by cold and expressed by all cell types in winter rye shoots. Physiol Plant 112: 359-371.

Phillips MA, Wildung MR, Williams DC, Hyatt DC, Croteau R (2003) cDNA isolation, functional expression, and characterization of (+)-α-pinene synthase and (-)-α-pinene synthase from loblolly pine (*Pinus taeda*): Stereocontrol in pinene biosynthesis. Arch Biochem Biophys 411: 267-276.

Piggott N, Ekramoddaoullah AKM, Liu JJ, Yu X (2004) Gene cloning of a thaumatin-like (PR-5) protein of western white pine (*Pinus monticola* D. Don) and expression studies of members of the PR-5 group. Physiol & Mol Plant Pathol 64: 1-8.

Pritchard JK, Rosenberg NA (1999) Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet 65: 220-228.

Quesada T, Gopal V, Cumbie WP, Eckert AJ, Wegrzyn JL (2010) Association mapping of quantitative disease resistance in a natural population of loblolly pine (*Pinus taeda* L.). Genet (Epub ahead of print).

R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available via http://www.r-project.org/. Accessed 11 Nov 2010.

Ralph SG, Park JY, Bohlmann J, Mansfield SD (2006a) Dirigent proteins in conifer defense: gene discover, phylogeny, and differential wound- and insect-induced expression of a family of DIR and DIR-like genes in spruce (Picea spp.). Plant Mol Biol 60: 21-40.

Ralph SG, Yueh H, Friedman M, Aeschliman D, Zeznik JA, et al. (2006b) Conifer defence against insects: microaaray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobe*) reveals large-scale changes of the host transcriptome. Plant Cell Environ 29: 1545-1570.

Ramanjulu S, Bartels D (2002) Drought- and desiccation-induced modulation of gene expression in plants. Plant Cell Environ 25: 141-151.

Ro DK, Arimura G, Lau SY, Piers E, Bohlmann J (2005) Loblolly pine abietadienol/abietadienal oxidase PtAO (CYP720B1) is a multifunctional, multisubstrate cytochrome p450. PNAS 22: 8060-8065.

Robatzek S, Somssich EI (2002) Targets of *At*WRKY6 regulation during plant senescence and pathogen defense. Genes Dev 16: 1139-1149.

Ryals JA, Neuenschwander UH, Willits MG, Molina A, Steiner HY, Hunt MD (1996) Systemic acquired resistance. Plant Cell 8: 1809-1819.

Sackton TB, Lazzaro BP, Clark AG (2010) Genotype and gene expression associations with immune function in *Drosophila*. PLoS Genet 6: 1-15.

Sairam RK, Tyagi A (2004) Physiology and molecular biology of salinity stress tolerance in plants. Curr Sci 86: 407-421.

Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, et al. (2002) DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration-and cold-inducible gene expression. Biochem Biophys Res Commun 290: 998-1009.

Sanchez P, Zabala MT, Grant M (2000) AtBI-1, a plant homologue of Bax Inhibitor-1, suppresses Bax-induced cell death in yeast and is rapidly upregulated during wounding and pathogen challenge. Plant J 21: 393-399.

Schmidt RA, Wilkinson RC, Moses CS, Broerman FS (1976) Drought and weevils associated with severe incidence of pitch canker in Volusia County, Florida. Ins Food and Agric Sci, Univ of Fla Prog Rep 76: 4 pp.

Schopmeyer CS (1939) Transpiration and physico-chemical properties of leaves as related to drought resistance in loblolly pine and shortleaf pine. Plant Physiol 14: 447-462.

Schultz RP (1997) Loblolly pine: The ecology and culture of loblolly pine (*Pinus taeda* L.). U.S. Dept of Agric, Forest Service, Washington, D.C.

Seo PJ, Lee AK, Xiang F, Park CM (2008) Molecular and functional profiling of Arabidopsis pathogenesis-related genes: insights into their roles in salt response of seed germination. Plant Cell Physiol 49: 334-344.

Shah J (2005) Lipids, lipases, and lipid-modifying enzymes in plant disease resistance. Annu Rev Phytopathol 43: 229-260.

Sharp RE, LeNoble ME (2002) ABA, ethylene and the control of shoot and root growth under water stress. J Exp Botany 53: 33-37.

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

Shinozaki K, Yamaguchi-Shinozaki K, Seki M (2003) Regulatory network of gene expression in the drought and cold stress responses. Curr Opin Plant Biol 6: 410-417.

Shinozaki K, Yamaguchi-Shinozaki K, Seki M (2007) Gene networks involved in drought stress response and tolerance. J Exp Biol 58: 221-227.

Slonim DK (2002) From patterns to pathways: gene expression data analysis comes of age. Nat Genet 32: 502-508.

Song Y, Wang L, Xiong L (2009) Comprehensive expression profiling analysis of OsIAA gene family in developmental processes and in response to phytohormone and stress treatments. Planta 299: 577-591.

Stahlberg A, Elbing K, Andrade-Garda JM, Sjogreen B, Forootan A, Kubista M (2008) Multiway real-time PCR gene expression profiling in yeast *Saccharomyces cerevisiae* reveals altered transcriptional response of ADH-genes to glucose stimuli. BMC Genomics 9:170-184.

Storey JD, Tibshirani R (2003) Statistical significance for genome-wide experiments. PNAS 100: 9440-9445.

Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, et al. (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. Sci 315: 848-853.

Staubach F, Teschke M, Voolstra CR, Wolf JBW, Tautz D (2009) A test of the netural model of expression change in natural populations of house mouse subspecies. Evol 64: 549-560.

Strauss SH, Lande R, Namkoong G (1992) Limitations of molecular-marker-aided selection in forest tree breeding. Can J For Res 22: 1050-1061.

Suzuki R, Shimodaira H (2006) Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinforma 22: 1540-1542.

Sweetlove LJ, Heazlewood JL, Herald V, Holzapffel R, Day DA, Leaver CJ, et al. (2002) The impact of oxidative stress on Arabidopsis mitochondria. Plant J 32: 891-904.

Szemenyei H, Hannon M, Long JA (2008) TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. Sci 319: 1384-1386.

Taji T, Ohsumi C, Iuchi S, Seki M, Kasuga M, et al. (2002) Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana. Plant J 29: 417-426.

Teskey RO, Bongarten BC, Cregg BM, Dougherty PM, Hennessey TC (1987) Physiology and genetics of tree growth response to moisture and temperature stress: an examination of the characteristic of loblolly pine (*Pinus taeda* L.). Tree Physiol 3: 41-61.

Tiwari SB, Wang XJ, Hagen G, Guilfoyle TJ (2001) Aux/IAA proteins are active repressors, and their stability and activity are modulated by auxin. Plant Cell 13: 2809-2822.

Tiwari SB (2004) Aux/IAA proteins contain a potent transcriptional respression domain. Plant Cell 16: 533-543.

