Tree Physiology 00, 1–13
doi:10.1093/treephys/tpw122

Research paper

Whole-transcriptome response to water stress in a California endemic oak, *Quercus lobata*

Paul F. Gugger[1,2,†], Juan Manuel Peñaloza-Ramírez[1,†], Jessica W. Wright[3] and Victoria L. Sork[1,4,5]

[1]Ecology and Evolutionary Biology, University of California, Los Angeles, CA 90095-7239, USA; [2]University of Maryland Center for Environmental Science, Appalachian Laboratory, Frostburg, MD 21532, USA; [3]Pacific Southwest Research Station, USDA Forest Service, Davis, CA 95618, USA; [4]Institute of the Environment and Sustainability, University of California, Los Angeles, CA 90095-1496, USA; [5]Corresponding author (vlsork@ucla.edu)

Received July 26, 2016; accepted December 12, 2016; handling Editor Jörg-Peter Schnitzler

Reduced water availability during drought can create major stress for many plant species. Within a species, populations with a history of seasonal drought may have evolved the ability to tolerate drought more than those in areas of high precipitation and low seasonality. In this study, we assessed response to water stress in a California oak species, *Quercus lobata* Née, by measuring changes in gene expression profiles before and after a simulated drought stress treatment through water deprivation of seedlings in a greenhouse setting. Using whole-transcriptome sequencing from nine samples from three collection localities, we identified which genes are involved in response to drought stress and tested the hypothesis that seedlings sampled from climatically different regions of the species range respond to water stress differently. We observed a surprisingly massive transcriptional response to drought: 35,347 of 68,434 contigs (52%) were differentially expressed before versus after drought treatment, of which 18,111 were down-regulated and 17,236 were up-regulated. Genes functionally associated with abiotic stresses and death were enriched among the up-regulated genes, whereas metabolic and cell part-related genes were enriched among the down-regulated. We found 56 contigs that exhibited significantly different expression responses to the drought treatment among the three populations (treatment × population interaction), suggesting that those genes may be involved in local adaptation to drought stress. These genes have stress response (e.g., WRKY DNA-binding protein 51 and HSP20-like chaperones superfamily protein), metabolic (e.g., phosphoglycerate kinase and protein kinase superfamily protein), transport/transfer (e.g., cationic amino acid transporter 7 and K^+ transporter) and regulatory functions (e.g., WRKY51 and Homeodomain-like transcriptional regulator). Baseline expression levels of 1310 unique contigs also differed among pairs of populations, and they were enriched for metabolic and cell part-related genes. Out of the large fraction of the transcriptome that was differentially expressed in response to our drought treatment, we identified several novel genes that are candidates for involvement in local adaptation to drought.

Keywords: drought, gene expression, local adaptation, *Quercus*, RNA-Seq.

Introduction

In plants, the capacity to respond to increases in temperature and drought conditions will be of particular importance if current climate trends continue (IPCC 2007, Diffenbaugh et al. 2015). Under most climate change scenarios, one of the main environmental stress factors will be increased drought (IPCC 2007, Dai 2013). In general, water availability is a major determinant of the distribution and abundance of plants. Understanding the mechanisms of how plants cope with water stress has been a central topic in plant physiology for decades (Stebbins 1952, Bray 1997). Many plants species have evolved specific adaptive mechanisms in response to the drought stress, exhibiting either

[†]These authors contributed equally to this paper.

drought escape or drought tolerance mechanisms (e.g., Savage and Cavender-Bares 2011). These mechanisms are complex, polygenic traits that involve cascades of responses, ranging from physiological changes to transcriptional regulation (Chaves et al. 2003). The key adaptive responses to drought in plants are reduced water loss, enhanced water uptake through roots and reduced radiation absorption. Physiological studies have shown that reduction in vegetative growth, stomatal closure and a decrease in the rate of photosynthesis are among the earliest responses to drought, protecting the plant from extensive water loss and oxidative stresses (Chaves et al. 2003). These drought response strategies and abilities can vary within species among different populations. For example, thick leaves and small stomata often found in oak species are characteristics that favor high water-use efficiency (Abrams 1990). Water-use efficiency, which is the ratio of the rate of photosynthesis to the rate of transpiration, is known to be important in local adaptation of oaks (*Quercus* spp.) (Ponton et al. 2002, Aranda et al. 2007, Ramírez-Valiente et al. 2009, Roussel et al. 2009).

With modern whole-transcriptome gene expression profiling (RNA-Seq) (Wang et al. 2009), it is feasible to study changes in gene expression controlled by water availability. Drought stress triggers many changes in gene expression that in turn affect a variety of physiological and metabolic processes, signal transduction, osmotic regulation and gene regulation in accordance with stress response (Shinozaki and Yamaguchi-Shinozaki 2007, Osakabe et al. 2014). Many of the genes involved in drought stress response have been identified as differentially expressed transcripts (Bray 2002, Ramanjulu and Bartels 2002, Osakabe et al. 2014). These genes can be categorized in two broad groups: those encoding stress response proteins versus regulatory proteins (Shinozaki and Yamaguchi-Shinozaki 2007). Stress response proteins include late embryogenesis abundant proteins or dehydrins, proteins involved in protecting photosynthesis, aquaporins that facilitate transmembrane water movement, lipid transfer proteins involved in cuticle biosynthesis, antioxidant proteins (e.g., glutathione-S-transferase) that reduce oxidative stress, carbohydrate transporters that adjust metabolism, osmolyte transporters to maintain osmotic balance, protein repair enzymes, proteases, protease inhibitors that block programmed cell death (e.g., Kunitz type) and other enzymes (Ramanjulu and Bartels 2002). Regulatory proteins of signal transduction and gene expression include transcription factors (e.g., WRKY, MYB, HD-ZIP, Myc-like bHLH), protein kinases, protein phosphatases, calmodulin-binding protein and those involved in abscisic acid biosynthesis (ABA), which can lead to stomatal closure, among other physiological phenomena (Seki et al. 2007). These types of proteins seem to be involved in immediate responses of plants to drought stress, but additional pathways can lead to long-term acclimation, damage control and repair, leaf senescence or programmed cell death (Osakabe et al. 2014). Furthermore, these responses and the molecules involved can vary within species depending on the genetic background of the individual, which may be related to local adaptation and different strategies or capacities for coping with drought (Villar et al. 2011, Lasky et al. 2014).

Quercus spp. (oaks) are known to display a wide range of adaptive variation for drought tolerance among species (Abrams 1990, Tyler et al. 2006) and among populations within species (Ponton et al. 2002, Roussel et al. 2009), but the molecular mechanisms underlying differences in tolerance are largely unknown. A few genes have been implicated in drought response in the European oak *Quercus petraea* based on a small expressed sequence tag (EST) study: betaine-aldehyde dehydrogenase, an ABA responsive transcription factor, an ABA-independent transcription factor, glutathione-S-transferase and a heat-shock cognate protein (Porth et al. 2005). In addition, a large EST microarray study of *Quercus robur* in response to long-term water stress showed that up-regulated genes encode protective proteins (e.g., late embryogenesis abundant, RCI2B, protein phosphatase 2CA), transcription factors (e.g., WRKY), transport proteins (e.g., inositol transporter 1 and lipid transfer proteins), leaf senescence proteins (ATP-dependent Cpl proteases and senescence-associated protein 21) and amylases (Spieß et al. 2012). Down-regulated genes were found to encode photosynthesis-related proteins. Finally, up to 20 genes with a broad range of functions were implicated in local adaptation to precipitation regimes in *Quercus lobata* Née in California based on extreme correlations of single-nucleotide polymorphisms with precipitation variables (Gugger et al. 2016, Sork et al. 2016). To assess the full array of genes that could be involved in drought adaptation, it is advantageous to use a whole transcriptome approach and avoid the ascertainment bias of studying only a few known genes. Additional knowledge of gene co-expression and the molecular pathways involved in adaptation of oaks to environmental changes will enhance our understanding of the molecular basis of adaptation and the capacity of oaks to recover under different frequencies, severities and durations of environmental stress (Spieß et al. 2012).

