TITLE

Genotype by environment interactions in gene regulation underlie 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

Key words: Genotype by environment interaction/ gene regulatory network/ evolution/ rewiring/ *Brachypodium distachyon*

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, yet the interactions between genotype and environment 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 in each combination of genotype and environmental treatment. 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 treatment, or in one genotype under treatment. We further find that some gene-gene interactions are conserved across conditions such that differential expression of one gene is apparently transmitted to a target gene. These variably detected edges cluster together in co-expression modules, suggestive of different constraints or selection strength acting on specific pathways. These variable features of gene regulatory networks may represent candidates modulate environmental response via genome editing, and suggest possible targets of evolutionary change in gene regulatory networks associated with phenotypic 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 may be controlled by a number of 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 may 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]. Molecular evolution studies or work 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].

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

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

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

genome averages, with a higher density of regulatory motifs[46]. An extension of this work also proposed that GxE genes might be associated with a greater number of *trans*-acting regulators than genes showing simple genotypic differential expression[47]. Previously, we showed that the number of genes connecting to, and the influence of genes showing GxE in expression are lower under drought stress but higher under cold stress compared to a random set of genes [48], again suggesting that the topology of GRNs could have been shaped by – or possibly shaped the efficacy of -- past selection in response to environmental conditions. However, that study, like most studies of environmentally dependent gene expression, assumed that co-expression relationships are invariant across genotypes or environmental treatments, which limits our ability to understand the evolution of complex gene expression phenotypes [42,49].

GxE is also observed in a large variety of species and its study plays important roles in fields as diverse as crop breeding and human health. However, studying GxE in plants facilitates highly replicated experiments with tightly controlled environments and genetically identical individuals. Here, we use two inbred natural accessions of the model grass *Brachypodium distachyon*[50] to study the physiological and transcriptomic patterns of GxE in response to soil water deficit. *B. distachyon* is closely related to economically significant grasses such as wheat, barley, and oat, and it has the advantages of small stature and small genome size[50,51]. *B. distachyon* populations are found throughout the Mediterranean where they most likely exhibit a winter annual strategy – germinating in the fall or spring and setting seed before the typically hot, dry summer months[52,53]. We selected two closely related accessions, Bd21 and Bd3-1 that have similar growth rates and life histories. Moreover, the low sequence divergence in the protein coding regions of these two accessions minimize challenges with mapping RNAsequencing reads. However, we previously documented GxE in leaf water and metabolite content between Bd21 and Bd3-1[52]. The comparison between two closely related varieties is analogous to modifying a crop variety with a small number of genome edits. We use highly replicated samples of factorial combinations of genotype and environment, harnessing variation from stochastic variance of gene expression and micro-environmental heterogeneity among replicates to reconstruct regulatory interactions[54,55]. We hypothesize that environmentally responsive, genotypically

149 specific, and genotype by environment interaction in transcript abundance and gene-
150 gene interactions are spatially clustered in the gene regulatory network(s) of *B.*
151 *distachyon*.
152
153 **RESULTS**
154 *Physiological responses to soil drying*
155 We first performed a drying experiment in which we measured a series of traits daily on
156 two inbred natural accessions as soil water content was experimentally decreased from
157 85% of field capacity to 50% of field capacity. These measurements suggest that both
158 accessions began reducing whole plant water usage starting on the fourth day of the
159 dry-down (Supplementary Figure 1A, t-test adjusted by Tukey, $p=0.048$) and thereafter
160 experienced a similar magnitude of soil drying stress with most assayed traits exhibiting
161 clear differences between control and drying treatments by Day 6. Linear mixed models
162 show that both accessions respond similarly along this dry-down, with the exception of
163 leaf glucose and fructose contents, which showed significant increase starting on Day 5
164 in Bd3-1 (z-test, $p<0.0001$) and Day 6 in Bd21 ($p<0.0001$; Supplementary Figure 1H
165 and Supplementary Figure 2B).

