Exploiting Differential Gene Expression and Epistasis to Discover Candidate Genes for Drought-Associated QTLs in *Arabidopsis thaliana*

John T. Lovell,a,b,1 Jack L. Mullen,b David B. Lowry,c Kedija Awole,b James H. Richards,d Saunak Sen,e Paul E. Verslues,f Thomas E. Juenger,a,g and John K. McKayb

a Department of Integrative Biology, University of Texas, Austin, Texas 78712
b Department of BioAgricultural Sciences and Pest Management, Colorado State University, Fort Collins, Colorado 80523
c Department of Plant Biology, Michigan State University, East Lansing, Michigan 48824
d Department of Land, Air, and Water Resources, University of California, Davis, California 95616
e Department of Epidemiology and Biostatistics, University of California, San Francisco, California 94143
f Institute of Plant and Microbial Biology, Academia Sinica, Taipei 115, Taiwan
g Institute of Cellular and Molecular Biology, University of Texas, Austin, Texas 78712

ORCID IDs: 0000-0002-8938-1166 (J.T.L.); 0000-0001-8222-7927 (J.H.R.); 0000-0003-4519-6361 (S.S.)

Soil water availability represents one of the most important selective agents for plants in nature and the single greatest abiotic determinant of agricultural productivity, yet the genetic bases of drought acclimation responses remain poorly understood. Here, we developed a systems-genetic approach to characterize quantitative trait loci (QTLs), physiological traits and genes that affect responses to soil moisture deficit in the TSUxKAS mapping population of *Arabidopsis thaliana*. To determine the effects of candidate genes underlying QTLs, we analyzed gene expression as a covariate within the QTL model in an effort to mechanistically link markers, RNA expression, and the phenotype. This strategy produced ranked lists of candidate genes for several drought-associated traits, including water use efficiency, growth, abscisic acid concentration (ABA), and proline concentration. As a proof of concept, we recovered known causal loci for several QTLs. For other traits, including ABA, we identified novel loci not previously associated with drought. Furthermore, we documented natural variation at two key steps in proline metabolism and demonstrated that the mitochondrial genome differentially affects genomic QTLs to influence proline accumulation. These findings demonstrate that linking genome, transcriptome, and phenotype data holds great promise to extend the utility of genetic mapping, even when QTL effects are modest or complex.

INTRODUCTION

Traits that drive adaptation in ecological and agricultural systems are typically affected by the allelic state at many loci, the environmental conditions, and the interaction of genes with the environment (Falconer and Mackay, 1996; Mackay, 2001). Elucidation of how genes and environments interact to produce complex phenotypes is a long-standing problem and “grand challenge” in modern biology as well as crop breeding (Araus et al., 2002, 2008; Rockman, 2012; Heslot et al., 2014). Among the most ecologically and agriculturally important environmental factors is variation in soil moisture availability, which has driven the evolution of morphological and physiological traits (Stebbins, 1952; Axelrod, 1972; Juenger, 2013) and directly affects agricultural productivity (Condon et al., 2004; Cattivelli et al., 2008; Richards et al., 2010).

Plants have evolved diverse, complex, and often highly inducible responses to soil moisture variation (Chaves et al., 2003; Chaves and Oliveira, 2004; Heschel and Riginos, 2005; Bogeat-Triboulot et al., 2007; Harb et al., 2010; Rosenthal et al., 2010; Pinheiro and Chaves, 2011; Mir et al., 2012). For example, drought-adapted genotypes may avoid cellular dehydration through plasticity of many traits, including stomatal conductance (Chater et al., 2011), root and shoot growth (MacMillan et al., 2008), leaf wilting or rolling (Kadioglu and Terzi, 2007), and phenology (Heschel and Riginos, 2005; Sherrard and Maherali, 2006). Upon exposure to drought, many plants accumulate high levels of the stress hormone abscisic acid (ABA). ABA-mediated signaling is important for the regulation of various drought-responsive traits, including stomatal conductance, gene expression (Cutler et al., 2010), and accumulation of the compatible solute proline.

Many plants accumulate high levels of proline upon exposure to drought. In *Arabidopsis thaliana*, expression of Δ^1^-PYRROLINE-5-CARBOXYLATE SYNTHETASE1 (P5CS1) is strongly induced by abiotic stress (Szabados and Savouré, 2010; Sharma et al., 2011). P5CS1 catalyzes the probable rate-limiting step in stress-induced proline biosynthesis. Conversely, expression of PROLINE DEHYDROGENASE1 (ProDH1), which encodes a mitochondrial-localized proline catabolism enzyme, is repressed by stress in many plant tissues. It is thought that induction of P5CS1 and repression of ProDH1 expression is important to suppress proline turnover and maximize the accumulation of proline for osmotic adjustment (Leprince et al., 2015). However, other studies have suggested that continued mitochondrial proline catabolism also contributes to drought resistance by balancing cellular redox status and maintaining a favorable ratio of oxidized versus reduced NADP (Sharma et al., 2011).
DNA sequence variation in both nuclear and cytoplasmic (mitochondrial and plastid) genomes may underlie variation in drought response, including proline and ABA accumulation. For example, biosynthesis of the carotenoid precursors of ABA as well as the rate-limiting caroteneoid cleavage reaction catalyzed by the 9-cis-epoxyketonoid dioxygenase family of enzymes occur in the chloroplast. While nuclear genes encode the enzymes involved in ABA biosynthesis (Milborrow, 2001; Finkelstein, 2013), the later steps of ABA biosynthesis occur in the cytoplasm. Likewise, proline catabolism occurs in mitochondria but is catalyzed by nucleo-encoded enzymes. It is not established whether proline or ABA metabolism can be influenced by sequence variation of mitochondrion- or plastid-encoded genes.

Quantitative trait locus (QTL) mapping and global gene expression analyses are useful methods to assess the genetic basis of traits involved in drought adaptation (McKay et al., 2008; Hall et al., 2010; Juenger et al., 2010; Des Marais et al., 2012; Schmalenbach et al., 2014; El-Soda et al., 2015). In particular, the genomic perturbation of experimental crosses utilizes recombination to break up linkage disequilibrium and allows causal inference of how variation at a given locus leads to phenotypic variation. One of the goals of QTL mapping is the identification of the polymorphisms underlying heritable physiological variation. While broadly utilized for this goal, linkage-based QTLs do not provide a framework to distinguish among candidate genes with further fine-mapping and/or reverse genetics (Rockman, 2012). The most common method to define candidates underlying a QTL is to search for physically proximate genes with annotations or gene ontology reflecting the trait of interest (Al-Shahrour et al., 2005). While sometimes successful in model organisms, this approach may inhibit the discovery of new genes or candidates in species without annotated reference genomes.

Recently, several studies have combined gene expression and phenotypic trait QTL mapping in experimental populations (Chen et al. 2010; reviewed in Hansen et al., 2008; Cubillos et al., 2012). Through analysis of colocalization between differentially expressed genes and phenotypic trait QTLs, it is possible to produce lists of candidate genes (Swamy et al., 2013); however, the researcher is often left with long and unwieldy lists of candidate genes and phenotypic trait QTLs (cis-eQTLs). We then applied our covariate method to assess the effects of each candidate gene on the local QTL peak (of each candidate gene) as an additive covariate.

We used this approach in the context of a large-scale QTL analysis of drought physiology in Arabidopsis. The experimental population consisted of recombinant inbred lines (RILs) derived from a cross between the KAS-1 and TSU-1 ecotypes. These accessions originated from environments with very different water availability and differed in their water use efficiencies (McKay et al., 2003). Furthermore, strong ecological differentiation between TSU and KAS has been documented in life history (Lovell et al., 2013) and survival in drought conditions (McKay et al., 2008). To test which traits respond to drought and the genetic loci underlying this response, we conducted a quantitative genetic analysis of 39 total phenotypic traits related to drought adaptation, 18 of which had significant multiple-QTL models. Candidate genes were defined separately for each QTL-phenotypic trait combination as those genes within the QTL confidence interval with cis-acting transcript abundance QTLs (cis-eQTLs). We then applied our covariate method to assess the effects of each candidate gene on the local QTL peak morphology. We demonstrated the utility of this approach by recovering known causal genes underlying QTLs and developed ranked lists of candidate genes for each individual QTL. Finally, we combined the candidate gene search with observed epistatic and additive effects to document how known proline metabolism genes interact with newly discovered effects of mitochondrial natural variation to regulate proline accumulation in response to drought.

RESULTS

Quantitative Genetics of Drought Physiology

Utilizing a population of 341 KAS-1 × TSU-1 RILs, we examined the effects of a progressive drought treatment on a series of traits, including transcript abundance, metabolites, physiology,
growth, and performance. Our experimental soil moisture treatment reduced soil water potential to −2 MPa (Figure 1). This low moisture level approximates water deficits often experienced by both wild and crop plant species.

