TITLE

Genetic Diversity In Gene Regulatory Interactions Underlies The Response To Soil Drying In The Model Grass *Brachypodium Distachyon*

SHORT TITLE

Gxe In Gene Regulation

AUTHORS

Jie Yun[1], Angela C. Burnett[2,a], Alistair Rogers[2,b], David L. Des Marais[1,c]

[1]Civil & Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

[2]Environmental & Climate Sciences Department, Brookhaven National Laboratory, Upton, NY 11973, USA

[a]Current Affiliation: Advanced Research + Invention Agency, Alan Turing Institute, British Library, 96 Euston Road, London, NW1 2DB, UK angie.burnett@aria.org.uk

[b]Current Affiliation: Climate & Ecosystem Sciences Division, Lawrence Berkeley National, Laboratory, Berkeley, CA 94720, USA alistair@lbl.gov

[c]Corresponding author: dldesmar@mit.edu

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (https://academic.oup.com/plcell/pages/General-Instructions) is: David L. Des Marais (dldesmar@mit.edu).

ABSTRACT

Gene expression is a quantitative trait under the control of genetic and environmental factors and their interaction, so-called GxE. Understanding the mechanisms driving GxE is fundamental for ensuring stable crop performance across environments, and for predicting the response of natural populations to climate change. Gene expression is regulated through complex molecular networks, however environmental and genotypic effects on genome-wide regulatory networks are rarely considered. In this study, we model genome-scale gene expression variation between two natural accessions of the model grass *Brachypodium distachyon* and their response to soil drying. We identified genotypic, environmental, and GxE responses in physiological, metabolic, and gene expression traits. We then identified gene regulation conservation and variation among conditions and genotypes, simplified as co-expression clusters found unique in or conserved across library types. Putative gene regulatory interactions are inferred as network edges with a graphical modelling approach, resulting in hypotheses about gene-gene interactions specific to -- or with higher affinity in -- one genotype, one environmental treatment, or in one genotype under treatment. We further find that some gene-gene interactions are conserved across conditions such differential expression is apparently transmitted to the target gene. These variably detected edges cluster together in co-expression modules, suggestive of different constraints or selection strength acting on specific pathways. We further applied our graphical modeling approach to identify putative, environmentally dependent regulatory mechanisms of leaf glucose content as an exemplar metabolite. Our study highlights an approach to identify variable features of gene regulatory networks and thereby identify key components for later genomic intervention to elucidate function or modulate environmental response. Our results also suggest possible targets of evolutionary change in gene regulatory networks associated with environmental plasticity.

INTRODUCTION

Gene expression represents a unique class of quantitative, complex traits. Like all quantitative traits, the transcript abundance of an individual gene can take a wide range of values, can show variation in those values among samples from a population, and this variation may be controlled by a number of both intrinsic and extrinsic factors [1–4]. And, because gene expression is a component of many quantitative traits such as morphology, life history, behavior, and physiology [5], using tools from quantitative genetics to study gene expression can provide insight into the molecular control of these higher-order traits [6].

Classically, variation in a quantitative trait can be modeled as a function of genetic differences among sampled individuals, of the effect of the environment on those individuals (i.e., phenotypic plasticity), and of the interaction between genotype and environment, GxE [7]. Recently, several studies have used expression QTL (eQTL) mapping to identify genetic loci that are associated with differences in the abundance of transcripts among natural accessions of plants [8–11]; these genetic variants describe heritable, constitutive differences in expression among genotypes. Abundant studies have likewise demonstrated the strong effects that differences in environment can have on the expression of single genes or suites of genes (e.g.,[12–15]).

Perhaps owing to the important role afforded phenotypic plasticity in plant ecology and evolution [16], hypotheses about its molecular basis pre-date modern genomic science [17] and remain central to studies of plant molecular biology, evolution, and biotechnology[18–24]. Recent studies assessing the prospects for biotechnology to enhance plant response to environmental stress often begin with transcriptomic, proteomic, or metabolomic assays of molecular responses to simulated stressors [25–29]. These -omics studies are predicated on the hypothesis that regulators of transcriptional activity offer attractive targets to manipulate suites of downstream molecular processes and, ultimately, organism-scale responses to stress (e.g., as described in [25,30]). However, identifying genes which can be manipulated to improve performance in the field while not also incurring a fitness penalty (so-called yield drag) remains challenging [31]. The molecular variants driving natural genetic variation in environmental response – GxE – represent a means to study how small genetic changes drive environmentally

responsive physiological and developmental traits [7]. Because such genetic variants are naturally occurring and have likely been exposed to generations of natural selection in wild populations, understanding their function *in vivo* may provide insight into how to design synthetic systems to improve stress tolerance without incurring yield drag. Genetic variants affecting GxE also represent the population-level variation on which natural selection might act to optimize fitness in response to climate change [32] and are thus of interest for predicting the future of natural plant communities.

GxE in plant gene expression has been reported in previous studies, most often at the level of individual genes or at global scale treating each gene as an independent variable [33,34]. Such an approach can be confounded by the nature of gene regulation, in which suites of genes are often under common regulatory control [35]; as a result, treating each gene as independent is difficult to justify statistically. Genomic features such as promoter and enhancer elements, post-transcriptional modification, and chromatin accessibility [36–40], add further complexity to co-variance among transcript abundance *in vivo*. Collectively, the interaction between genes at the transcriptional level can be conceptualized as a gene regulation network (GRN), within which there are dependencies of expression among genes and the possible flow of information through a cell [41,42]. The topology of GRNs implies some regulators directly regulate or are "upstream" of many genes and may therefore exhibit pleiotropic effects when perturbed. GRNs, and the sub-networks, modules, or communities that comprise them, may be seen as an integrated molecular phenotype that allows for identifying transcriptional associations with physiological, metabolic, or developmental functions, thereby improving biological interpretability [43].

The dependency of gene co-expression and the resulting topology of GRNs on environmental variation is poorly understood [44]. By extension, we have a limited understanding of how environmentally responding networks evolve and might be successfully manipulated using biotechnology to improve crop resilience. For example, genes exhibiting GxE in their expression might be clustered into modules, as it is hypothesized that environmental pressure results in the evolution of modules such that organismal sub-systems can function independently [45]. A study on *Caenorhabditis elegans* found that genes exhibiting GxE in their expression have longer promoters than

125 genome averages, with a higher density of regulatory motifs [46]. An extension of this
126 work also proposed that GxE genes might be associated with a greater number of *trans*-
127 acting regulators than genes showing simple genotypic differential expression [47].
128 Previously, we showed that the number of genes connecting to and the influence of
129 genes showing GxE in expression are lower under drought stress but higher under cold
130 stress compared to a random set of genes [48], again suggesting that the topology of
131 GRNs could have been shaped by – or possibly shaped the efficacy of -- past selection
132 in response to environmental conditions. However, that study, like most studies of
133 environmentally dependent gene expression, assumed that co-expression relationships
134 are invariant across genotypes or environmental treatments, which limits our ability to
135 understand the evolution of complex gene expression phenotypes [42,49].
136 Here, we use two inbred natural accessions of the model grass *Brachypodium*
137 *distachyon* [50] to study the physiological and transcriptomic patterns of GxE in response
138 to soil water deficit. *B. distachyon* is closely related to economically significant grasses
139 such as wheat, barley, and oat, and it has the advantages of small stature and small
140 genome size [50,51]. Moreover, as a natural species with no history of artificial selection, it
141 may exhibit higher phenotypic plasticity as compared to domesticated crops selected in
142 stable environments with high resource input [7]. Indeed, we previously documented GxE
143 in leaf water and metabolite content in *B. distachyon* response to soil drying[52]. *B.*
144 *distachyon* populations are found throughout the Mediterranean where they most likely
145 exhibit a winter annual strategy – germinating in the fall or spring and setting seed
146 before the typically hot, dry summer months[52,53]. We use highly replicated samples of
147 factorial combinations of genotype and environment, harnessing variation from
148 stochastic variance of gene expression and micro-environmental heterogeneity among
149 replicates to reconstruct regulatory interactions [54,55].
150
151 **RESULTS**
152 *Physiological responses to soil drying*
153 We first performed a drying experiment in which we measured a series of traits daily on
154 two inbred natural accessions as soil water content was experimentally decreased from
155 85% of field capacity to 50% of field capacity. These measurements suggest that both

156 accessions began reducing whole plant water usage starting on the fourth day of the
157 dry-down (Figure 1A, t-test adjusted by Tukey, p=0.048) and thereafter experienced a
158 similar magnitude of soil drying stress with most assayed traits exhibiting clear
159 differences between control and drying treatments by Day 6. Carbon assimilation
160 (Figure 1C; p= 0.0379) and stomatal conductance (Figure 1D; p= 0.0265) were
161 significantly lower in droughted plants starting on day 5. Total plant fresh weight started
162 to decrease in the droughted plants on day 6 (Figure 1E; p-value<0.0001) and exhibited
163 significantly lower dry weights (Figure 1F; p-value=0.0018) on day 6. Dry root:shoot
164 ratio is slightly higher in drought on day 7 (Figure 1G; p-value= 0.0143). Leaf hydraulic
165 potential is marginally more negative in droughted plants (Figure 1B; p-value=0.09) on
166 day 6 and is significantly more negative under drought on Day 7 (p-value<0.0001).
167 Relative chlorophyll content exhibits lower but non-significant values under drought
168 (Figure 1J; F-test, p-value=0.08). Linear mixed models show that both accessions
169 respond similarly along this dry-down, with the exception of leaf glucose and fructose
170 contents, which showed significant increase starting on Day 5 in Bd3-1 (z-test,
171 p<0.0001) and Day 6 in Bd21 (p<0.0001; Figure 1H and Supplementary Figure 1).

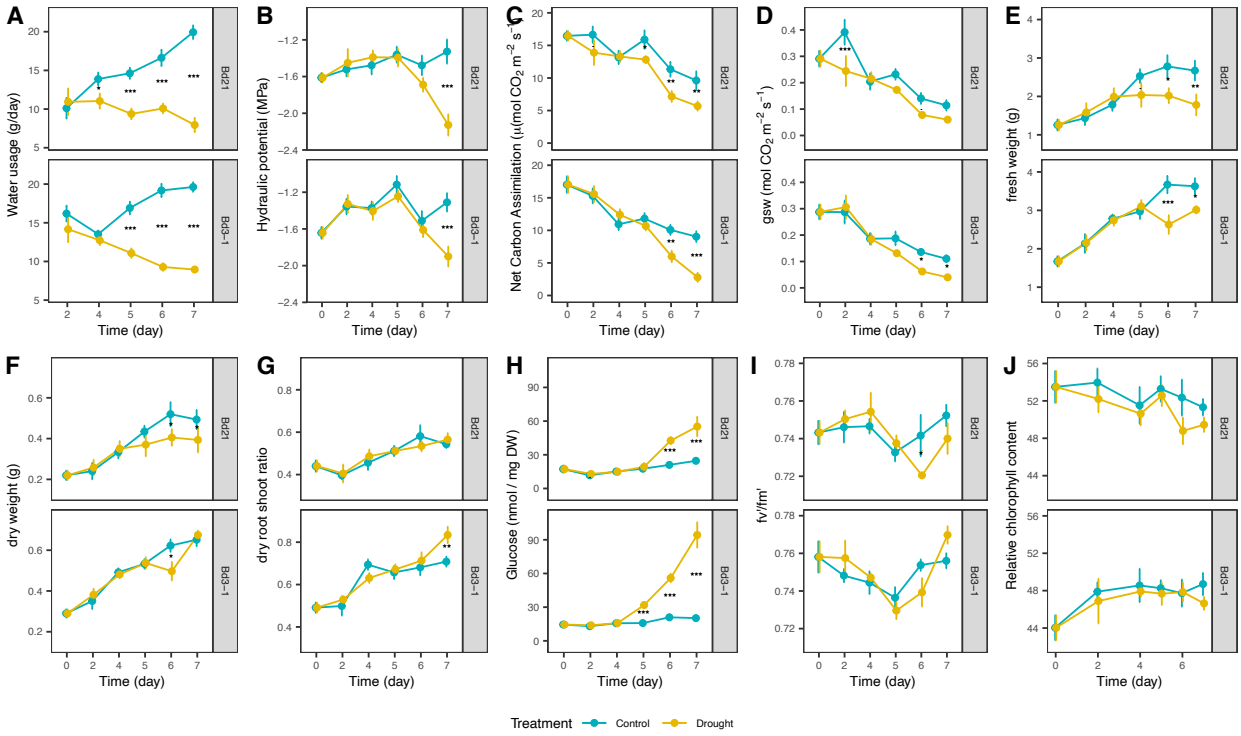


Figure 1. Comparison of physiological and metabolic traits between control and dry-down in two accessions from Day 0 to Day 7. (Generalized) Linear models were determined based on AIC score. Random-effects components of the models were based on likelihood ratio tests. Differences of means shown here with asterisks, were determined using contrasts between the two treatments at each time point in each genotype by Tukey's test. N = 9 for each point.

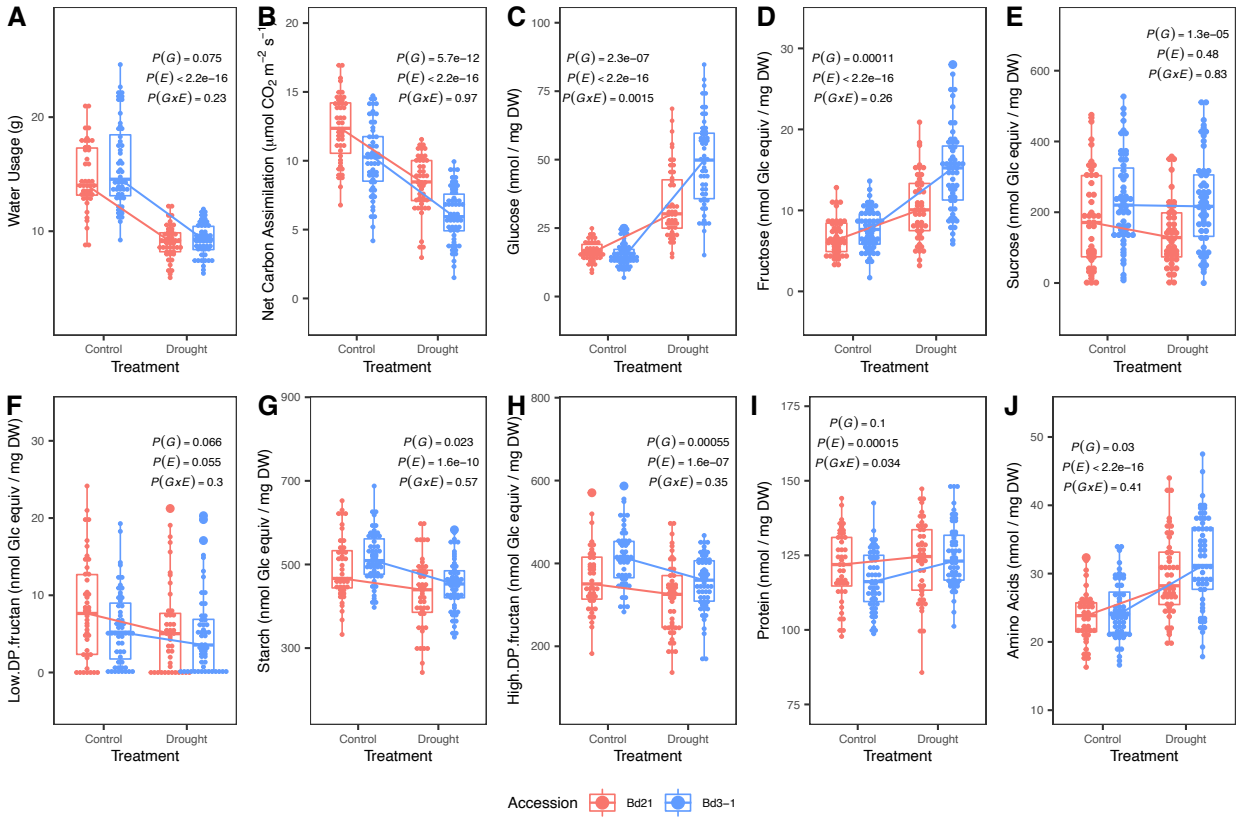


Figure 2. Leaf blade traits of Bd21 and Bd3-1 on Day 6. The p-values of effect terms were determined using (generalized) linear (mixed) models, by removing each term and performing likelihood ratio test (anova and Chisq test, with maximum likelihood mode if models have random factors). DW = Dry Weight of tissue. Glc Equiv = glucose equivalents. N = 48-58 for each combination of accession and treatment.

We next performed a more highly replicated experiment (n=48-58 for each accession in each treatment) focused on contrasts between the two accessions at soil water contents of 85% and 55%. These plants constitute the sampling for all following analyses. Consistent with the trend observed before, tissue water content decreases under drought stress in both genotypes, as does the leaf photosynthesis rate (Figure 2).