Tognolli M, Penel C, Greppin H, Simon P (2002) Analysis and expression of the class III peroxidase large gene family in *Arabidopsis thaliana*. Gene 288: 129-138.

VanGuilder HD, Vrana KE, Freeman WM (2008) Twenty-five years of quantitative PCR for gene expression analysis. BioTechniques, 44: 619-626.

Veyrieras JB, Camus-Kulandaivelu L, Gouesnard B, Manicacci D, Charcosset A (2007) Bridging genomics and genetic diversity: linkage disequilibrium structure and association mapping in maize and other cereals. Crop Sci 47: S60-S71.

Wakamiya I, Newton RJ, Johnston JS, Price HJ (1993) Genome size and environmental factors in the genus *Pinus*. Am J Botany 80: 1235-1241.

Walley JW, Dehesh K (2010) Molecular mechanisms regulating rapid stress signaling networks in *Arabidopsis*. J Integr Plant Biol 52: 354-359.

Wallis C, Eyles A, Chorbadjian R, Gardener M, Hansen R, et al. (2008) Systemic induction of phloem secondary metabolism and its relationship to resistance to a canker pathogen in Austrian pine. New Phytologist 177: 767-778.

Wang H, Azuaje F, Bodenreider O, Dopazo J (2004) Gene expression correlation and gene ontology-based similarity: an assessment of quantitative relationships. IEEE Symp Comput Intell Bioninforma Comput Biol Proc, pp 25-31.

Watkinson JI, Sioson AA, Vasquez-Robinet C, Shukla M, Kumar D, et al. (2003) Photosynthetic acclimation is reflected in specific patterns of gene expression in drought-stressed loblolly pine. Plant Physiol. 133: 1702-1716.

Wegener A, Gimbel W, Werner T, Hani J, Ernst D, Sandermann H Jr (1997) Molecular cloning of ozone-inducible protein from *Pinus sylvestris* L. with high sequence similarity to vertebrate 3-hydroxy-3-methylglutaryl-CoA-synthase. Biochimica et Biophysica Acta 1350: 247-252.

Wilkie GS, Dickson KS, Gray NK (2003) Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. Trends Biochem Sci 28: 182-188.

Wingfield MJ, Hammerbacher A, Ganley RJ, Steenkamp ET, Gordon TR, et al. (2008) Pitch canker caused by *Fusarium circinatum*—a growing threat to pine plantations and forests worldwide. Australasian Plant Pathol 37: 319-334.

Winnett SM (1998) Potential effects of climate change on U.S. forests: a review. Clim Res 11: 39-49.

Wu H, Echt CS, Popp MP, Davis JM (1997) Molecular cloning, structure and expression of an elicitor-inducible chitinase gene from pine trees. Plant Mol Biol 33: 979-987.

Wu H, Michler CH, LaRussa L, Davis JM (1999) The pine *Pschi4* promoter directs wound-induced transcription. Plant Sci 142: 199-207.

Xie DX, Feys BF, James S, Nieto-Rostro M, Turner JG (1998) COI1: an Arabidopsis gene required for jasmonate-regulated defense and fertility. Sci 280: 1091-1094.

Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA (2003) Role of anthocyanidin reductase, encoded by BANYULS in plant flavanoid biosynthesis. Sci 299: 396-399.

Xiong L, Lee H, Ishitani M, Zhu JK (2002) Regulation of osmotic stress-responsive gene expression by the LOS6/ABA1 locus in Arabidopsis. J Biol Chem 277: 8588-8596.

Xu Z, Escamilla-Trevino L, Zeng L, Lalgondar M, Bevan D, Winkel B, et al. (2004) Functional genomic analysis of Arabidopsis thaliana glycoside hydrolase family 1. Plant Mol Biol 55: 343-367.

Yan J, He C, Zhang H (2003) The BAG-family proteins in *Arabidopsis thaliana*. Plant Sci 165: 1-7.

Yang Y, Shah J, Klessig DF (1997) Signal perception and transduction in plant defense responses. Genes Dev 11: 1621-1639.

Yang L, Tang R, Zhu J, Liu H, Mueller-Roeber B, et al. (2008) Enhancement of stress tolerance in transgenic tobacco plants constitutively expressing AtIpk2beta, an inositol polyphosphate 6-/3-kinase from Arabidopsis thaliana. Plant Mol Biol 66: 329-343.

Yanhui C, Xiaoyuan Y, Kun H, Meihua L, Jigang L, et al. (2006) The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. Plant Mol Biol 60: 107-124.

Yu J, Pressoir G, Briggs WH, Bi IV, Yamasaki M, et al. (2005) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38: 203-208.

Yuan J, Zhu M, Lightfoot DA, Iqbal MJ, Yang JY, et al. (2008) In silico comparison of transcript abundances during *Arabidopsis thaliana* and Glycine max resistance to *Fusarium virguliforme*. BMC Genomics 9: S6.

Zhu J, Verslues PE, Zheng X, Lee BH, Zhan X, et al. (2005) HOS10 encodes an R2R3-type MYB transcription factor essential for cold acclimation in plants. PNAS 102: 9966-9971.

APPENDIX A

Table A.1 Genes included in all expression analyses. Genes are listed with their Unigene number, primer sequences, and the study that they were originally identified in.