We measured 39 total phenotypic traits. Twelve phenotypic traits were measured in both well-watered and reduced water potential treatments (Table 1). These were related to growth, biomass partitioning, as well as key drought physiology traits. We also calculated and mapped QTLs for plasticity, which can also be considered a quantitative character (Falconer and Mackay, 1996). To assess plasticity for phenotypic traits measured in both environments, we calculated the RIL-specific difference between quantile normalized breeding values in wet and dry conditions. We quantified two-leaf transition responses to drought stress only in the dry environment: change in leaf width (rolling) and leaf length (wilting). Finally, we examined days to flowering (FT) for each RIL, using previously reported data (Lovell et al., 2013).

As expected, drought stress had a substantial effect on most measured phenotypic traits (Table 1). Growth traits responded particularly strongly. Leaf area was reduced by ∼50% in the dry treatment, and shoot fresh mass in the dry treatment was only 25% of that in the wet treatment (Figures 1B and 1C). For several phenotypic traits, including growth rate, the degree of plasticity was strongly positively correlated with treatment-specific breeding values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D). Physiological traits also exhibited strong stress responses. Leaf tissue concentrations of the compatible solute, proline, and carbon isotope ratio (δ13C), were measured in both environments. We quantified δ13C values (Figure 1D).

### Multiple-QTL Modeling

To define genomic regions associated with drought physiology, we implemented a stepwise model selection approach (Manichaikul et al., 2009) within the QTL mapping package R/qtl for all 39 traits (Broman et al., 2003). QTL mapping was conducted on the breeding values for each trait within each environment (n_wet = 13, n_dry = 14) and the plasticity estimate for each of the 12 traits measured in both environments.

Single- and multiple-QTL modeling revealed 36 significant QTLs across 18 of the 39 traits (Table 3, Figure 2; Supplemental Table 2). Strong QTL peaks were found for several traits on proximate chromosome 4 (Chr4) and distal Chr2. The Chr4 QTL was previously cloned as FRIGIDA for WUE (wet) and FT (Lovell et al., 2013); however, this marks the first documentation of colocализed QTLs for proline, root mass ratio (RMR), WUE plasticity, and wilting phenotypes at the same region. The QTL hot spot on Chr2 was centered on an extremely strong peak for proline (dry treatment, 24.9% variance explained) but also included single QTLs for WUE (dry), water content (dry), leaf rolling (dry), and proline plasticity. While correlations among traits with colocализed QTLs were generally weak, several pairs of traits were strongly correlated, including FT-WUE (wet) (r = 0.34, P < 0.0001; Supplemental Figure 2) and proline (dry)-WC (dry) (r = −0.35, P < 0.0001; Supplemental Figure 2B).

Aside from the FRIGIDA region, WUE QTLs also colocализed with those for FT on distal Chr4, proximate Chr5, and distal Chr5 (Figure 2; Supplemental Figure 3). In general, allelic variation underlying each QTL peak produced a parallel response of WUE and FT, as indicated by similar QTL profiles across the genome for each trait (Supplemental Figure 3A). For three-quarters of the pleiotropic QTLs, the KAS allele conferred later FT and higher WUE (Supplemental Figure 3B). However, for the pleiotropic FT/WUE QTL on distal Chr5, the KAS allele was associated with earlier FT and lower WUE (Supplemental Figure 3B).

We found plasticity QTL for proline, WUE, leaf area, and shoot growth rate (Table 3). The QTL peaks for these traits largely aligned with the strongest QTL peaks for either the wet (WUE) or dry main-effect phenotype. In these cases, we observed “weak” interactions, where the magnitude but not the direction of the QTL effect changed with the environment.

Since this population used reciprocal crosses, we were able to evaluate the effects of the cytoplasmic genomes. Adding cytoplasm as a factor improved model fit and was therefore retained in multiple QTL models for 12 of the phenotypic traits (Supplemental Table 2). In particular, there was a strong additive effect of cytoplasm on proline (10.6% of the total variance; Supplemental Tables 2 and 3; Figures 3A and 3B); however, for other phenotypic traits, cytoplasm generally explained little of the total variance (Supplemental Table 2). We tested the significance of QTL-cytoplasm epistasis post-hoc by iteratively fitting an interaction term between cytoplasm and each QTL in the final multiple-QTL model for each phenotypic trait where cytoplasm was retained in the model (Supplemental Table 4). This interaction was significant, improved model fit and was added to the model for two QTLs: plasticity of growth rate QTL 3@60 and WUE (dry) 3@58 (QTL identifier follows: phenotypic trait “QTL” Chr@CM position; Supplemental Figure 4). Interestingly, the WUE QTL 3@58 broadly colocализed with proline QTL 3@44, another QTL with a strong, albeit additive, effect of cytoplasm (Figure 3B).

QTL-QTL epistasis was found in only two phenotypic traits, proline and FT (Table 3; Supplemental Table 3). Proline was strongly increased in lines with KAS alleles at QTL 3@44, but TSU alleles at the main effect QTL 2@74 (Figure 3C). The latest FT phenotype was conferred by KAS alleles at both FT QTLs 4@3 and 4@62.

### Candidate Gene Analysis

Our candidate gene discovery method used a three-step approach to determine the effect of gene expression on the peak
QTL LOD score. First, all genes with significant gene expression polymorphism were extracted (for each QTL interval) by including only genes with significant cis-eQTLs (mapped in a recent study using the same gene expression data set; Lowry et al., 2013). Second, we ran QTL scans in which expression of each gene was iteratively added (and subsequently removed) as an additive covariate to the previously generated multiple-QTL model. This allowed us to determine the relative effect of transcript abundance of each gene on the focal QTL peak height. Finally, we ran a permutation test (10,000 permutations) for all genes in each QTL to determine significance of the effect (Figure 4).

The WUE QTL 4@4 was previously cloned, and the phenotypic variation resulted from DNA sequence variation that caused an expression polymorphism at FRIGIDA. Within the WUE QTL 4@4 confidence interval, there were 92 genes. Our candidate selection approach returned FRIGIDA (AT4G00650) as the strongest candidate gene for WUE. We tested the significance of the estimates by permuting the gene expression covariate data and rerunning the QTL scans. This test resulted in an empirical P = 0.0057 for FRIGIDA (Figures 4A and 4B). For the pleiotropic FT QTL 4@3, FRIGIDA was the second strongest candidate. While not the strongest in the list, FRIGIDA had a significant effect on the FT QTL 4@3 LOD score (P = 0.01).

The proline QTL 2@74 contained 239 genes, many more than the WUE QTL 4@4. This QTL colocalized with a well-documented sequence polymorphism in the proline biosynthesis gene, P5CS1 (AT2G39800), described by Kesari et al. (2012). We ran the same covariate screening process with the proline QTL 2@74, with the additive effect of cytoplasm included in the model. Incorporation of cytoplasm effects into candidate selection resulted in three genes with significant gene expression covariate effects, including P5CS1 (Figures 4C and 4D, Table 4; Supplemental Table 5; P > 0.001). These results provided a proof of concept that our method could discover causal genes in moderately large QTL regions.

To explore the utility of this method, we conducted covariate candidate gene scans for all QTLs with intervals spanning <25 centimorgans (cM), including 20 main effect QTLs and two epistatic QTL regions. For the two epistatic QTLs, with significant interactive effects but small main effects (FT QTL 4@63 and proline QTL 3@50), gene expression polymorphism was defined by the interactive effect between gene expression at the local and interacting QTLs. We then ranked these lists of potential candidates by the relative proportion of the LOD score absorbed by the gene expression covariate (Table 4; Supplemental Table 5). Overall, we screened 652 genes, 169 of which had an empirical q-transformed P < 0.1. Mean gene expression was highly elevated across significant candidate genes relative to those without significant effects (t_{427.5} = 9.87, P < 0.0001).

Several of the candidate genes produced from our methods had particularly interesting gene annotations. The strongest candidate for the FT QTL5@15 was AT5G17880 (CSA1), which

\[ r = 0.87 \] for the wet treatment, and \[ r = 0.06 \] for the dry treatment.

---

Figure 1. Physiological Effects of the Experimental Drought on Leaf Growth.

(A) Soil moisture release curve. Closed circles represent measurements of soil water potential across a range of soil water content. The open circle shows mean soil moisture at the end of the dry-down treatment. These data were used to estimate water potential for the dry-down treatment.

(B) Isolated overhead image of a genotype (KT207) in drought (left) and well-watered (right) conditions. The images have the same scale and color adjustments.

(C) Reaction norms for shoot fresh mass under different water availability treatments.