Leaf metabolite content was affected by soil drying. Leaf starch content decreased on Day 6 (Figure 2G; Chisq test and p<1.6e-10). Sucrose shows no treatment effects in the leaf (Figure 2E; p=0.48). The degradation of starch and the catabolism of sucrose might contribute to the accumulation of glucose and fructose, both of which are strongly elevated in our Day 6 samples (Figure 2C, D; both p<2.2e-16). However, leaf glucose increases to a larger degree in Bd3-1 than in Bd21, shown as a significant GxE response (Figure 2C, p=0.0015). Another important group of

199 carbohydrates in cool-season Pooid grasses like *Brachypodium* are fructans, which may
200 serve as a short-term storage carbohydrate when sucrose exceeds demand [56]. Here,
201 we assayed high-degree polymerized fructan, which showed detectable decreases in
202 leaves on Day 6 (Figure 2H; $p<1.6e-7$), similar to what we observed for starch. Low-
203 degree polymerized fructan also decrease slightly (Figure 2F; $p= 0.055$), similar to
204 sucrose. In general, most carbohydrates show a genotypic effect with higher values in
205 Bd3-1. Amino acids per unit leaf mass was higher in droughted plants on Day 6 (Figure
206 2J; $p<2.2e-16$). Total protein content increased in Bd3-1 and marginally increased in
207 Bd21 (Figure 2I; $p=0.00015$), showing a GxE response ($p=0.034$).
208
209 *Factors affecting the transcript abundance of individual genes*

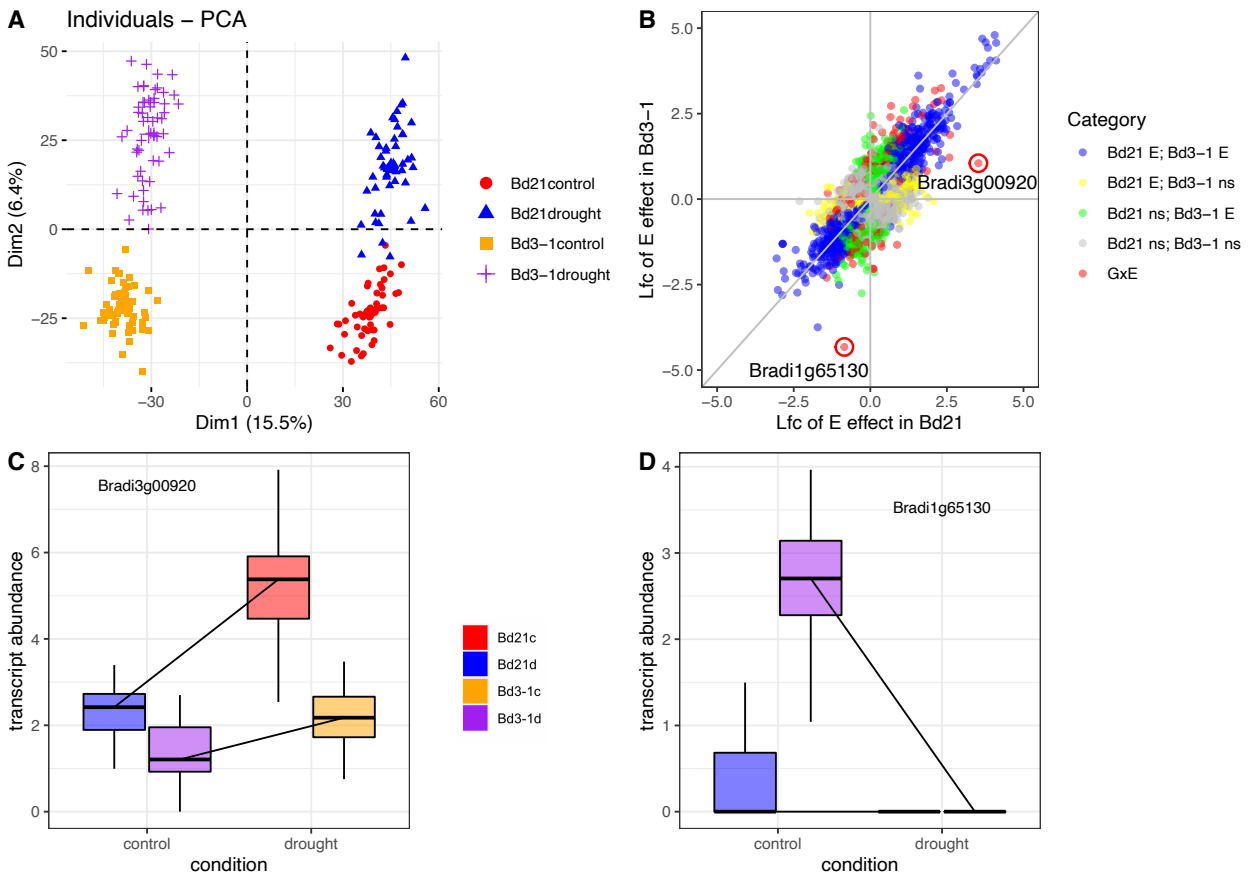


Figure 3. Gene expression diversity. **A)** PCA shows clustering of RNASequencing library by accession and environmental treatment. **B)** Factors affecting the transcript abundance of individual genes. Each point represents a gene model in the *B. distachyon* Bd21 genome. Lfc = Log-fold change of transcript abundance in response to soil drying. Colors indicate the significance of treatment effect on Bd21 and Bd3-1 samples; genes with a significant GxE effect are highlighted in red and two genes discussed in the text are circled. Boxplots show two genes with large GxE effect: **C)** Bradi3g00920 and **D)** Bradi1g65130, gene expression reaction norm based on log2(norm_count+1). (Red Bd21d, blue Bd21c, orange Bd3-1d, purple Bd3-1c)

221

222 The transcripts of 29,740 gene models were identified in our RNASequencing data,
223 representing 91.4% of those in the Bd21 v3.0 reference. Principle component analysis
224 shows clear grouping of samples by library type (Figure 3A), in which the first PC
225 separates the two genotypes and explains 15.5% of total sample variance and the
226 second PC separates the two treatments and explains 6.4% of total variance.

227 26,388 genes passed our filtering criteria and are included in the following
228 analysis of by-gene differential expression. Complete DGE gene list is available in
229 Supplementary Data Set 1. At an FDR=0.05, 4725 genes are constitutively higher, and
230 4087 genes are constitutively lower in Bd21 as compared to Bd3-1 (i.e., show a
231 significant genotypic effect). 3111 genes are constitutively higher, and 2832 genes are
232 lower in drought than control (environmental effect); among these, 3066 genes exhibit
233 both significant genotype and environment effects (hereafter "G+E genes"). In Bd21,
234 4024 and 3983 genes are up- or down-regulated, respectively, under drought. In Bd3-1,
235 3997 and 4010 genes are up- or down-regulated, respectively, under drought (Figure
236 3B). 538 genes exhibit significant GxE interaction at FDR = 0.1 (Figure 3B; hereafter
237 "GxE genes"). Among these GxE genes, 107 are up-regulated and 66 are down-
238 regulated in both accessions but with differing magnitudes of response in the two
239 accessions; most such genes exhibit more dramatic changes in Bd3-1 (Figure 3B). 135
240 genes are uniquely up-regulated and 99 uniquely down-regulated in response to drying
241 in Bd3-1 (so-called "conditional neutrality;"[7]); similarly, 39 genes are uniquely up -
242 regulated and 60 uniquely down-regulated in Bd21. Only 21 genes are expressed in
243 different directions in response to the drought treatment in the two accessions.

244 Of the two genes showing most striking GxE effect, Bradi3g00920 showed a 3.52
245 log-fold increase under drought stress in Bd21 (Wald test corrected by BH, p_{adj} = 5.55e-
246 68) but increases only 1.05 log-fold increase in Bd3-1 (p_{adj}=3.08e-5). This leads to an
247 additional 2.48 lfc in addition to genotypic and environmental effect, named GxE lfc
248 (p_{adj}=1.20e-12, Figure 3C). The gene is annotated to be Leucine rich repeat family
249 protein. Similarly, Bradi1g65130 showed a significant 4.32 log-fold decrease in Bd3-1 in
250 response to drying (p_{adj}=2.90e-8) but only 0.84 log-fold decrease in Bd21 (p_{adj}=0.13,

251 Figure 3D), with a 3.48 GxE lfc. The gene is predicted to be a cyclase/dehydrase family
252 protein.
253 There are 40 annotated transcription factors among the 538 GxE genes,
254 including a predicted MYB-related transcription factor with the largest GxE Lfc of -1.5
255 among them. GO terms enriched among GxE genes include processes related to
256 chlorophyll, starch and organic substances, amino acids, translation, and protein folding
257 processes, and response to stimulus (Supplementary Data Set 2).
258
259 *Variation in gene-gene interaction within co-expression modules*
260 The high replication used in our RNA sequencing experiment (N = 48-58, subsampled
261 to 48 samples for each library type) affords us the power to estimate co-expression
262 relationships independently in each of the four library types (Bd21 control watering,
263 Bd21 drought, Bd3-1 control, Bd3-1 drought) using WGCNA. Genes in a module might
264 represent a group of genes commonly activated by a set of regulatory factors and
265 sharing molecular functions [43]. Depending on the parameter settings implemented in
266 WGCNA we identified 14-24 modules in Bd21 control, 14-18 modules in Bd21 drought,
267 14-16 modules in Bd3-1 control, and 15-20 modules in Bd3-1 drought (Figure 4).
268 We first identify modules that overlap among the four library types using a high
269 dimensional Fisher's exact test [57], testing the hypothesis that a set of modules identified
270 within or between library types share significantly more genes than expected. These
271 cliques of overlapping module are interrogated for four general patterns. In the first
272 pattern, we reason that a clique comprised of one module from each of the four library
273 types may represent gene co-expression relationships that are conserved across
274 genotype and treatment (Figure 4, teal polygons). We find 40 such modules comprising
275 20 overlapping cliques, among which 36 modules show significant GO enrichments.
276 Clique 1 is comprised of four modules that have common GO terms of circadian rhythm.
277 Clique 8 is comprised of a chain of seven partially overlapping cliques enriched for
278 similar core cellular processes including protein translation, photosynthesis, electron
279 transport chain, photosystem, and glycogen metabolic process (Table 1). Some
280 modules in conserved cliques are highly correlated with physiological traits. In
281 particular, Bd3-1d21 in Clique 8-6 is correlated with daily water usage (p-value=4.4e-5)

282 while Bd3-1d4 in Clique 8-5 is correlated with both daily water usage (p-value=1.5e-4),
283 and tissue water content (p-value=1.9e-4). Both modules are enriched with GO terms
284 associated with photosynthesis, and Bd3-1d12 in Clique 8-8 is correlated with leaf
285 glucose content (p-value=6.4e-5) and partially overlaps with a Clique 8-7 enriched for
286 glycogen metabolic process.

287 The second pattern is cliques comprised of two modules from the same genotype
288 but with different treatments (Figure 4, yellow polygons). We find three such cliques,
289 which may represent modules that are only coherent in one genotype. For example,
290 Clique 8-9 solely comprises modules in Bd21 and is enriched with functions such as
291 response to wounding and phenylpropanoid biosynthetic processes (including *B.*
292 *distachyon* orthologs of C4H, PAL, 4CL and F3H, F3'H and an R2R3 MYB transcription
293 factor which regulates phenylpropanoid biosynthesis in other species, see
294 Supplementary Data Set 3-4). Clique 8-12 is comprised solely of modules in Bd3-1 and
295 is enriched in chlorophyll and pigment metabolic process and terpenoid biosynthetic
296 processes; it contains five NAC regulators/domain related genes, as well as a
297 chlorophyll A-B binding protein and a putative senescence-inducible chloroplast stay-
298 green protein.

299 The third type of clique is treatment specific, composed of two overlapping
300 modules from different genotypes but the same environmental condition (Figure 4,
301 green polygons). We identified one such case, Clique 3. Both modules in this clique are
302 correlated with leaf glucose content (p-value=9.4e-8 and 1.6e-6 for Bd21d and Bd3-1d,
303 respectively). Although the two modules have no GO terms in common, Bd21d13 is
304 enriched with a GO term of response to water and contains two RAB18 orthologs
305 (Responding to ABA), an ortholog of rice ASR5 (abscisic acid-, stress- and ripening-
306 induced gene), and one IAR4. Many genes in Bd21d13 -- including the two RAB18
307 genes -- have very low expression in the control condition.

308 Finally, a clique may be comprised of modules either from a single library type or
309 from three library types and excluding the fourth library type; we interpret this pattern to
310 represent modules exhibiting GxE (Figure 4, pink or orange polygons). There are three
311 cliques which are unique to a single module in a single library type. Among these,
312 Clique 2 is comprised solely by a module identified in Bd21d and, although not enriched

313 for any GO terms, the module is highly correlated with glucose content (p-value=1.1e-
314 5). The annotations of several genes in this clique relate to abscisic acid (ABA)
315 signaling and regulation. Moreover, the proximal promoters of genes in this module are
316 enriched with a predicted bZIP motif (p-value = 2.6e-7). Clique 5 comprises only one
317 module from Bd21c and is enriched with GO term of phosphatidylethanolamine
318 biosynthesis process. Clique 7 is identified in Bd21c, enriched with GO terms of
319 glutathione metabolic process and response to chemical. We also identified three
320 cliques comprised of modules from three library types with the fourth condition
321 unrepresented. Among these, Clique 8-11 lacks a Bd3-1d module while the three
322 included modules have GO terms in common related to cell wall biogenesis. Many of
323 the genes in these modules are not included in any reconstructed module in Bd3-1d.

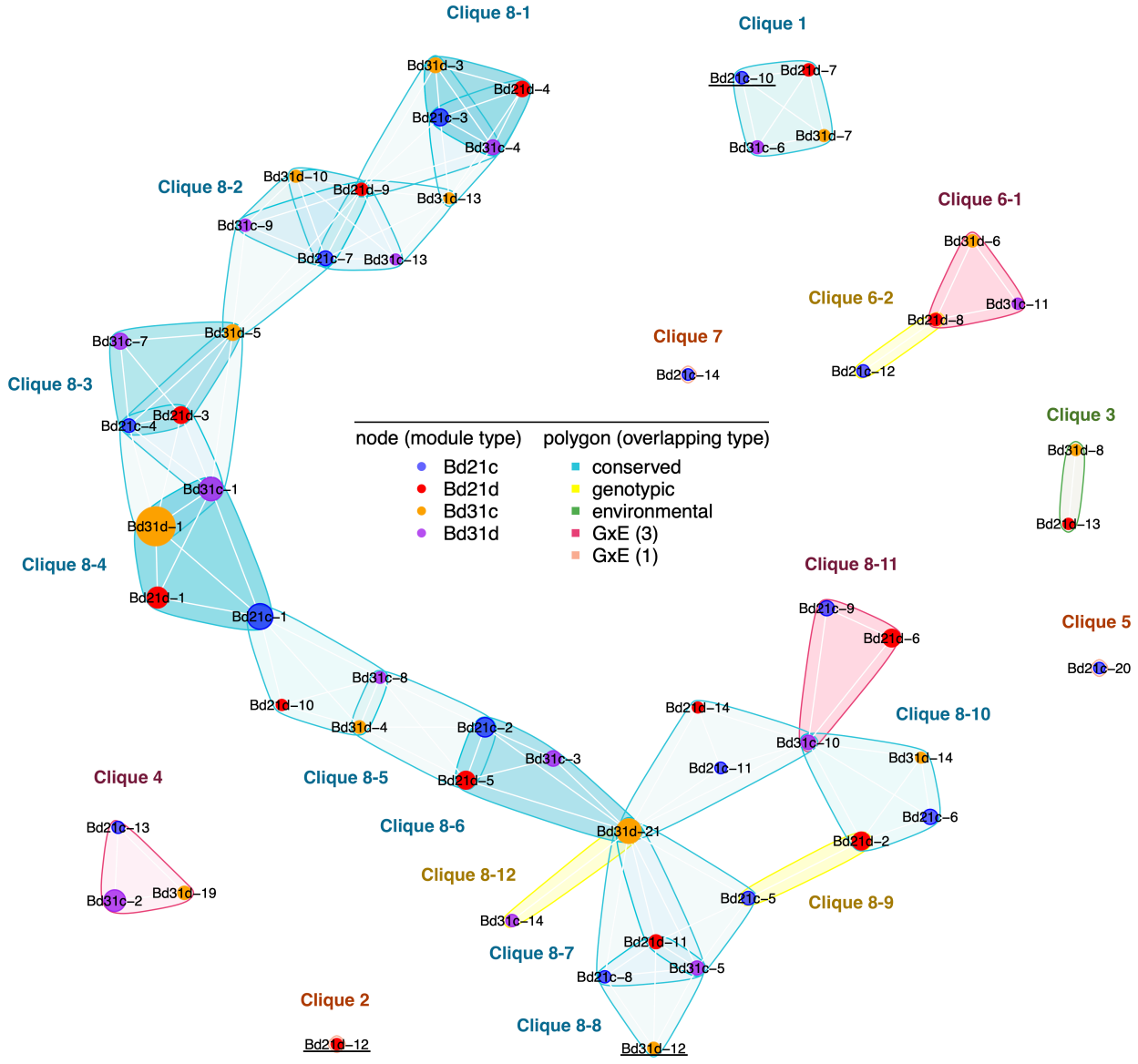


Figure 4. Conservation of gene co-expression modules among genotypes and treatments. Nodes represent modules identified in each of the four library types, with the size of the node corresponding to the number of genes in each module (red, Bd21d; blue, Bd21c; orange, Bd3-1d; purple, Bd3-1c). Clusters of modules enclosed by a polygon -- cliques -- show significant gene overlap using the Fisher's Exact Test(Data are available in Supplementary Data Set 5); polygon color denotes the pattern of module conservation among libraries (teal, four-module stable overlap; pink, three-module overlap; green, two-module of same environment overlap; yellow, two-module of same genotype overlap; and orange, module uniquely identified in a single library type) with the transparency of the color proportional to the number of genes in common. Underlined module names are shown as inferred network in Figure 5.

338 Table 1. Common GO terms at lowest level among modules in overlapping groups
339 (based on all module versions FDR=0.1) and correlated traits (based on moderate size
340 module version, in specific library type, only p<0.001 are reported, not corrected by
341 multiple test). DGE enrichment in modules are shown (based on constraint module size
342 version, only p<0.001 are reported, not corrected by multiple test, and the percentage is
343 shown if > 75%). Cliques with no common GO terms or correlated traits identified are
344 not shown. Complete lists are available in Supplementary Data Set 6-8.