In this study, we take one approach to understanding plant gene expression response to water stress in *Q. lobata*, a widely distributed California oak found across a range of climate conditions. The distribution of this species is water-limited (McLaughlin and Zavaleta 2012), but its range includes a tremendous precipitation gradient (Sork et al. 2010). Thus, *Q. lobata* offers an excellent system to study tolerance and local adaptation in natural populations. Different populations might respond differently to projected increases in drought due to genetic differences, as well as differences in the severity of drought across the distribution. In this study, we assess the transcriptional response to drought stress by measuring gene expression profiles across the entire transcriptome. Specifically, we examine what genes are involved in response to drought stress and test the hypothesis that seedlings sampled from different climatic regions of

the species range respond to water stress differently, consistent with local adaptation.

Materials and methods

Sample design

The seed sources came from different maternal lines in each of three populations with contrasting climatic environments from across California: Malibu Creek State Park (MC); Middle Creek Campground, Mendocino National Forest (MK); and Springville (SV) (Figure 1 and Table S1 available as Supplementary Data at *Tree Physiology* Online). Springville is an open oak savanna ecosystem with the highest mean temperature and temperature seasonality and the lowest mean annual precipitation of the three sites, contributing to it having the highest climatic water deficit (CWD), which is an integrated measure of water availability including temperature, rainfall and soil characteristics (Flint et al. 2013). Middle Creek Campground is an open-canopy, mixed-oak woodland with the lowest CWD, highest precipitation and moderate temperature seasonality. Malibu Creek State Park is a mixture of oak savanna and mixed-oak woodland with CWD nearly as high as SV, the lowest temperature seasonality and low precipitation. In October 2011, acorns were collected from these three sites as part of a trial for a large provenance study of valley oak planted at the Institute of Forest Genetics, USDA Forest Service, Placerville, CA (Delfino Mix et al. 2015). Acorns were planted in Sunshine #4 (Sungro, Agawam, MA, USA) aggregate plus soil mix in Steuwe and Sons D40 pots (25 cm depth × 6.4 cm diameter) (Tangent, OR, USA) and watered twice per week and as needed to maintain moist soil. On 23 August 2012, mature leaf tissue samples (fourth node from top) of nine seedlings (three from each of three source populations) were collected and frozen immediately between slabs of dry ice for RNA isolation to be used as control samples before treatment. To simulate a water stress condition, these same seedlings were deprived of water for 15 days, after which (7 September 2012) we collected tissue on dry ice for RNA isolation.

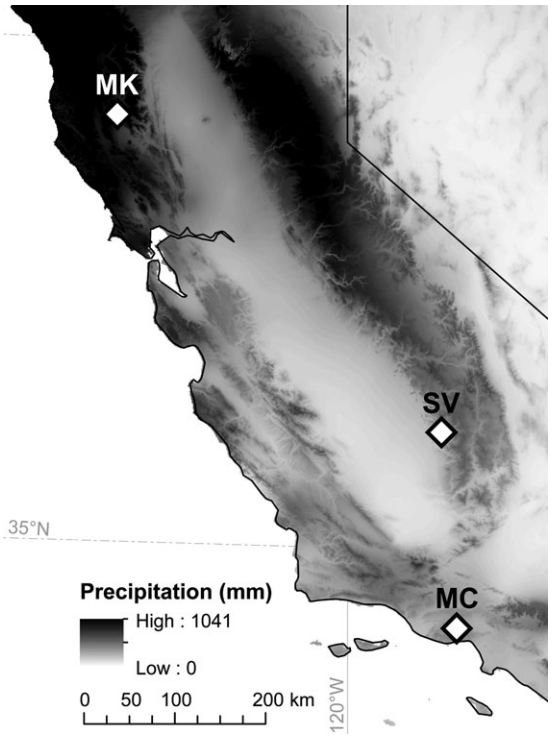

Figure 1. Acorn sampling sites in California that were used for greenhouse water stress experiment. Precipitation gradient shown as grayscale. MC, Malibu Creek State Park; SV, Springville; MK, Middle Creek.

Library preparation and sequencing

Total RNA was isolated with a prewash protocol (http://openwetware.org/wiki/Conifer_RNA_prep) and RNeasy Mini Kit (Qiagen, Hilden, Germany). The quality and quantity of the total RNA was determined using a BioAnalyzer 2100 with RNA Nano chip (Agilent Technologies, Santa Clara, CA, USA). We then prepared cDNA from poly-A purified mRNA for 50-bp, single-end sequencing using the Illumina TruSeq RNA library preparation kit. Each sample was barcoded using standard Illumina adaptors 1–12 to allow up to 12 samples to be pooled in each of two lanes of sequencing on an Illumina HiSeq 2000 v3 at the Broad Stem Cell Research Center, UCLA. A given before–after pair was sequenced on the same lane, and sample pairs drawn from different populations were stratified across the two sequencing lanes (see Table S1 available as Supplementary Data at *Tree Physiology* Online).

Processing and mapping Illumina reads

The quality and contamination levels of RNA-Seq reads generated by the Illumina HiSeq 2000 were ascertained using FastQC. Low quality reads, primer/adapter contamination, and long repetitive sequences were filtered using tools on the UCLA Galaxy server (http://galaxy.hoffman2.idre.ucla.edu/). Filtered reads were aligned against all contigs in the reference transcriptome available for *Q. lobata* (Cokus et al. 2015) using BWA, the Burrows–Wheeler Aligner with default settings (Li and Durbin 2009, 2010). Burrows–Wheeler Aligner uses an index built with the Burrows–Wheeler transformation that allows for fast searching and also reports a meaningful mapping quality score that can be used to discard alignments that are not well supported due to a high number of mismatches or low mapping quality. The numbers of reads mapped to each contig were generated using Samtools 0.1.19 'Gene to Counts' (idxstats) (Li et al. 2009) and used for subsequent analyses. Approximately 23 k of 83 k contigs in the *Q. lobata* reference transcriptome (Cokus et al. 2015) have gene models assigned to them, of which 9.4 k were assigned functional annotations based on *Arabidopsis* orthologs (Swarbreck et al. 2008) and 19 k were assigned annotations based on matches to the Pfam protein family database (Finn et al. 2014).

Identification of differentially expressed genes

Differential expression before versus after drought, differences in baseline expression levels among populations, as well as differential response to drought by different populations were analyzed using gene lengths and count data in DESeq2 1.6.3 (Anders and Huber 2010, Love et al. 2014) as implemented in R 3.1.2 (R Development Core Team 2014) and the R Bioconductor project (http://www.bioconductor.org). DESeq2 implements a negative binomial generalized linear model utilizing size-factor normalized gene expression data (Anders and Huber 2010). Our final model included 'treatment' (before vs after drought), sequencing lane, library prep batch, population and population × treatment interaction as predictors (see Table S1 available as Supplementary Data at *Tree Physiology* Online) and expression data for all contigs as response variables. We ran this model with Wald tests and extracted the relevant contrasts (e.g., before versus after, and interactions). Contigs with low expression across most samples for a given contrast are automatically discarded from testing in DESeq2. Resulting *P*- values were adjusted for multiple testing using the Benjamini and Hochberg (1995) false discovery rate (FDR) method, and 0.05 was used as a threshold for significance. Based on regularized log-transformed expression count data, principal components analysis and hierarchical clustering with heatmaps were used to visualize the results.

To arrive at our final DESeq2 model (above), we considered simpler models that included expression data only at contigs with gene models or excluded 'lane' and 'library prep' terms. The results were highly similar, thus we have preferred to include all contigs and follow the common practice of accounting for batch effects. Furthermore, the DESeq2 modeling framework does not allow for a term for 'individual' to be included, thus we assessed the potential for bias in our model by estimating the intra-individual correlation based on mixed models with individual, lane, library prep batch, population and treatment as random effects, as implemented in the 'variancePartition' package in R 3.3.1 using regularized log-transformed expression data. The median variance explained for the 'individual' term across all genes, which can be interpreted as the intra-individual correlation, is equal to zero (see Figure S1 available as Supplementary Data at *Tree Physiology* Online), suggesting that individual variation is unlikely to create spurious results favoring our hypotheses.