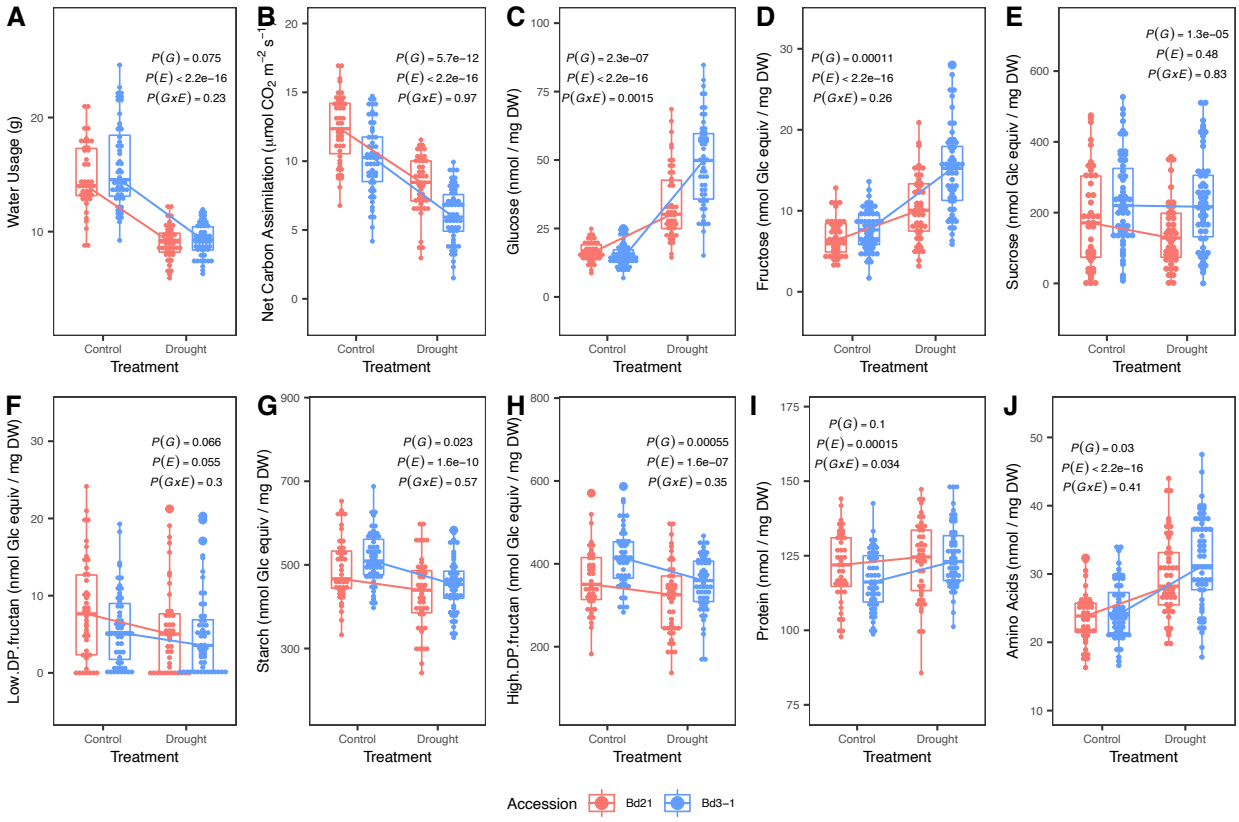


Figure 1. Leaf blade traits of Bd21 and Bd3-1 on Day 6. The p-values of effect terms were determined using (generalized) linear (mixed) models, by removing each term and performing likelihood ratio test (anova and Chisq test, with maximum likelihood mode if models have random factors). DW = Dry Weight 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%, corresponding to Day 6 of the dry-down process. These plants constitute the sampling for all following analyses.

Leaf metabolite content was affected by soil drying. Leaf starch content decreased on Day 6 (Figure 1G; Chisq test and p=1.6e-10). Sucrose shows no treatment effects in the leaf (Figure 1E; 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 1C, 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 1C, p=0.0015). Another important group of carbohydrates in cool-season Pooid grasses like *Brachypodium* are fructans, which may