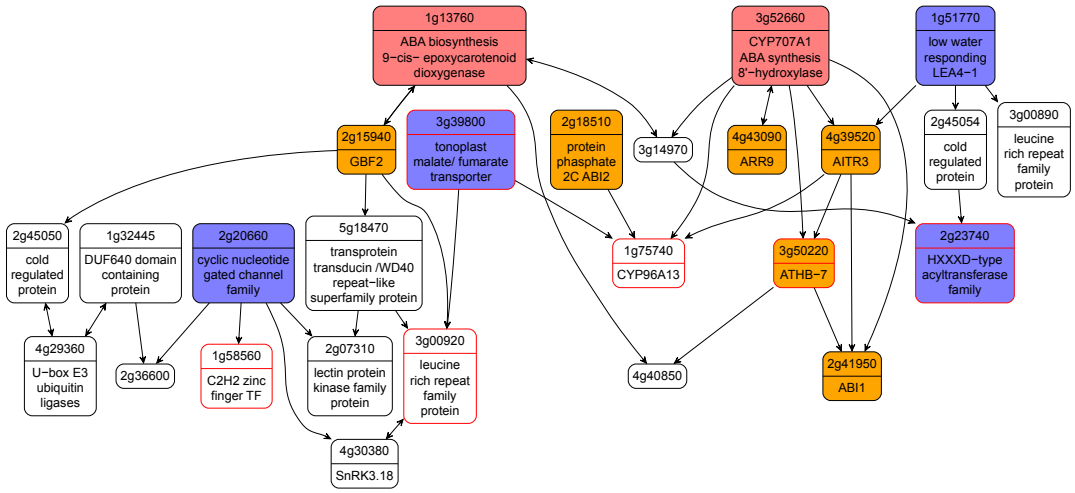
Name (# of included modules)	Common GO	Correlated Trait (p<1e-03)	DGE Enrichment (p<1e-03; percentage when >75%)
Clique 1 (4)	circadian rhythm		Bd3-1c6: E
Clique 2 (1)		glucose (1.1e-05)	E (97%),GxE
Clique 3 (2)		Bd21d13: Glucose (9.4e-08) Bd3-1d8: Glucose (1.6e-06)	Bd21d13: E (100%); Bd3-1d8: E (92%),GxE
Clique 4 (3)		Bd3-1d19: leaf water content (8.5e-04)	Bd21c13: G (76%); Bd3-1c2: E,G,GxE; Bd3-1d19: G (97%)
Clique 5 (1)	phosphatidylethanolamine biosynthetic process		
Clique 7 (1)	glutathione metabolic process; alternative respiration; response to chemical		GxE
Clique 8-1 (4)	ribosomal small subunit assembly; translation; photosynthetic electron transport in photosystem II; photorespiration; ATP synthesis coupled proton transport; reductive pentose-phosphate cycle; electron transport coupled proton transport, respiratory electron transport chain		
Clique 8-2 (4)	electron transport chain, cellular nitrogen compound biosynthetic process, cellular respiration		
Clique 8-3 (4)	translation		
Clique 8-4 (4)	cytoplasmic translation; regulation of actin polymerization or depolymerization; ribosomal large subunit biogenesis, protein maturation		
Clique 8-5 (4)	Photorespiration; reductive pentose-phosphate cycle	Bd3-1d4: Leaf water content (1.5e-04) daily water usage (1.9e-04)	Bd21c2: E,G; Bd21d5: E (82%);
Clique 8-6 (4)	response to light stimulus; photosynthesis, light harvesting in photosystem I; photosystem II stabilization; pigment biosynthetic process	Bd3-1d21: daily water usage (4.4e-05)	Bd21c2: E,G; Bd21d5: E (82%); Bd3-1c3: E; Bd3-1d21: E,G,GxE
Clique 8-7 (4)	glycogen metabolic process		Bd21c8: E,G; Bd21d11: E,G,GxE; Bd3-1c5: E,G,GxE; Bd3-1d21: E,G,GxE
Clique 8-8 (4)		Bd3-1d12: Glucose (6.4e-05)	Bd21c8: E,G; Bd21d11: E,G,GxE; Bd3-1c5: E,G,GxE; Bd3-1d12: E,GxE
Clique 8-9 (2)	response to wounding; phenylpropanoid biosynthetic process; regulation of defense response; monocarboxylic acid metabolic process		Bd21c5: G,GxE; Bd21d2: E,G,GxE
Clique 8-11 (3)	plant-type primary cell wall biogenesis; plant-type secondary cell wall biogenesis; cellulose biosynthetic process		Bd21c9: E,G (100%); Bd21d6: E,G (95%); Bd3-1c10: E,G (82%)
Clique 8-12 (2)	terpenoid biosynthetic process; chlorophyll metabolic process; pigment metabolic process	Bd3-1d21: daily water usage (4.4e-05)	Bd3-1c14: E,G (76%) Bd3-1d21: E,G,GxE

345

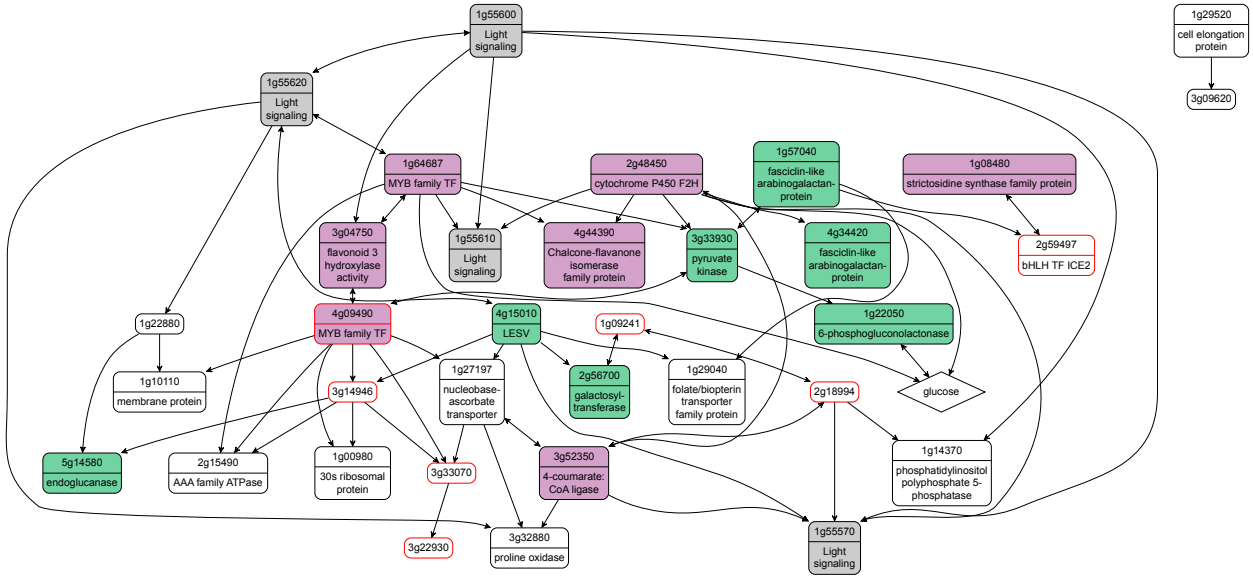
346 *Inferring regulatory interactions within modules*

347 The preceding analyses demonstrate that the gene composition of some co-expression
348 modules vary as a function of genotype, treatment, and GxE, while other modules are
349 conserved across genotypes and treatments. We next assess a possible cause of
350 module-level variation: variation in regulatory interactions among genes in modules. We
351 use graphical models to infer the conditional independence of genes to estimate causal
352 relationships among them, here working with the simplifying assumption that the
353 regulator(s) of a focal gene is in the same module as that gene. We focus on three
354 exemplar modules. Since the regulation is based on estimates of transcript abundance,
355 gene "interaction" might represent either direct transcription factor-target regulation or
356 indirect regulation such as arising from through system-level phenomena such as
357 metabolism or cell cycle similar as described [58].

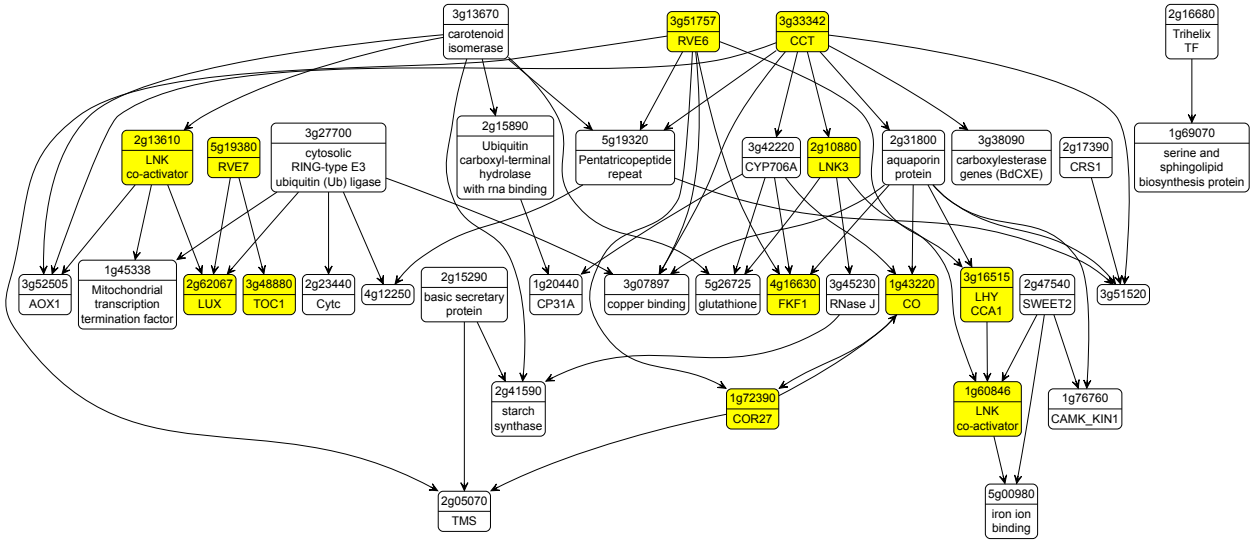
358 **A**



359
360 **B**



361
362 **C**



363

364 Figure 5. Inferred directed interactions within modules. **A)** Module 12 in Bd21d, **B)**
365 Module 14 in Bd3-1d and **C)** Module 10 in Bd21d inferred using the UTI-GSP method.
366 Nodes are genes and edges are inferred causal interactions. Edges with two arrows
367 indicate either the direction of interaction is uncertain or that the interaction is inferred to
368 be in both directions. Second row of each node shows annotation of orthologs in
369 Arabidopsis or rice. Glucose is included as a node in in module Bd3-1d12 as an
370 example of a module with high correlation between leaf glucose content and the module
371 eigengene (Table 1). Genes exhibiting a significant pattern of GxE in our DGE analyses
372 are highlighted with red outline.

373
374 Bd21d12 is the sole module in Clique 2, and many of its genes relate to ABA
375 biosynthesis and signaling. Two of these are predicted to act upstream many other
376 genes in the module (Figure 5A red fill): Bradi1g13760, an ortholog of AtNCED9, a 9-
377 cis-epoxycarotenoid dioxygenase, and Bradi3g52660, orthologous to AtCYP707A1, with
378 inferred ABA 8'-hydroxylase activity. The module also contains genes whose orthologs
379 are ABA-induced or ABA-dependent regulators (Figure 5A orange fill), including
380 Bradi4g39520 (orthologous to a transcription repressor, AtAITR3), Bradi3g50220
381 (orthologous to ABA-responsive AtATHB-7), Bradi2g15940 (orthologous to AtGBF2, an
382 ABA regulated bZIP transcription factor), and Bradi2g41950 and Bradi2g18510
383 (orthologous to AtABI1 and AtABI2, respectively). Other genes are associated with
384 drought response in Arabidopsis (Figure 5A blue fill), including stomatal movement
385 (Bradi2g20660, Bradi3g39800), wax biosynthesis (Bradi2g23740), and a late embryo
386 abundant (LEA) protein (Bradi1g51770). The gene with strongest GxE effect in our
387 dataset, Bradi3g00920 (Figure 3C), has the highest eigenvalue in this module.
388 Module 12 in Bd3-1d is correlated with leaf glucose content and is a member of
389 Clique 8-8 which share three modules with Clique 8-7 enriched for glycogen metabolic
390 processes (Figure 5B). Bradi1g55600 is inferred to be upstream of many other genes in
391 this module (Figure 5B grey fill); its Arabidopsis ortholog is an early light-induced protein
392 (ELIP2) and regulates chlorophyll synthesis [59]. Paralogs of Bradi1g55600 are also
393 present in this module. Downstream in the inferred network, there are genes whose
394 Arabidopsis orthologs encode enzymes in sugar and related metabolic pathways
395 (Figure 5B green fill), including Bradi3g33930 (pyruvate kinase), Bradi1g22050 (6-
396 phosphogluconolactonase in OPP pathway), Bradi5g14580 (endoglucanase in cell wall
397 degradation), Bradi2g56700 (galactosyltransferase in proteoglycan synthesis), and
398 Bradi1g18650 (4-alpha-glucanotransferase in starch degradation). Above, we reported
399 that leaf glucose content is correlated with this module; when we include glucose as a
400 node in the module, we find that it is predicted to directly interact with Bradi1g22050, the
401 6-phosphogluconolactonase. Bd3-1d12 also contains Bradi4g15010, whose Arabidopsis
402 ortholog is predicted to have starch synthesizing and degrading activity (LESV).
403 Additional genes in this module have predicted functions in secondary metabolism

404 (Figure 5B purple fill), including orthologs of an Arabidopsis MYB family transcription
405 factor involved in flavonol biosynthesis (Bradi1g64687), flavonoid 3'-hydroxylase
406 (Bradi3g04750), flavanone 2-hydroxylase (Bradi2g48450) and a 4-coumarate: CoA
407 ligase in the phenylpropanoid pathway (Bradi3g52350). Notably, several GxE genes
408 with unknown functions are connected in this module.
409 Finally, we report on module Bd21d 10 in Clique 1, that is conserved in all four
410 library types and is enriched in genes relate to circadian processes (Figure 5C, yellow
411 fill). There are five predicted MYB TFs in this module, including Bradi3g51757
412 (orthologous to AtRVE6 and OsRVE4), Bradi2g62067 (orthologous to OsLUX),
413 Bradi5g19380 (orthologous to OsMYB511, AtRVE7), and Bradi3g16515 (orthologous to
414 OsLHY/CCA1), Bradi1g29680 (orthologous to OsLHY in rice and AtRVE2/ ERP1 in
415 Arabidopsis). Additional genes in this module are orthologous to CONSTANS (CO) and
416 related proteins in Arabidopsis. Studies suggest genes in circadian-controlled pathways
417 show conserved oscillation over 24 hour intervals [60], suggesting that stochastic variance
418 among individuals may have provided the signal for reconstructing these conserved
419 interactions in our data.
420
421 *Genes predicted to interact show similar or different patterns of differential gene*
422 *expression*
423 Genes in modules are enriched with certain types of DGEs as shown in Table 1.
424 However, within a module, gene expression variation between two genes connected by
425 an edge in our graphs (Figure 5) can be dependent on: genotype (G change, e.g.,
426 regulator and target gene show no DGE and G DGE respectively, see Supplementary
427 File 1 for testing method and illustrative examples), environment (E change, e.g.,
428 regulator and target gene show no DGE and E DGE respectively), genotype by
429 environment (GxE change, e.g., regulator and target gene show no DGE and GxE DGE
430 respectively) or no effect (e.g., regulator gene and target gene both show no DGE). We
431 summarize the different types in the inferred gene regulation network in modules in
432 Supplementary Table 1 and Supplementary Figure 2 (based on modules in Figure 5).
433 For example, in module Bd21d12 that only exists in Bd21 under drought condition, of 41
434 total edges, there are two edges showing no effect of expression changes, five edges

435 showing G changes, 11 edges showing E changes, 12 edges showing G+E changes,
436 11 edges showing GxE changes. While in Bd21 specific module Bd21c5, of 62 total
437 edges, there are 23 edges showing no DGE changes, 25 edges showing G changes,
438 nine edges showing E changes, eight edges showing G+E changes, and no edges
439 showing GxE changes.
440
441 *Genetic and Environmental Effects on Interactions among Genes*
442 An edge between nodes in our inferred gene regulation network represents a
443 hypothesis about the regulatory interaction between two genes. The high biological
444 replication of our RNASequencing across four library types enables us to fit linear
445 models to explore the effects of genotype, environment, and their interaction on the
446 expression relationship between pairs of genes. There are four patterns of interest in
447 our dataset.
448 First, the regulatory interaction between two genes may be present under one or
449 some library types (combinations of genotype and treatment) but absent in others; this
450 pattern would be diagnosed by identifying significant slope change such that the
451 correlation between these two genes is no longer significant. One possible explanation
452 for this pattern is environmentally dependent regulatory interaction (Figure 6B). In ABA
453 module Bd21d12, described above, the regression of Bradi4g40850 (unknown) on
454 upstream genes shows a significant E effect of slope from gene Bradi3g50220 (ATHB-7
455 which is regulated in an ABA-dependent manner). Similarly, in drought specific module
456 Bd21d13, the regression of Bradi1g37410 (which encodes a dehydrin family protein) on
457 upstream genes suggests that Bradi1g44480 (dehydrogenase E1 component domain
458 containing protein) regulates Bradi1g37410 only in drought (Supplementary Figure 3A).
459 Across modules, we see clustering of such context-dependent edges as reflected in
460 module Bd21d13 in clique 3 – reconstructed only under drought conditions – containing
461 32 edges of which 14 are environmentally dependent (44%) (Supplementary Table 1).
462 A regulatory interaction between two genes also may be observed in only one of
463 the two genotypes (Figure 6D). For example, in Bd21 specific module Bd21c5, the
464 regression of Bradi1g15840 (putative allene oxide cyclase in jasmonic acid synthesis)
465 on upstream genes suggest Bradi3g44990 (putative cell-wall invertase) regulates

466 Bradi1g15840 only in Bd21 (Figure 6D). Consistent with the observation that in Clique
467 8-9 specific to Bd21, we find that Bd21c5 has 12 edges showing a significant effect of
468 genotype among its among 65 edges (18%) (Supplementary Table 1).
469 We might also observe differences among the slopes of the library types such
470 that a regression is only significant in a single library type, indicative of GxE for a given
471 pairwise regulatory interaction (Figure 6A). In module Bd21d12 discussed above, the
472 regression of Bradi3g00920 (Leucine Rich Repeat family protein) on upstream genes
473 suggests that Bradi2g15940 (GBF2) interacts with Bradi3g00920 with a significant GxE
474 effect of the slope (Figure 6A). Module Bd21d12 is found only in Bd21 under drought
475 stress and has many genes associated with ABA signaling.
476 In the second general pattern, the interaction between two genes is observed to
477 change in magnitude between library types; this might be caused if the binding
478 interaction between a transcription factor and a regulatory motif is stronger in one
479 genotype than a second genotype (Figure 6E). Such an interaction would appear as a
480 change in the slope of the regression of target gene on its putative regulator but without
481 a complete loss of the correlation. As one example, Bradi4g20520 (unknown function) is
482 predicted to regulate Bradi3g19730 (jacalin-like lectin domain containing protein) in
483 Bd21c6 in all library types but with a greater slope in the Bd21 libraries than in Bd3-1
484 libraries.
485 Third, the baseline expression of genes might be different across library types
486 (Figure 6G), possibly due to the presence of cofactors or the general resource and
487 energy availability of the cell; this type of regulation may be detected as a difference in
488 intercept among library types in our multiple regression analysis. In module Bd21c10,
489 Bradi3g45230 (RNA-metabolising metallo-beta-lactamase) regression on genes
490 including Bradi2g10880 (LNK3) shows the regulation is similar in all library types, but
491 the baseline in Bd3-1 is constitutively higher in both treatments. This pattern is observed
492 as a higher intercept in Bd3-1 libraries which leads to genetically dependent expression
493 of the downstream gene. Such intercept differences as a function of environment,
494 genotype, or GxE are widely found among the pairwise interactions identified in our
495 data.