(D) Dependence of the change in leaf area during the treatment on total leaf area. For the wet treatment, \( r = 0.87 \), and for the dry treatment, \( r = 0.06 \).
Table 1. Variance Component Estimates and Summary Statistics for the Phenotypic Traits Measured

<table>
<thead>
<tr>
<th>Phenotype (Units or Calculation)</th>
<th>Abbreviation</th>
<th>Environment</th>
<th>Mean</th>
<th>SD</th>
<th>VarEnv (%)</th>
<th>VarRILxEnv (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growth Leaf area (cm²)</td>
<td>LA</td>
<td>Wet, dry, plast.</td>
<td>18.40</td>
<td>9.160</td>
<td>66</td>
<td>1</td>
</tr>
<tr>
<td>Growth rate (LA_{harvest}/LA_{pretreatment})</td>
<td>GR</td>
<td>Wet, dry, plast.</td>
<td>5.44</td>
<td>7.660</td>
<td>72</td>
<td>NA</td>
</tr>
<tr>
<td>Relative GR [ln(LA)-ln(GR-LA)]</td>
<td>RGR</td>
<td>Wet, dry, plast.</td>
<td>0.07</td>
<td>0.990</td>
<td>71</td>
<td>NA</td>
</tr>
<tr>
<td>Leaf wilting (%)</td>
<td>Wet, dry, plast.</td>
<td>3.16</td>
<td>2.350</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Leaf rolling (%)</td>
<td>Dry</td>
<td>3.93</td>
<td>2.280</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Biomass Shoot fresh mass (g)</td>
<td>SFM</td>
<td>Wet, dry, plast.</td>
<td>0.68</td>
<td>0.490</td>
<td>72</td>
<td>2</td>
</tr>
<tr>
<td>Shoot dry mass (g)</td>
<td>SDM</td>
<td>Wet, dry, plast.</td>
<td>0.06</td>
<td>0.026</td>
<td>36</td>
<td>0</td>
</tr>
<tr>
<td>Root dry mass (g)</td>
<td>RDM</td>
<td>Wet, dry, plast.</td>
<td>0.01</td>
<td>0.004</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Shoot:root ratio (SDM/RDM)</td>
<td>SR</td>
<td>Wet, dry, plast.</td>
<td>6.02</td>
<td>1.860</td>
<td>42</td>
<td>0</td>
</tr>
<tr>
<td>Root mass ratio [RDM/(RDM+SDM)]</td>
<td>RMR</td>
<td>Wet, dry, plast.</td>
<td>0.15</td>
<td>0.033</td>
<td>46</td>
<td>0</td>
</tr>
<tr>
<td>Grav. water content (SDM/SFM)</td>
<td>WC</td>
<td>Wet, dry, plast.</td>
<td>87.80</td>
<td>5.540</td>
<td>83</td>
<td>2</td>
</tr>
<tr>
<td>Physiology Abscisic acid conc. (µmol/g SDM)</td>
<td>ABA</td>
<td>Wet, dry, plast.</td>
<td>5.92</td>
<td>5.960</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>ABA (aqueous) (µmol/g SDM)*</td>
<td></td>
<td>Wet, dry, plast.</td>
<td>0.72</td>
<td>1.100</td>
<td>35</td>
<td>0</td>
</tr>
<tr>
<td>Water use efficiency (δ¹³C)</td>
<td>WUE</td>
<td>Wet, dry, plast.</td>
<td>−29.80</td>
<td>0.890</td>
<td>76</td>
<td>1</td>
</tr>
<tr>
<td>Proline conc (µmol/g SFM)</td>
<td>Proline</td>
<td>Wet, dry, plast.</td>
<td>93.60</td>
<td>102.500</td>
<td>85</td>
<td>4</td>
</tr>
<tr>
<td>Flowering time (days)</td>
<td>FT</td>
<td>Wet</td>
<td>23.70</td>
<td>4.580</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Genotype effects cannot be estimated for the growth rate traits because a single replicate was measured within each treatment. Environmental contributions cannot be estimated for the single environment phenotypes, wilting, rolling, and FT. Those traits that were not used for QTL mapping are indicated with an asterisk. SDM, shoot dry mass; SFM, shoot fresh mass.

Table 2. Genetic Correlations among Traits

<table>
<thead>
<tr>
<th>LA</th>
<th>SFM</th>
<th>SDM</th>
<th>RDM</th>
<th>RMR</th>
<th>ABA</th>
<th>WUE</th>
<th>Proline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dry</td>
<td>Wet</td>
<td>Dry</td>
<td>Wet</td>
<td>Dry</td>
<td>Wet</td>
<td>Dry</td>
<td>Wet</td>
</tr>
<tr>
<td>Wet</td>
<td>0.25*</td>
<td>0.91*</td>
<td>0.19*</td>
<td>0.85*</td>
<td>0.18*</td>
<td>0.74*</td>
<td>0.10</td>
</tr>
<tr>
<td>Dry</td>
<td>0.22*</td>
<td>0.89*</td>
<td>0.28*</td>
<td>0.59*</td>
<td>0.20</td>
<td>0.55*</td>
<td>0.04</td>
</tr>
<tr>
<td>Wet</td>
<td>0.19*</td>
<td>0.90*</td>
<td>0.16*</td>
<td>0.80*</td>
<td>0.15</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td>Dry</td>
<td>0.23*</td>
<td>0.69*</td>
<td>0.25</td>
<td>0.68*</td>
<td>0.16</td>
<td>0.20</td>
<td>0.03</td>
</tr>
<tr>
<td>SDM Wet</td>
<td>0.23*</td>
<td>0.71*</td>
<td>0.08</td>
<td>0.19</td>
<td>0.15</td>
<td>0.05</td>
<td>0.12*</td>
</tr>
<tr>
<td>Dry</td>
<td>0.38*</td>
<td>0.61*</td>
<td>0.18</td>
<td>0.07</td>
<td>0.06</td>
<td>0.11</td>
<td>0.02</td>
</tr>
<tr>
<td>RDM Wet</td>
<td>0.33*</td>
<td>0.37*</td>
<td>0.20</td>
<td>0.34*</td>
<td>0.05</td>
<td>0.03</td>
<td>0.26*</td>
</tr>
<tr>
<td>Dry</td>
<td>0.30*</td>
<td>0.59*</td>
<td>0.16</td>
<td>0.29*</td>
<td>0.05</td>
<td>0.05</td>
<td>0.48*</td>
</tr>
<tr>
<td>RMR Wet</td>
<td>0.37*</td>
<td>0.05</td>
<td>0.13</td>
<td>0.07</td>
<td>0.31*</td>
<td>0.02</td>
<td>0.17*</td>
</tr>
<tr>
<td>Dry</td>
<td>−0.26</td>
<td>0.15</td>
<td>0.21</td>
<td>0.35*</td>
<td>−0.10</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>ABA Wet</td>
<td>0.43*</td>
<td>−0.05</td>
<td>0.00</td>
<td>0.08</td>
<td>0.11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry</td>
<td>−0.04</td>
<td>0.02</td>
<td>0.05</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WUE Wet</td>
<td>0.21*</td>
<td>0.17*</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry</td>
<td>0.02</td>
<td>0.29*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proline Wet</td>
<td>−0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The asterisk (and bold) indicates a significant effect (P < 0.05). Data for ABA levels are ln-transformed to improve normality. Abbreviations are defined in Table 1.
Table 3. Multiple-QTL Model Statistics

<table>
<thead>
<tr>
<th>Trait</th>
<th>Treatment</th>
<th>Formula</th>
<th>%Var</th>
<th>P Value</th>
<th>pLOD</th>
</tr>
</thead>
<tbody>
<tr>
<td>RGR</td>
<td>Wet</td>
<td>y ~ cytoplasm + 3@054</td>
<td>7.754</td>
<td>0.016</td>
<td>2.418</td>
</tr>
<tr>
<td>WUE</td>
<td>Wet</td>
<td>y ~ cytoplasm + 3@018 + 3@065 + 4@04 + 4@42 + 5@037 + 5@089</td>
<td>34.301</td>
<td>&lt;0.001</td>
<td>8.621</td>
</tr>
<tr>
<td>FT</td>
<td>Wet</td>
<td>y ~ cytoplasm + 1@084 + 4@03 + 4@062 + 5@15 + 5@72</td>
<td>70.998</td>
<td>&lt;0.001</td>
<td>63.283</td>
</tr>
<tr>
<td>LA</td>
<td>Dry</td>
<td>y ~ 1@032 + 3@090</td>
<td>3.974</td>
<td>0.121</td>
<td>0.493</td>
</tr>
<tr>
<td>RDM</td>
<td>Dry</td>
<td>y ~ cytoplasm + 3@33</td>
<td>14.674</td>
<td>0.281</td>
<td>0.438</td>
</tr>
<tr>
<td>RMR</td>
<td>Dry</td>
<td>y ~ 3@19 + 4@01</td>
<td>35.884</td>
<td>0.029</td>
<td>1.419</td>
</tr>
<tr>
<td>GR</td>
<td>Dry</td>
<td>y ~ cytoplasm + 3@04</td>
<td>1.965</td>
<td>0.360</td>
<td>0.818</td>
</tr>
<tr>
<td>RGR</td>
<td>Dry</td>
<td>y ~ 1@12 + 3@30</td>
<td>6.003</td>
<td>0.043</td>
<td>0.207</td>
</tr>
<tr>
<td>Wilt</td>
<td>Dry</td>
<td>y ~ 4@02</td>
<td>7.251</td>
<td>0.006</td>
<td>0.469</td>
</tr>
<tr>
<td>Roll</td>
<td>Dry</td>
<td>y ~ cytoplasm + 2@74</td>
<td>9.765</td>
<td>0.005</td>
<td>2.259</td>
</tr>
<tr>
<td>WUE</td>
<td>Dry</td>
<td>y ~ 2@074 + 3@058 + 5@084</td>
<td>6.222</td>
<td>0.088</td>
<td>1.118</td>
</tr>
<tr>
<td>ABA</td>
<td>Dry</td>
<td>y ~ 2@16 (+2@16*cytoplasm)</td>
<td>4.349</td>
<td>0.034</td>
<td>1.324</td>
</tr>
<tr>
<td>Proline</td>
<td>Dry</td>
<td>y ~ cytoplasm + 2@074 + 3@050 + 4@03 + 2@74 + 3@050</td>
<td>44.165</td>
<td>&lt;0.001</td>
<td>21.949</td>
</tr>
<tr>
<td>WC</td>
<td>Dry</td>
<td>y ~ cytoplasm + 2@87</td>
<td>7.731</td>
<td>0.017</td>
<td>0.130</td>
</tr>
<tr>
<td>LA</td>
<td>Plasticity</td>
<td>y ~ cytoplasm + 1@03 + 3@027</td>
<td>10.124</td>
<td>0.011</td>
<td>0.327</td>
</tr>
<tr>
<td>GR</td>
<td>Plasticity</td>
<td>y ~ cytoplasm + 3@04 + 4@060 (+4@060*cytoplasm)</td>
<td>5.065</td>
<td>0.151</td>
<td>1.056</td>
</tr>
<tr>
<td>WUE</td>
<td>Plasticity</td>
<td>y ~ cytoplasm + 4@04</td>
<td>12.718</td>
<td>0.001</td>
<td>0.178</td>
</tr>
<tr>
<td>Proline</td>
<td>Plasticity</td>
<td>y ~ cytoplasm + 2@074</td>
<td>18.199</td>
<td>&lt;0.001</td>
<td>7.629</td>
</tr>
</tbody>
</table>