Gene ontology enrichment analyses

We performed singular enrichment analyses using agriGO (Du et al. 2010) to identify functional classes of genes based on Plant Gene Ontology (GO) slim (Ashburner et al. 2000) annotations drawn from the *Arabidopsis* Information Resource (TAIR10, http://www.arabidopsis.org; Swarbreck et al. 2008) that are overrepresented in the lists of significantly differentially expressed genes relative to a background of all genes that were tested in DESeq2. Specifically, we performed Fisher's exact tests for GO terms that occurred in at least 10 genes and corrected for multiple testing using the FDR method of Benjamini and Yekutieli (2001). A threshold of 0.05 was used to assess significance based on the FDR-adjusted *P*- values. This analysis was necessarily restricted to the set of 9.4 k oak contigs for which *Arabidopsis* orthologs had been assigned (Cokus et al. 2015).

Weighted gene co-expression network analysis

To construct a gene co-expression network, we deployed WGCNA 1.51 (Langfelder and Horvath 2008) using biweight midcorrelation among normalized gene expression values based on the regularized log-transformation as implemented in DESeq2. We considered only genes with expression level counts ≥ 10 in at least 90% of libraries. Groups of genes with similar expression patterns ('modules') were identified based on hierarchical clustering (soft threshold power = 7, minimum module size = 100, 0.1 ≤ mergeCutHeight ≤ 0.3). The module eigengenes (i.e., the first principal component of the module) were associated with the main drought stress treatment (as dummy variable) and geographic and climatic variables from the sample site of origin using linear regression (Zhang and Horvath 2005). Climate data included growing season growing degree-days above 5 °C derived from a spline model (Rehfeldt 2006), and CWD (a measure of evaporative demand exceeding soil moisture), actual evapotranspiration, mean maximum temperature (T_{max}) and mean minimum temperature (T_{min}) derived from the California Basin Characterization Model 1950–1980 averages (see Table S1 available as Supplementary Data at *Tree Physiology* Online) (Flint et al. 2013). All WGCNA settings were set at their default values unless previously mentioned. An advantage of the WGCNA approach is that it can identify groups of putatively interacting genes independent of preexisting annotation information.

Results

Gene expression and GO enrichment

In total, nine individuals in three populations of *Q. lobata* were sequenced successfully across both experimental conditions (before and after drought stress) (see Table S1 available as Supplementary Data at *Tree Physiology* Online). A range of 6.5–24.7 million (mean = 13 million) 50-base reads were obtained for each treatment of each biological sample after filtering (NCBI BioProject <TBD>). Of these, 77% of filtered RNA-Seq reads mapped onto the *Q. lobata* reference transcriptome (unaligned reads were discarded).

Surprisingly, 35,347 of 68,434 contigs with sufficient expression levels for testing (52%) were significantly differentially expressed before versus after drought treatment (Figure 2a and Table S2 available as Supplementary Data at *Tree Physiology* Online). Of these, 18,111 were down-regulated and 17,236 were up-regulated in response to drought (Table 1, Figure 2 and Table S2 available as Supplementary Data at *Tree Physiology*

Online). Of these genes, 91% had greater than twofold change in expression before versus after the drought treatment. This large response is also apparent in a principal components analysis showing that samples collected before treatment cluster together and those after treatment cluster together (Figure 3). Genes up-regulated in response to the drought treatment are involved especially in responses to stimulus, response to stress, death and various nuclear cellular components (Figure 4a and Figure S2 available as Supplementary Data at *Tree Physiology* Online). Down-regulated genes primarily have functions related to photosynthesis, metabolic processes and various cell parts (Figure 4b and Figure S3 available as Supplementary Data at *Tree Physiology* Online). Given the very large number of genes that responded to the treatment, we also performed an additional enrichment analysis restricted to the 3000 most significant genes. The results were largely the same (not shown). A number of candidate genes identified in other studies are also differentially expressed in ours (see Discussion).

Fifty-six contigs had different expression responses to the treatment depending on the population (population × treatment interaction) (Tables 1 and 2, and Figure 2b). A dendrogram and heatmap generated based on regularized log-transformed expression levels reveal two clusters of genes with similar interaction patterns (Figure 5). The clusters seem unrelated to any functional type of gene (e.g., metabolic genes are in both clusters). These 56 genes include eight with *Arabidopsis* orthologs: a WRKY transcription factor (m01oak15323cc-t01.1, AT5G64810), an ATPase family gene (m01oak16888Ci-t01.1, AT1G67120), HSP20-like chaperones superfamily protein (m01oak35262cF-t01.1, AT4G10250), phosphoglycerate kinase (m01oak07301CC-t01.1, AT1G79550), protein kinase superfamily protein (m01oak13924cC-t01.1, AT2G23200) and SPFH/Band 7/PHB domain-containing membrane-associated protein family (m01oak14227Cc-t01.1, AT2G03510). Of the remaining 49 genes, 21 had no known function and 28 contained Pfam domains with a broad range of molecular functions, especially related to stress response, regulation of metabolism, molecule transport or transfer, and nucleic acid binding and signaling. Due to the small number of TAIR-annotated genes, we were unable to statistically test for enrichment of GO functions.

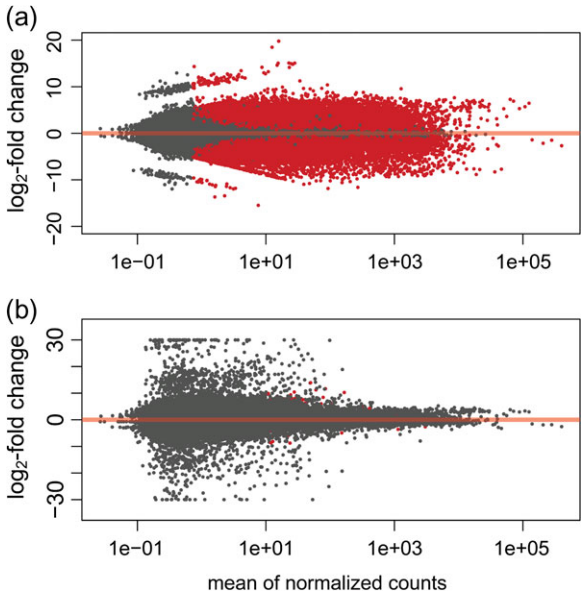

Figure 2. Log_2-fold gene expression change versus normalized expression level for (a) test of before versus after the drought stress treatment and (b) test of the drought treatment by population interaction. Red points are statistically significant after FDR adjustment.

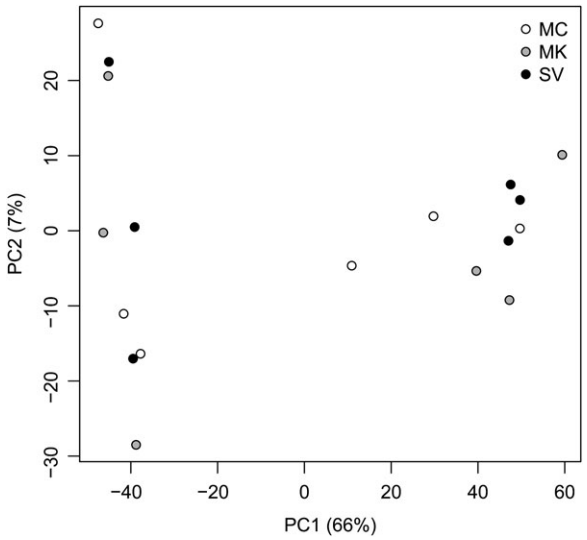

Figure 3. Principal components analysis of regularized log-transformed gene expression data for the 500 most variable genes colored by sample site. All samples before drought treatment are clustered on the left and after are clustered on the right. MC, Malibu Creek State Park; SV, Springville; MK, Middle Creek.