496 The final pattern is no difference in the pairwise regulation between genes
497 among library types but an apparent transmission of a regulatory response between the
498 genes. This pattern is manifested as neither a significant slope change nor an intercept
499 effect in the regression, and a common differential expression between putative
500 regulator and its target gene. One such case is Bradi5g18277 (serine-type
501 endopeptidase) and Bradi1g53680 (zinc transporter 3 precursor) in module Bd31c16
502 (Figure 6F). In module Bd31d12, GxE genes are found more likely to share a regulatory
503 interaction with other genes exhibiting GxE than non GxE genes in the module (t-test, p-
504 value = 0.013; see Supplementary Figure 3B for an example).
505 Finally, we use our approach to study the transcriptional control of leaf glucose
506 content, which shows a significant GxE effect in our assays (Figure 1H) and is highly
507 correlated with four module eigengenes (Bd21d12, Bd21d13, Bd3-1d8, and Bd3-1d12;
508 Table 1). Nine genes are inferred to regulate glucose when it is included as a node in
509 these modules. Among these, we find that Bradi4g39520 only exhibits an interaction
510 with glucose under drought stress (Figure 6C); the glucose content is same between
511 two genotypes and not dependent on the expression level of Bradi4g39520 under
512 control conditions. Bradi4g39520 is an ortholog of AITR3 in Arabidopsis, ABA-induced
513 transcription repressor that acts as feedback regulator in ABA signaling.
514

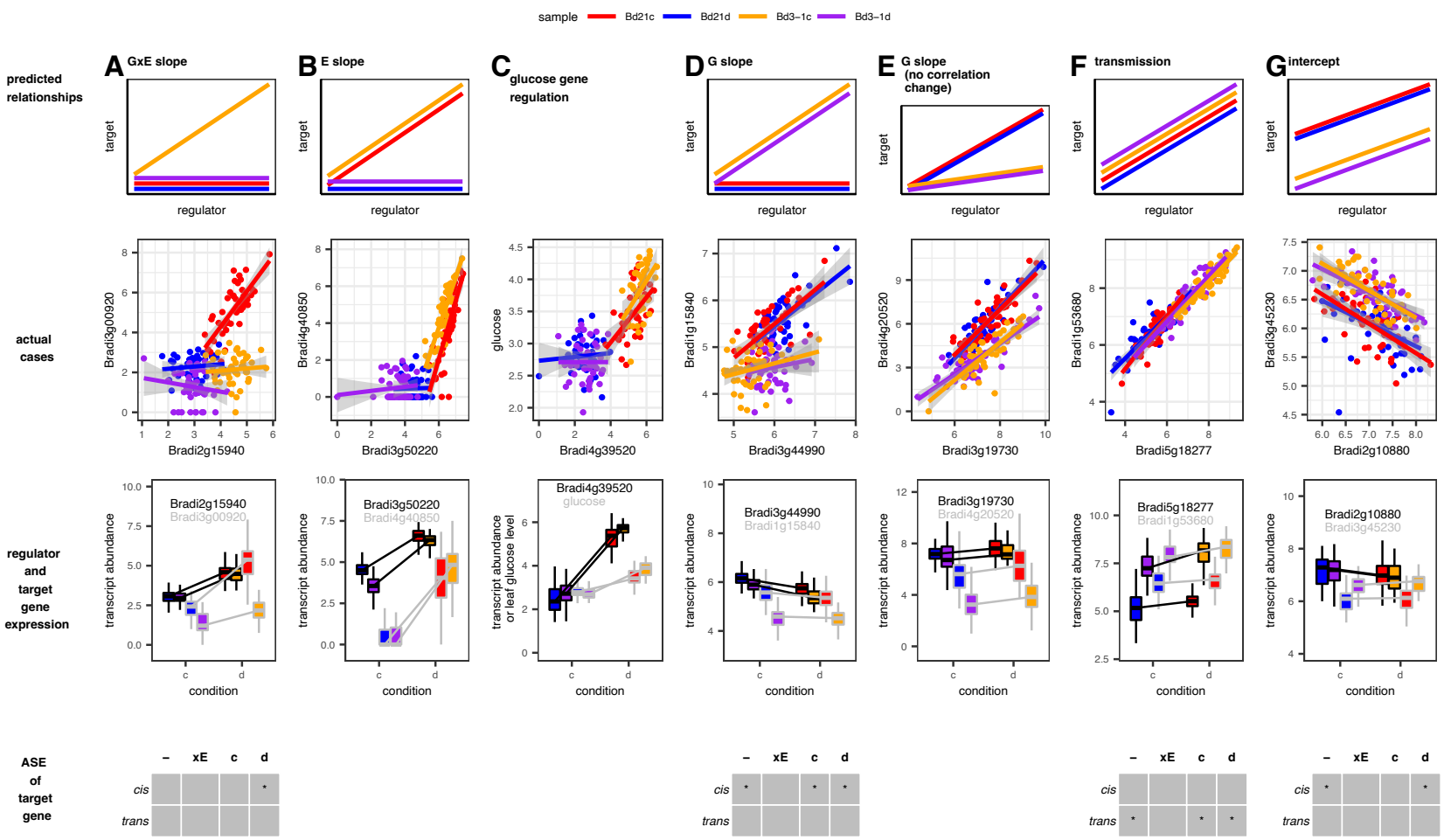


Figure 6. Predicted and actual examples of regulations between pairs of genes selected. **A-B, D-G)** show regulation slope changes, transmission, and regulation intercept changes. These are illustrated with linear regression between regulator and target, the expression level of each gene, and allele specific expression (ASE) types. ASE types include constitutive regulation (-) and environment specific (xE) detected in full model (see method), or regulation detected in control (c) and drought (d) condition. P-value <0.1 and not corrected by multiple tests. **C)** regression of log-transformed leaf glucose content on upstream identified gene Bradi4g39520 as a putative regulator of leaf glucose content. Transcript abundances are log2 (norm_count+1).

523 *Identifying allele-specific expression in target genes*
524 We next used allele specific expression assays to infer whether changes in regulatory
525 interactions attributed to genetic or GxE effects in transcript abundance arise from
526 factors acting in *cis* or in *trans*. In the above examples of the genotype-dependent slope
527 between the regulation from Bradi3g44990 to Bradi1g15840 (Figure 6D), Bradi1g15840
528 is found to be *cis*-regulated (Supplementary Figure 4B). In the example of GxE
529 regulation of Bradi3g00920 by Bradi2g15940 (Figure 6A), we identified Bradi3g00920 is
530 regulated by a *cis*- acting factor under drought stress (Supplementary Figure 4A).
531 Moreover, in the example of an apparent transmission from Bradi5g18277 to
532 Bradi1g53680 (Figure 6F), we observed that transcript variation of Bradi1g53680 is
533 controlled by a *trans*- acting variant (Supplementary Figure 4C) as expected. As for the
534 example of Bradi3g45230 being regulated by Bradi2g10880 with an effect of genotype
535 on the intercept of the regression (Figure 6G), Bradi3g45230 is found to be *cis*-
536 regulated (Supplementary Figure 4D). It might be an example of variation in an
537 enhancer or similar cis-regulatory element leading to base level expression difference.
538
539 **DISCUSSION**
540 Gene expression has long been a focus of both evolutionary and functional genetic
541 studies. Here, we aimed to integrate these two perspectives by studying the
542 environmental and genetic control of regulatory interactions between genes using highly
543 replicated RNASequencing libraries of *Brachypodium distachyon* accessions exposed
544 to moderate soil drying and control conditions.
545 Here, we extend past work showing that the expression of individual genes can
546 vary as a function of genotype and environmental treatment to demonstrate variation in
547 interactions between genes. One example highlighted throughout the analyses in this
548 study is Bradi3g00920, which shows a pattern of GxE in its gene expression. By
549 developing a directed graphical network for the module in which the gene is located, we
550 infer that Bradi3g00920 is relatively downstream of several other genes putatively
551 regulating it. Multiple regression of its gene expression suggests that Bradi3g00920 is
552 regulated by Bradi2g19540 only under drought in Bd21. We infer that this pattern
553 suggests a possible cause of the observed GxE in gene expression for Bradi3g00920.

554 Further, allele-specific expression reveals that this gene is associated with a *cis*-
555 regulated expression pattern.
556 We next consider how a differentially expressed gene may transmit its
557 expression pattern through its interactions with other genes. Here, we refer to the
558 correlated pattern of differential expression as “transmission.” As shown in the example
559 in Figure 6F, an inferred target gene (Bradi1g53680) shows the same dependency on a
560 putative regulator (Bradi5g18277) across all library types. Indeed, the target gene is
561 found to be *trans*- regulated, suggesting that the genetic variant(s) causing the
562 differential expression is in the direct regulator or upstream of it. It suggests the
563 interpretation of *trans*- acting loci in QTL studies might include what we here call gene
564 expression transmission and not a protein coding change in the trans-acting regulator.
565 The transmission of the expression is also a direct example showing certain genetic
566 variants can affect more than one genes, and that the expression of individual genes
567 should not be considered independent variables.
568 Variation in gene-gene interactions observed in modules could lead to a module
569 not being reconstructed in certain combinations of genotype and environment. For
570 example, 44% of the inferred edges in module Bd21d13 exhibit environmentally
571 dependent regulation and, perhaps as a result, this module is found only under drought
572 stress. In contrast, in modules that have fewer pairwise gene-gene regulatory variations,
573 or when gene expression states are transmitted faithfully from upstream nodes in the
574 network, these modules may be more likely observed under all sampled conditions.
575 Genetic variants affecting expression, whether constitutive differences between
576 varieties (in our study, those with a significant G effect) or GxE, may be acted on by
577 selection if they ultimately have an effect on phenotype. Past work addressed the
578 possibility of spatial separation across a network between a genetic variant and its
579 effect[35,61], or between the perception of an environmental stimulus and the downstream
580 transcriptional response[62]. The network context of a genetic variant is important for
581 evolutionary studies because the number and functions of genes which interact with a
582 variable gene influences the pleiotropy of that genetic variant[63]. Recent work considered
583 the topology and/or pleiotropy of gene expression in the context of the evolution of
584 plasticity[48,64]; however, most studies do not account for the prospect of pleiotropic

585 effects of these genetic variants and, as a result, variation in the network topology itself.
586 We specifically consider such regulatory interaction variation here. The large
587 RNASequencing dataset reported here also allowed us to estimate gene co-expression
588 independently in each combination of genotype and environmental treatment, thus
589 relaxing the assumption of an invariant network topology across conditions used
590 previously[48,64,65]. We note that other approaches to account for network structure, such
591 as mapping covariance or principle components as traits in QTL analysis [3,11,66] may
592 detect similar regulatory structure.
593 Changes in gene-gene interactions are found to be clustered in some modules,
594 suggesting a spatial pattern of regulatory variation in networks and supporting the
595 hypothesis that modularity allows for reduced pleiotropic effects of mutations[63]. A
596 related hypothesis is that certain regulatory pathways are under weaker purifying
597 selection, thus leading to accumulation genetic of variation in a subset of modules.
598 Consistent with this latter hypothesis, we observed that variable interactions are more
599 common in modules not associated with essential biological pathways (e.g., translation
600 and ATP synthesis processes). The small overlap between variable modules and the
601 chain structure in Figure 4 is representative of these patterns. Previously, we observed
602 a similar pattern in which individual genes exhibiting expression GxE were localized at
603 the periphery of a gene co-expression network under drought[48]. The pattern we observe
604 here is similar, but now at the level of gene-gene regulatory interactions.
605 While we have highlighted a number of cases where the interactions among
606 genes vary as a function of genotype, environment, and GxE, we do observe several
607 modules that, while conserved among library types, are enriched for genes that respond
608 to stress. Clique 8-6, for example (Figure 4), is comprised of one module from each of
609 the four library types but is significantly enriched for genes that individually show a
610 significant response to soil drying stress in our study. These gene lists are also enriched
611 for GO terms associated with the light reactions of photosynthesis. Thus, while the gene
612 interaction network of the pathway forms a system whose interactions are constant
613 across various conditions and genotypes, this system is nonetheless environmentally
614 responsive. The mechanism of this response -- whether via active regulation or more
615 systems-level metabolic differences -- cannot be determined with this type of data but

616 we highlight this difference from the environmentally dependent gene-gene interactions
617 that we observe in, e.g., Clique 3 (Figure 4).
618 Our study also revealed a functional connection between cellular and whole-plant
619 traits to specific gene regulatory features. First, modules associated with traits or
620 enriched with GO terms represent hypotheses about the transcriptional control of key
621 drought responses. However, even in modules representing housekeeping pathways,
622 fewer than 50% of genes have annotations; our network-based approach therefore
623 represents to discover gene function, particularly for genes in the overlapping
624 intersection between modules. Second, certain traits, like leaf glucose content, might be
625 controlled by several pathways, as we found to be associated with four modules (Table
626 1). We also show how these associations may change in different environments as
627 some modules only exist under drought stress.
628 Collectively, our work highlights how natural selection may modify gene
629 regulation networks to have new functions or environmental responses and may also
630 shed light on how we can modify traits through breeding. First, a gene's expression can
631 be modified at various steps of its regulation, e.g., at its direct regulator or further
632 upstream. Conversely, modification of a gene's expression to be responsive to an
633 environmental factor will likely require accounting for gene-gene interactions; our study
634 shows one way to identify such targets. Second, regulatory variation or environmentally
635 responsive features seem to accumulate in specific modules. Variants whose effects
636 are restricted to such modules may have fewer pleiotropic effects and thereby be more
637 likely to be retained in populations. Third, it might be possible to change plasticity to
638 certain stressors without having pleiotropic effects in other environments; modules with
639 many GxE edges but few G edges might be fruitful targets to such avoid trade-offs. On
640 the other hand, targeting modifications to modules with E edges but few GxE edges
641 may prevent rank-changing interactions when breeding for a range of environments[19,20].
642 While our results suggest how natural or artificial selection may shape, or be shaped by,
643 physiological and regulatory responses to the environment, how these factors directly
644 contribute to drought tolerance and lifetime fitness cannot be determined in the chamber
645 growth experiments like those performed here. Future studies need to build the

646 connection from the current transient gene expression and regulation characterization to
647 fitness in the natural environment.
648
649 **METHODS**
650 **Experimental growth conditions and drought treatments**
651 The experiment protocol was adapted from a study by Des Marais et al.[67] Two
652 accessions of *Brachypodium distachyon* (Bd21, Bd3-1) seeds were placed at 4 °C for
653 seven days to synchronize seedling germination. After that, the seeds were planted in
654 Deepot D16H pots (Stuewe&Sons) filled with Profile porous ceramic rooting media
655 (Profile Products, Buffalo Grove, IL). The dry weight of each pot (DW) and its field
656 capacity for water (FC) were recorded. FC was determined by saturating a pot with
657 water and allowing it to drain under gravity overnight and calculating the weight
658 difference from DW; this was used as the basis for the subsequent controlled soil dry
659 down. Plants were grown in growth chambers with 50% humidity, 500 μmol photons m^{-2}
660 s^{-1} light intensity, and 25°C day/20°C night (12h/12h) and bottom watered with tap water
661 (pH=5-6) every other day throughout the experiment, supplemented with DYNA-GRO
662 GROW Liquid Plant Food 7-9-5 (Dyna-Gro, Richmond, CA, USA) over the first 32 days.
663 On day 33 after sowing, control plants were watered to 85% FC every night while dry-
664 down plants were allowed to decrease their soil water content to 50% FC over 7 days
665 (5% decrease each day). The plants were divided into two groups with Group 1 used to
666 characterize plant responses along the dry-down process, and Group 2 used for gene
667 expression and metabolite content.
668 On days 0 (85%FC), 2 (75%FC), 4 (65% FC), 5 (60% FC), 6 (55% FC), and 7
669 (50% FC) nine individuals of each genotype from each treatment (control and drying)
670 were used for gas exchange measurements at 3–5 hours after lights-on, non-
671 photochemical quenching (NPQ) at 5-6 hours after lights-on, and destructive sampling
672 for leaf hydraulic potential, leaf metabolites, and total plant biomass at 6-8 hours after
673 lights-on. Our preliminary analyses determined that gas exchange and leaf water
674 potentials of droughted plants began to deviate from water controls on days 5 and 6 of
675 the dry-down; accordingly, we selected day 6 for more detailed analysis. 48-58
676 replicates of leaf tissue in each genotype and treatment were measured for gas