185 serve as a short-term storage carbohydrate when sucrose exceeds demand [56]. Here,
186 we assayed high-degree polymerized fructan, which showed detectable decreases in
187 leaves on Day 6 (Figure 1H; p=1.6e-7), similar to what we observed for starch. Low-
188 degree polymerized fructan also decrease slightly (Figure 1F; p= 0.055), similar to
189 sucrose. In general, most carbohydrates show a genotypic effect with higher per-dry-
190 mass values in Bd3-1. Amino acids per unit leaf mass was higher in droughted plants
191 on Day 6 (Figure 1J; p<2.2e-16). Total protein content increased in Bd3-1 and
192 marginally increased in Bd21 (Figure 1I; p=0.00015), showing a GxE response
193 (p=0.034).
194
195 *Factors affecting the transcript abundance of individual genes*

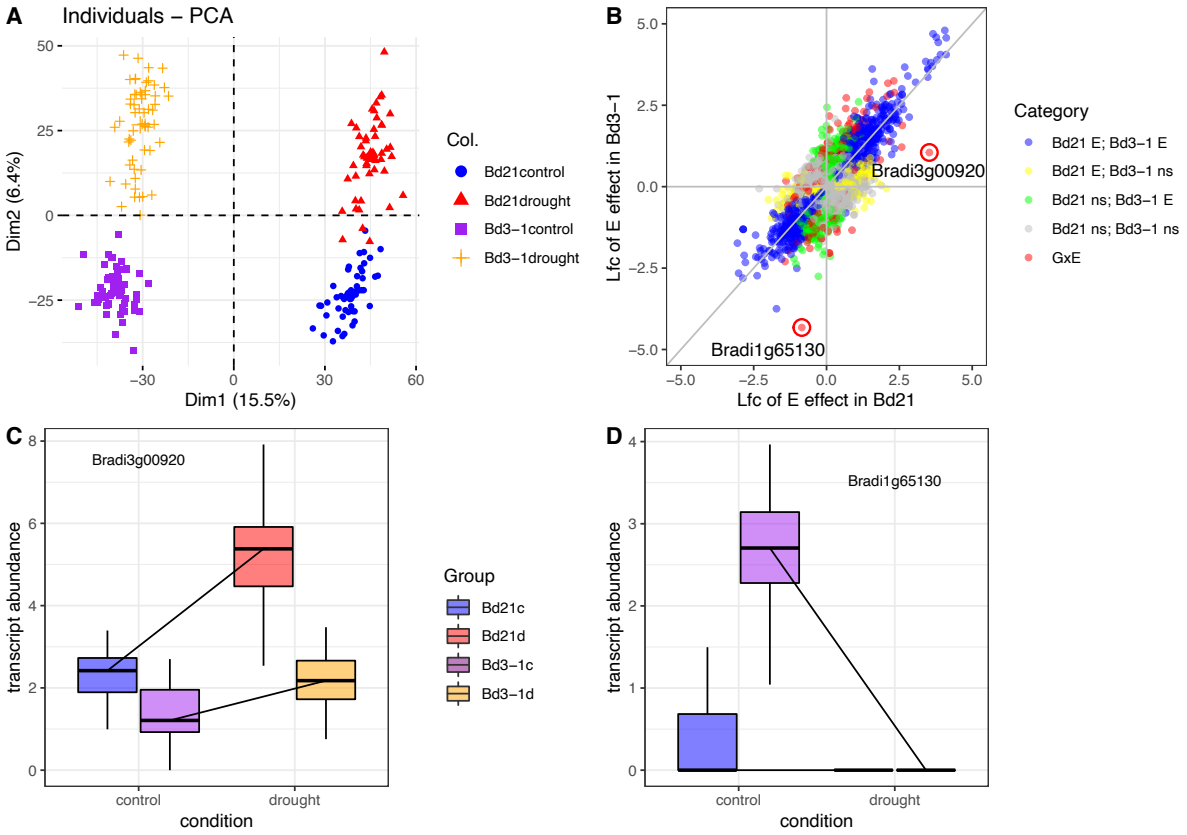


Figure 2. Gene expression diversity. **A)** PCA shows clustering of RNASequencing library by accession and environmental treatment. **B)** Factors affecting the transcript abundance of individual genes. Each point represents a gene model in the *B. distachyon* Bd21 genome. Lfc = Log-fold change of transcript 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 the two genes plotted in **C and D** 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).