Degrees of freedom, percentage of variance explained by the model, and P values derived from $\chi^2$ tests were generated by fitting the QTL model with ANOVA. The penalized whole-model LOD score ($pLOD$) was derived from stepwise model selection, where models that increase $pLOD$ are retained. Significant interactions between the cytoplasm and QTLs were indicated in the formula; statistics for the interactions are reported in Supplemental Table 4.

Distal Chr5 contained a pleiotropic QTL for FT, WUE (wet), and WUE (dry). The effects of this QTL on WUE in both conditions were parallel, indicating that a single constitutive locus may have caused the QTL. Of the 24 unique significant candidates among these three traits, one gene was found in all three: AT5G55180 encoding a plasma membrane-localized O-glycosyl hydrolase that had not been annotated for any physiological attributes. However, members of the glycosyl hydrolase family of proteins are involved in a diverse array of physiological functions in plants, including drought responses (Bray, 2004), signaling, and development (Minic, 2008), which presents the possibility of pleiotropic gene action across phenological and physiological traits.

Isolation of the Quantitative Trait Nucleotides in Cytoplasmic Genomes

Since the multiple-QTL model for proline (dry) was strongly influenced by an additive effect of cytoplasm (>10% variance explained), we explored DNA sequence variation in the cytoplasmic genomes. We downloaded high-depth resequencing of both parental lines (Lowry et al., 2013). Sequence comparisons revealed that the chloroplast genomes of TSU and KAS had identical DNA sequences. However, 16 single-nucleotide polymorphisms (SNPs) existed between the mitochondrial genomes. Of these SNPs, four were in gene-coding regions of ATMG00050 (unknown function), ATMG00070 (NADH DEHYDROGENASE SUBUNIT 9 [NAD-9]), ATMG00510 (NAD-7), and ATMG00710 (unknown function) (Supplemental Table 6). Interestingly, across a diverse panel of Arabidopsis accessions (218 genotypes from the Salk Institutes 1001 genome resequencing project, http://signal.salk.edu/atg1001/index.php), NAD-9 has two major haplotypes (I, 40%; II, 43%) that exist at relatively equal frequencies, a departure from neutral expectations that instead suggests balancing selection at the locus (Supplemental Figure 7). The other three candidate genes have single dominant haplotypes with a small proportion of independent, rare mutations (Supplemental Figure 7).

We verified the presence of the NAD-9 SNP through capillary sequencing. The TSU-1 (a high proline accumulator) NAD-9 allele differed from the allele common to KAS-1 (a low proline accumulator) and the Columbia-0 reference. This result was consistent with an effect of the KAS cytoplasm genotype in the RIL population (Figures 3A and 3B). While SNPs in introns and untranslated regions were found within several other mitochondrial genes, the NAD-9 SNP is a missense mutation that caused valine to be substituted for phenylalanine in the 19th codon.

DISCUSSION

Drought adaptation involves multivariate and often correlated evolution of physiological, developmental, and life history phenotypes. Underlying these physiological responses are diverse patterns of sequence and gene expression variation. Comparisons of gene expression and physiological traits have revealed a complex genetic basis of drought responses (Liu et al., 1998; Shinozaki et al., 2003; Des Marais et al., 2012). By exploiting the causal connections between environmental variation and the genotype-phenotype map, here, we present candidate genes for constitutive and plastic responses to soil moisture reduction in the context of QTL mapping.

Genetic Architecture of Drought-Responsive Traits

Our soil moisture reduction treatment imposed drought stress on the TSU-KAS RIL mapping population (Figures 1A and 1B).
and induced physiological responses across many phenotypic traits (Table 1). For example, all lines increased proline concentration in drought conditions. Increased proline may contribute to osmotic adjustment and cellular redox balance (Szabados and Savouré, 2010; Verslues and Sharma, 2010), traits that may confer improved cellular dehydration tolerance.

Even though our experimental design was able to detect loci that explained as little as 1.8% of the total phenotypic variance (Table 3), for the majority of constitutive traits and plasticity estimates, we found no QTLs. The low number and small-effect sizes of QTLs determined in our analysis were indicative of a genomic architecture of drought-responsive traits that was decisively polygenic. While loci of small effect were the most common observation, there were several genomic regions that explained a very large proportion of phenotypic variance. In particular, the FT QTL 4@3 (FRIGIDA), RMR QTL 3@18, and proline QTL 2@74 (P5CS1) explained 59, 19.7, and 25% of the total variation, respectively (Table 3). In concert with observed (and unobserved) small-effect loci, the presence of these large-effect loci, and several moderate effect-size QTLs (e.g., RMR QTL4@1 and root mass QTL 3@33) provided evidence for an exponential distribution of allelic effects on potentially adaptive traits (Orr, 1998). This pattern has been observed in other recent physiological QTL mapping studies (Ågrena et al., 2013; Joseph et al., 2013a).

Genetic Correlations Underlie Pleiotropic QTLs

Correlations among drought acclimation responses can directly affect the fitness (yield) of genotypes when challenged with low soil water potentials. For example, mild early season drought may simultaneously select for cellular dehydration avoidance through stomatal closure (Heschel et al., 2002) and reduced growth rates (Schmalenbach et al., 2014). Alternatively, strong late season drought may select for drought escape through early flowering and fast growth while soil water conditions are favorable (Meyre et al., 2001; Heschel and Riginos, 2005; Sherrard and Maherali, 2006). Genetic correlations among many of these traits have been observed both within (McKay et al., 2003; Lovell et al., 2013) and among (Angert et al., 2009) species, providing further evidence that selection acts on both the plasticity of and correlations among drought-associated phenotypes (Endler, 1986).

In our population, a “dehydration avoidance” drought adaptive strategy was conferred by increased water use efficiency, decreased growth rate, increased water foraging through root growth,

Figure 2. Mapping Positions of Significant QTLs.

QTL point estimates (filled circles) and accompanying drop 1.5 (solid colored lines) confidence intervals for all phenotypes with significant multiple QTL models. Phenotypes in red were collected in the drought treatment. Blue-labeled phenotypes were from the “wet” treatment, and purple traits are plasticity estimates. The two focal regions for candidate gene method validation are highlighted in green.

Figure 3. Interaction Plots for Proline QTLs 2@74 (Best Candidate P5CS1), 3@44 (ProDH), and the Effect of Cytoplasmic Variation.

Allelic means ± se are plotted.
we found evidence for parallel responses where allelic effects for unique significance and phenology. QTLs for FT and WUE colocalized for 4/7 of the traits, especially between WUE for many QTLs (Supplemental Table 2).