677 exchange and photosynthesis at 1.5-5 hours after lights-on on day 6 and harvested for
678 RNA extraction and metabolites assay between 6-7.5 hours after lights-on (equivalent to
679 mid-day). Sampling of individuals was randomized to avoid confounding time-of-day
680 harvest with genotype or treatment.
681
682 **Physiological and metabolic measurements**
683 Net CO_2 assimilation rate (A), and stomatal conductance (g_{sw}) of the youngest fully
684 expanded leaf of a selected tiller were determined using a Li-6800 (LI-COR
685 Biosciences, Lincoln, NE). The LiCor chamber was set to a flow speed of 200 μmol s^{-1},
686 humidity at 50%, CO_2 concentration at 400 μmol mol^{-1}, fan speed of 10,000 rpm, and
687 photosynthetic photon flux density (PPFD) at 1000 μmol (m^{-2} s^{-1}). Leaves were left for 5
688 min to stabilize in the chamber before each reading. Photochemical quenching (qP),
689 non-photo-chemical quenching (qN), and Nonphotochemical chlorophyll fluorescence
690 quenching (NPQ), PSII maximum efficiency of light adapted leaves (fv'/fm') were
691 measured on leaves at noon; Parameters follow Li-6800 protocol of Determining PSII
692 efficiency. Leaf chlorophyll content was estimated by taking the average of two leaves
693 measured with MultispeQ V2.0 prior to lights-on in the growth chamber.
694 The leaf water potential was measured by the Scholander Pressure Chamber
695 Model 670 (PMS Instrument Company, Albany, OR) on the youngest fully expanded
696 leaf of sampled plants.
697 Above-ground tissue was excised at soil level with a razor blade. Roots were
698 cleaned with tap water to gently remove growth media and dried with tissue paper.
699 Fresh tissue biomass was recorded before samples were dried for 48 h at 80°C and
700 then weighed for dry biomass.
701
702 **Statistical analyses**
703 For the traits collected during dry-down (Supplementary Data Set 9), most traits were
704 fitted into linear mixed models using lm and lmer [68] in R[69]. Glucose, fructose were fitted
705 with generalized linear mixed models with Gamma distribution in glm and glmer [68], with
706 family=Gamma(link= "inverse"). Accession (two levels), Treatment (two levels), Harvest
707 day (six levels), and their interactions, were modeled as fixed factors and freeze dry

708 session (four levels) and assay batches (five levels) for metabolites were modeled as
709 random factors. Model selection was first performed on the fixed-effect parts of each
710 model, based on maximum likelihood Akaike information criterion (AIC) using the
711 dredge function in the MuMin package [70]. Random-effects factors of each model are
712 based on maximum likelihood of models using the anova function. Models are shown in
713 Supplementary Table 2. Throughout the dry-down process, the pairwise significance
714 test between treatments for each accession on each harvesting day (Figure 1 and
715 Supplementary Figure 1) was estimated with a t-test (or Z-test for glm, glmer) adjusted
716 by Tukey using emmeans function [71] in R based on the model selected. Outliers were
717 removed before model fitting. The results are visualized using ggplot2 and ggpubr[72,73].
718 Similarly, leaf trait data collected on Day 6 (Supplementary Data Set 10) were fit
719 with (generalized) linear mixed models using lm, lmer, glm, or glmer, as appropriate,
720 with fixed factors of Accession, Treatment, and their interaction, and random factors of
721 growth batch, freeze dry sessions, and assay batch for metabolites (Figure 2). The p-
722 values shown are based on likelihood ratio test (Chisq test) of dropping each term using
723 anova function.
724
725 **Transcriptome sampling and processing**
726 Leaf tissue flash frozen on liquid nitrogen was ground using an MP FastPrep-24™
727 Classic bead beating grinder with CoolPrep™ 24 x 2mL Adapter (MP Biomedicals)
728 loaded with dry ice. Approximately 100mg of tissue was used for RNA extraction using
729 Spectrum Plant Total RNA Kit (Sigma-Aldrich, USA) according to the manufacturer
730 instructions, followed by DNase treatment (On-Column DNase I Digestion Set, Sigma-
731 Aldrich, USA). The quality and quantity of the RNA samples were determined both by
732 absorbance measurements (Nanodrop 1000, NanoDrop Technologies) and with a Tape
733 Station (4200 TapeStation System, Agilent). 25ng of RNA were processed for High-
734 Throughput 3' Digital Gene Expression [74,75] by the MIT BioMicro Center, pooled and
735 then sequenced across three flowcells on an Illumina Novaseq (NovaSeq6000
736 Sequencing System, Illumina) generating, on average, 3.9±1.3 million raw reads per
737 sample.

738 Raw sequence reads were trimmed by Trimmomatic [76] and Cutadapt [77] to
739 remove adapters and poly(A) tails. All reads were mapped to the *B. distachyon* Bd21
740 primary transcriptome sequence
741 (Bdistachyon_556_v3.2.transcript_primaryTranscriptOnly, downloaded from
742 Phytozome) and the annotated sequence of the whole Bd21 plastome, which is
743 available at Genbank (LT558597 accession). Read alignments from Bowtie2 [78] were
744 corrected by RNA-Seq by Expectation-Maximization (RSEM) [79] to account for non-
745 uniquely aligned reads. Bowtie2 alignment is done with seedlength -L 25 and mismatch
746 allowance -n 2 which allows mismatches of Bd3-1 read alignments. Alignment with --
747 estimate-rspd specifies the reads position distributions that are highly 5' or 3'- biased [79].
748 Reads with the same unique molecular identifier and aligned to the same genes were
749 removed to reduce PCR biases during library generation [74,75] prior to adjustment in
750 RSEM. RNASeq reads from Bd21 and Bd3-1 mapped to the Bd21 references at
751 77.0±1.7% and 78.1±1.8 %, respectively, suggesting no genotypic effect on mapping.
752 Reads which did not map to the genome during this first round were mapped onto
753 ribosomal RNA, after which the total mapping rates increased to ~96%.
754
755 **Differential gene expression (DGE)**
756 RNASeq reads were normalized using Relative Log Expression normalization
757 implemented by DESeq2 [80]; these normalized data with $\log_2$(Count+1) transformation
758 are used for downstream analyses. The grouping is visualized with PCA using
759 factoextra package in R.[81] In DGE step, genes with >0 counts in more than 10 samples
760 are kept. DESeq2 fit the transcript count of each gene into the generalized linear model
761 Transcript Count ~ Genotype + Treatment (+ Genotype * Treatment) or Transcript
762 Count ~ Treatment, stratified by genotype or Transcript Count ~ genotype, stratified by
763 treatment. Each gene is fitted into a negative binomial distribution with mean μ and
764 variance α. Results in Figure 3 were visualized with ggplot in R. [72,73]
765
766 **Identifying patterns of variation in clusters of gene-gene interactions**
767 For this and following analyses, we removed transcripts exhibiting low variance (<0.1)
768 among replicates in each library type; in total, 16,207 genes are included. Co-

769 expression modules in each library type (Bd21 Control, Bd21 Drought, Bd3-1 Control,
770 Bd3-1 Drought) were generated using Weighted Gene Co-expression Network Analysis
771 (WGCNA)[82]. For WGCNA, the starting correlation matrix is unsigned, as both negative
772 and positive correlations are of interest in this study. Weight power=5 is added to the
773 correlation matrix to achieve a scale-free adjacency matrix. The topology similarity
774 matrices in the Bd21c, Bd3-1d, and Bd3-1c data sets are scaled to have a similar
775 distribution to Bd21d, which we chose arbitrarily as the reference library type.
776 Dendrograms were generated from topology similarity matrices, and modules were
777 generated by cutting at three parameter settings: (1) moderate size modules, used for
778 most analysis below, high dissimilarity (height= NULL(0.99) with a dynamic hybrid
779 algorithm (minClusterSize =30, deepsplit=2, pamStage = FALSE, mergeheight=0.25);
780 (2) conservative (smaller) modules, used for causal inference analyses discussed
781 below, cutHeight = 0.984, minClusterSize = 15, deepSplit = 3, pamStage = FALSE,
782 mergeheight=0.25; and (3) liberal (larger) modules, cutHeight = NULL(0.99),
783 minClusterSize = 30, deepSplit = 1, pamStage = FALSE, mergeheight=0.25. The three
784 different settings are used to better understand the effect of threshold choice within
785 each library type. In brief, our WGCN analyses recovered 15, 24, and 14 modules in
786 Bd21c in the three settings, respectively; 14, 18, and 14 modules in Bd21d in the three
787 settings, respectively; 14, 16, and 14 modules in Bd3-1c in three settings, respectively;
788 and 15, 20, and 15 modules in Bd3-1d in the three settings, respectively. Gene
789 membership is summarized in Supplementary Data Set 9.
790 We next addressed metrics of module preservation across our four library types,
791 Bd21c, Bd21d, Bd3-1c, and Bd3-1d. First, we assessed the degree to which the lists of
792 genes in a given module are conserved among our two genotypes and two
793 environmental conditions using a test for membership overlap. We required that module
794 overlap must be significant in all three parameter settings (moderate, conservative, and
795 liberal, described above, as shown in Figure 4, three modules with unstable overlapping
796 are not shown). To examine the overlap in gene module membership among library
797 types we used the SuperExact test [57] and a p-value cut-off of with $< 10^{-12}$ with the
798 alternative hypothesis that the size intersect between two groups or among more
799 groups is larger than expected by chance (Supplementary Data Set 5). Hypergraphs

representing module overlap networks -- here, termed “cliques” for subnetworks -- were visualized using the R package igraph [83,84].

Second, we tested for enrichment of transcripts exhibiting significant DGE in our co-expression modules using Fisher’s Exact Tests [57]. Significant enrichments (p-value <0.001 uncorrected by test number) are reported in Supplementary Data Set 6 and summarized in Table 1. Here the enrichment is reported for modules estimated from our conservative parameter settings to be consistent with following network inference analysis.

Third, biological roles of each module were inferred using GO term enrichment, by inspecting gene annotations, and by testing for module correlation with the sampled physiological and metabolic traits. Module GO enrichment is conducted on using the PANTHER classification[85] and implemented in rbioapi[86], with Fisher’s exact statistical overrepresentation tests of GO biological process. Only the significantly enriched GO terms at FDR = 0.1 and the lowest-rank GO categories are reported. Here the enrichment is reported if positive in modules of any of the parameter settings (Supplementary Data Set 7). For common GO term comparison among modules in an overlapping clique, all GO categories are considered (summarized in Table 1). Common genes and annotations in overlapping cliques are list in Supplementary Data Set 3-4. Gene annotations are from MapMan (https://mapman.gabipd.org/mapmanstore) based on the Phytozome database. Correlations between module eigengenes (i.e., their first principal component) and traits was performed in each library type. The correlation p-values are reported in Supplementary Data Set 11 and summarized in Table 1 based on moderate criteria modules, uncorrected for the number of tests.

Finally, we tested the hypothesis that genes within modules share common DNA motifs corresponding to known transcription factor binding sites. All possible 6-mer sequences were tested for enrichment in the -500 to +200 upstream window regions[87] relative to each gene’s predicted transcriptional start site in the Bd21 reference genome, in the module using AME in MEME suite[88]. The AME parameters were: --verbose 1 --scoring avg --method fisher --hit-lo-fraction 0.25 --evalue-report-threshold 10.0, with the control set based on all the genes. The Arabidopsis DAPSeq database [89] was used as a

830 reference to interpret enriched motif sequence identified. Here the enrichment is
831 reported for moderate criteria modules.
832
833 **Causal network inference in modules**
834 Methods such as WGCNA, based on correlations in transcript abundance, cannot
835 provide information about the direction of regulatory interactions. Our high replication
836 provides the power to model genes as random variables and infer conditional
837 independence of gene expression to represent possible causality. Here, we use the
838 Unknown-Targeted Interventional Greedy Sparest Permutation (UT-IGSP) algorithm [90]
839 to infer regulatory causality, representing a group of genes as a directed acyclic graph
840 (DAG). We use all genes in a given WGCNA module – estimated from one library type -
841 - as priors to reduce latent variables in the model, relying on the simplifying assumption
842 that there are no edges among genes between modules. We include both samples that
843 the module is inferred from (here, representing "observational" data) as well as samples
844 from all other library types (here, representing "interventional data"), representing the
845 assumption that most of the gene regulation structure is the same between library
846 types. In these analyses, we use the conservative module criteria (most restrictive
847 inclusion in the WGCNA analyses) and limit the analysis to modules with number of
848 genes fewer than 70% of samples in the specific library type. When testing for modules
849 that are conserved within a treatment or within a genotype, modules with as many as 70
850 genes can be used because more samples are available for inference. We use stability
851 selection [91] as implemented in Belyaeva et al [92] to minimize the effect of our selected
852 hyperparameters. Edges with a selection probability higher than upper 0.95 (or 0.97 for
853 bigger conserved modules to keep sparse edges, and 0.95 for modules with glucose as
854 a node) quantile are kept. DAGs were exported from R using the rgraphviz and
855 visualized using Graphviz Visual Editor [93].
856
857 **Identifying patterns of variation in gene-gene interactions**
858 We extended our analysis of gene expression responses using a regression approach
859 that explicitly models the magnitude and direction of putative interaction between genes
860 based on their transcript abundances. For a putative regulator and target pair identified

861 in our DAGs, we regress the target gene transcript abundance on the regulator
862 transcript abundance, and test whether the slope and intercept coefficients of
863 regression changes as a function of genotype or environmental treatment. When there
864 are multiple inferred regulators for a target gene, a multivariate linear regression is
865 applied.

866 For the hypothesis that gene x_i $(i = 1, \ldots, n)$ regulates gene y, to infer the
867 intercept or slope differences between genotypes, treatment, or their interactions, we
868 modeled their ($\log_2$ -transformed) gene expression x_i and y as:

869 $$y = \gamma_0 + \gamma_g g + \gamma_e e + \gamma_{ge} g * e + \sum_{i=1}^{n} \left(\beta_{0,i} + \beta_{g,i} g + \beta_{e,i} e + \beta_{ge,i} g * e\right) * x_i$$

870 Here g, e are used to differentiate the dataset used to train the parameters (i.e., g, e are
871 "0,0", "0,1", "1,0", "1,1" for Bd21 control, Bd21 drought, Bd3-1 control and Bd3-1 drought
872 respectively). Thus, consistent with the DGE test using DESeq2, above, a significant p-
873 value of γ_g, β_g suggest a genotypic effect on regression intercept and slopes,
874 γ_e, β_e suggest an environmental effect on intercept and slopes, and γ_{ge}, β_{ge} suggest a
875 GxE effect of intercept and slopes. Only genes with regression R^2 greater than >0.5 are
876 considered, as estimated with the Car package in R. [94] P-value of <0.05 is reported to
877 be significant without multiple test correction. Note that if the transcript abundances are
878 standardized before fitting our regression models, the slope differentiation test should
879 closely match the results of CILP tests introduce by Lea et al[66] which we include as
880 Supplementary Figure 5.
881

882 **Allele specific expression**

883 We used allele-specific expression to assess whether GxE in transcript abundances
884 were caused by cis- or trans-acting variants. Bd21 served as a pollen donor on
885 emasculated Bd3-1 flowers, resulting in ten F1 hybrid seeds. To confirm successful
886 outcrossing in this largely selfing species, we extracted DNA from leaves (DNeasy Plant
887 Pro Kit; Qiagen, MD, US) and tested for a known length polymorphism on a 3% agarose
888 gel (data not shown). We performed a dry-down experiment on these F1s and extracted
889 leaf RNA, as above. 100ng of RNA for each of 10 samples along with 12 samples (three
890 replicated of each genotype and condition) from the initial experiment were aliquoted for

891 sequencing. For DGE, and to identify SNPs between alleles, we generated libraries
892 using NEB Ultra II Directional RNA with Poly(A) Selection (New England Biolabs,
893 Ipswich, MA). Libraries were sequenced on an Illumina Novaseq generating 34.1 ± 13.7
894 million raw reads per sample.
895 To generate a reference diploid transcriptome, RNAseq reads of each Bd3-1
896 sample were aligned to the Bd21 reference genome (Bdistachyon_556_v3.2) based on
897 a RNAseq short variant discovery pipeline [95], with STAR alignment [96] of two pass and
898 alignSJoverhangMin=30 and alignSJDBoverhangMin=10, variants calling GATK [97] and
899 the flag of dont-use-soft-clipped-bases. The variants called in each sample were then
900 filtered to find SNPs and in/dels with homozygous alternative genotype mapping. To
901 reduce false discovery only variants identified in more than half of the replicates of each
902 genotype were kept. The synthesized Bd21 or Bd3-1 vcf together with Bd21 genome
903 and annotation gff was then used to generate *de novo* Bd21 or Bd3-1 genome (only
904 exons are corrected) using g2gtools [98], and a diploid genome with emase [99]. The reads
905 from all the samples were mapped to this diploid transcriptome using bowtie[100] and a
906 count matrix was estimated from alignments using emase (an alternate pipeline
907 employing bowtie2 followed by RSEM provides similar results; not shown) to account for
908 the multi-mapping particularly accounting for allelic expression. Isoforms are not
909 differentiated in this study.
910 Genes with different counts between two alleles are considered to have
911 sequence variants but sparse differences in a gene might not accurately differentiate
912 two alleles. Accordingly, we included 3963 genes that have more than >1 count of the
913 allele in one accession and <20% total count of the allele in the opposing accession in
914 all samples, modified from Gould et al [101]. The RNASeq count matrix was normalized
915 using RLE in DESeq2, as above. All genes (>0 in more than 10 samples, including
916 allele-indifferentiable genes) are included in the normalization step with allele-
917 distinguishable genes shown as allelic expression and other genes as total expression
918 to ensure the normalization factors are calculated based on most genes for the
919 comparable of genes with allelic variants. Possible read-mapping bias was assessed as
920 in Lovell et al. [102] shown in Supplementary Figure 6. PCA of samples are shown in
921 Supplementary Figure 7.