The transcripts of 29,740 gene models were identified in our RNASequencing data, representing 91.4% of those in the Bd21 v3.0 reference genome. Principle component analysis shows clear grouping of samples by library type (Figure 2A), in which the first PC separates the two genotypes and explains 15.5% of total sample variance and the second PC separates the two treatments and explains 6.4% of total variance.

26,388 genes passed our filtering criteria and are included in the following analysis of by-gene differential expression (Supplementary Data Set 1). At an FDR=0.05, 4725 genes are constitutively higher, and 4087 genes are constitutively

215 lower in Bd21 as compared to Bd3-1 (i.e., show a significant genotypic effect). 3111
216 genes are constitutively higher, and 2832 genes are lower in drought than control
217 (environmental effect); among these, 3066 genes exhibit both significant genotype and
218 environment effects. In Bd21, 4024 and 3983 genes are up- or down-regulated,
219 respectively, under drought. In Bd3-1, 3997 and 4010 genes are up- or down-regulated,
220 respectively, under drought (Figure 2B). 538 genes exhibit significant GxE interaction at
221 FDR = 0.1 (Figure 2B; hereafter "GxE genes"). Among these GxE genes, 107 are up-
222 regulated and 66 are down-regulated in both accessions but with differing magnitudes
223 of response in the two accessions (Figure 2B). 135 genes are uniquely up-regulated
224 and 99 uniquely down-regulated in response to drying in Bd3-1 (so-called "conditional
225 neutrality;"[7]); similarly, 39 genes are uniquely up -regulated and 60 uniquely down-
226 regulated in Bd21. Only 21 genes are expressed in different directions in response to
227 the drought treatment in the two accessions.
228 There are 40 annotated transcription factors among the 538 GxE genes,
229 including a predicted MYB-related transcription factor with the largest GxE effect (-1.5
230 log-fold change). GO terms enriched among GxE genes include processes related to
231 chlorophyll, starch and organic substances, amino acids, translation, and protein folding
232 processes, and response to stimulus (Supplementary Data Set 2).
233
234 *Variation in gene-gene interaction within co-expression modules*
235 We next test the hypothesis that groups of putatively co-regulated genes show
236 genotypic, environmental, or GxE effects. The high replication used in our RNA
237 sequencing experiment (N = 48-58, subsampled to 48 samples for each library type)
238 affords us the power to estimate co-expression relationships independently in each of
239 the four library types (Bd21 control watering, Bd21 drought, Bd3-1 control, Bd3-1
240 drought) using Weighted Gene Co-expression Network Analysis (WGCNA)[57]. Genes in
241 a module might represent a group of genes commonly activated by regulatory factors
242 and sharing cellular functions[43]. Depending on the two parameter settings implemented
243 in WGCNA we identified 14-24 modules in Bd21 control, 14-18 modules in Bd21
244 drought, 14-16 modules in Bd3-1 control, and 15-20 modules in Bd3-1 drought (Figure
245 3).

246 We first identify modules that overlap among the four library types using a high
247 dimensional Fisher's exact test[58], testing the hypothesis that a set of modules identified
248 within or between library types share significantly more genes than expected. We
249 interrogated these "cliques" of overlapping modules for four general patterns. In the first
250 pattern, we reason that a clique comprised of one module from each of the four library
251 types may represent gene co-expression relationships that are conserved across
252 genotype and treatment (Figure 3, teal polygons). We find 40 such modules comprising
253 20 overlapping cliques, among which 36 modules show significant GO enrichments.
254 Clique 1 is comprised of four modules that have common GO terms of circadian rhythm.
255 Clique 8 is comprised of a chain of nine partially overlapping cliques enriched for core
256 cellular processes following a spatial sequence, including electron transport chain,
257 cellular respiration, protein translation, photosynthesis light reaction, and glycogen
258 metabolic process (Table 1). Some modules in conserved cliques are highly correlated
259 with physiological traits. In particular, Bd3-1d21 in Clique 8-6 is correlated with daily
260 water usage (p-value=4.4e-5) while Bd3-1d4 in Clique 8-5 is correlated with both daily
261 water usage (p-value=1.5e-4), and tissue water content (p-value=1.9e-4). Both modules
262 are enriched with GO terms associated with photosynthesis. Bd3-1d12 in Clique 8-8 is
263 correlated with leaf glucose content (p-value=6.4e-5) and partially overlaps with Clique
264 8-7 which is enriched for glycogen metabolic process.
265 The second pattern is a clique comprised of two modules from the same
266 genotype and both treatments (Figure 3, yellow polygons). We find three such cliques,
267 which may represent genes that are only co-regulated in one genotype. For example,
268 Clique 8-9 solely comprises modules in Bd21 and is enriched with functions such as
269 response to wounding and phenylpropanoid biosynthetic processes (including *B.*
270 *distachyon* orthologs of genes encoding C4H (Bradi2g31510), PAL (Bradi5g15830),
271 4CL (Bradi3g52350) and F3H (Bradi2g54090, Bradi5g19250, Bradi2g54090),
272 F3'H(Bradi4g31380), PPO (Bradi2g52260) and an R2R3 MYB transcription factor
273 (Bradi2g11676) which regulates phenylpropanoid biosynthesis in other species;
274 Supplementary Data Sets 3-4). Clique 8-12 is comprised solely of modules in Bd3-1 and
275 is enriched in chlorophyll and pigment metabolic processes and terpenoid biosynthetic
276 processes; it contains five NAC regulators/domain related genes (Bradi1g17480,