2012), an observation consistent with the effects of the TSU allele (Sherrard and Maherali, 2006; Wilczek et al., 2009; Banta et al., 2012). Ability usually favors rapid cycling and drought escape strategies exist. In annual plants such as Arabidopsis, greater moisture availability throughout the growing season, such as Northwest India, may favor a dehydration-avoidant strategy (Heschel et al., 2002; Blum, 2005; Kesari et al., 2012). Capillary sequencing revealed a functional polymorphism at P5CS1, where intronic sequence variation yielded a reduced function allele in KAS, which was nearly identical to the reduced function allele of P5CS1 previously described for the ecotype “Sha” (Kesari et al., 2012). Aside from FRIGIDA and P5CS1, we were able to define candidate genes for all other narrow QTLs. Several of these were annotated to have similar effects as those shown in our physiological assays. For example, we found CSA1 (CONSTITUTIVE SHADE AVOIDANCE1) as a candidate for FT QTL 5@70. Furthermore, the much wider, but colocated WUE QTL 5@37 also had CSA1 as a strong candidate (J.T. Lovell, unpublished data). CSA1 responds directly to shade and red/far red light ratios (Supplemental Figure 5), altering life history and vegetative growth structure (Faigón-Soverna et al., 2006). There is significant physiological crosstalk between shade avoidance, drought physiology, and phenology (Maliakal candidate genes for any QTL.

We tested the effects of 652 genes across all QTLs with intervals narrower than 25 cM. Transcript abundance of 169 of these genes significantly (Q-transformed P_{10000permutations} < 0.1) affected the LOD score of the overlying QTL. Interestingly, these 169 represent a highly expressed subset of the total 25,662 genes with expression data.

We were able to recover the top two priori candidates for the main FT and proline QTLs, FRIGIDA and P5CS1, respectively. FRIGIDA is a vernalization-responsive transcription factor that affects flowering and pleiotropically drives variation in WUE (Lovell et al., 2013). In our RIL population, lines with the low-expression TSU alleles flowered earlier and had lower WUE, a drought escape life history strategy. Alternatively, KAS alleles were associated with dehydration avoidance physiology through increased WUE and FT. These effects were mediated by stomatal conductance and other upstream physiological traits (Lovell et al., 2013). P5CS1 catalyzes the rate-limiting step in proline biosynthesis. The induction of P5CS1 gene expression and subsequent increased P5CS1 protein abundance is required for high levels of proline accumulation (Kesari et al., 2012). Capillary sequencing revealed a functional polymorphism at P5CS1, where intron-sequence variation yielded a reduced function allele in KAS, which was nearly identical to the reduced function allele of P5CS1 previously described for the ecotype “Sha” (Kesari et al., 2012).

Candidate Genes for Drought Adaptation

One of the goals of genetic mapping is to discover the regions, interactions, and ultimately the genes that underlie physiological variation. Many approaches permit inference of potential candidate genes underlying QTLs, including CIT (Schadt et al., 2005), differential expression analyses (Drake et al., 2006; Farber et al., 2009), and partial regressions (Bing and Hoeschele, 2005); however, it is difficult to rank or infer the effect of each without significant additional data, such as reverse genetics experiments. By combining gene expression data with genetic mapping approaches, we presented a method to define and rank sets of candidate genes for any QTL.

We found strong evidence of a genetic basis for correlations that confered drought-adaptive syndromes, especially between WUE and phenology. QTLs for FT and WUE colocated for 4/7 of the unique significant loci across the traits. Additionally at each locus, we found evidence for parallel responses where allelic effects for each trait produced phenotypic vectors of similar, positive orientation (Supplemental Figure 3). Even where the sign of the effects reversed (e.g., QTL 5@70), the correlation conferred by alleles at this QTL remained positive. This effect, which antagonized the degree of physiological differentiation between TSU and KAS, increased the strength of the drought adaptive trade-off between dehydration avoidance and drought escape in the RIL population. Finally, the genome-wide correlation between FT and WUE QTL locations suggests that pleiotropy (at the QTL level) between these traits is not unique to the FRIGIDA locus (Lovell et al., 2013).

and delayed flowering. Dehydration-avoidant phenotypic values and QTL effects were typical of KAS alleles, while TSU alleles conferred a drought escape strategy. Habitats with consistently limited precipitation throughout the growing season, such as Northwest India where the KAS genotype originated (McKay et al., 2008), may favor a dehydration-avoidant strategy (Heschel et al., 2002; Blum, 2005; Schmalenbach et al., 2014). By contrast, TSU originates from Southwestern Japan, where more mesic growing season conditions exist. In annual plants such as Arabidopsis, greater moisture availability usually favors rapid cycling and drought escape strategies (Sherrard and Maherali, 2006; Wilczek et al., 2009; Banta et al., 2012), an observation consistent with the effects of the TSU allele for many QTLs (Supplemental Table 2).

We found strong evidence of a genetic basis for correlations that conferred drought-avoidance syndromes, especially between WUE and phenology. QTLs for FT and WUE colocaled for 4/7 of the unique significant loci across the traits. Additionally at each locus, we found evidence for parallel responses where allelic effects for each trait produced phenotypic vectors of similar, positive orientation (Supplemental Figure 3). Even where the sign of the effects reversed (e.g., QTL 5@70), the correlation conferred by alleles at this QTL remained positive. This effect, which antagonized the degree of physiological differentiation between TSU and KAS, increased the strength of the drought adaptive trade-off between dehydration avoidance and drought escape in the RIL population. Finally, the genome-wide correlation between FT and WUE QTL locations suggests that pleiotropy (at the QTL level) between these traits is not unique to the FRIGIDA locus (Lovell et al., 2013).

Candidate Genes for Drought Adaptation

One of the goals of genetic mapping is to discover the regions, interactions, and ultimately the genes that underlie physiological variation. Many approaches permit inference of potential candidate genes underlying QTLs, including CIT (Schadt et al., 2005), differential expression analyses (Drake et al., 2006; Farber et al., 2009), and partial regressions (Bing and Hoeschele, 2005); however, it is difficult to rank or infer the effect of each without significant additional data, such as reverse genetics experiments. By combining gene expression data with genetic mapping approaches, we presented a method to define and rank sets of candidate genes for any QTL.

We tested the effects of 652 genes across all QTLs with intervals narrower than 25 cM. Transcript abundance of 169 of these genes significantly (Q-transformed P_{10000permutations} < 0.1) affected the LOD score of the overlying QTL. Interestingly, these 169 represent a highly expressed subset of the total 25,662 genes with expression data.

We were able to recover the top two priori candidates for the main FT and proline QTLs, FRIGIDA and P5CS1, respectively. FRIGIDA is a vernalization-responsive transcription factor that affects flowering and pleiotropically drives variation in WUE (Lovell et al., 2013). In our RIL population, lines with the low-expression TSU alleles flowered earlier and had lower WUE, a drought escape life history strategy. Alternatively, KAS alleles were associated with dehydration avoidance physiology through increased WUE and FT. These effects were mediated by stomatal conductance and other upstream physiological traits (Lovell et al., 2013). P5CS1 catalyzes the rate-limiting step in proline biosynthesis. The induction of P5CS1 gene expression and subsequent increased P5CS1 protein abundance is required for high levels of proline accumulation (Kesari et al., 2012). Capillary sequencing revealed a functional polymorphism at P5CS1, where intron-sequence variation yielded a reduced function allele in KAS, which was nearly identical to the reduced function allele of P5CS1 previously described for the ecotype “Sha” (Kesari et al., 2012).

Aside from FRIGIDA and P5CS1, we were able to define candidate genes for all other narrow QTLs. Several of these were annotated to have similar effects as those shown in our physiological assays. For example, we found CSA1 (CONSTITUTIVE SHADE AVOIDANCE1) as a candidate for FT QTL 5@15. Furthermore, the much wider, but colocated WUE QTL 5@37 also had CSA1 as a strong candidate (J.T. Lovell, unpublished data). CSA1 responds directly to shade and red/far red light ratios (Supplemental Figure 5), altering life history and vegetative growth structure (Faigón-Soverna et al., 2006). There is significant physiological crosstalk between shade avoidance, drought physiology, and phenology (Maliakal
### Table 4. List of the Top 10 Significant Candidate Genes for Each QTL Ranked by LOD Effect