Table 1. Counts of differentially expressed genes based on DESeq2 during a 15-day simulated drought experiment for three populations of *Q. lobata* sampled in California.

	Total contigs	Tested in DESeq2	Differentially expressed	Up-regulated	Down-regulated
Before vs after ('treatment')	83,645	68,434	35,347	17,236	18,111
Before vs after (TAIR only)	9431	9234	6768	3129	3639
Population × treatment	83,645	36,254	56		
Malibu vs Middle Creek	83,645	32,292	537		
Malibu vs Springville	83,645	28,277	767		
Springville vs Middle Creek	83,645	40,324	329		

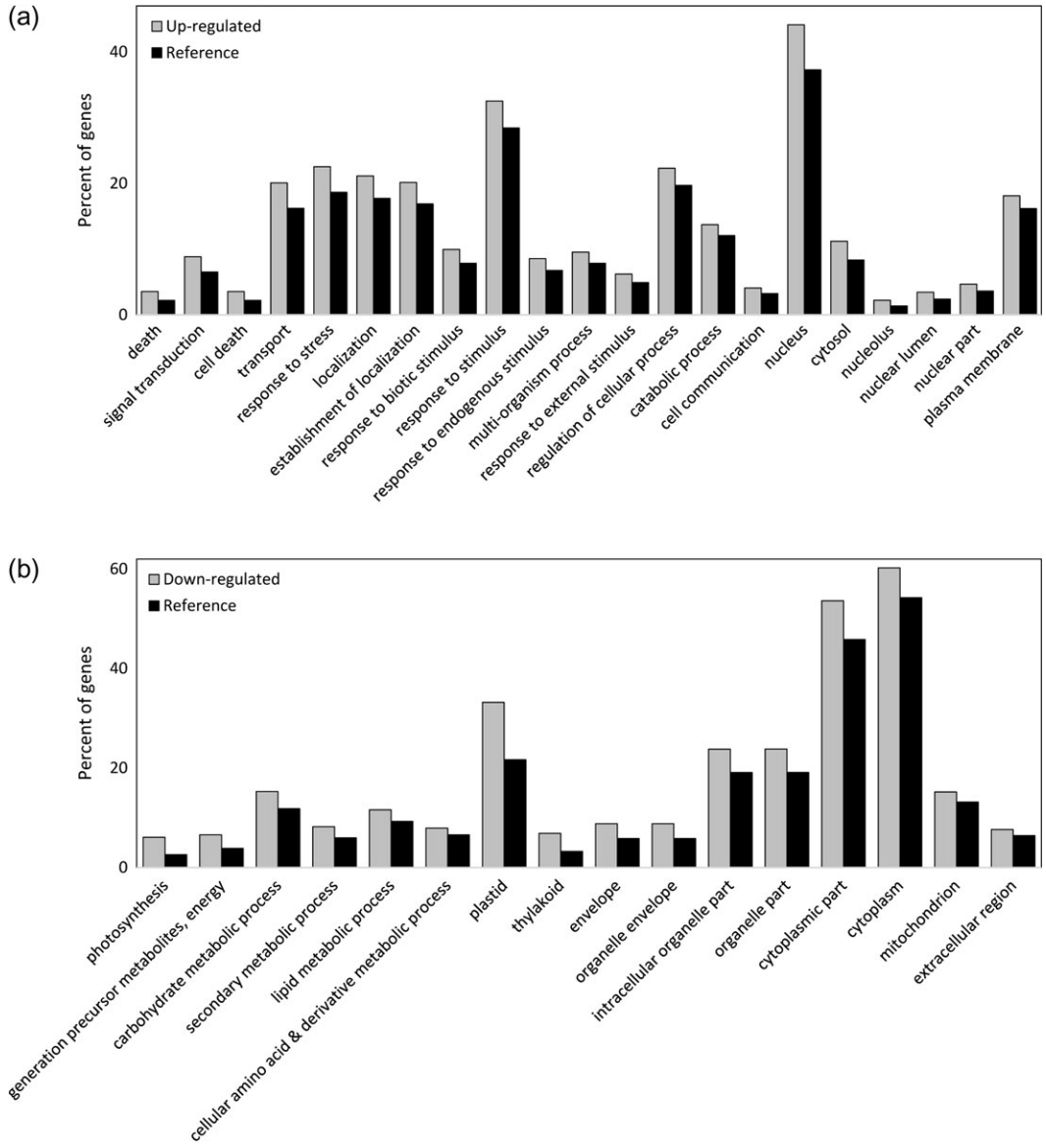

Figure 4. Results of singular enrichment analyses based on Plant GO slim functional categories for genes that were significantly (a) up-regulated and (b) down-regulated in response to the drought treatment.

Baseline differences in gene expression levels among populations were also apparent: 530 genes differed in expression among MC vs MK, 752 genes among MC vs SV and 310 genes among SV vs MK (Table 1 and Table S3 available as Supplementary Data at *Tree Physiology* Online). Gene Ontology enrichment in these genes was primarily in categories unrelated to stress, such as photosynthesis, metabolism and cellular parts (Figure 6 and Figure S4 available as Supplementary Data at *Tree Physiology* Online).

Weighted gene co-expression network analysis

Weighted gene co-expression network analysis based on 17,601 genes with sufficient expression revealed two large modules of co-expressed genes (Figure 7). The first ('grey') module is composed of 3270 genes, and its eigengene has moderate, insignificant correlations ($0.1 < |r| < 0.37$; $P > 0.1$) with the climate of origin and virtually no correlation with the drought treatment (Figure 8, and Figures S5 and S6 available as Supplementary Data at *Tree Physiology* Online). The second ('black') module is composed of 14,331 genes whose eigengene is strongly associated with the drought treatment ($r = 0.96$, $P < 7 \times 10^{-10}$). Thus, the genes can be broken down as one very large group that responded to the drought treatment synchronously and another fairly large group that did not respond to the drought but may have modest differences in baseline expression level among sample sites or individuals. The coarse resolution of clusters may be due to relatively low sample size.

Table 2. Genes or contigs with significant differential expression for the drought treatment × population interaction.