Gene name and putative function	Unigene	Primer sequence	Reference
ALDH7B4 Aldehyde dehydrogenase/ oxidoreductase	Pta. 8865	F: 5'-ACGCTGTTAGGACCACTGCATA-3' R: 5'-AGAATGCACCCACCCTGAGATT-3'	Kotchoni and Bartels (2003)
PtMYB12 MYB transcription factor	Pta. 15187	F: 5'-ACATCTTTCCACCACGAGTCCT-3' R: 5'-TGTGGATTATCGCAGCACGA-3'	Chapter II
PtNAP Non-intrinsic ABC protein 4	Pta. 10827	F: 5'-GAATTGGACTGGAAATCATCGC-3' R: 5'-CATTAGGTGGGGGAGGTAGAGG-3'	Chapter II
CAB Ubiquitin protein ligase/zinc ion binding	Pta. 6152	F: 5'-CAGTCCAGAGTCCAGACACCA-3' R: 5'-TGCCTCTGGCCATACTAGATGA-3'	Kanyuka et al. (2003)
cMyc C-Myc binding protein	Pta. 111354	F: 5'-CTGTCGATGGGTTGATGGATTT-3' R: 5'-GGATTGCTTGCATCTCATGAAA-3'	Chapter II
GRAS GRAS family transcription factor	Pta. 7913	F: 5'-GGAGTAGGCAGCAACGCTTT-3' R: 5'-GCCAACCGAGGAGCAGAAT-3'	Chapter II
IAA2 Auxin-induced transcription factor	Pta. 11474	F: 5'-CACGATCGAGACCAAATCAAGA-3' R: 5'-TGAACCTTGCGATAATGAACGT-3'	Chapter II
IAA3 Auxin-induced transcription factor	Pta. 8823	F: 5'-AGCCAATGACCCTCCAAGAAT-3' R: 5'-GACAGGTGGCCATCCAACA-3'	Chapter II
IPK Inositol polyphosphate kinase	Pta. 21742	F: 5'-AACGTTTCAATGCCACCCTAA-3' R: 5'-CACCTGGAAGGCAGACAAGAA-3'	Yang et al. (2008)
MYND Zinc finger (MYND type) family protein	Pta. 11048	F: 5'-CCGGATGTTACGATTTTGGATT-3' R: 5'-CGTCCAGCGCGTGAATTT-3'	Chapter II
PtWRKY2 WRKY transcription factor	Pta. 12744	F: 5'-ACATACACCGTCTGCCTGCTT-3' R: 5'-CAGATCTGTGAAGAGCCACGAA-3'	Chapter II
PtMYB13 MYB transcription factor	Pta. 18610	F: 5'-GCTTCCGCGTTTCAGATCAG-3' R: 5'-TGCTGCGATGTGGGAAGAG-3'	Zhu et al. (2005)
PtPAL Phenylalanine ammonia-lyase	Pta. 2030	F: 5'-CGCCGAACAGCATAACCAG-3' R: 5'-GATGGCCTCGGCAGATTTT-3'	Chapter II
TTA7 Flavanoid 3'-monooxygenase	Pta. 15557	F: 5'-ATGCCCCATCACTCGCTCTA-3' R: 5'-GGTTGTCCCGAGCTTCAAATAC-3'	Chapter II
ABA1 Zeaxanthin epoxidase	Pta. 4440	F: 5'-TGTTGTGCACTGGAGGATTCA-3' R: 5'-ATGTGGGCAAACAAAGCCAA-3'	Xiong et al. (2002)
PR-1 Pathogenesis-related anti-fungal protein	Pta. 15491	F: 5'-GCTGCGCTCAAGCTCAATG-3' R: 5'-CTACATAGTTGCCAGGCGGATC-3'	Lee et al. (2007)

Table A.1 continued.

Gene name and putative function	Unigene	Primer sequence	Reference
ATT1 CYP86A subfamily cytochrome P450	Pta. 16436	F: 5'-GATCTGCCTGGGTAAGGACTTG-3' F: 5'- GATGGCGCAGTAGAATTGCA -3'	Kosma et al. (2009)
KCS4 3-ketoacyl-CoA synthase	Pta. 3828	F: 5'-ATGGAGCACTCTGAGCTTTCG-3' F: 5'-AAACCCGATCGCTCCAGTATC-3'	Kosma et al. (2009)
GPAT6 Glycerol-3-phosphate acyltransferase	Pta. 3703	F: 5'-GCAACATTCCCATCCCAGAG-3' R: 5'-TTCCCTTGACGACCAGCTTG-3'	Kosma et al. (2009)
LACS6 Long-chain acyl-CoA synthetase	Pta. 16871	F: 5'-TTCCGTTGGCACACATTTATG-3' R: 5'-GGTAAAATCCCATGGCAGCTC-3'	Kosma et al. (2009)
LACS9 Long-chain acyl-CoA synthetase	Pta. 7897	F: 5'-TCAAGGTGGACGAGAGAGGAA-3' R: 5'-CAAGGCAGCCATCAGGATG-3'	Kosma et al. (2009)
CER1 *Eceriferum* wax biosynthetic gene	Pta. 2916	F: 5'-GGGCTGATTTACCTTGCGATT-3' R: 5'-TCCTTGTGAAACGCACGATG-3'	Kosma et al. (2009)
CER4 Fatty acyl-CoA reductase	Pta. 17144	F: 5'-GATGTGGCCTTGGAAACGAA-3' R: 5'-TTGGGATTTGTCGCAGCAT-3'	Kosma et al. (2009)
WBC11 ATPase/fatty acid transporter	Pta. 23058	F: 5'-ATGCTGGTTGCTGGATTTTTTC-3' R: 5'- GGGATAGCGCCAAACTGGTT -3'	Kosma et al. (2009)
WBC12 ABC transporter	Pta. 19540	F: 5'-TTGGAGGAAGAATTCAAGGGAA-3' R: 5'-GGGCATTGCTGTAAGGAACC-3'	Kosma et al. (2009)
WIN1 Ethylene responsive transcription factor	Pta. 10578	F: 5'-TCCACGCTCACTGCCAAAC-3' R: 5'-CTTCTTCGATGAGCCCGATC-3'	Kosma et al. (2009)
ADR1 Activated disease resistance protein	Pta. 9635	F: 5'-CGGTTCAACAACATCCAGTCG-3' R: 5'-CTTGGACCACTGCTGTGATTTG-3'	Chini et al. (2004)
EB9D 14-3-3 protein	Pta.11581	F: 5'-GTTTAAGACAGGCGCCGAGA-3' R: 5'-CAAGTCTGCGCAGGATATTGC-3'	Lapointe et al. (2001)
CYPA Cytochrome p450 monooxygenase	Pta.26	F: 5'-TTGATGGCCAAACAACTGCTCG-3' R: 5'-ATGCAGTGATCGCCATTCCCAT-3'	Ro et al. (2005)
CYPB Cytochrome p450 monooxygenase	Pta.6578	F: 5'-TTGTTGTGTCCGTGGATCCA-3' R: 5'-AGGCAATTTCAGGCCAACTATC-3'	Ro et al. (2005)
CYPC Cytochrome p450 monooxygenase	Pta.15878	F: 5'-GCCTCATGTTCAAGACTTCGTG-3' R: 5'-CCAAAATTACACGGGAAACCAT-3'	Ro et al. (2005)
CYPD Cytochrome p450 monooxygenase	Pta.4611	F: 5'-TTGCTGCTCGTGACACAACA-3' R: 5'-TGCAAGCATCAGCATGTACAAG-3'	Ro et al. (2005)
AOS Allene oxide synthase	Pta.5467	F: 5'-AACAGAATCCCCGACCGTTAG-3' R: 5'-ATGATCGCTCGGCTGTTTGT-3'	Lawrence et al. (2006) Matos et al. (2008)
ATAF-1 NAC transcription factor	Pta.19461	F: 5'-CTGGCAAACCAGCTCATTCAG-3' R: 5'-CAACTTGGGTCTGGACATTGTG-3'	Lu PL et al. (2007)

Table A.1 continued.