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>QTL</th>
<th>%Var</th>
<th>Candidate Genes (Ordered by P Value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA.plast</td>
<td>1@3</td>
<td>(-) 3.13</td>
<td>AT1G655490, FAS1</td>
</tr>
<tr>
<td>RGR.dry</td>
<td>1@12</td>
<td>(-) 3.30</td>
<td>AT2G03140, AT2G04380**, ALDH6B2, AT2G04170*, UGT773B5, AT2G05915</td>
</tr>
<tr>
<td>LA.dry</td>
<td>1@32</td>
<td>(-) 3.48</td>
<td></td>
</tr>
<tr>
<td>FT.wet</td>
<td>1@84</td>
<td>(+) 3.96</td>
<td>ACX5*, AT2G36460, CTF2A, GLTP1*</td>
</tr>
<tr>
<td>ABA.dry</td>
<td>2@16</td>
<td>(-) 5.16</td>
<td>AT2G38800*, P5CS1, PROT1*</td>
</tr>
<tr>
<td>Roll.dry</td>
<td>2@74</td>
<td>(-) 6.23</td>
<td>AT2G36460, CTF2A, GLTP1*</td>
</tr>
<tr>
<td>Proline.dry</td>
<td>2@74</td>
<td>(+) 13.6</td>
<td>AT2G38800*, P5CS1, PROT1*</td>
</tr>
<tr>
<td>Proline_plast</td>
<td>2@74</td>
<td>(+) 3.6</td>
<td>AT4G24430, emb2076, AT3G43430, AT3G43670, ATIVD, CYP77A5P</td>
</tr>
<tr>
<td>WUE.dry</td>
<td>2@74</td>
<td>(+) 3.6</td>
<td>AT4G24430, emb2076, AT3G43430, AT3G43670, ATIVD, CYP77A5P</td>
</tr>
<tr>
<td>WC.dry</td>
<td>2@87</td>
<td>(+) 3.34</td>
<td>MAC3B, AT2G43210, AT2G46220, IQD14, AT2G46100**, AT2G38800, AT2G32150, AT2G42490, AT2G32160</td>
</tr>
<tr>
<td>GR.dry</td>
<td>3@87</td>
<td>(-) 4.56</td>
<td>CBSDFUCH1, RPL18AC*, AT3G18530, AT3G14595</td>
</tr>
<tr>
<td>GR_plast</td>
<td>3@87</td>
<td>(-) 4.38</td>
<td>CBSDFUCH1, RPL18AC*, AT3G18530, AT3G14595</td>
</tr>
<tr>
<td>WUE.wet</td>
<td>3@91</td>
<td>(+) 19.7</td>
<td>AT3G31230, AT3G28080, AT3G27250, ATCSLC04, AT3G30300*, AT3G26670*, PRODH, CAF1-9*, ATMYB30</td>
</tr>
<tr>
<td>RMR.dry</td>
<td>3@91</td>
<td>(+) 19.7</td>
<td>AT3G31230, AT3G28080, AT3G27250, ATCSLC04, AT3G30300*, AT3G26670*, PRODH, CAF1-9*, ATMYB30</td>
</tr>
<tr>
<td>LA.dry</td>
<td>3@27</td>
<td>(-) 4.18</td>
<td>AT3G18750, AT3G25240, BRT1 (UGT84A2)</td>
</tr>
<tr>
<td>RGR_plast</td>
<td>3@27</td>
<td>(-) 4.18</td>
<td>AT3G18750, AT3G25240, BRT1 (UGT84A2)</td>
</tr>
<tr>
<td>RDM.dry</td>
<td>3@32</td>
<td>(+) 19.2</td>
<td>AT4G24250, AT4G22990, AT4G24050, ATCSLC04, AT3G26670*, PRODH, CAF1-9*, ATMYB30</td>
</tr>
<tr>
<td>Proline.dry</td>
<td>3@50</td>
<td>(-) 4.91</td>
<td>AT4G24250, AT4G22990, AT4G24050, ATCSLC04, AT3G26670*, PRODH, CAF1-9*, ATMYB30</td>
</tr>
<tr>
<td>RGR.wet</td>
<td>3@54</td>
<td>(-) 6.69</td>
<td>AT3G44430, emb2076, AT3G43430, AT3G43670, ATIVD (IVD), ATML03 (ML03)*, AT3G45555</td>
</tr>
<tr>
<td>WUE.dry</td>
<td>3@58</td>
<td>(-) 3.76</td>
<td>HR4, ABC2 Homolog 1 (ATATH1)<em>, scpl48, ALDH2B4**, AT3G47580, IVD, AT3G53730</em>, ATEXLA1, AT3G51470**, CSR1</td>
</tr>
<tr>
<td>WUE.wet</td>
<td>3@65</td>
<td>(-) 5.14</td>
<td>HR4, ABC2 Homolog 1 (ATATH1)<em>, scpl48, ALDH2B4**, AT3G47580, IVD, AT3G53730</em>, ATEXLA1, AT3G51470**, CSR1</td>
</tr>
<tr>
<td>LA.dry</td>
<td>3@90</td>
<td>(+) 3.30</td>
<td>ATSTE24, SAM-2*</td>
</tr>
<tr>
<td>RMR.dry</td>
<td>4@81</td>
<td>(+) 15.9</td>
<td>AT4G00740, FRI, ECA2</td>
</tr>
<tr>
<td>Wilt.dry</td>
<td>4@82</td>
<td>(-) 4.01</td>
<td>AT4G00740, FRI, ECA2</td>
</tr>
<tr>
<td>FT.wet</td>
<td>4@83</td>
<td>(-) 59.1</td>
<td>AT4G01130, SAM-2*, ATSTE24, AT4G00270, AT4G02540, ML01</td>
</tr>
<tr>
<td>Proline.dry</td>
<td>4@83</td>
<td>(-) 3.46</td>
<td>AT4G01130, SAM-2*, ATSTE24, AT4G00270, AT4G02540, ML01</td>
</tr>
<tr>
<td>WUE.wet</td>
<td>4@84</td>
<td>(-) 7.09</td>
<td>FRI</td>
</tr>
<tr>
<td>WUE.plast</td>
<td>4@84</td>
<td>(+) 3.47</td>
<td>FRI</td>
</tr>
<tr>
<td>WUE.wet</td>
<td>4@84</td>
<td>(+) 2.99</td>
<td>FRI</td>
</tr>
<tr>
<td>GR.plast</td>
<td>4@84</td>
<td>(-) 4.82</td>
<td>CRK22, CRK21, AT4G22990, AT4G24050, ISU1, AT4G21910, ATSBT3.12, TOM1*</td>
</tr>
<tr>
<td>FT.wet</td>
<td>4@86</td>
<td>(-) 5.19</td>
<td>CRK22, CRK21, AT4G22990, AT4G24050, ISU1, AT4G21910, ATSBT3.12, TOM1*</td>
</tr>
<tr>
<td>FT.wet</td>
<td>5@15</td>
<td>(-) 1.83</td>
<td>CSA1, CHS3, AT5G18950, AT5G16890, AT5G17680, PAT1(TRP1), GDH1**, ATCBR*</td>
</tr>
<tr>
<td>WUE.wet</td>
<td>5@37</td>
<td>(-) 3.49</td>
<td>ATATG18F, AT5G55180, AT5G54710, AT5G54720, ORA*, ATM2</td>
</tr>
<tr>
<td>FT.wet</td>
<td>5@72</td>
<td>(+) 4.24</td>
<td>ATATG18F, AT5G55180, AT5G54710, AT5G54720, ORA*, ATM2</td>
</tr>
<tr>
<td>WUE.dry</td>
<td>5@84</td>
<td>(+) 3.53</td>
<td>XYL4**, AGL62, AT5G61660, ABA1, LECRK110, AT5G62350, DARS, AT5G55180, KCA2</td>
</tr>
<tr>
<td>WUE.wet</td>
<td>5@89</td>
<td>(+) 6.33</td>
<td>AT5G54710, AT5G55180, AT5G54720*, ARF2*, ATATG18F, AT5G63020, AT5G53700, SNRK2-3, AT5G60160**, AT5G61600</td>
</tr>
</tbody>
</table>

The QTL-specific percentage of variance explained is presented and is preceded by the direction of the QTL effect. Positive values indicate a higher mean of the TSU allele. Candidate genes for QTL with confidence intervals that spanned >25 cM are not listed. Bold font indicates those genes discussed in the text. Genes without DNA sequence or with neither DNA sequence nor protein divergence between TSU and KAS alleles are marked by one asterisk or two asterisks, respectively.
et al., 1999; Schmitt et al., 2003), raising the possibility that CSA1 pleiotropically affects phenological and physiological traits in the TSUxKAS population, a hypothesis that requires further testing.

Several studies have found that single QTL peaks can fractionate into multiple linked peaks, each caused by separate, but linked polymorphisms (Studer and Doebley, 2011; Johnson et al., 2012). To examine this possibility, we screened for QTLs with multiple strong candidate genes with correlated expression patterns. Several genes within the ABA 2@15 QTL returned strong covariate effects. Of these, two genes displayed gene expression phenotypes that were highly correlated, AT2G03140 and AT2G04380. To determine the relative strength of each candidate, we analyzed DNA sequence and protein polymorphism within the genes and compared TAIR-10 gene annotations. Only AT2G03140 had any sequence polymorphism between the TSU and KAS parental accessions (Supplemental Table 6). While gene function can easily be affected by polymorphism outside of the coding region, genes that are conserved between the widely divergent genomes of TSU and KAS may be less likely to contribute to quantitative genetic divergence than those with sequence, and especially protein, divergence (Lowry et al., 2013).

It is important to note that this approach does not definitively document functional effects of loci. Instead, genes such as AT2G03140, which lack functional annotation for drought response or ABA biosynthesis, necessitate further functional genetic analyses. With that said, AT2G03140 is a highly interesting candidate gene as it encodes a putative chloroplast-localized protein with similarity to CAAX amino terminal proteases involved in membrane anchoring of proteins (Choy et al., 1999). As the early steps of ABA biosynthesis occur within the chloroplast (Endo et al., 2008; Cutler et al., 2010; Lee et al., 2013) and involve lipid-soluble carotenoids and membrane-associated enzymes (Milborrow, 2001; Seo and Koshiba, 2002), the connection of AT2G03140 to ABA metabolism is plausible and highly promising for further analysis.