277 Bradi1g63600, Bradi3g59380, Bradi4g02060, Bradi4g44000) as well as a chlorophyll A-
278 B binding protein (Bradi4g31257) and a putative senescence-inducible chloroplast stay-
279 green protein (Bradi4g36060).

280 The third type of clique is treatment specific, composed of two overlapping
281 modules from different genotypes but the same environmental condition (Figure 3,
282 green polygons). We identified one such case, Clique 3. Both modules in this clique are
283 correlated with leaf glucose content (p-value=9.4e-8 and 1.6e-6 for Bd21 and Bd3-1,
284 respectively). Although the two modules have no GO terms in common, Bd21d13 is
285 enriched with GO terms including response to water and contains two RAB18 orthologs
286 (Bradi1g37410, Bradi3g43870), one rice ASR5 ortholog (Bradi4g24650), and one IAR4
287 (Bradi1g44480). As expected, many genes in Bd21d13 -- including the two RAB18
288 genes -- have very low expression in the control condition.

289 Finally, a clique may be comprised of modules either from a single library type or
290 from three library types and excluding the fourth library type; we interpret this pattern to
291 represent modules exhibiting GxE (Figure 3, pink or orange polygons). There are three
292 cliques which are unique to a single module in a single library type. Among these,
293 Clique 2 is comprised solely by a module identified in Bd21 drought and, although not
294 enriched for any GO terms, the module is highly correlated with leaf glucose content (p-
295 value=1.1e-5). The annotations of several genes in this module relate to abscisic acid
296 (ABA) signaling and regulation. Moreover, the proximal promoters of genes in this
297 module are enriched with a predicted bZIP motif (p-value = 2.6e-7). Clique 5 comprises
298 only one module from Bd21c and is enriched with GO term of phosphatidylethanolamine
299 biosynthesis process. Clique 7 is identified in Bd21 control watering, enriched with GO
300 terms of glutathione metabolic process and response to chemical. We also identified
301 three cliques comprised of modules from three library types with the fourth condition
302 unrepresented. Among these, Clique 8-11 lacks a Bd3-1d module and the three
303 included modules have GO terms in common related to cell wall biogenesis. Many of
304 the genes in Clique 8-11's modules are not included in any reconstructed module in
305 Bd3-1d.