Gene name and putative function	Unigene	Primer sequence	Reference
PSCHI4 Extracellular chitinase	Pta.3891	F: 5'-GAGTGGCGGGATATGGAATG-3' R: 5'-CACAGGCACAATAAGCGACG-3'	Wu et al. (1997) Wu et al. (1999)
CPK3 Calcium-dependent protein kinase	Pta.15829	F: 5'-CCGCCGATTTGACTAGGACGAT-3' R: 5'-GGAGTTATGCACCGGGACTTGA-3'	Mori et al. (2006)
PtGSTU18 Classs tau glutathione *S*-transferase	Pta.16094	F: 5'-AAGATCTTTCCATGCATGCGA-3' R: 5'-GAAAGCAGTGGAAGAGAGCGTT-3'	Dixon et al. (2002)
CslA1 Cellulose synthase-like protein	Pta.205	F: 5'-TGGTTCTGGTGCAGATTCCCAT-3' R: 5'-ATTGGAGCAGCTTGTGGGCTTT-3'	Liepman et al. (2007) Hematy et al. (2009)
GLX1 Glyoxylase	Pta.11449	F: 5'-TGCAGAGGCTCTTCGGGTAGTT-3' R: 5'-AACCAGGGCCAATACCAGGGAT-3'	Sairam and Tyagi (2004)
HMG-CoA Hydroxymethylglutaryl-CoA-synthase	Pta.4795	F: 5'-AACCCTTTGCTGGCTTGTCTGA-3' R: 5'-AAGGTTTCTCAGCAGCTTGCGA-3'	Wegener et al. (1997)
HPL Hydroperoxide lyase	Pta. 18452	F: F'-TGTATCGTGAACCGAAATCCC-3' R: R'-GCGTTAAAAAGTGGCTTGCG-3'	Feussner and Wasternack (2002)
PtEMB1 Late embryogenesis abundant protein	Pta.6938	F: F'-GCCAGAGACGCTGTGAAACAAA-3' R: F'-AAATTACCCGAAGCCGAAAGAGGG-3'	Dong and Dunstan (1999)
PtEMB2 Late embryogenesis abundant protein	Pta.7648	F: 5'-TGAAACCCTCCCCTGTACTCG-3' R: 5'-ACCTGCCGTCTTTGCCTTTAG-3'	Dong and Dunstan (1999)
PtEMB3 Late embryogenesis abundant protein	Pta.49	F: 5'-GGCGGAGGGAATACAAAGAGA-3' R: 5'-GACTGGATTCCGGAGGATCAC-3'	Dong and Dunstan (1999)
PtEMB4 Late embryogenesis abundant protein	Pta. 2123	F: 5'-TACTCTCGGATCCCTCCTGGTA-3' R: 5'-CGACTATAAAGCATCGCGAGC-3'	Dong and Dunstan (1999)
(-)-α-pinene synthase Monoterpene synthase	Pta.117	F: 5'-GCTCTGGTTCGCCTCAAATGAC-3' R: 5'-ACTTTGGCTTCTTGCATTGCGT-3'	Phillips et al. (2003)
(+)-α-pinene synthase Monoterpene synthase	Pta.116	F: 5'-AGTTATTGGACCGAAAAAGGCA-3' R: 5'-TCAACTCAACTGCCTTGGGC-3'	Phillips et al. (2003)
α-terpineol-synthase Monoterpene synthase	Pta. 117	F: 5'-GTGGATCGCCACTGGTTTTC-3' R: 5'-CTTGGAGAACGGGAAAGTTAGC-3'	Phillips et al. (2003)
PX-1 Class III secretory peroxidase	Pta.1699	F: 5'-TGACCCAACATTGGACACCA-3' R: 5'-ATGGCACCGACGACAATCA-3'	Marjamaa et al. (2006) Tognolli et al. (2002)
PX-2 Class III secretory peroxidase	Pta.5586	F: 5'-GCCTAAAATTCGCCACACAAA-3' R: 5'-TTTAAGGTGACAGCGCTCATCA-3'	Marjamaa et al. (2006) Tognolli et al. (2002)
PX-3 Class III secretory peroxidase	Pta.18	F: 5'-TGTGAACGCTCTGCCAACTC-3' R: 5'-AGTTGCCCGTCATTGGAGTC-3'	Marjamaa et al. (2006) Tognolli et al. (2002)
PX-4 Class III secretory peroxidase	Pta.3054	F: 5'-GCAGGCCCAAATGCTAATTC-3' R: 5'-TCAAGTTGAGGCAGCCTGC-3'	Marjamaa et al. (2006) Tognolli et al. (2002)

Table A.1 continued.

Gene name and putative function	Unigene	Primer sequence	Reference
PX-5 Class III secretory peroxidase	Pta.20723	F: 5'-CTTCACTTTCACGATTGCTTCG-3' R: 5'-GACAACGCTACGTTTACCGGA-3'	Marjamaa et al. (2006) Tognolli et al. (2002)
PR-10 Pathogenesis-related family 10 protein	Pta.13322	F: 5'-AGATGGTGTCAGGGACTGCAA-3' R: 5'-ACTATGGAACGCCACGGTGA-3'	Liu and Ekramoddoullah (2004)
SamCMT Jasmonic acid carboxyl methyltransferase	Pta.18552	F: 5'-ACGGAAGGCTATTTCCAGACAA-3' R: 5'-CTTCACTGGTTTTCCCAGGTTC-3'	Seo et al. (2001)
STR_SYN Jasmonate-responsive alkaloid biosynthesis enzyme	Pta.16305	F: 5'-CTGCTTGTGGTTGGTCCTGA-3' R: 5'-GATGGGGTTCCTTTGAGGTTC-3'	Pauw and Memelink (2004)
PtWRKY WRKY transcription factor	Pta.15442	F: 5'-TGTTGATCAAACCGAAGCCA-3' R: 5'-TCCGAGGCTCCCATGATAAG-3'	Robatzek and Somssich (2002)
Avr-cf9 Avr9/Cf-9 rapidly elicited LRR domain-containing protein	Pta.295	F: 5'-TCATGATGGAACGAGGTAAAG-3' R: 5'-CATCATTCCAGAGGCAACAATC-3'	Li and Asiegbu (2004)
PtBAG-1 Class I Bcl-2 associated athanogene	Pta.2860	F: 5'-GATGTTCTGAAGGTTCGCAATG-3' R: 5'-AGTGGTGACAACAAAGTGGGAA-3'	Yan et al. (2003)
PtBAG-4 Class I Bcl-2 associated athanogene	Pta.3028	F: 5'-GGCAATTTCTGCAGTCAAGGCT-3' R: 5'-AGGCAGCTGTGAATGGTGGAA-3'	Kabbage and Dickman (2008)
BI-1 Bcl-2 asociated (Bax)-inhibitor protein	Pta.5096	F: 5'-TTTGGAATATTGGCGGTCTCC-3' R: 5'-GGCTTAAGTCCATCCCTGCTT-3'	Sanchez et al. (2000)
HDZ32 Homeodomain/leucine-zipper protein	Pta.5267	F: 5'-TGCGTGGTGTTGTAGCATCC-3' R: 5'-ACTGGGTTCACATCTTGGGC-3'	Ohashi-Ito and Fukuda (2003)
ERD3 Early-responsive to dehydration protein	Pta.3290	F: 5'-TGGCGCCCAGAGATAATCAT-3' R: 5'-TCCTGCAATGCTGGGTATCA-3'	Shionzaki et al. (2003)
ERF1 Ethylene-responsive transcription factor	Pta.13277	F: 5'-AAGGGTTGGACGCCGTTAAC-3' R: 5'-TAAATCCTCGGCAGGAGACG-3'	Shinozaki et al. (2003)
PtGPX1 Glutathione peroxidase	Pta.2451	F: 5'-TGCATTCCCTTGCAACCAAT-3' R: 5'-AAAGGAATTTGCATGCACCAG-3'	Milla et al. (2003)
PtGPX2 Glutathione peroxidase	Pta.1107	F: 5'-ACAGTCAAGGACATCCGTGGTA-3' R: 5'-TTTGCTCATTGTCAACGTTGC-3'	Milla et al. (2003)
PtGPX3 Glutathione peroxidase	Pta.7951	F: 5'-GCCTTCCCATGCAATCAGTT-3' R: 5'-CGGAAAAGATCTGCACTCGC-3'	Milla et al. (2003)
PT1 Phenylcoumaran benzylic ether reductase	Pta.10991	F: 5'-AATCAGGGCAGTGGACGATC-3' R: 5'-TCCTGCCAACACTCTGTCCTT-3'	Gang et al. (1999)
Sesqui-TPS Sesquiterpene synthase	Pta.18448	F: 5'-AAAGAATGCCGTCGCTAGCTCA-3' R: 5'-TCCCGAACAACATTCTGCAGCA-3'	Martin et al. (2004)