Quercus gene or contig	*Arabidopsis* TAIR or Pfam ID	Symbol	Protein or domain name	General function
m01oak00470CT-t01.1	D:PF02969, D:PF00125	TAF Histone	TATA box binding protein associated factor (TAF), Core histone H2A/H2B/H3/H4	Metabolic
m01oak01092CC-t01.1	F:PF02225, D:PF00082	PA Peptidase_S8	PA domain, Subtilase family	Metabolic
m01oak03031cT-t01.1	D:PF01419	2 × Jacalin	Jacalin-like lectin domain	
m01oak03252cf			Unknown	
m01oak05037Cz			Unknown	
m01oak05736Cz			Unknown	
m01oak05832cm			Unknown	
m01oak06276cf-t01.1	F:PF01764	Lipase_3	Lipase (class 3)	Metabolic
m01oak06325cT-t01.1	F:PF03767	Acid_phosphat_B	HAD superfamily, subfamily IIIB (Acid phosphatase)	Metabolic
m01oak06417jC-t01.1	F:PF03169	OPT	OPT oligopeptide transporter protein	Transport/transfer
m01oak07301CC-t01.1	AT1G79550	PGK	Phosphoglycerate kinase	Metabolic
m01oak07996cC			Unknown	
m01oak08414cC-t01.1	F:PF03492	Methyltransf_7	AM-dependent carboxyl methyltransferase	Transport/transfer
m01oak09862cf-t01.1	D:PF13561, F:PF08659, D:PF00106	adh_short_C2, KR, adh_short	Enoyl-(Acyl carrier protein) reductase, KR domain, short chain dehydrogenase	Metabolic
m01oak12499Cz			Unknown	
m01oak12526cC-t01.1	D:PF00210	Ferritin	Ferritin-like domain	Ion homeostasis
m01oak12585Ji			Unknown	
m01oak13134Ct-t01.1	D:PF00717	Peptidase_S24	Peptidase S24-like	
m01oak13430cC-t01.1	AT5G58900		Homeodomain-like transcriptional regulator	Transcription factor
m01oak13475Cz			Unknown	
m01oak13924cC-t01.1	AT2G23200		Protein kinase superfamily protein	Metabolic
m01oak14227Cc-t01.1	AT2G03510		SPFH/Band 7/PHB domain-containing membrane-associated protein family	
m01oak15142CC-t01.1	r:PF00560	LRR_1	Leucine Rich Repeat	Protein binding
m01oak15323cc-t01.1	AT5G64810	WRKY51	WRKY DNA-binding protein 51	Stress response, transcription factor
m01oak16669CC			Unknown	
m01oak16888Ci-t01.1	AT1G67120		ATPases; nucleotide binding; ATP binding; nucleoside-triphosphatases; transcription factor binding	Transcription factor, stress response
m01oak17139CC-t01.1	F:PF01657	2 × Stress-antifung	Salt stress response/antifungal	Stress response, pathogen response
m01oak17816cC-t01.1	AT3G10600	CAT7	Cationic amino acid transporter 7	Transport/transfer
m01oak18171JT-t01.1	D:PF00931, r:PF00560	NB-ARC, LRR_1	NB-ARC domain, Leucine Rich Repeat	Nucleid acid binding, signaling, pathogen response
m01oak18430Ci-t01.1	D:PF01738, D:PF12695	DLH, Abhydrolase_5	Dienelactone hydrolase family, α/β hydrolase family	Metabolic
m01oak18580cM-t01.1, -t01.2	F:PF05978, F:PF00083, F:PF07690	UNC-93, Sugar_tr, MFS_1	On channel regulatory protein UNC-93, Sugar (and other) transporter, Major Facilitator Superfamily	Transport/transfer
m01oak20011Cz			Unknown	
m01oak20039cf-t01.1	F:PF01657, D:PF00069, D:PF07714	2 × Stress-antifung, Pkinase, Pkinase_Tyr	Salt stress response/antifungal, Protein kinase domain, Protein tyrosine kinase	Stress response, metabolic
m01oak20950Cz			Unknown	
m01oak22392Cz			Unknown	
m01oak22555cC-t01.1	F:PF00954, D:PF01453, D:PF00069, D:PF07714	S_locus_glycop, B_lectin, Pkinase, Pkinase_Tyr	S-locus glycoprotein family, D-mannose binding lectin, Protein kinase domain, Protein tyrosine kinase	Metabolic
m01oak22836cC-t01.1	r:PF00560	3 × LRR_1	Leucine Rich Repeat	
m01oak23201CC-t01.1	F:PF03552	Cellulose_synt	Cellulose synthase	Metabolic
m01oak23892Jz			Unknown	
m01oak26417cf-t01.1	F:PF01145	Band_7	SPFH domain/Band 7 family	
m01oak26582Cz			Unknown	

(*Continued*)

Table 2. (Continued)

Quercus gene or contig	*Arabidopsis* TAIR or Pfam ID	Symbol	Protein or domain name	General function
m01oak27349Cz			Unknown	
m01oak29202Cz			Unknown	
m01oak29709cF-t01.1	F:PF01582	TIR	TIR domain	Response to stimulus, signaling
m01oak31397Jz			Unknown	
m01oak32376Cz			Unknown	
m01oak34915jC-t01.1	F:PF08276, F:PF00954, D:PF01453, D:PF00069, D:PF07714	PAN_2, S_locus_glycop, B_lectin, Pkinase, Pkinase_Tyr	PAN-like domain, S-locus glycoprotein family, D-mannose binding lectin, Protein kinase domain, Protein tyrosine kinase	Metabolic, RNA binding
m01oak35262cF-t01.1	AT4G10250	ATHSP22.0	HSP20-like chaperones superfamily protein	Stress response
m01oak35742Cz			Unknown	
m01oak39454sf			Unknown	
m01oak46055cC-t01.1	F:PF08387, D:PF00646	FBD, F-box	FBD, F-box domain	Protein binding
m01oak50236Jt-t01.1	D:PF14392, D:PF14111, D:PF0337,2 D:PF13966, F:PF00078, D:PF13456	zf-CCHC_4, DUF4283, Exo_endo_phos, zf-RVT, RVT_1, RVT_3	Zinc knuckle, Domain of unknown function (DUF4283), Endonuclease/Exonuclease/phosphatase family, zinc-binding in reverse transcriptase, Reverse transcriptase (RNA-dependent DNA polymerase), Reverse transcriptase-like	Metabolic, nucleotide binding
m01oak50609Cz			Unknown	
m01oak51147Cl-t01.1	D:PF12850, D:PF00149	Metallophos_2, Metallophos	Calcineurin-like phosphoesterase superfamily domain, Calcineurin-like phosphoesterase	Metabolic
m01oak64020cf-t01.1	F:PF02705	K_trans	K^+ potassium transporter	Transport/transfer
m01oak66628cF-t01.1	F:PF03140	DUF247	Plant protein of unknown function	

Discussion

Massive response to drought

A striking result of this study is the massive gene expression changes observed in response to the drought treatment: 35,347 of 68,434 contigs tested, or 52%, were differentially expressed, most of which (91%) exhibited greater than twofold change in expression level (Figure 2a and Table S2 available as Supplementary Data at *Tree Physiology* Online). This result was even more drastic when considering only the highest quality contigs with assigned *Arabidopsis* orthologs: 6768 of 9234 genes or 73% (Table 1). Even in attempting to define subsets of genes with similar co-expression patterns, we found that genes could only be grouped into a large set that responded synchronously to the treatment (most genes) versus those that did not (Figure 7). Several studies that used similar approaches have reported responses involving up to 20–37% of the genes analyzed (Cohen et al. 2010, Bhardwaj et al. 2015), but most report <15% (e.g., Kreps et al. 2002, Utsumi et al. 2012, Yates et al. 2014). The extremely large gene expression responses to short-term drought provides many candidate genes for involvement in oak response to water stress, but it also likely includes other types of associated responses.

There are at least four potential non-mutually exclusive explanations for the large response to drought. One possibility is that gene expression varies dramatically through time, largely unrelated to water stress, and thus a large component of the apparent response stems from developmental or other changes that occurred within individuals during the 15-day interval pre- and post-treatment samples. Time-series experiments often identify many genes whose expression differs through time (usually in relation to development or stress), but typically the response is not as large as that observed here (e.g., Breeze et al. 2011, Bechtold et al. 2016). In addition, we attempted to minimize effects of development by collecting tissue from leaves of similar maturity at the same time of day. A second explanation is that the treatment was so severe that the plants were in the process of dying and shutting down expression altogether. In this scenario, we would expect the plants to show bias toward down-regulation of all genes after treatment; however, we observe similar fractions of genes up- and down-regulated (Table 1 and Figure 2, and see Table S2 available as Supplementary Data at *Tree Physiology* Online). A third, more likely, explanation is that our drought treatment was severe enough that pathways unrelated or indirectly related to drought, such as those related to leaf senescence, were initiated. Indeed, several leaves began browning around the time of collection, death-related genes were up-regulated after the treatment, and a number of non-drought related genes were differentially expressed (Figure 4), although this latter observation is common in RNA-Seq differential expression studies. Fifteen days without water occurs frequently in nature, and seedlings must routinely survive such stresses. However, the importance of acclimation in