922 We employed a pipeline based on DESeq2 as introduced by Lovell et al [102].
923 Briefly, the transcript abundance of a target gene is fit into full model
924 $log_2 q_{ij} = \beta_0 + \beta_T T_i + \beta_A A_i + \beta_G G_i + \beta_{AG} A_i * G_i + \beta_{TG} T_i * G_i + \beta_{AT} T_i * A_i + \beta_{AGT} A_i * T_i * G_i$
925 where q_{ij} is gene expression of sample i for gene j, T_i is treatment (control: 0, drought:
926 1), A_i is allele type (Bd21: 0, Bd3-1: 1), G_i is generation (F1: 0, Parents: 1). Accordingly,
927 β_A tests for cis-acting regulation; β_{AG} tests for trans-acting; β_{AT} tests for cis-by-
928 environment-acting; and β_{AGT} tests for trans-by-environment-acting. Here, the approach
929 was also modified to perform tests in each environmental condition (as a "half model").
930 To be specific, each test of cis/trans is done separately in each condition because of our
931 limited power and low expression of the gene in samples in some conditions; GxE
932 cannot be directly tested in this analysis. P-values <0.1 is reported to be significant
933 without multiple test correction. Results are summarized in Supplementary Data Set 12.
934
935 **Accession numbers**
936 RNASequencing data for DGE and network analysis and ASE from this article can be
937 found in the GenBank data libraries under GSE267880 and GSE267878. All source
938 code used in this study is available online
939 (https://github.com/jyjiey/GxE_in_gene_regulation_network).

940

941 **SUPPLEMENTARY DATA**
942 **Supplementary Data Set 1.** Gene DGE list
943 **Supplementary Data Set 2.** GxE DGE gene GO enrichment
944 **Supplementary Data Set 3.** Module common genes
945 **Supplementary Data Set 4.** Module common gene annotation sorted
946 **Supplementary Data Set 5.** Module overlap p-value and intersect gene number
947 **Supplementary Data Set 6.** Module G,E,GxE DGE gene enrichment p-value
948 **Supplementary Data Set 7.** Module GO enrichment
949 **Supplementary Data Set 8.** Module traits correlation p-value
950 **Supplementary Data Set 9.** Traits data along dry-down
951 **Supplementary Data Set 10.** Traits data on day6 matching RNA samples

952 **Supplementary Data Set 11.** Gene module membership
953 **Supplementary Data Set 12.** ASE type
954 **Supplementary Figure 1.** Other physiological traits during dry-down in addition to
955 Figure 1
956 **Supplementary Figure 2.** Inferred directed interactions within selected modules with
957 gene DGE and edge DGE changes labeled
958 **Supplementary Figure 3.** Examples of regulations between pairs of genes selected in
959 addition to Figure 6
960 **Supplementary Figure 4.** ASE of the genes in Figure 6
961 **Supplementary Figure 5.** Heatmap comparing CILP and correlation variation from
962 regression method
963 **Supplementary Figure 6.** Mean reads of Bd21 allelic count vs. mean reads of Bd3-1
964 allelic count across all samples
965 **Supplementary Figure 7.** PCA of allelic gene expression
966 **Supplementary Table 1.** Percentage of E, G, G+E and GxE DGE change and slope
967 change among edges identified in small modules with inferred directed interactions
968 **Supplementary Table 2.** Models selected for physiological traits during dry-down
969 reported in Figure 1 and Supplementary Figure 1
970 **Supplementary File 1.** Direct test of DGE change between two genes
971

972 **ACKNOWLEDGEMENTS**

973 This work was supported by a grant from the MIT Jameel Water and Food Systems Lab
974 (JWAFS) and NSF CAREER Award (IOS 2239070) to DLD. JY was supported, in part,
975 by a graduate research fellowship from MIT JWAFS. ACW & AR were supported
976 through the United States Department of Energy contract no. DE-SC0012704 to
977 Brookhaven National Laboratory. We acknowledge the MIT SuperCloud and Lincoln
978 Laboratory Supercomputing Center for providing (HPC, database, consultation)
979 resources that have contributed to the research results reported within this paper. We
980 also acknowledge the computational support of Massachusetts Green High-
981 Performance Computer the Commonwealth Computational Cloud for Data Driven
982 Biology (C3DDB).

983

984 **AUTHOR CONTRIBUTIONS**

985 DLD and JY designed the research; JY, ACB, AR and DLD performed research; JY and
986 DLD design the analytic and computational tools, analyzed data. DLD and JY wrote the
987 paper with contributions from AR and ACB.

988

989 **REFERENCE**

990 1. Hill, M. S., Vande Zande, P. & Wittkopp, P. J. Molecular and evolutionary processes
991 generating variation in gene expression. *Nature Reviews Genetics* 1–13 (2020)
992 doi:10.1038/s41576-020-00304-w.
993 2. Wittkopp, P. J. & Kalay, G. Cis-regulatory elements: molecular mechanisms and
994 evolutionary processes underlying divergence. *Nat Rev Genet* **13**, 59–69 (2012).
995 3. Veltsos, P. & Kelly, J. K. The quantitative genetics of gene expression in Mimulus
996 guttatus. *PLOS Genetics* **20**, e1011072 (2024).
997 4. Huang, W., Carbone, M. A., Lyman, R. F., Anholt, R. R. H. & Mackay, T. F. C.
998 Genotype by environment interaction for gene expression in Drosophila
999 melanogaster. *Nat Commun* **11**, 5451 (2020).
1000 5. National Academies of Sciences, E. *Next Steps for Functional Genomics:*
1001 *Proceedings of a Workshop*. (2020). doi:10.17226/25780.
1002 6. Rockman, M. V. Reverse engineering the genotype–phenotype map with natural
1003 genetic variation. *Nature* **456**, 738–744 (2008).
1004 7. Des Marais, D. L., Hernandez, K. M. & Juenger, T. E. Genotype-by-Environment
1005 Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic
1006 Environment. *Annual Review of Ecology, Evolution, and Systematics* **44**, 5–29
1007 (2013).
1008 8. Lowry, D. B. *et al.* Expression Quantitative Trait Locus Mapping across Water
1009 Availability Environments Reveals Contrasting Associations with Genomic Features
1010 in Arabidopsis. *The Plant Cell* **25**, 3266–3279 (2013).
1011 9. Lovell, J. T. *et al.* The genomic landscape of molecular responses to natural drought
1012 stress in Panicum hallii. *Nat Commun* **9**, 5213 (2018).

1013 10. Josephs, E. B. *et al.* The Evolutionary Forces Shaping Cis- and Trans-Regulation of
1014 Gene Expression within a Population of Outcrossing Plants. *Molecular Biology and*
1015 *Evolution* **37**, 2386–2393 (2020).
1016 11. Josephs, E. B., Lee, Y. W., Stinchcombe, J. R. & Wright, S. I. Association mapping
1017 reveals the role of purifying selection in the maintenance of genomic variation in
1018 gene expression. *Proceedings of the National Academy of Sciences* **112**, 15390–
1019 15395 (2015).
1020 12. Varoquaux, N. *et al.* Transcriptomic analysis of field-droughted sorghum from
1021 seedling to maturity reveals biotic and metabolic responses. *Proceedings of the*
1022 *National Academy of Sciences* **116**, 27124–27132 (2019).
1023 13. Winkelmüller, T. M. *et al.* Gene expression evolution in pattern-triggered immunity
1024 within Arabidopsis thaliana and across Brassicaceae species. *The Plant Cell* (2021)
1025 doi:10.1093/plcell/koab073.
1026 14. Wu, T.-Y. *et al.* Evolutionarily conserved hierarchical gene regulatory networks for
1027 plant salt stress response. *Nat. Plants* **7**, 787–799 (2021).
1028 15. Reynoso, M. A. *et al.* Evolutionary flexibility in flooding response circuitry in
1029 angiosperms. *Science* **365**, 1291–1295 (2019).
1030 16. Bradshaw, A. D. Evolutionary Significance of Phenotypic Plasticity in Plants. in
1031 *Advances in Genetics* (eds. Caspari, E. W. & Thoday, J. M.) vol. 13 115–155
1032 (Academic Press, 1965).
1033 17. Smith, H. Signal perception, differential expression within multigene families and the
1034 molecular basis of phenotypic plasticity. *Plant, Cell & Environment* **13**, 585–594
1035 (1990).
1036 18. Fox, R. J., Donelson, J. M., Schunter, C., Ravasi, T. & Gaitán-Espitia, J. D. Beyond
1037 buying time: the role of plasticity in phenotypic adaptation to rapid environmental
1038 change. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**,
1039 20180174 (2019).
1040 19. Xiong, W. *et al.* Increased ranking change in wheat breeding under climate change.
1041 *Nat. Plants* **7**, 1207–1212 (2021).

1042 20. Choquette, N. E. *et al.* Environment-specific selection alters flowering-time plasticity
1043 and results in pervasive pleiotropic responses in maize. *New Phytologist* **238**, 737–
1044 749 (2023).
1045 21. Levis, N. A. & Pfennig, D. W. Evaluating 'Plasticity-First' Evolution in Nature: Key
1046 Criteria and Empirical Approaches. *Trends Ecol Evol* **31**, 563–574 (2016).
1047 22. Rispail, N. *et al.* Multi-Environmental Trials Reveal Genetic Plasticity of Oat
1048 Agronomic Traits Associated With Climate Variable Changes. *Front. Plant Sci.* **9**,
1049 (2018).
1050 23. Des Marais, D. L. & Juenger, T. E. Pleiotropy, plasticity, and the evolution of plant
1051 abiotic stress tolerance. *Annals of the New York Academy of Sciences* **1206**, 56–79
1052 (2010).
1053 24. Remold, S. K. & Lenski, R. E. Pervasive joint influence of epistasis and plasticity on
1054 mutational effects in Escherichia coli. *Nat Genet* **36**, 423–426 (2004).
1055 25. Wei, S. *et al.* A transcriptional regulator that boosts grain yields and shortens the
1056 growth duration of rice. *Science* **377**, eabi8455 (2022).
1057 26. Da Ros, L. *et al.* Multi-omics atlas of combinatorial abiotic stress responses in
1058 wheat. *The Plant Journal* **116**, 1118–1135 (2023).
1059 27. Shu, J. *et al.* Transcriptomic, proteomic, metabolomic, and functional genomic
1060 approaches of Brassica napus L. during salt stress. *PLOS ONE* **17**, e0262587
1061 (2022).
1062 28. Xu, K. *et al.* Proteomic analysis reveals the molecular mechanism underlying the
1063 cold acclimation and freezing tolerance of wheat (Triticum aestivum L.). *Plant Sci*
1064 **318**, 111242 (2022).
1065 29. Amiour, N. *et al.* The use of metabolomics integrated with transcriptomic and
1066 proteomic studies for identifying key steps involved in the control of nitrogen
1067 metabolism in crops such as maize. *Journal of Experimental Botany* **63**, 5017–5033
1068 (2012).
1069 30. Zhang, H., Zhu, J., Gong, Z. & Zhu, J.-K. Abiotic stress responses in plants. *Nat Rev*
1070 *Genet* **23**, 104–119 (2022).
1071 31. Passioura, J. B. Translational research in agriculture. Can we do it better? *Crop*
1072 *Pasture Sci.* **71**, 517–528 (2020).

1073 32. Nicotra, A. B. *et al.* Plant phenotypic plasticity in a changing climate. *Trends Plant*
1074 *Sci* **15**, 684–692 (2010).
1075 33. Des Marais, D. L. *et al.* Physiological Genomics of Response to Soil Drying in
1076 Diverse Arabidopsis Accessions[W][OA]. *Plant Cell* **24**, 893–914 (2012).
1077 34. Shin, J. H. *et al.* Transcriptomic changes due to water deficit define a general
1078 soybean response and accession-specific pathways for drought avoidance. *BMC*
1079 *Plant Biol* **15**, 26 (2015).
1080 35. Liu, X., Li, Y. I. & Pritchard, J. K. Trans Effects on Gene Expression Can Drive
1081 Omnigenic Inheritance. *Cell* **177**, 1022-1034.e6 (2019).
1082 36. Control of Transcription Factor Activity. in *Plant Genes, Genomes and Genetics*
1083 149–160 (John Wiley & Sons, Ltd, 2015). doi:10.1002/9781118539385.ch10.
1084 37. Transcription Factors Interpret cis-Regulatory Information. in *Plant Genes, Genomes*
1085 *and Genetics* 133–148 (John Wiley & Sons, Ltd, 2015).
1086 doi:10.1002/9781118539385.ch9.
1087 38. Making mRNAs – Control of Transcription by RNA Polymerase II. in *Plant Genes,*
1088 *Genomes and Genetics* 121–131 (John Wiley & Sons, Ltd, 2015).
1089 doi:10.1002/9781118539385.ch8.
1090 39. Marand, A. P., Eveland, A. L., Kaufmann, K. & Springer, N. M. cis-Regulatory
1091 Elements in Plant Development, Adaptation, and Evolution. *Annu Rev Plant Biol* **74**,
1092 111–137 (2023).
1093 40. Jones, D. M. & Vandepoele, K. Identification and evolution of gene regulatory
1094 networks: insights from comparative studies in plants. *Current Opinion in Plant*
1095 *Biology* **54**, 42–48 (2020).
1096 41. Alon, U. Network motifs: theory and experimental approaches. *Nat Rev Genet* **8**,
1097 450–461 (2007).
1098 42. Lachowiec, J., Queitsch, C. & Kliebenstein, D. J. Molecular mechanisms governing
1099 differential robustness of development and environmental responses in plants. *Ann*
1100 *Bot* **117**, 795–809 (2016).
1101 43. Thompson, D., Regev, A. & Roy, S. Comparative analysis of gene regulatory
1102 networks: from network reconstruction to evolution. *Annu Rev Cell Dev Biol* **31**,
1103 399–428 (2015).

1104 44. Sun, Y. & Dinneny, J. R. Q&A: How do gene regulatory networks control
1105 environmental responses in plants? *BMC Biology* **16**, 38 (2018).
1106 45. Wagner, G. P., Pavlicev, M. & Cheverud, J. M. The road to modularity. *Nature*
1107 *Reviews Genetics* **8**, 921–931 (2007).
1108 46. Grishkevich, V. *et al.* A genomic bias for genotype–environment interactions in C.
1109 elegans. *Molecular Systems Biology* **8**, 587 (2012).
1110 47. Grishkevich, V. & Yanai, I. The genomic determinants of genotype × environment
1111 interactions in gene expression. *Trends Genet* **29**, 479–487 (2013).
1112 48. Des Marais, D. L., Guerrero, R. F., Lasky, J. R. & Scarpino, S. V. Topological
1113 features of a gene co-expression network predict patterns of natural diversity in
1114 environmental response. *Proceedings of the Royal Society B: Biological Sciences*
1115 **284**, 20170914 (2017).
1116 49. Luscombe, N. M. *et al.* Genomic analysis of regulatory network dynamics reveals
1117 large topological changes. *Nature* **431**, 308–312 (2004).
1118 50. Scholthof, K.-B. G., Irigoyen, S., Catalan, P. & Mandadi, K. K. Brachypodium: A
1119 Monocot Grass Model Genus for Plant Biology. *The Plant Cell* **30**, 1673–1694
1120 (2018).
1121 51. Draper, J. *et al.* Brachypodium distachyon. A New Model System for Functional
1122 Genomics in Grasses. *Plant Physiol* **127**, 1539–1555 (2001).
1123 52. Des Marais, D. L., Lasky, J. R., Verslues, P. E., Chang, T. Z. & Juenger, T. E.
1124 Interactive effects of water limitation and elevated temperature on the physiology,
1125 development and fitness of diverse accessions of Brachypodium distachyon. *New*
1126 *Phytologist* **214**, 132–144 (2017).
1127 53. López-Alvarez, D. *et al.* Environmental niche variation and evolutionary
1128 diversification of the Brachypodium distachyon grass complex species in their native
1129 circum-Mediterranean range. *American Journal of Botany* **102**, 1073–1088 (2015).
1130 54. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic Gene
1131 Expression in a Single Cell. *Science* **297**, 1183–1186 (2002).
1132 55. Cortijo, S., Bhattarai, M., Locke, J. C. W. & Ahnert, S. E. Co-expression Networks
1133 From Gene Expression Variability Between Genetically Identical Seedlings Can
1134 Reveal Novel Regulatory Relationships. *Front. Plant Sci.* **11**, (2020).

1135 56. Livingston, D. P., Hincha, D. K. & Heyer, A. G. Fructan and its relationship to abiotic
1136 stress tolerance in plants. *Cell. Mol. Life Sci.* **66**, 2007–2023 (2009).
1137 57. Wang, M., Zhao, Y. & Zhang, B. Efficient Test and Visualization of Multi-Set
1138 Intersections. *Sci Rep* **5**, 16923 (2015).
1139 58. Xiong, Y. *et al.* Glucose–TOR signalling reprograms the transcriptome and activates
1140 meristems. *Nature* **496**, 181–186 (2013).
1141 59. Tzvetkova-Chevolleau, T. *et al.* The light stress-induced protein ELIP2 is a regulator
1142 of chlorophyll synthesis in Arabidopsis thaliana. *The Plant Journal* **50**, 795–809
1143 (2007).
1144 60. Global Profiling of Rice and Poplar Transcriptomes Highlights Key Conserved
1145 Circadian-Controlled Pathways and cis-Regulatory Modules. *PLOS ONE* **6**, e16907
1146 (2011).
1147 61. Marchand, G. *et al.* Bridging physiological and evolutionary time-scales in a gene
1148 regulatory network. *New Phytol* **203**, 685–696 (2014).
1149 62. Chevin, L.-M., Leung, C., Le Rouzic, A. & Uller, T. Using phenotypic plasticity to
1150 understand the structure and evolution of the genotype-phenotype map. *Genetica*
1151 **150**, 209–221 (2022).
1152 63. Wagner, G. P. & Zhang, J. The pleiotropic structure of the genotype–phenotype
1153 map: the evolvability of complex organisms. *Nature Reviews Genetics* **12**, 204–213
1154 (2011).
1155 64. Laitinen, R. A. E. & Nikoloski, Z. Genetic basis of plasticity in plants. *J Exp Bot* **70**,
1156 739–745 (2019).
1157 65. Hämälä, T., Gorton, A. J., Moeller, D. A. & Tiffin, P. Pleiotropy facilitates local
1158 adaptation to distant optima in common ragweed (Ambrosia artemisiifolia). *PLOS*
1159 *Genetics* **16**, e1008707 (2020).
1160 66. Lea, A. *et al.* Genetic and environmental perturbations lead to regulatory
1161 decoherence. *eLife* **8**, e40538 (2019).
1162 67. Des Marais, D. L., Razzaque, S., Hernandez, K. M., Garvin, D. F. & Juenger, T. E.
1163 Quantitative trait loci associated with natural diversity in water-use efficiency and
1164 response to soil drying in Brachypodium distachyon. *Plant Science* **251**, 2–11
1165 (2016).

68. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* **67**, 1–48 (2015).
69. R Core Team (2022). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
70. Bartoń, K. MuMIn: Multi-Model Inference. (2023).
71. Lenth, R. V. *et al.* emmeans: Estimated Marginal Means, aka Least-Squares Means. (2024).
72. ggplot2 Based Publication Ready Plots. https://rpkgs.datanovia.com/ggpubr/.
73. ggplot2: Elegant Graphics for Data Analysis | SpringerLink. https://link.springer.com/book/10.1007/978-3-319-24277-4.
74. Xiong, Y. *et al.* A Comparison of mRNA Sequencing with Random Primed and 3′-Directed Libraries. *Scientific Reports* **7**, 14626 (2017).
75. Soumillon, M., Cacchiarelli, D., Semrau, S., Oudenaarden, A. van & Mikkelsen, T. S. Characterization of directed differentiation by high-throughput single-cell RNA-Seq. *bioRxiv* 003236 (2014) doi:10.1101/003236.
76. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, btu170.
77. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
78. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357–359 (2012).
79. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
80. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).
81. Kassambara, A. & Mundt, F. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. (2020).
82. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
83. Csárdi, G. & Nepusz, T. The igraph software package for complex network research. in (2006).

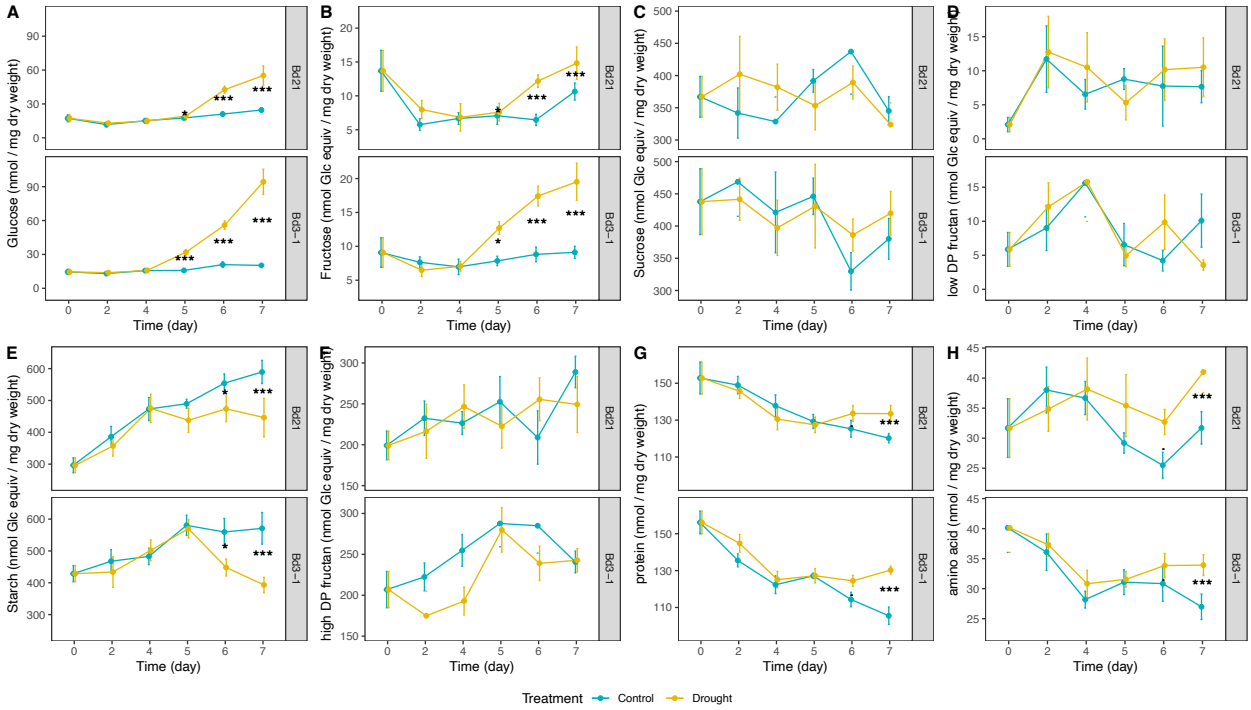
1197 84. Csárdi, G. *et al.* igraph: Network Analysis and Visualization. (2023).
1198 85. Thomas, P. D. *et al.* PANTHER: A Library of Protein Families and Subfamilies
1199 Indexed by Function. *Genome Res.* **13**, 2129–2141 (2003).
1200 86. Rezwani, M., Pourfathollah, A. A. & Noorbakhsh, F. rbioapi: user-friendly R interface
1201 to biologic web services' API. *Bioinformatics* **38**, 2952–2953 (2022).
1202 87. Ksouri, N. *et al.* Tuning promoter boundaries improves regulatory motif discovery in
1203 nonmodel plants: the peach example. *Plant Physiology* **185**, 1242–1258 (2021).
1204 88. McLeay, R. C. & Bailey, T. L. Motif Enrichment Analysis: a unified framework and an
1205 evaluation on ChIP data. *BMC Bioinformatics* **11**, 165 (2010).
1206 89. O'Malley, R. C. *et al.* Cistrome and Epicistrome Features Shape the Regulatory
1207 DNA Landscape. *Cell* **165**, 1280–1292 (2016).
1208 90. Squires, C., Wang, Y. & Uhler, C. Permutation-Based Causal Structure Learning
1209 with Unknown Intervention Targets. Preprint at
1210 https://doi.org/10.48550/arXiv.1910.09007 (2020).
1211 91. Meinshausen, N. & Bühlmann, P. Stability Selection. *Journal of the Royal Statistical*
1212 *Society Series B: Statistical Methodology* **72**, 417–473 (2010).
1213 92. Belyaeva, A. *et al.* Causal network models of SARS-CoV-2 expression and aging to
1214 identify candidates for drug repurposing. *Nat Commun* **12**, 1024 (2021).
1215 93. Rgraphviz. *Bioconductor* http://bioconductor.org/packages/Rgraphviz/.
1216 94. Fox, J. *et al.* car: Companion to Applied Regression. (2023).
1217 95. RNAseq short variant discovery (SNPs + Indels). *GATK*
1218 https://gatk.broadinstitute.org/hc/en-us/articles/360035531192-RNAseq-short-
1219 variant-discovery-SNPs-Indels- (2023).
1220 96. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15
1221 (2013).
1222 97. Auwera, G. A. V. der & O'Connor, B. D. *Genomics in the Cloud: Using Docker,*
1223 *GATK, and WDL in Terra*. (O'Reilly Media, Inc., 2020).
1224 98. g2gtools documentation. https://churchill-lab.github.io/g2gtools/.
1225 99. Raghupathy, N. *et al.* Hierarchical analysis of RNA-seq reads improves the accuracy
1226 of allele-specific expression. *Bioinformatics* **34**, 2177–2184 (2018).

100. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* **10**, R25 (2009).

101. Gould, B. A., Chen, Y. & Lowry, D. B. Gene regulatory divergence between locally adapted ecotypes in their native habitats. *Molecular Ecology* **27**, 4174–4188 (2018).

102. Lovell, J. T. *et al.* Drought responsive gene expression regulatory divergence between upland and lowland ecotypes of a perennial C4 grass. *Genome Res* **26**, 510–518 (2016).

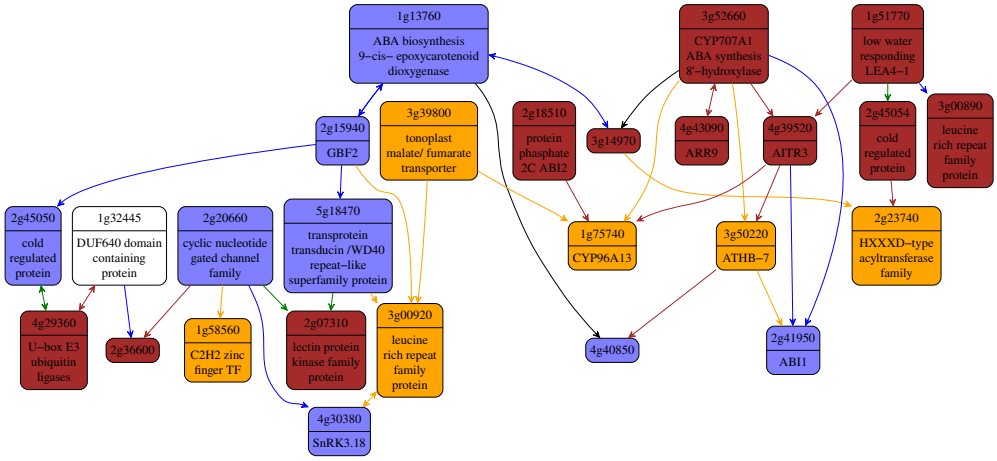
SUPPLEMENTARY INFORMATION



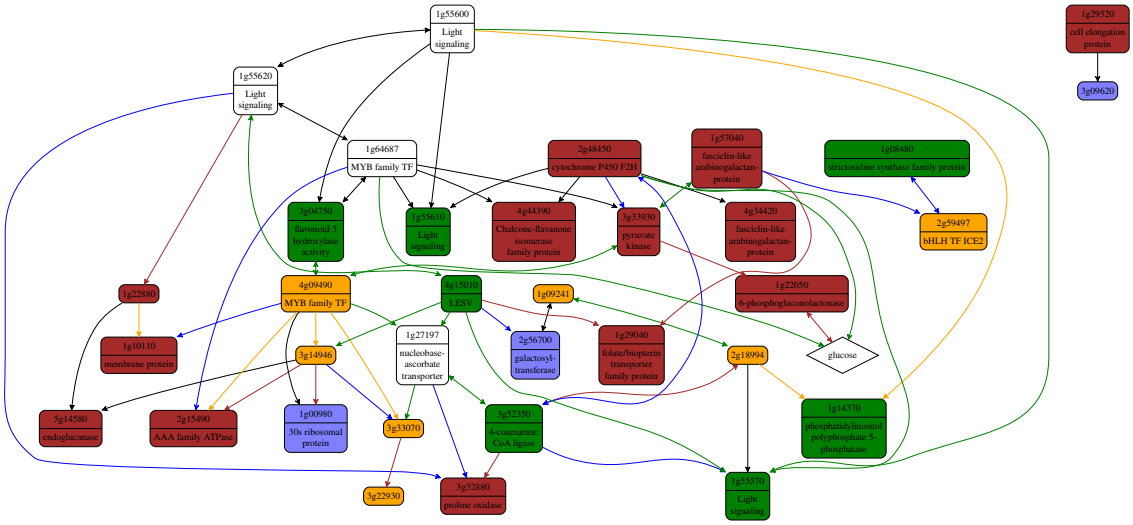
Supplementary Figure 1. Other physiological traits during dry-down in addition to Figure 1. Comparison of physiological and metabolic traits between control and dry-down in two accessions from Day 0 to Day 7. (Generalized) Linear models were determined based on AIC score. Random-effects components of the models were based on

1246 likelihood ratio tests. Based on the selected model, differences of means shown here
1247 with asterisks, were determined using contrasts between the two treatments at each
1248 time point in each genotype by Tukey's test. N = 9 for each point.
1249

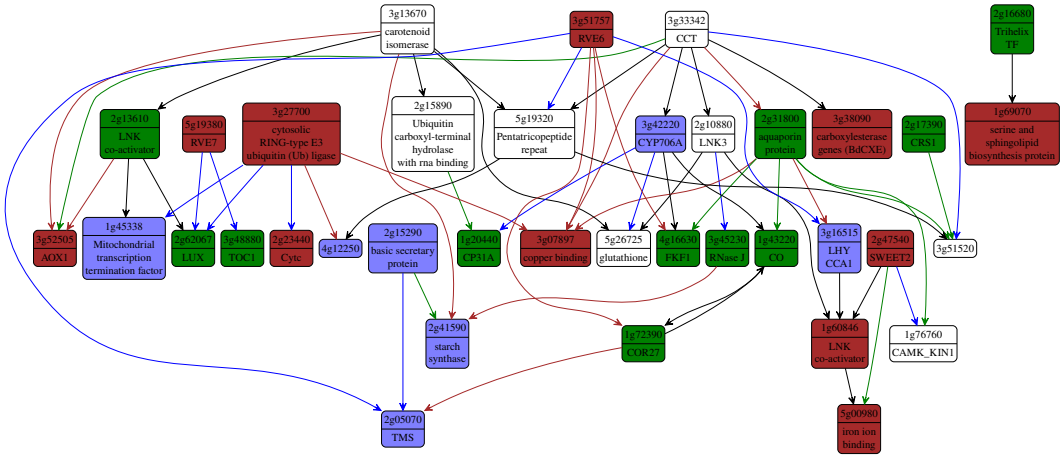
A


1g13760
ABA biosynthesis 9−cis− epoxycarotenoid dioxygenase
3g52660
CYP707A1 ABA synthesis 8'−hydroxylase
1g51770
low water responding LEA4−1
2g15940
GBF2
3g39800
tonoplast malate/ fumarate transporter
2g18510
protein phasphate 2C AB12
3g14970
4g43090
ARR9
4g39520
AITR3
2g45054
cold regulated protein
3g00890
leucine rich repeat family protein
2g45050
cold regulated protein
1g32445
DUF640 domain containing protein
2g20660
cyclic nucleotide gated channel family
5g18470
transprotein transducin /WD40 repeat−like superfamily protein
1g75740
CYP96A13
3g50220
ATHB−7
2g23740
HXXXD−type acyltransferase family
4g29360
U−box E3 ubiquitin ligases
2g36600
1g58560
C2H2 zinc finger TF
2g07310
lectin protein kinase family protein
3g00920
leucine rich repeat family protein
4g40850
2g41950
ABI1
4g30380
SnRK3.18

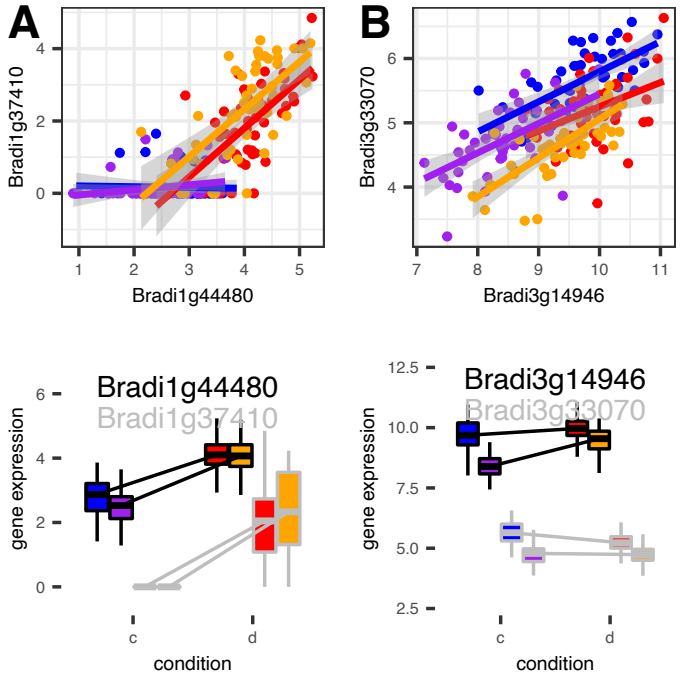
B


1g55600
Light signaling
1g29520
cell elongation protein
1g55620
Light signaling
1g64687
MYB family TF
2g48450
cytochrome P450 F2H
1g57040
fasciclin-like arabinogalactan-protein
1g08480
strictosidine synthase family protein
3g07750
flavonoid 3 hydroxylase activity
1g55610
Light signaling
4g44390
Chalcone-flavanone isomerase family protein
3g33930
pyruvate kinase
4g34420
fasciclin-like arabinogalactan protein
2g59497
bHLH TF ICE2
1g22880
4g09490
MYB family TF
4g15010
LESV
1g09241
1g22050
6-phosphogluconolactonase
1g10110
membrane protein
1g27197
nucleobase-ascorbate transporter
2g56700
galactosyl-transferase
1g29040
folate/biopterin transporter family protein
3g14946
2g18994
glucose
5g14580
endoglucanase
2g15490
AAA family ATPase
1g00980
30s ribosomal protein
3g33070
3g52350
4-coumarate-CoA ligase
1g14370
phosphatidylinositol polyphosphate 5-phosphatase
3g22930
3g32880
proline oxidase
1g55570
Light signaling