Table A.1 continued.

Gene name and putative function	Unigene	Primer sequence	Reference
PtMLO2 Mlo-family calmodulin binding protein	Pta.7890	F: 5'-GCACTTGTGACACAGATGGGTT-3' R: 5'-GCAGACATCCAATGCAGTGAAG-3'	Buschges et al. (1997)
PtMYB6 MYB transcription factor	Pta.11841	F: 5'-AGCTCCAAAGCCATGTCTGC-3' R: 5'-AAGGTAACGGCATTGAATCCC-3'	Du et al. (2009)
bZIP bZIP domain transcription factor	Pta.18030	F: 5'-AATGAACTGAGGTCGGCAGTG-3' R: 5'-CTGGTTGAAGGTGTCATGGGA-3'	Kesarwani et al. (2007)
PtPDIR1 Dirigent-like protein	Pta.9601	F: 5'-AAGAATGCGGCTCTGGAACA-3' R: 5'-TATTTCGGTTTGCTCGAGGC-3'	Ralph et al. (2006a)
CHS Naringenin-chalcone synthase	Pta.15022	F: 5'-GTGCGTGCGTCCACTTCATA-3' R: 5'-ATGTTCAACCAGCGGAGAGG-3'	Dixon and Paiva (1995)
PR-2 β-1,3-glucanase	Pta.1332	F: 5'-CCTTCTATCCCGCCACCAAT-3' R: 5'-AAGCACGACGTACGGATCGT-3'	Ryals et al. (1996)
PR-3 Basic chitinase	Pta.24081	F: 5'-ACAACCCCGACATTGTTGCT-3' R: 5'-TTCTGGATGACCGCCCAAT-3'	Loopstra and Sathyan (2004)
PR-5 Thaumatin-like protein	Pta.13241	F: 5'-AAGGAAGCTGCAGCACTGGT-3' R: 5'-GGACAGTCTCCTACAACGCTGG-3'	Piggott et al. (2004)
PTI4 Ethylene-response element binding factor	Pta.17765	F: 5'-AAGGGCGCTAGAGTTTGGCTT-3' R: 5'-ATGACCGTGCCGCTTACAAGAT-3'	Fujimoto et al. (2000)
RAP2.1 AP2 domain-containing	Pta.4168	F: 5'-AGCACGTAAGGAGCAGCAGAG-3' R: 5'-ACCGTCCTCTATTCAGCGGAG-3'	Sakuma et al. (2002)
RAP2.4 AP2 domain-containing	Pta.4168	F: 5'-TTCATGGCAGCAGCCTGTC-3' R: 5'-CCACTGGGGTAAATGGGTTG-3'	Sakuma et al. (2002)
ABI1 Serine/threonine phosphatase 2C	Pta.8814	F: 5'-ATGGCAAGGATTGAGGCAGCA-3' R: 5'-CTGGCAATGTCAAGGGCCATT-3'	Merlot et al. (2001)
BALDH Betaine aldehyde dehydrogenase	Pta.1047	F: 5'-ATAATGCTGACTTGCGCAACC-3' R: 5'-GGGTTAATTGTGCACAGCCAA-3'	Kotchoni and Bartels (2003)
CHI Chalcone isomerase	Pta.15647	F: 5'-TCAATCGCAGATCGGGTTTC-3' R: 5'-ACTACTGTGCTCAACGGCGTG-3'	Dixon and Paiva (1995)
PtMYB9 MYB domain transcription factor	Pta.1692	F: 5'-ATGCCAGATCGGGCTTACAGAT-3' R: 5'-AGACACGGTTGCATGAACAGCT-3'	Loopstra and Sathyan (2004)
OXR1 2OG-Fe(II) oxygenase family protein	Pta.5273	F: 5'-AGCTGTTGAGTCTGTTCGCTGA-3' R: 5'-AGGCCATGAATTCGATCCGCA-3'	Lee et al. (2005)
HSP17.4-CIII Class III heat shock protein	Pta.6499	F: 5'-TGCCAATGTCGAGGGCATTTCT-3' R: 5'-TTCCGCCTGCCATGAAGTTCAA-3'	Nishizawa et al. (2006)
NAC1 NAC domain transcription factor	Pta.17081	F: 5'-ACGAACTGGGTAATGCACGAGT-3' R: 5'-ACTGCAAGGATGAATGGGTGGT-3'	Hu et al. (2006)
NCED 9-*cis*-epoxycarotenoid dioxygenase	Pta.607	F: 5'-ACGGATTCCACGGCACATTCAT-3' R: 5'-TGTCTTCCGCCCTTCTTGCTTT-3'	Iuchi et al. (2001)

Table A.1 continued.