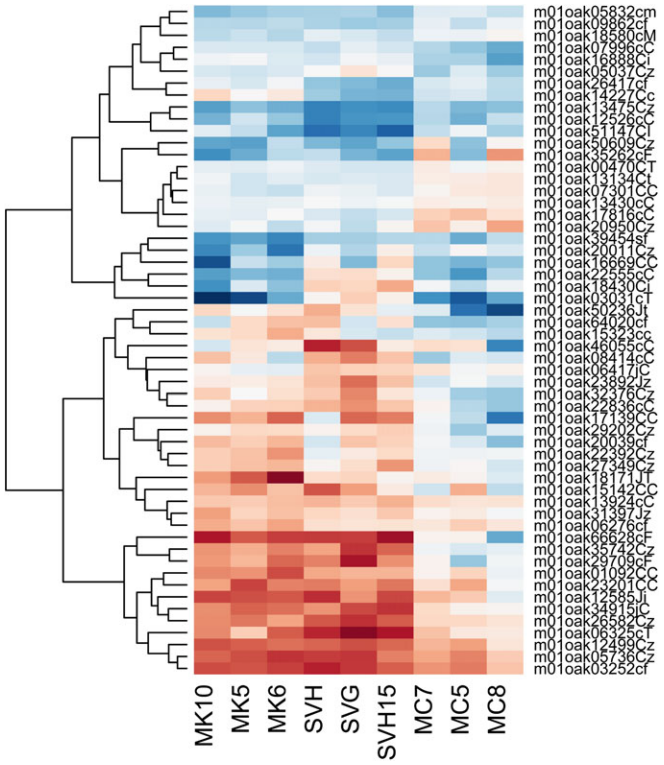

Figure 5. Heatmap and hierarchical clustering of $\log_2$-fold changes in gene expression in response to the drought treatment for the 56 genes showing significant drought treatment by population interaction terms from DESeq2 analysis. The dendrogram clusters genes with similar expression change patterns. The color scale reflects the $\log_2$-fold change in gene expression in response to drought, ranging from down-regulated (blue) to up-regulated (red). MC, Malibu Creek State Park; SV, Springville; MK, Middle Creek.

stress response is widely reported in the plant physiology literature (Yordanov et al. 2000), and thus it is possible that this seemingly realistic drought treatment imposed on unacclimated greenhouse plants was 'felt' severely. A fourth possible explanation is that such a massive transcriptomic response is normal in oaks, which would indicate a flexible and dynamic response to the drought-prone environments that oaks often inhabit. In this case, we might expect that other species that occupy water-limited environments might also exhibit large responses. Alternatively, large responses may be a feature of all trees or some other aspect of the natural history of oaks. However, we do not observe either trend in other studies (Kreps et al. 2002, Cohen et al. 2010, Utsumi et al. 2012, Yates et al. 2014, Bhardwaj et al. 2015), and the response does not seem to be large (3%) in other oaks under long-term water stress (Spieß et al. 2012). The distinction between the latter two hypotheses for explaining the massive response (severe treatment versus real response) in valley oak can be better understood in future studies using data collected from seedlings in nature or additional greenhouse studies allowing for acclimation.

The massive response precludes us from pinpointing the precise molecular mechanisms of drought stress response, but the observed enrichment for a number of gene functional categories related to abiotic stress responses are similar to those found in *Arabidopsis* and other plants (Figure 4). Among those genes are many that are commonly implicated in drought response in oaks and other species, such as late embryogenesis abundant proteins, glutathione-S-transferase, zeaxanthin epoxidase (involved in ABA biosynthesis), chitinase family proteins, inositol monophosphatase family proteins, dehydration-induced proteins, several proteins that interact with ABA under drought, heat-shock proteins, lipid transfer proteins, and various transcription factors including WRKY32, WRKY57, MYB homeodomain-like superfamily proteins and bHLH nucleotide binding proteins (see Table S2 available as Supplementary Data at *Tree Physiology* Online) (Porth et al. 2005, Seki et al. 2007, Shinozaki and Yamaguchi-Shinozaki 2007, Spieß et al. 2012). Consistent with research in other oaks, we further find that senescence-associated and cell death-associated proteins are up-regulated and photosynthesis-related proteins are down-regulated (Spieß et al. 2012). In addition, genes whose patterns of nucleotide polymorphism implicate them in adaptation to local climate in *Q. lobata* (Gugger et al. 2016) are also enriched among the large set of genes responding to the drought treatment (40 of 50 climate-related genes from Gugger et al. are found in differentially expressed genes from this study; hypergeometric test: $P < 10^{-4}$), lending further support to their role in climate stress adaptation.

Local adaptation to drought

We find initial evidence of local adaptation to drought in the different responses of the different sampled populations to the drought stress treatment. Specifically, the expression response of 56 genes to the drought treatment depended on population of origin (treatment × population interaction) (Table 1, and Figures 2b and 4). Based on limited available information, these genes are not known to be involved in local climate adaptation in oaks (Gugger et al. 2016, Sork et al. 2016) nor *Arabidopsis* (Hancock et al. 2011), but have plausible functions for such involvement. These genes primarily function in metabolic processes, stress response, transport/transfer of other molecules, signaling and transcription regulation. In particular, HSP20-like chaperones superfamily protein is known to be involved in stress response and 'memory' (Stief et al. 2014), and other heat-shock proteins have been associated with water stress response in other oaks (Porth et al. 2005). In addition, phosphoglycerate kinase is involved in energy metabolism under stress (Larkindale and Vierling 2008) and protein kinase superfamily protein is involved in metabolic processes. Given the prevalence of molecules involved in metabolism (13 of 29 ascribed general functions), we hypothesize that local adaptation to drought in oaks may include different abilities to metabolize or different ways of altering metabolic activities in the face of drought. These different metabolic responses may be mediated by some of the

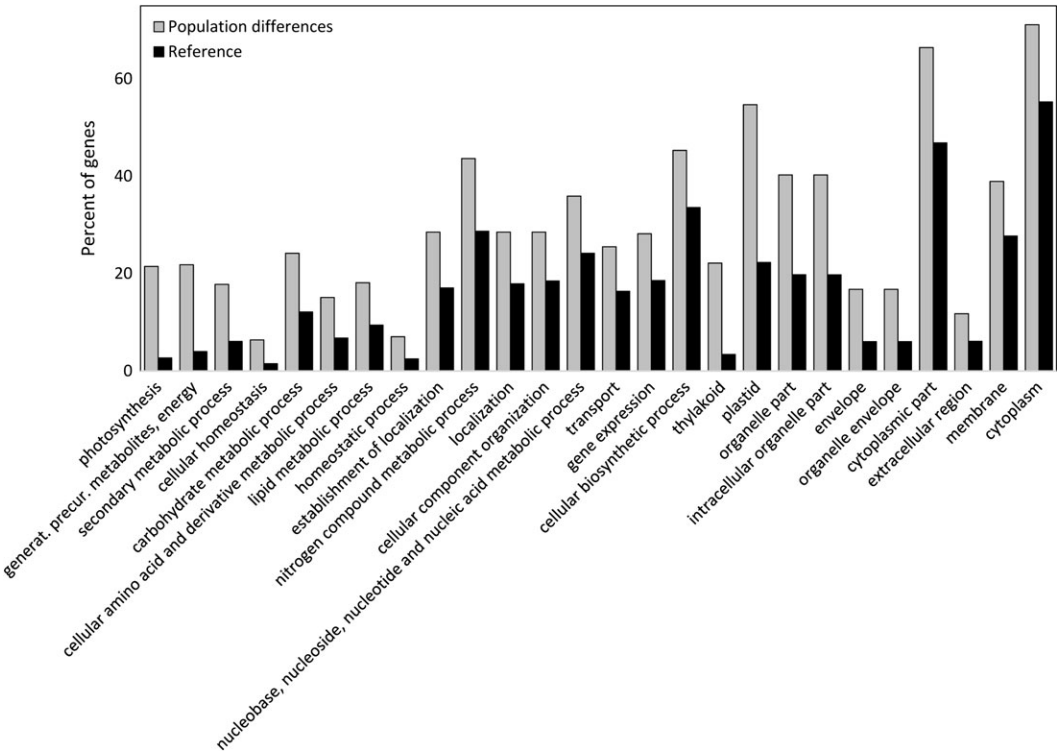

Figure 6. Results of singular enrichment analyses based on Plant GO slim functional categories for genes that were differentially expressed among sample sites.