**Documentation of Genetic Networks through QTL Mapping**

Our candidate gene discovery method made use of the details of genetic architecture, through incorporation of epistasis, additive effects of cytoplasm (and other covariates of interest), and environmental interactions. For example, this approach validated the effects of P5CS1 and permitted inference of potential candidate quantitative trait nucleotides in the cytoplasmic genome. However, as with any other candidate gene selection approach, our method provides hypotheses and does not document functional variation of candidate genes.

Our analysis of the parental genomes (sequences were published in Lowry et al., 2013) revealed that the chloroplast genomes of TSU and KAS were identical and only 16 SNPs existed in the mitochondrial genome. These results differed from published TSU-1 (SRX246442) and KAS-1 (SRX246466) sequences (http://www.ncbi.nlm.nih.gov/sra; Joseph et al., 2013b), where many more cytoplasmic SNPs were documented. However, despite having the same name, the KAS-1 and TSU-1 on the short read archive are not closely related to the KAS-1 and TSU-1 genomes that represent the parents of our mapping population. For example, over half of the published sequenome SNPs that are polymorphic in our mapping population are monomorphic between NCBI TSU-1 and KAS-1 (http://naturalvariation.org/hapmap). It is important to note that, while well suited for the characterization of SNPs, the short read sequencing and reference-based alignment used by Lowry et al. (2013) to sequence the mapping parents TSU and KAS may be unable to detect genomic rearrangements. As large-scale rearrangements are characteristic of the Arabidopsis mitochondrial genome (Davila et al., 2011), it is possible that these sequence variants are augmented by other undetected polymorphisms.

Proline concentration was strongly affected by cytoplasmic variation in our mapping population, indicating that sequence variation in the cytoplasmic genomes affected quantitative variation of adaptive traits. While the plastid genomes were monomorphic, there were 16 mitochondrial SNPs between the TSU and KAS parents of our mapping population, four of which were genic, including two in genes encoding NADH dehydrogenase subunits (NAD-7 and NAD-9). The observation that two of the mitochondrial polymorphisms were in genes for NADH dehydrogenase subunits is consistent with proposals that proline accumulation is tightly related to cellular redox status and that proline catabolism in the mitochondria is important in drought resistance (Sharma et al., 2011). For example, Szabados and Savouré (2010) and Verslues and Sharma (2010) found that proline metabolism is connected to oxidation/reduction status, and Sharma et al. (2011) showed that mitochondrial catabolism of proline is required to maintain growth under low water potential. We observed that p5cs1-4, which is blocked in stress-induced proline accumulation, was associated with strongly upregulated expression of a number of genes for NAD(P)H-dehydrogenases as well as additional genes related to mitochondrial respiration (P.E. Verslues, unpublished data).

To determine which of the NAD genes was the most likely candidate, we conducted capillary sequencing and downloaded sequence data from 218 natural accessions. Interestingly, while the SNP within NAD-7 was within an intron, the NAD-9 SNP was a missense mutation. Furthermore, there was evidence for historical balancing selection at NAD-9, but not NAD-7. While neutral evolution should yield many low-frequency haplotypes, and directional selection would reduce the number of variants, historical balancing selection should yield multiple haplotypes at elevated frequencies without a single dominant haplotype (reviewed in Nielsen, 2005). While three of the four candidate genes in the mitochondrial genome had a single dominant haplotype, the two main haplotypes of NAD-9 were maintained at >40% across a sample of 218 accessions. These results are consistent with the findings of Joseph et al. (2013b), who demonstrated that the mitochondrial genes for the NADH dehydrogenase complex harbor many more sequence polymorphisms than expected by neutral evolution. Given the nonsynonymous nature of the SNP, and evidence of historical balancing selection, it is possible that variation in NAD-9 affects proline catabolism in the mitochondria, a process that has consequences for redox balance and growth during drought.

To assay the effects of epistasis, we incorporated gene expression patterns of interacting loci (or covariates) into the candidate selection approach. Since many epistatic loci lacked strong additive effects, we fit a model where the expression of each gene underlying the epistatic QTL was a function of the local genotype and gene expression of the interacting QTL. Candidate determination of the proline epistatic locus QTL3@44 (with P5CS1) revealed a strong candidate: AT3G30775, ProDH1. These loci and...
patterns, and protein structure in hand, it would be possible to cal-
tain of sequence and expression data to bolster our candidate
tical information that may improve the extensibility of our approach.
Finally, it is important to note that the genotype-phenotype cas-
cade operates in many fashions. Our approach can discover only
a distinct subset where transcript abundance is the causal phenon-
type underlying physiological traits. In particular, we expect our ap-
proach to be limited by the strength of correlations between
physiological and gene expression phenotypes. For example, if
genomes are expressed at a similar rate, but alternative splicing or RNA
sequence polymorphism causes protein and trait variation, we ex-
pact to have little power to detect signals of connections between
candidate genes and QTLs. However, sequence-based gene ex-
pression quantification methods, such as RNA-seq, provide addi-
tional information that may improve the extensibility of our approach.
For example, with information on alternative splicing, methylation
patterns, and protein structure in hand, it would be possible to cal-
culate a measure of gene functionality. As such, functionality, and not
simply transcript abundance, could be used as a covariate in our method.
While the data presented here come from microarray technology and do not permit such inference, we expect a combi-
nation of sequence and expression data to bolster our candidate
gene approach in future analyses.

METHODS

Plant Materials and Growth Conditions

Seed of 341 RILs from reciprocal crosses between Arabidopsis thaliana ac-
cessions KAS (Kas-1; CS903) and TSU (Tsu-1; CS1640), along with the pa-
rents were sown on fritted clay (Profile Products) in 2.5-inch pots in duplicate
in each of two blocks. The Tsu-1 x Kas-1 mapping population is publically
available through the ABRC (ID: CS97026). Seeds were planted in a ran-
domized complete block design, and then the pots were refrigerated at 4°C in
darkness for 6 d to cold-stratify the seeds prior to commencement of a 12-h
photoperiod in two Conviron ATC60 growth chambers (Controlled Environ-
ments), at 23°C and 40% humidity during the day and 18°C and 50% humidity
during the dark period. Light intensity was ~330 μmol m⁻² s⁻¹. After 4 weeks
of growth, half of the plants were given a drought treatment, while the others
remained fully watered. Two replicates of each RIL were randomly assigned
to each treatment.

The drought treatment consisted of a slow decrease in soil moisture con-
tent over the course of 1 week. The treatment was imposed at the level of the
flat (tray of 32 plants) and randomized within each chamber. Each day, all
pots assigned to the drought treatment were weighed, and water was added to
individual pots to bring them up to the target gravimetric water content.
The target water content decreased each day, in the following series: 100, 90,
80, 70, 60, 45, and 40% of saturation. We had previously calculated the soil
moisture release curve for fritted clay: 40% soil moisture content relates to
approximately ~2 MPa soil water potential (Figure 1A).

Phenotypic Analyses

At the end of the drought treatment, photographs were taken of each
plant, and the shoots were excised at the hypocotyls and weighed to
obtain shoot fresh mass. The shoots were then freeze-dried and their dry
mass was measured. In a subset of 240 plants, root tissue was collected
by rinsing away the fritted clay. Root tissue was then freeze-dried for dry
mass determination.

Photographs of the plants were taken and used to calculate leaf area by
summing pixels comprising the rosette image using the image processing
software Scion Image (Scion). For half of the plants, a photograph was also
taken prior to the onset of the dry-down treatment, so that we could calculate
growth in leaf area during the treatment and relative growth rate.

ABA was assayed with a Phytodetek enzyme-linked immunosorbent assay
kit from Agdia. Samples were prepared and measured according to the
protocol from Agdia. Each sample and eight standards were run in duplicate on
32-well Phytodetek plates. A BioTek PowerWave HT spectrophotometer was
used to quantify the absorbances (at 450 nm), which were fit to the standards
of each plate using a logistic equation. To reduce residual variance caused by
fresh mass variance among and within treatments, we performed all additional
analyses on ABA concentrations standardized by the dry mass of the rosette.

Leaf tissue from each plant was crushed and lyophilized to quantify δ¹³C
using a dual-inlet mass spectrometer at the Stable Isotope facility at Uni-
viersity of California, Davis. Proline concentration was assayed by an acid
rhynhydrin assay adapted to 96-well plates (Bates et al., 1973; Verslues, 2010).

In addition to traits measured in both environments, we collected several
environment-specific phenotypes. Flowering time for each line was measured in
a separate experiment (Lovell et al., 2013) and reanalyzed here to make
comparisons with all other phenotypes. Plants in the drought treatment were
photographed both at the onset and the conclusion of the drought treatment.
Fully expanded leaf characteristics were compared between the two time-
points to determine the degree of rolling (lwidth) and wilting (llength).