Gene name and putative function	Unigene	Primer sequence	Reference
PtAN1 AN1-like zinc finger protein	Pta.10959	F: 5'-CGAAACTCCTAACCGGTGCTT-3' R: 5'-CAAATGTCGGTGCGGAAATAC-3'	Mukhopadyay et al. (2004)
P5CDH 1-pyrroline-5-carboxylate dehydrogenase	Pta.11660	F: 5'-GATCCACGCATGACCCTTTT-3' R: 5'-TTGATCTGAAAGGGCGTGTG-3'	Deuschle et al. (2004)
RPK1 Protein kinase	Pta.5851	F: 5'-TATGCCAGCACAGGAATGCTGA-3' R: 5'-AGAAGCCCTGTTGATTGTGGCA-3'	Ramanjulu and Bartels (2002)
COI1 Ubiquitin protein ligase	Pta.2657	F: 5'-AAGAGCATGCGACAGAGACCA-3' R: 5'-TTGAGGCTGAGATGCCTGCTAT-3'	Xie et al. (1998)
ANR Anthocyanidin reductase	Pta.8026	F: 5'- CTGGGCGTTTTTCATGTCG-3' R: 5'-ATGATGTGATCAAGCCAGCG-3'	Xie et al. (2003)
ARF ADP-ribosylation factor	Pta.12493	F: 5'-GGTCTCGATAATGCCGGAAA-3' R: 5'-TGGTGGTCACGCATCCTACA-3'	Gillingham and Munro (2007)
BGLU1 Glycosyl hydrolase family 1 protein	Pta. 17476	F: 5'-CGGAATATTTCCACCGACGA-3' R: 5'-ACAGGGAATTGCTCGTCAGC-3'	Xu et al. (2004)
PtCPN10 Mitochondrial chaperonin	Pta.11459	F: 5'-CATATCAGTTGGGCCTGGTTCT-3' R: 5'-TGTCAAGGAGGGTGACCATGT-3'	Sweetlove et al. (2002)
ERECTA Transmembrane receptor protein kinase	Pta. 3189	F: 5'-CCAGTGTTGGCAGGTCACAA-3' R: 5'-TTCAGGGACGAGCCAAGAGT-3'	Masle et al. (2005)
PtGOLS1 Galactinol synthase	Pta. 7539	F: 5'-CTTCGCGGAGCAGGATTTAT-3' R: 5'-TCCCTTGGCCTACAATCTGGT-3'	Taji et al. (2002)
PtGSTF9 Class phi glutathione s-transferase	Pta. 14237	F: 5'-AGAGCAAATACTTGGCAGGCGA-3' R: 5'-ACACAATACCTCGTCACGGCTT-3'	Dixon et al. (2002)
IMPA Importin, alpha isoform	Pta.6960	F: 5'-AGCAACAACCCAATTCCGAA-3' R: 5'-ATGCAGGTGTTGTTCCTCGC-3'	Miyamoto et al. (2004)
PtOMT *o*-methyltransferase	Pta. 11924	F: 5'-TTGATCTGGTAATGGTGGCGCA-3' R: 5'-TTGGAAGGCGGGTTTAGTCGTT-3'	Goujon et al. (2003)
LP3 ABA/water deficit inducible protein	Pta. 447	F: 5'-AATTGGGTGGACTGGGAACTG-3' R: 5'-CGCATCGAAGAAGGATCCAG-3'	Chang et al. (1996)
OXR2 2OG-Fe(II) oxygenase family protein	Pta. 8739	F: 5'-AATCTGGGTCAGCTGCCTCA-3' R: 5'-ACCAGATCTTGGACTGGAGGG-3'	Yuan et al. 2008
P5CR Pyrroline-5-carboxylate reductase	Pta. 4812	F: 5'-TTGCCATAGAGGCATTGGCTGA-3' R: 5'-TTGCATCATCACTGGCTTCCCA-3'	Deuschle et al. (2004)
Pt31 Class I small heat shock protein	Pta.9363	F: 5'-AACAAGCAGGGCGTAACGTC-3' R: 5'-TAGCGGTGACAGGAACGTGA-3'	Loopstra and Sathyan (2004)
Pt38 Hypothetical protein	Pta. 11066	F: 5'-TACGGGTCGAAGCCCAATC-3' R: 5'-TGGGACATTCCACACCAGG-3'	Loopstra and Sathyan (2004)
PtAldh 3-chloroallyl aldehyde dehydrogenase	Pta. 966	F: 5'-GTTATGGCTTGGCAGCAGGA-3' R: 5'-TTACTCGGGCATTACGTGTCG-3'	Loopstra and Sathyan (2004)

Table A.1 continued.

Gene name and putative function	Unigene	Primer sequence	Reference
PtGRP Glycine-rich protein	Pta. 17	F: 5'-GCGAACGAACTTGTTGAAAGG-3' R: 5'-GTGTGGTACAGTTCGGGAAGG-3'	Loopstra and Sathyan (2004)
PtGTP GTP-binding protein	Pta. 13969	F: 5'-AGGGCTATTGCAGTTCGCTTC-3' R: 5'-GAGTTGCAGGGTATCGGGACT-3'	Loopstra and Sathyan (2004)
PtIP Inorganic pyrophosphatase	Pta. 13088	F: 5'-TATGCGCTCAAGGTAATTGCG-3' R: 5'-CGTTCTATTCCTGCAGGGGAG-3'	Loopstra and Sathyan (2004)
PtLEA2 Late embryogenesis abundant protein	Pta. 7849	F: 5'-GCCAGTATGATGGACAAAGCG-3' R: 5'-GGAAAAACCCTCCGCAGAG-3'	Loopstra and Sathyan (2004)
PtLEA3a Late embryogenesis abundant protein	Pta. 11747	F: 5'-AAGGGTTGGGCTAGGTGGAA-3' R: 5'-CGCATTTATCTGTTTCCAGGGA-3'	Loopstra and Sathyan (2004)
PtLEA3c Late embryogenesis abundant protein	Pta. 1596	F: 5'-TCCATTCCCACTGCCTACACT-3' R: 5'-ATTGTCTTGTGTCCCGTTTTCA-3'	Loopstra and Sathyan (2004)
AOC Abietadienol/abietadienal oxidase/cytochrome p450	Pta. 8786	F: 5'-TTGACAGGGCGTCTCGTGAAT-3' R: 5'-AGGAGAGGTGGCTTCTGCAGTA-3'	Ro et al. (2005)

APPENDIX B

Table B.1 Average gene expression and expression range between the highest and lowest expressing individuals in the NCSU association population calculated for each gene. Values are expressed in ($\Delta\Delta C_T$, cycle threshold value). Negative numbers represent higher average expression relative to the calibrator individual. Each cycle is equal to a 2-fold difference in expression.