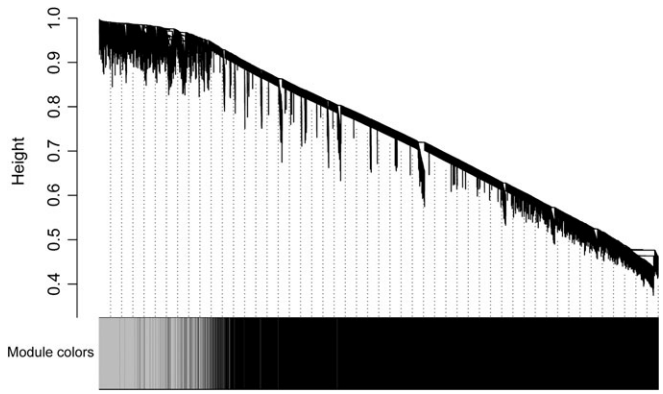

Figure 7. Hierarchical clustering dendrogram of regularized log-transformed gene expression and inferred modules of co-expressed genes (gray versus black) based on WGCNA.

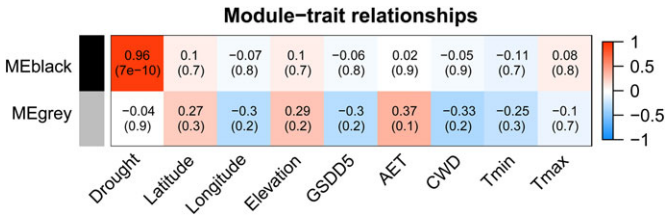

Figure 8. Heatmap of correlations (*P*- value in parentheses) of module eigengenes with the drought treatment, sample site coordinates and climate variables at the sample sites. Black and gray modules are as in Figure 7. GSDD5, growing season growing degree-days above 5 °C; AET, actual evapotranspiration; CWD, climatic water deficit; T_{min}, mean annual minimum temperature; and T_{max}, mean annual maximum temperature.

transcription factors among the list. These include: WRKY51, which is involved in regulating stress response to reduce reactive oxygen species and cell death (Rushton et al. 2010, Gao et al. 2011); an ATPase family gene involved in transcription regulation and stress response, possibly in an auxin signaling pathway (He et al. 2005); a homeodomain-like transcriptional regulator; or the proteins containing domains involved in signaling, such as m01oak50236Jt-t01.1 (Zn knuckle), m01oak29709cF-t01.1 (TIR domain) and m01oak18171JT-t01.1 (NB-ARC domain and leucine-rich repeats). Therefore, we speculate that, collectively, these molecules may be related to the commonly observed differences in water-use efficiency among populations of oaks (Aranda et al. 2007, Roussel et al. 2009).

Generally, we observed two clusters of genes that have similar drought treatment by population expression responses (Figure 5): one (lower half) corresponding to up-regulation in response to drought in MK and SV and modest up-regulation to down-regulation in the MC, and another (upper half) containing a number of population-specific patterns (Figure 5). Very few of these genes (e.g., m01oak15323cc, WRKY51; m01oak14227Cc, SPFH/Band7/PHB domain-containing membrane-associated protein

family) seem to follow a pattern that would suggest that high water stress sources (SV, MC) differ in response from low water stress sources (MK) (see Table S1 available as Supplementary Data at *Tree Physiology* Online), nor do they suggest a coast–interior divide or north–south divide as has been observed in genetic studies (Gugger et al. 2013, 2016). However, three seed sources are not sufficient to make robust inferences. These patterns of co-expression raise the possibility that tight interactions exist among the molecules in each cluster. Unfortunately, none of the eight oak genes assigned *Arabidopsis* orthologs are known or predicted to interact in *Arabidopsis* (Geisler-Lee et al. 2007), and such information is entirely lacking for oaks.

Baseline levels of expression in 1310 genes differed among populations in this study. These results do not inherently indicate anything about drought response, but fixed population differences in expression levels can contribute to local adaptation more generally. These genes tended to be involved in metabolism and cell parts (Figure 6), further implicating metabolic differences in population differentiation. These genes were not enriched among genes in the WGCNA 'grey' module, whose eigengene has moderate but insignificant correlations with the sample sites and their climates and has no correlation with the drought treatment (Figure 8). Nonetheless, each result independently offers lists of candidate genes for understanding differentiation among populations of *Q. lobata*.

In conclusion, this study provides initial evidence that seedlings from different localities and environments respond differently in gene expression, consistent with predictions based on local adaption. Given these differential responses to the treatment among populations and the different baseline expression levels among populations, we expect that different regional populations will have different responses and different capacities to respond to rapid climate change. This idea is consistent with inferences from ecological niche models suggesting differential responses among different regions (Sork et al. 2010) and a long history of provenance tests in trees suggesting local adaptation to environment and different tolerances for novel environments (Rehfeldt et al. 1999, Savolainen et al. 2007).

Acknowledgments

We thank Annette Delfino Mix for critical help with the greenhouse experiment. We are grateful to the following people for assistance with sampling and valuable discussions: Ana Albarrán-Lara, Lawren Sack, Megan Bartlett, Kimberly Crispin, Grace John, Pamela Thompson, Stephanie Steele, Keith Gaddis, Jianli Zhao, Jinming Chen, Sergio Nigenda-Morales and Rodrigo Méndez-Alonso. We also acknowledge the University of California Reserve System and the California Department of Parks and Recreation for access to some sample sites; the Institute of Forest Genetics, Pacific Southwest Research Station, Placerville, CA for rearing seedlings used in this experiment; the Broad Stem Cell Research Center at UCLA for its sequencing facilities; and the computational services available through the Hoffman2 Shared Cluster provided by the UCLA Institute for Digital Research and Education's Research Technology Group. Any use of trade, product or firm names is for descriptive purposes only and does not imply endorsement by the US Government.

Supplementary Data

Supplementary Data for this article are available at *Tree Physiology* Online.

Conflict of interest

None declared.

Funding

J.M.P.-R. was supported by a postdoctoral fellowship from UC MEXUS-CONACYT; V.L.S. received support for research and P.F.G. from UCLA; and J.W.W. received support from the Pacific Southwest Research Station, USDA Forest Service, which also funded the acorn collection and tree culturing.

References

Abrams MD (1990) Adaptations and responses to drought in *Quercus* species of North America. Tree Physiol 7:227–238.

Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11:R106.

Aranda I, Pardos M, Puértolas J, Jiménez MD, Pardos JA (2007) Water-use efficiency in cork oak (*Quercus suber*) is modified by the interaction of water and light availabilities. Tree Physiol 27:671–677.

Ashburner M, Ball CA, Blake JA et al. (2000) Gene Ontology: tool for the unification of biology. Nat Genet 25:25–29.

Bechtold U, Penfold CA, Jenkins DJ et al. (2016) Time-series transcriptomics reveals that AGAMOUS-LIKE22 affects primary metabolism and developmental processes in drought-stressed *Arabidopsis*. Plant Cell 28:345–366.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 57:289–300.

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29:1165–1188.

Bhardwaj AR, Joshi G, Kukreja B et al. (2015) Global insights into high temperature and drought stress regulated genes by RNA-Seq in economically important oilseed crop *Brassica juncea*. BMC Plant Biol 15:9.

Bray EA (1997) Plant responses to water deficit. Trends Plant Sci 2: 48–54.

Bray EA (2002) Classification of genes differentially expressed during water-deficit stress in *Arabidopsis thaliana*: an analysis using microarray and differential expression data. Ann Bot (Lond) 89:803–811.

Breeze E, Harrison E, McHattie S et al. (2011) High-resolution temporal profiling of transcripts during *Arabidopsis* leaf senescence reveals a distinct chronology of processes and regulation. Plant Cell 23: 873–894.

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

Cohen D, Bogeat-Triboulot M-B, Tisserant E et al. (2010) Comparative transcriptomics of drought responses in *Populus*: a meta-analysis of genome-wide expression profiling in mature leaves and root apices across two genotypes. BMC Genomics 11:630.