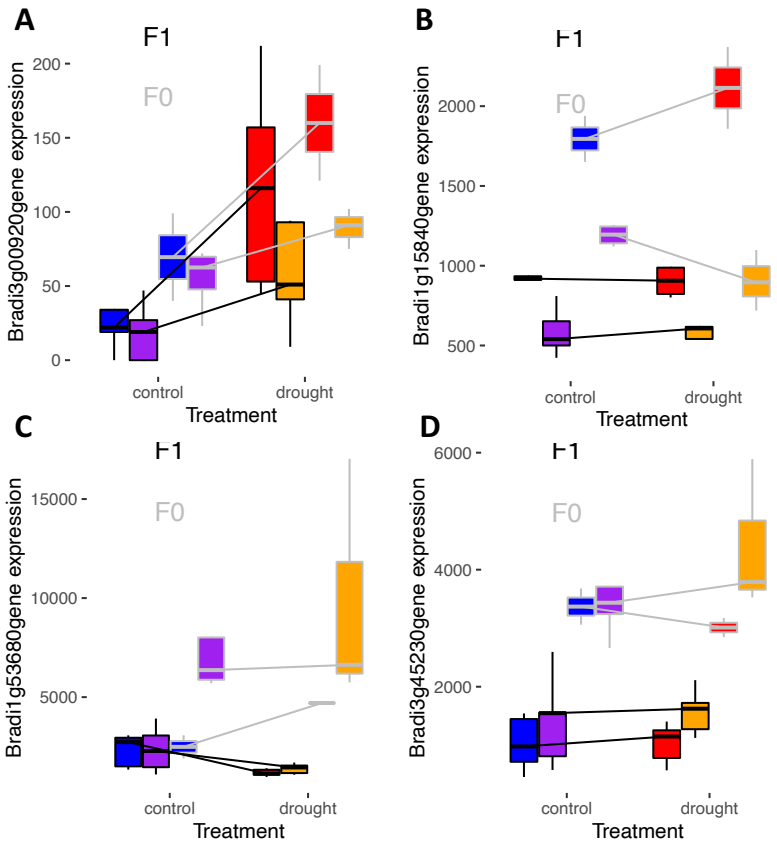
C


3g13670
carotenoid isomerase
3g51757
RVE6
3g33342
CCT
2g16680
Trihelix TF
2g13610
LNK co-activator
5g19380
RVE7
3g27700
cytosolic RING-type E3 ubiquitin (Ub) ligase
2g15890
Ubiquitin carboxyl-terminal hydrolase with rna binding
5g19320
Pentatricopeptide repeat
3g42220
CYP706A
2g10880
LNK3
2g31800
aquaporin protein
3g38090
carboxylesterase genes (BdCXE)
2g17390
CRS1
1g69070
serine and sphingolipid biosynthesis protein
3g52505
AOX1
1g45338
Mitochondrial transcription termination factor
2g62067
LUX
2g43880
TOC1
2g23440
Cytc
4g12250
2g15290
basic secretary protein
1g30440
CP31A
3g07897
copper binding
5g26725
glutathione
1g16630
FKF1
1g45230
RNase J
1g43220
CO
3g16515
LHY CCA1
2g47540
SWEET2
3g51520
2g41590
starch synthase
1g72390
COR27
1g60846
LNK co-activator
1g76760
CAMK_KIN1
2g05070
TMS
5g00980
iron ion binding

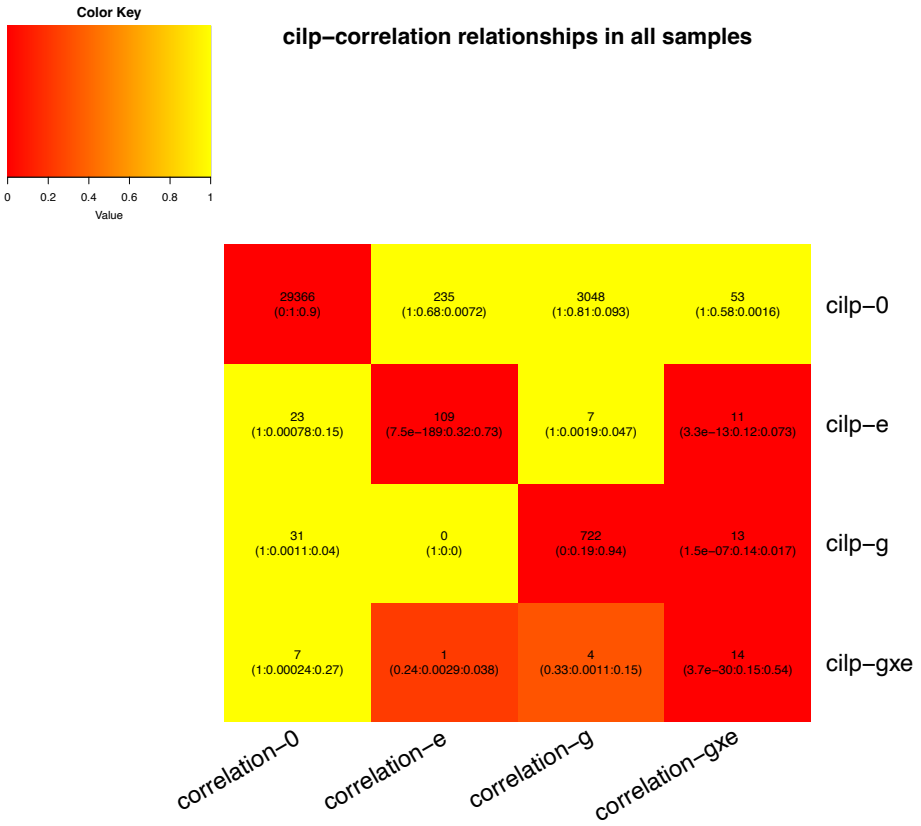
Supplementary Figure 2. Inferred directed interactions within selected modules with gene DGE and edge DGE changes, corresponding to **A)** Module 12 in Bd21d, **B)** Module 14 in Bd3-1d and **C)** Module 10 in Bd21d shown in Figure 5, with node colors are based on gene DGE and edge colors are based on edge DGE changes (blue: E, green: G, brown: G+E, orange: GxE, black/white: none).



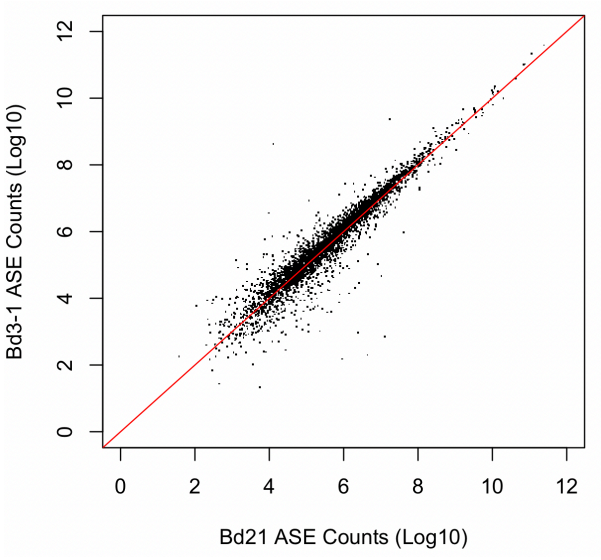
Supplementary Figure 3. Examples of regulations between pairs of genes selected in addition to Fig.6 **A)** shows regulation E slope changes, and **B)** shows transmission of GxE DGE. They are illustrated with linear regression between regulator and target, their expression levels based on log2(norm_count+1). Bd21c, Bd21d, Bd3-1c, Bd3-1d are colored with blue, red, purple, orange, respectively.



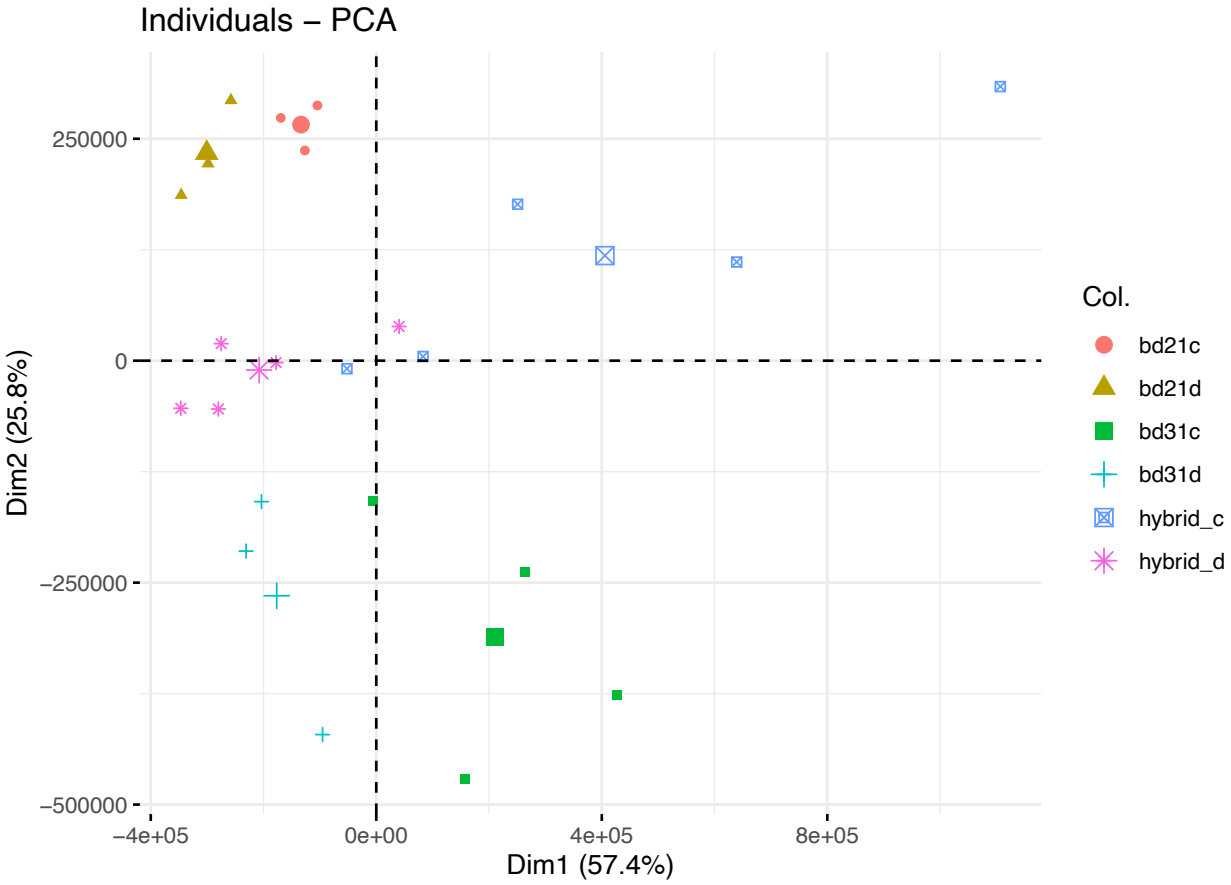
Supplementary Figure 4. ASE of the genes in Figure 6 (Bradi3g00920, Bradi1g15840, Bradi1g53680, Bradi3g45230). F1 is expression in hybrid and F0 is expression in parents. Values are normalized count. Bd21c, Bd21d, Bd3-1c, Bd3-1d are colored with blue, red, purple, orange, respectively.



Supplementary Figure 5. Heatmap comparing CILP and correlation variation from regression method based on gene pairs with correlation >0.75 in any of the library types. The number represent the # of the cases found in the intersect. The three number in the parenthesis are p-value of fisher's exact test, the proportion of the genes in the intersect in the correlation category, and the proportion of the genes in the intersect in the CILP category. The color is based on p-value.



Supplementary Figure 6. mean reads of Bd21 allelic count vs. mean reads of Bd3-1 allelic count across all samples of allele specific genes suggest no alignment bias to Bd21 or Bd3-1



1289 Supplementary Figure 7. PCA of allelic gene counts colored based on samples
1290 genotype and treatment (after normalization)
1291
1292 Supplementary Table 1. Percentage of E, G, G+E and GxE DGE change, or slope
1293 change identified in edges in small module (<40 genes) inferred gene regulation
1294 network (e.g., Figure 5).

		edge DGE change				edge rewiring %			
module no.	total edges	G	E	G+E	GxE	G	E	G+E	GxE
bd21c_13	67	25	4	10	6	10%	12%	3%	9%
bd21c_14	52	19	8	5	0	8%	17%	0%	6%
bd21c_17	37	16	4	6	0	5%	5%	3%	5%
bd21c_21	30	8	4	12	0	10%	20%	3%	7%
bd21c_22(part of 9)	22	5	3	3	0	5%	23%	9%	5%
bd21c_23(part of 10)	25	5	3	10	0	4%	8%	4%	0%
bd21d_12	41	5	11	12	11	34%	15%	5%	12%
bd21d_13	32	2	11	7	5	44%	3%	6%	6%
bd21d_16 (part of 5)	23	6	9	6	1	17%	4%	4%	17%
bd31d_12	68	15	16	11	6	10%	9%	0%	1%
bd31d_14	51	17	3	8	0	10%	4%	4%	4%
bd31d_18(part of 21)	32	4	10	10	2	13%	19%	9%	9%
bd31d_19	108	44	15	18	4	7%	24%	5%	6%
bd31c_13	79	8	4	0	0	8%	5%	3%	13%
bd21c_5	65	25	9	8	0	5%	18%	2%	9%
bd31c_10	55	11	19	10	0	7%	9%	2%	11%
bd31c_16	27	13	1	3	0	15%	19%	0%	15%

1295
1296 Supplementary Table 2. Models used for physiological traits during dry-down reported in
1297 both Figure 1 and Supplementary Figure 1. (Generalized) Linear models were
1298 determined based on AIC score. Random-effects components of the models were
1299 based on likelihood ratio tests.

No.	Trait	Model selected
1	Water usage	Accession +Treatment+harvest.day+ Accession*harvest.day+ Treatment*harvest.day
2	Hydraulic potential	Accession +Treatment+harvest.day+ Treatment*harvest.day
3	Net photosynthesis rate	Accession +Treatment+harvest.day+ Treatment*harvest.day
4	gsw	Treatment+harvest.day+ Treatment*harvest.day

5	Fresh weight	Accession +Treatment+harvest.day+ Treatment*harvest.day
6	Dry weight	Accession +Treatment+harvest.day+ Treatment*harvest.day
7	Dry root shoot ratio	Accession +Treatment+harvest.day+ Treatment*harvest.day
8	Glucose	Accession + Treatment +harvest.day+ Accession* Treatment +Treatment*harvest.day
9	fv'/fm'	Accession
10	Relative chlorophyll content	Accession +harvest.day
11	Fructose	Accession + Treatment +harvest.day+ Accession* harvest.day + Treatment*harvest.day+(1| DW.Plate)
12	Sucrose	Accession
13	Low DP fructan	-
14	Starch	Accession + Treatment +harvest.day+ Accession* harvest.day + Treatment*harvest.day+(1| DW.Plate..)
15	High DP fructan	harvest.day
16	Protein	Accession + Treatment +harvest.day+Treatment*harvest.day
17	Amino acid	Treatment +harvest.day+Treatment*harvest.day

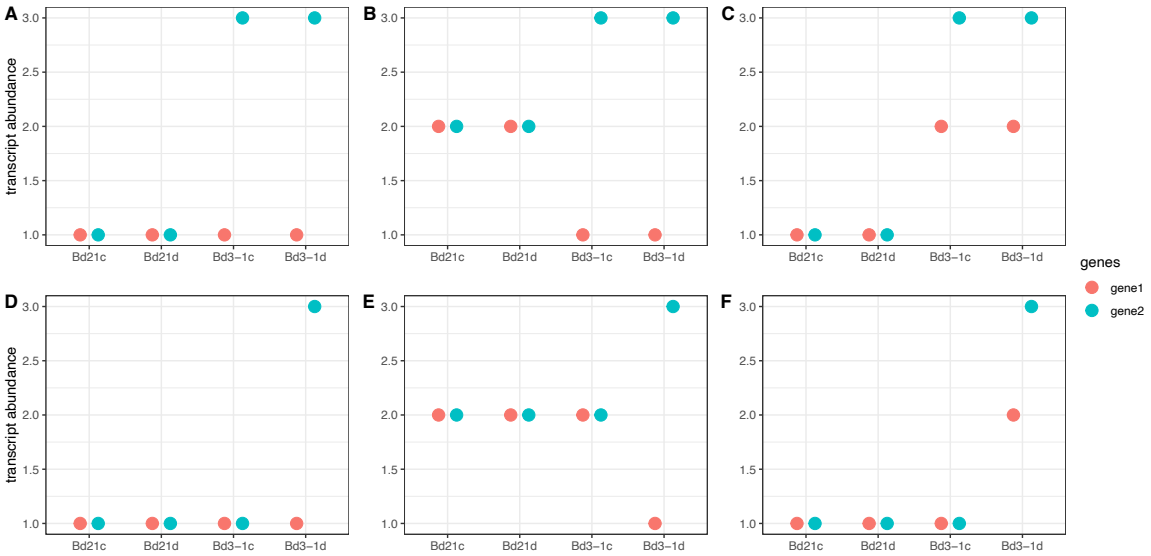
Supplementary File 1 Direct test of DGE change between two genes

Direct test of DGE change between two genes

We fit pairs of genes in regressions to formally test whether DGE changes occur between them. To be specific, a vector of log2 -transformed transcript abundance for two correlated genes were fit with the following model:

$$x = gene * g * e$$

where x is a combined vector of gene expression of two genes, and variable $gene$ $(1,2)$ differentiates which the gene of gene expression value x. Thus, the regression coefficients of $gene * g$ which differ significantly from 0 suggest that effects of genotype on transcript abundances differ between the two genes, and similarly coefficients of $gene * e$ which differ significantly from 0 suggest that effects of treatment environment on transcript abundances differ between the two genes, and significant coefficients for $gene * g * e$ suggests a GxE effect on transcript abundance. P-values <0.001 is reported to be significant without multiple test corrections. Results are shown in Supplementary Figure 4. This method avoids the indirect comparison by DGEs of two genes, but generally consistent with it (see Supplementary Figure 4).



Examples illustrating gene-gene interaction. **A-C)** DGE G changes; **E-G)** DGE GxE changes. In each type of variation, the three plots show examples of different responses (**A,D**), direction (**B,E**), and sensitivity(**C,F**), respectively. Red and blue dots show two gene expression levels respectively, in four library types.