Quantitative Genetic Analyses

The phenotypic data set was analyzed with a linear mixed model, with RIL as
a fixed effect and treatment, and the genotype-treatment interaction as random
effects. These models and variance component estimates were calculated
using PROC MIXED in the SAS software package (SAS Institute). Least square
means of trait values were estimated for each RIL, and genetic correlations
among traits were calculated as the standard Pearson pairwise correlations.
Phenotypes were in general very normal; however, FT and ABA were both
marginally skewed. Quantile normalizations of these traits did not strongly
influence our QTL analysis (for detailed comparisons, see Supplemental Methods),
so we opted to map the raw breeding values (Supplemental Data Set 1A).

A linkage map for this population was described previously (McKay et al.,
2008). To this map we added eight additional simple sequence length poly-
merism markers and an additional 276 single nucleotide polymorphism
markers, based on Sequenom technologies (Supplemental Data Set 1B; For
Ooijen, 2006) with the Kosambi mapping function, for a total of 450 markers.

Genotype probabilities were calculated for each locus and a set of
pseudo-markers were placed in any region with a >1 cM gap in the map. The
kosambi algorithm with an error probability of 0.01 was used to infer genotype
probabilities. QTL mapping was performed using the Haley-Knott regression
algorithm implemented in the R/qtl package within the R statistical computing
environment (Broman et al., 2003; Broman and Sen, 2009). We developed
multiple-QTL models via a penalized stepwise model selection approach
(Manichaikul et al., 2009) where terms were included at
multiple-QTL models via a penalized stepwise model selection approach
(Manichaikul et al., 2009) where terms were included at
Manichaikul et al., 2009) where terms were included at
Manichaikul et al., 2009) where terms were included at
degree of rolling (lwidth) and wilting (llength).

To achieve increased accuracy in our estimates of QTL peak means and
breadth, we calculated confidence intervals (1.5 LOD drop) for each QTL
point estimate separately by varying the position of the focal QTL while

Soil Moisture QTLs in Arabidopsis 11 of 15

breadth, we calculated confidence intervals (1.5 LOD drop) for each QTL
point estimate separately by varying the position of the focal QTL while

breadth, we calculated confidence intervals (1.5 LOD drop) for each QTL
point estimate separately by varying the position of the focal QTL while

breadth, we calculated confidence intervals (1.5 LOD drop) for each QTL
point estimate separately by varying the position of the focal QTL while
controlling for all other terms in the model. We also conducted multiple-QTL modeling using cytoplasm as an additive covariate and conducted post-hoc tests for cytoplasm-QTL interactions for all QTLs in those models with evidence for additive effects of cytoplasm.

It is important to note that we defined the significance of all covariate affects (e.g., gene expression and cytoplasm-QTL interactions) post-hoc. That is, the base multiple-QTL model was defined following Manichaikul et al. (2009). The QTL positions and interactions from this model were fixed for all other analyses. To determine significance of additional effects, we added and then removed a single term. Significance was determined by comparing the fit of the original and more complex QTL models.

All scripts and pipelines to conduct the QTL and covariate analyses have been posted on github: https://github.com/jtlovell/r-QTL_functions.

Candidate Gene Analysis

We downloaded gene expression and DNA sequence data from Lowry et al. (2013), who used Affymetrix atSNPtile1.0 arrays (Zhang et al., 2007) to map eQTLs for the T SUxKAS population from RNA extracted from the experiment presented here. Gene expression data was available for a 104 RIL subset of our mapping population. To identify candidate genes, we conducted a three-step protocol that combined this gene expression data with QTL mapping results and covariates.

The first step of our candidate gene identification approach was to define a list of candidate genes that had significant gene expression polymorphism for each phenotypic trait QTL interval. For each QTL confidence interval for each trait, we extracted all genes that fell between the maximum and minimum physical positions of all markers within the interval. This approach was necessary because there are many small rearrangements throughout the population relative to the Columbia reference genome. Additionally, there is a single large inversion on Chr4 (Supplemental Figure 8). If the QTL interval was so narrow that it only included a single marker, the interval was expanded to the nearest bounding markers. For the majority of QTLs that had simple additive effects, we defined gene expression polymorphism as those genes with cis-eQTLs (Lowry et al., 2013). For other QTLs with strong additive effects of cytoplasm, differentially expressed genes were further culled to only those with significantly different expression between cytoplasmic backgrounds using fixed effects ANOVA. Finally, for QTLs with primarily epistatic effects, candidates were determined as those that have gene expression polymorphism that is significantly associated with gene expression of the primary candidate at the interacting QTL. In the latter two cases, significance was assessed with q-value estimation in the R package “q-value” (Dabney and Storey, 2014).

The second step of our analysis was to rank the candidate genes by their effect on the local phenotypic trait QTL. To accomplish this, we extracted RI L-specific gene expression values for all candidate genes. These expression values were iteratively added to the original QTL model (which may include several QTLs, cytoplasmic covariates, and epistasis) as a single additive term, or in the case of epistasis, as an interactive covariate with the epistatic QTL. To make comparisons with identical patterns of missing data, the entire genotype, gene expression, and phenotype matrices were culled so there was no missing data. This reduced the peak LOD scores for all QTLs, but made comparisons among models possible. Gene expression covariates that explained residual variance had higher LOD scores (negative difference) and those that were correlated with the phenotypic trait breeding values decreased the LOD score of the focal peak (Supplemental Figure 1). Therefore, we took the difference between LOD scores at the QTL point estimate in the original multiple-QTL model lacking the expression covariate and the new model with a gene-expression covariate as the estimated effect of that gene. We then obtained a ranked list for each QTL, where the strongest candidate gene had the most positive covariate effect.

The last step was to determine the significance of a subset of genes with the strongest covariate effect. To accomplish this, we permuted the gene expression data and reran the covariate scan 10,000 times and reported the LOD difference at the QTL point estimate. The number of permuted observations with a greater difference than the empirical data/ \( n_{perm} \) was used as our empirical P value.

Accession Numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: Locus:2127013, AT4G00650, FRIGIDA; Locus:504954491, ATMGO0070, NAD9; Locus:2063907, AT2G39800, P5CS1; Locus:2089706, AT3G30775, PRODH1; Locus:2161710, AT5G51680, O-GLYCOSYL HYDROLASE FAMILY 17 PROTEIN; Locus:2170333, AT5G17880, CSA1; and Locus:2056891, AT2G03140, \( \alpha/\beta \)-HYDROLASE SUPERFAMILY PROTEIN.

Supplemental Data

Supplemental Figure 1. Visualization of the concept of the covariate scan approach.

Supplemental Figure 2. Correlation of phenotypes with colocalized QTL on proximate Chr4 and distal Chr2.

Supplemental Figure 3. Effect of allelic variation on the correlation between WUE and FT.

Supplemental Figure 4. Cytoplasmic interactions with genomic QTLs.

Supplemental Figure 5. Validation of the allelic effect of CSA1 using NILs.

Supplemental Figure 6. Hierarchical clustering of the covariance of all genes within each narrow QTL interval.

Supplemental Figure 7. Haplotyping of the four genes that contained SNPs in the mitochondrial genome.

Supplemental Figure 8. Comparison of the physical position (bp) for all TAIR10 gene models with the mapping position in cM.

Supplemental Table 1. Phenotypic correlations between plasticity and mean breeding values for all measured phenotypic traits.

Supplemental Table 2. Summary statistics for all terms in each QTL model.

Supplemental Table 3. T statistics for the additive effect of cytoplasm.

Supplemental Table 4. Significance of cytoplasm epistasis on each QTL.

Supplemental Table 5. Significance, effect, and divergence of each candidate gene in each narrow QTL.

Supplemental Table 6. List of cytoplasmic SNPs between TSU and KAS.

Supplemental Data Set 1A. Complete phenotypic trait data.

Supplemental Data Set 1B. Complete genotype matrix.
Supplemental Methods. Additional information pertaining to the analytical pipeline for candidate gene analyses, QTL methods, and materials.

ACKNOWLEDGMENTS

Our work was supported by National Science Foundation Grants DEB 0420111 and DEB 0618347 to J.K.M., T.E.J., and J.H.R. and an Academia Sinica Career Development Award to P.E.V. J.T.L. was supported by NSF PRFB IOS-1402393. We thank A. Heiliger for assistance with phenotypic trait data collections and B. Gibson for laboratory assistance. K. Broman provided helpful feedback on the QTL analyses. Comments from D. Kliebenstein, J. Weber, D. Sloan, S. Schwartz, L. Milano, and two anonymous reviewers greatly improved earlier versions of the article.

AUTHOR CONTRIBUTIONS


Received February 8, 2015; revised March 13, 2015; accepted April 1, 2015; published April 14, 2015.

REFERENCES


Exploiting Differential Gene Expression and Epistasis to Discover Candidate Genes for Drought-Associated QTLs in *Arabidopsis thaliana*

John T. Lovell, Jack L. Mullen, David B. Lowry, Kedija Awole, James H. Richards, Saunak Sen, Paul E. Verslues, Thomas E. Juenger and John K. McKay

*Plant Cell*; originally published online April 14, 2015; DOI 10.1105/tpc.15.00122

This information is current as of June 27, 2017