Cokus SJ, Gugger PF, Sork VL (2015) Evolutionary insights from *de novo* transcriptome assembly and SNP discovery in California white oaks. BMC Genomics 16:552.

Dai A (2013) Increasing drought under global warming in observations and models. Nat Clim Change 3:52–58.

Delfino Mix A, Wright JW, Gugger PF, Liang CF, Sork VL (2015) Establishing a range-wide provenance test in valley oak (*Quercus lobata* Née) at two California sites. Proceedings of the Seventh Oak Symposium: Managing Oak Woodlands in a Dynamic World. USDA Forest Service, Pacific Southwest Research Station, General Technical Report PSW-GTR-251, Albany, CA, pp 413–424.

Diffenbaugh NS, Swain DL, Touma D (2015) Anthropogenic warming has increased drought risk in California. Proc Natl Acad Sci 112: 3931–3936.

Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res 38:W64–W70.

Finn RD, Bateman A, Clements J et al. (2014) Pfam: the protein families database. Nucleic Acids Res 42:D222–D230.

Flint L, Flint A, Thorne J, Boynton R (2013) Fine-scale hydrologic modeling for regional landscape applications: the California Basin Characterization Model development and performance. Ecol Proc 2:25.

Gao Q-M, Venugopal S, Navarre D, Kachroo A (2011) Low oleic acid-derived repression of jasmonic acid-inducible defense responses requires the WRKY50 and WRKY51 proteins. Plant Physiol 155: 464–476.

Geisler-Lee J, O'Toole N, Ammar R, Provart NJ, Millar AH, Geisler M (2007) A predicted interactome for *Arabidopsis*. Plant Physiol 145: 317–329.

Gugger PF, Ikegami M, Sork VL (2013) Influence of late Quaternary climate change on present patterns of genetic variation in valley oak, *Quercus lobata* Née. Mol Ecol 22:3598–3612.

Gugger PF, Cokus SJ, Sork VL (2016) Association of transcriptome-wide sequence variation with climate gradients in valley oak (*Quercus lobata*). Tree Genet Genomes 12:1–14.

Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG, Toomajian C, Roux F, Bergelson J (2011) Adaptation to climate across the *Arabidopsis thaliana* genome. Science 334: 83–86.

He X-J, Mu R-L, Cao W-H, Zhang Z-G, Zhang J-S, Chen S-Y (2005) AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development. Plant J 44:903–916.

IPCC (2007) Climate change 2007: the physical science basis: summary for policymakers. Cambridge University Press, Cambridge, UK.

Kreps JA, Wu Y, Chang H-S, Zhu T, Wang X, Harper JF (2002) Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. Plant Physiol 130:2129–2141.

Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9:559.

Larkindale J, Vierling E (2008) Core genome responses involved in acclimation to high temperature. Plant Physiol 146:748–761.

Lasky JR, Des Marais DL, Lowry DB, Povolotskaya I, McKay JK, Richards JH, Keitt TH, Juenger TE (2014) Natural variation in abiotic stress responsive gene expression and local adaptation to climate in *Arabidopsis thaliana*. Mol Biol Evol 31:2283–2296.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760.

Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics 26:589–595.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Subgroup GPDP (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25: 2078–2079.

Love M, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.

McLaughlin BC, Zavaleta ES (2012) Predicting species responses to climate change: demography and climate microrefugia in California valley oak (*Quercus lobata*). Glob Chang Biol 18:2301–2312.

Osakabe Y, Osakabe K, Shinozaki K, Tran L-SP (2014) Response of plants to water stress. Front Plant Sci 5:86.

Ponton S, Dupouey J-L, Bréda N, Dreyer E (2002) Comparison of water-use efficiency of seedlings from two sympatric oak species: genotype × environment interactions. Tree Physiol 22:413–422.

Porth I, Koch M, Berenyi M, Burg A, Burg K (2005) Identification of adaptation-specific differences in mRNA expression of sessile and pedunculate oak based on osmotic-stress-induced genes. Tree Physiol 25:1317–1329.

R Development Core Team (2014) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

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

Ramírez-Valiente JA, Lorenzo Z, Soto A, Valladares F, Gil L, Aranda I (2009) Elucidating the role of genetic drift and natural selection in cork oak differentiation regarding drought tolerance. Mol Ecol 18: 3803–3815.

Rehfeldt G (2006) A spline model of climate for the western United States. USDA Forest Service, Port Collins, OH.

Rehfeldt GE, Ying CC, Spittlehouse DL, Hamilton DA Jr (1999) Genetic responses to climate in *Pinus contorta*: niche breadth, climate change, and reforestation. Ecol Monogr 69:375–407.

Roussel M, Le Thiec D, Montpied P, Ningre N, Guehl J-M, Brendel O (2009) Diversity of water use efficiency among *Quercus robur* genotypes: contribution of related leaf traits. Ann. For Sci 66:408–408.

Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010) WRKY transcription factors. Trends Plant Sci 15:247–258.

Savage JA, Cavender-Bares JM (2011) Contrasting drought survival strategies of sympatric willows (genus: *Salix*): consequences for coexistence and habitat specialization. Tree Physiol 31:604–614.

Savolainen O, Pyhajarvi T, Knurr T (2007) Gene flow and local adaptation in trees. Ann Rev Ecol Evol Syst 38:595–619.

Seki M, Umezawa T, Urano K, Shinozaki K (2007) Regulatory metabolic networks in drought stress responses. Curr Opin Plant Biol 10: 296–302.

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

Sork VL, Davis FW, Westfall R, Flint A, Ikegami M, Wang H, Grivet D (2010) Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata* Née) in the face of climate change. Mol Ecol 19:3806–3823.

Sork VL, Squire K, Gugger PF, Steele S, Levy ED, Eckert AJ (2016) Landscape genomic analysis of candidate genes for climate adaptation in a California endemic oak, *Quercus lobata* Née (Fagaceae). Am J Bot 103:33–46.

Spieß N, Oufir M, Matušíková I, Stierschneider M, Kopecky D, Homolka A, Burg K, Fluch S, Hausman J-F, Wilhelm E (2012) Ecophysiological and transcriptomic responses of oak (*Quercus robur*) to long-term drought exposure and rewatering. Environ Exp Bot 77:117–126.

Stebbins GL Jr (1952) Aridity as a stimulus to plant evolution. Am Nat 86:33–44.

Stief A, Altmann S, Hoffmann K, Pant BD, Scheible W-R, Bäurle I (2014) *Arabidopsis* miR156 regulates tolerance to recurring environmental stress through SPL transcription factors. Plant Cell 26: 1792–1807.

Swarbreck D, Wilks C, Lamesch P et al. (2008) The *Arabidopsis* Information Resource (TAIR): gene structure and function annotation. Nucleic Acids Res 36:D1009–D1014.

Tyler CM, Kuhn B, Davis FW (2006) Demography and recruitment limitations of three oak species in California. Quart Rev Biol 81:127–152.

Utsumi Y, Tanaka M, Morosawa T et al. (2012) Transcriptome analysis using a high-density oligomicroarray under drought stress in various genotypes of cassava: an important tropical crop. DNA Res 19: 335–345.

Villar E, Klopp C, Noirot C, Novaes E, Kirst M, Plomion C, Gion JM (2011) RNA-Seq reveals genotype-specific molecular responses to water deficit in eucalyptus. BMC Genomics 12:538.

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57–63.

Yates S, Swain M, Hegarty M, Chernukin I, Lowe M, Allison G, Ruttink T, Abberton M, Jenkins G, Skot L (2014) *De novo* assembly of red clover transcriptome based on RNA-Seq data provides insight into drought response, gene discovery and marker identification. BMC Genomics 15:453.

Yordanov I, Velikova V, Tsonev T (2000) Plant responses to drought, acclimation, and stress tolerance. Photosynthetica 38:171–186.

Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 41:17.