Gene	Avg. C_T Avg. Range	Gene	Avg. C_T Avg. Range	Gene	Avg. C_T Avg. Range	Gene	Avg. C_T Avg. Range
ABI1	-0.252 3.39	**CYPC**	-0.050 7.43	**PtOMT**	0.473 3.93	**Pt38**	-1.88 4.66
ADR1	-0.133 4.11	**CYPD**	0.101 4.60	**LP3**	0.080 4.28	**PtAldh**	-0.610 2.68
ANR	-0.284 7.14	**EB9D**	0.089 7.91	**PT1**	0.811 3.10	**PtGRP**	-0.967 5.33
AOC	1.21 4.10	**ERD3**	-0.108 4.52	**PtMLO2**	-0.184 6.37	**PtGTP**	-0.431 3.77
AOS	0.360 5.02	**ERECTA**	-0.218 4.19	**PtMYB9**	0.140 2.39	**PTI4**	0.347 7.15
ARF	-0.212 3.56	**ERF1**	0.621 4.95	**PtMYB6**	-0.165 4.00	**PtIP**	0.326 3.72
ATAF-1	0.194 10.0	**PtEMB1**	-0.623 6.00	**NAC1**	0.273 5.59	**PtLEA2**	-1.01 4.92
Avr-Cf9	0.090 2.84	**PtEMB2**	-0.764 6.22	**NCED**	-1.11 7.59	**PtLEA3a**	-0.268 6.94
BALDH	0.022 2.57	**PtEMB3**	-0.793 5.22	**OXR1**	0.065 5.40	**PtLEA3c**	-0.259 7.35
PtBAG-1	0.428 3.58	**PtEMB4**	3.72 13.2	**OXR2**	0.732 3.65	**PX-1**	1.03 3.53
PtBAG-4	-0.310 3.08	**GLX1**	-0.167 5.40	**(-)pin-TPS**	0.480 6.00	**PX-2**	1.56 7.53
BGLU1	-0.111 5.44	**PtGPX1**	-0.374 4.34	**(+)pin-TPS**	0.578 7.34	**PX-3**	4.04 9.42
BI-1	0.126 3.38	**PtGPX2**	0.242 3.44	**P5CDH**	-0.207 3.27	**PX-4**	0.725 5.94
bZip	-0.469 5.96	**PtGPX3**	-0.048 8.04	**P5CR**	-0.071 3.28	**PX-5**	0.848 5.44
CHI	0.788 3.98	**PtGOLS1**	0.061 9.29	**PtAN1**	0.003 3.31	**RAP2.1**	0.140 4.00
CHS	1.02 7.19	**PtGSTF9**	1.18 3.03	**PtPDIR1**	0.197 6.98	**RAP2.4**	0.512 4.24
COI1	-0.672 4.08	**PtGSTU18**	0.577 6.38	**PR-2**	-0.538 7.09	**RPK1**	0.374 4.20
CPK3	-0.272 5.22	**HDZ32**	-0.048 8.04	**PR-3**	0.945 7.92	**SamCMT**	-0.518 13.4
PtCPN10	-1.38 5.66	**HMGCoA**	0.265 4.35	**PR-5**	-0.738 7.89	**Sesqui-TPS**	0.295 7.02

Table B.1 continued.

Gene	Avg. C_T Avg. Range	Gene	Avg. C_T Avg. Range	Gene	Avg. C_T Avg. Range	Gene	Avg. C_T Avg. Range
CslA1	0.499 6.74	**HPL**	-1.14 6.59	**PR-10**	1.90 10.1	**STR_SYN**	0.123 4.85
CYPA	0.534 5.06	**HSP17.4CIII**	-0.739 4.76	**PSCHI4**	-0.158 7.71	**trp-TPS**	1.09 6.76
CYPB	0.944 6.10	**IMPA**	-0.579 3.84	**Pt31**	0.122 9.96	**PtWRKY**	-0.399 6.51

APPENDIX C

Table C.1 Partial correlations among genes controlling for treatment groups. The correlations listed were significant ($p<0.05$) and considered strong correlations ($R^2>0.5$).

Partial correlations, ($R^2>0.5$)											
	ABA1		ALDH7B4		CYPD		cMyc		CAB		GRAS
PtWRKY2	0.591	CYPB	0.583	CslA1	0.594	IPK	0.569	CPK3	0.567	AOS	0.566
PR-1	0.571	AtNAP4	0.533	PtPDIR1	0.526	IAA2	0.543	PR-10	0.522	EB9D	0.544
HMGCoA	0.515	PtEMB2	0.524	PX-1	0.520			PtMYB13	0.518	HDZ32	0.511
										MYND	0.510
Partial correlations, ($R^2>0.5$)											
	PR-1		PtMYB13		ERECTA		BGLU1		ARF		ERD3
PR-10	0.673	PtEMB3	0.582	Pt38	0.667	ERECTA	0.741	ERECTA	0.656	PtAN1	0.740
PR-2	0.557	PtEMB1	0.555	PtIP	0.665	ADR1	0.525	PtLEA2	0.655	NCED	0.676
SamCMT	0.557	AOS	0.526	PtLEA2	0.635	PtAldh	0.611	Pt38	0.636	PtLEA3a	0.586
		PtEMB2	0.525	PtAldh	0.564	ABI1	0.590	PtCPN10	0.631	COI1	0.586
		STR_SYN	0.521	ABI1	0.561	BI-1	0.533	PtIP	0.623	PtOMT	0.586
				HSP17.4	0.554	PtIP	0.530	PtAldh	0.619	BI-1	0.554
				PtOMT	0.554	Pt38	0.524	BGLU1	0.619	ABI1	0.708
						OXR1	0.508	P5CR	0.554		
Partial correlations, ($R^2>0.5$)											
	TTA7		PtGOLS1		P5CR		Pt38		BI-1		PtMYB9
NCED	0.539	(+)pinTPS	0.571	Pt38	0.661	PtLEA2	0.831	PtMBY9	0.675	ERF1	0.667
PtAN1	0.506	AOC	0.524	PtLEA2	0.636	PtBAG-4	0.611	PtAN1	0.617	OXR1	0.638
		PtGPX1	0.517	PtBAG-4	0.517	PtIP	0.549	COI1	0.588	PR-3	0.540
		STR_SYN	0.517	BI-1	0.516	OXR1	0.512	OXR1	0.540	PtAN1	0.538
				PtGSTF9	0.534	BI-1	0.507	NCED	0.528	NCED	0.506
						PtAldh	0.505	RPK1	0.528		

Table C.1 continued.

Partial correlations, (R^2>0.5)											
	Avr-Cf9		BALDH		PtLEA2		COI1		ERF1		PtLEA3a
RPK1	0.586	PtMYB9	0.740	PtBAG-4	0.586	PtAN1	0.769	OXR1	0.622	ABI1	0.679
BI-1	0.554	PtBAG-4	0.676	NCED	0.576	ANR	0.608	PR-3	0.586	COI1	0.610
PtAN1	0.514	ERF1	0.653	PtLEA3C	0.529	BZip	0.557	PR-2	0.550	PtAN1	0.529
		OXR1	0.578			PtEMB3	-0.552				
Partial correlations, (PR>0.5)											
	OXR1		HSP17.4		NCED		PtAN1		PR-2		RPK1
PR-2	0.655	P5CDH	0.547	PtAN1	0.627	ANR	0.560	PR-3	0.741	ANR	0.507
PR-3	0.638	RPK1	0.519	PtLEA3c	0.502	bZip	0.523	HMGCoA	0.579	NAC1	0.536
		NAC1	0.509					PR-5	0.504		
		PtBAG-1	0.535								
Partial correlations, (PR>0.5)											
	AOC		trp-TPS		PtGPX2		HDZ32		PtMLO2		PtMYB6
(+)pinTPS	0.842	(+)pinTPS	0.684	PtEMB2	0.595	BZip	0.614	PtMYB6	0.700	BZip	0.625
(-)pinTPS	0.825	(-)pinTPS	0.604	PTI4	0.595	RAP2.1	0.587	PtEMB1	0.607	(+)pinTPS	0.560
CYPB	0.737	AOC	0.573	PtEMB3	0.581	PtMYB6	0.537	(+)pinTPS	0.592	PSCHI4	0.533
STR_SYN	0.661			CPK3	0.576	AOS	0.523	AOC	0.558	RAP2.1	0.520
PtEMB1	0.587			PT1	0.559			PX-1	0.543	RAP2.4	0.516
CslA1	0.552			ATAF-1	0.553			CYPB	0.539	PtPDIR1	0.511
HPL	0.516			PtGSTU18	0.547			GLX1	0.518	HPL	0.509
GLX1	0.505							PtPDIR1	0.510	GLX1	0.501
PX-2	0.501							PSCHI4	0.503		
Partial correlations, (PR>0.5)											
	PX-1		PX-5		RAP2.1		CslA1		(+)pinTPS		PX-2
CslA1	0.523	PtWRKY	0.607	RAP2.4	0.652	HPL	0.542	PSCHI4	0.534	STR_SYN	0.518

Table C.1 continued.

Partial correlations, (PR>0.5)											
	PtEMB1		PtGSTU18		GLX1		PtEMB3		HMG CoA		HPL
STR_SYN	0.773	PtEMB2	0.789	PtEMB1	0.700	PtEMB2	0.631	PX-2	0.587	PtEMB1	0.550
PtEMB2	0.728	STR_SYN	0.775	PX-2	0.607	STR_SYN	0.617	SamCMT	0.531	(-)pinTPS	0.632
SamCMT	0.661	PtEMB3	0.681	HPL	0.589	SamCMT	0.570			(-)pinTPS	0.514
PR-10	0.569	SamCMT	0.643	STR_SYN	0.543	PtEMB1	0.538			PX-2	0.508
PtEMB4	0.543	PtEMB1	0.625	PtEMB2	0.540	PtEMB4	0.520				
PX-2	0.538	PR-10	0.576								
PSCHI4	0.521	OXR2	0.556								
Partial correlations, (PR>0.5)											
	PR-10		CYPC		PtEMB2		PX-4		PtEMB4		PSCHI4
SamCMT	0.701	PR-10	0.502	STR_SYN	0.832	PX-5	0.729	SamCMT	0.626	PX-2	0.535
PSCHI4	0.609	PR-3	0.532	PR-10	0.587	RAP2.1	0.560	PtEMB2	0.520	SamCMT	0.512
STR_SYN	0.586			SamCMT	0.579	PtWRKY	0.541			STR_SYN	0.504
				PSCHI4	0.503	AOS	0.450				
Partial correlations, (PR>0.5)											
	CYPB		PtCPN10		ABI1		CYPA		PtBAG-4		bZip
STR_SYN	0.845	PtLEA2	0.647	COI1	0.768	(+)pinTPS	0.748	HSP17.4	0.676	RAP2.1	0.708
PtGSTU18	0.740	Pt38	0.619	PtAN1	0.736	(-)pinTPS	0.730	PtMYB9	0.637	RAP2.4	0.635
PtEMB1	0.736	PtIP	0.555	BI-1	0.704	trp-TPS	0.669	P5CDH	0.607		
PtEMB2	0.723	PtBAG-4	0.537	NCED	0.657	AOC	0.651	ERF1	0.543		
PR-10	0.577	OXR2	0.526	PtMYB9	0.637	PtGPX1	0.634	OXR2	0.508		
SamCMT	0.568			ANR	0.626	PtMLO2	0.558				
PtEMB3	0.563			RPK1	0.552	PtMYB6	0.534				
				CHI	0.508						

Table C.1 continued.

Partial correlations, (PR>0.5)											
	PtWRKY2		PtGPX1		AtNAP4		AOS		ATAF1		CPK3
PR-1	0.689	PtMLO2	0.795	PtMYB13	0.739	GLX1	0.703	CPK3	0.715	PtEMB2	0.811
PtEMB3	0.575	PtMYB6	0.742	PtEMB1	0.627	PX-2	0.668	PtEMB4	0.641	PtEMB1	0.723
PtWRKY	0.531	(+)pinTPS	0.661	CAB	0.616	PtEMB1	0.625	PtEMB2	0.640	PtGSTU18	0.720
ATAF-1	0.525	PSCHI4	0.600	GRAS	0.588	EB9D	0.616	CYPC	0.604	STR_SYN	0.674
PR-10	0.512	AOC	0.593	CYPB	0.553	PtEMB2	0.586	HMG CoA	0.571	CYPB	0.646
PtEMB4	0.508	PtEMB1	0.592	PtGPX1	0.535	HPL	0.565	PtEMB3	0.569	PR-10	0.640
CPK3	0.504	PtGPX3	0.571	PtGOLS1	0.530	HMG CoA	0.518	SamCMT	0.559	CYPC	0.633
PtGSTU18	0.503	(-)pinTPS	0.559	PtEMB2	0.517	CPK3	0.511	PtEMB1	0.556	SamCMT	0.619
PR-2	0.502	GLX1	0.541	CMyc	0.511			PR-10	0.537	PtEMB4	0.570
PR-3	0.501	trp-TPS	0.537	STR_SYN	0.503			PtGSTU18	0.529	PtEMB3	0.557
		STR_SYN	0.528					PR-3	0.524		
		CYPB	0.520								
		PX-1	0.515								
		PX-2	0.507								
Partial correlations, (PR>0.5)											
	SamCMT		(-)pinTPS		EB9D						
STR_SYN	0.588	(+)pinTPS	0.855	GLX1	0.689						
		PX-2	0.517	PtEMB1	0.585						

VITA

Name: Candace Marie Seeve

Address: 2123 TAMU, College Station, TX 77843

Email Address: cseeve@tamu.edu

Education: B.S., Cell & Molecular Biology, The University of Hawai'i at Hilo, 2006