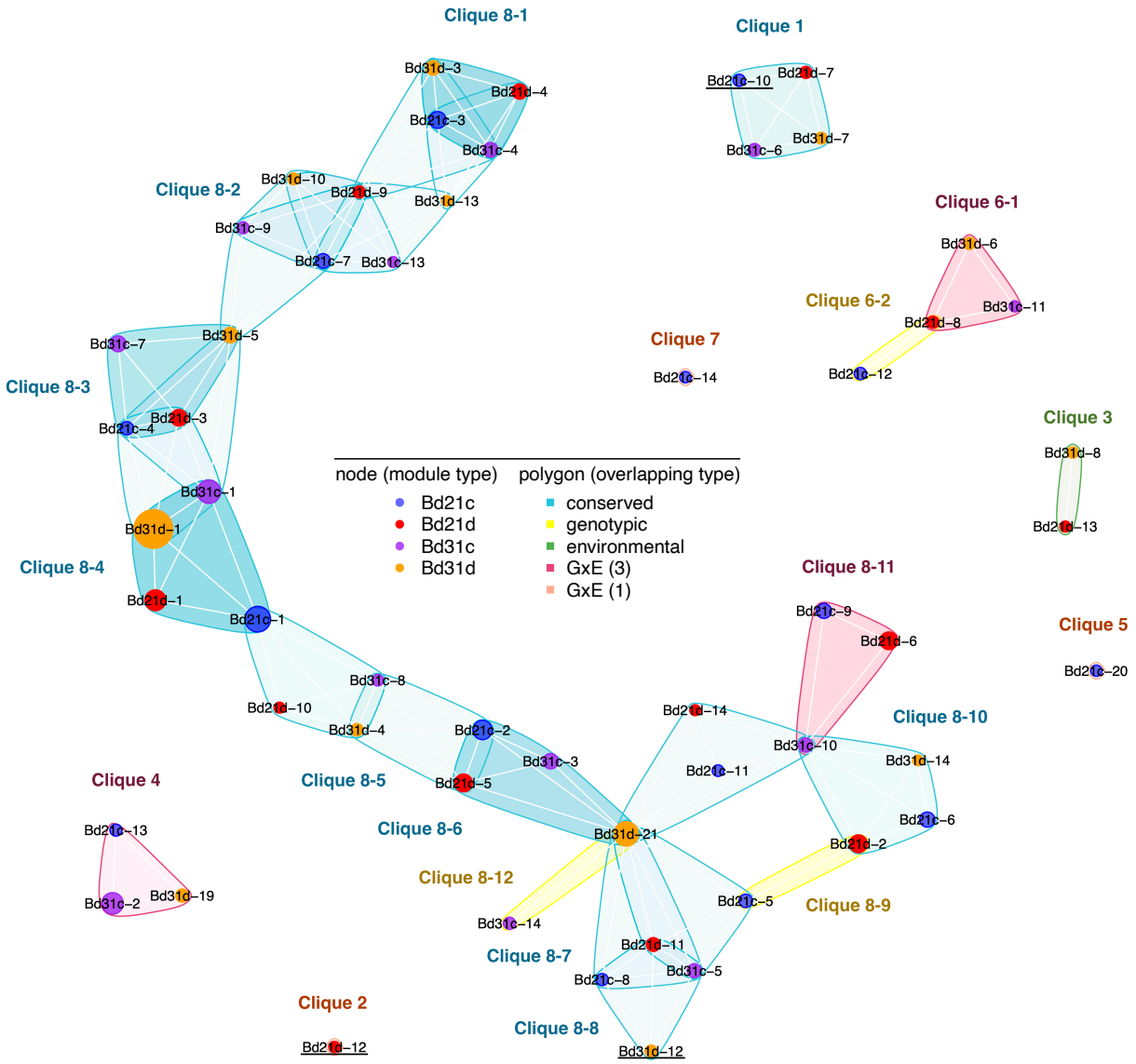


Figure 3. Conservation of gene co-expression modules among genotypes and treatments. Nodes represent modules identified in each of the four library types, with the size of the node corresponding to the number of genes in each module (red, Bd21d; blue, Bd21c; orange, Bd3-1d; purple, Bd3-1c). Clusters of modules enclosed by a polygon -- cliques -- show significant gene overlap using the high-dimensional 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 4.

319 Table 1. Common GO terms at lowest level among modules in overlapping groups
320 (based on all module versions FDR=0.1) and correlated traits (based on moderate size
321 module version, in specific library type, only p<0.001 are reported, not corrected by
322 multiple test). DGE enrichment in modules are shown (based on constraint module size
323 version, only p<0.001 are reported, not corrected by multiple test, and the percentage is
324 shown if > 75%). Cliques with no common GO terms or correlated traits identified are
325 not shown. Complete lists are available in Supplementary Data Set 6-8.
326

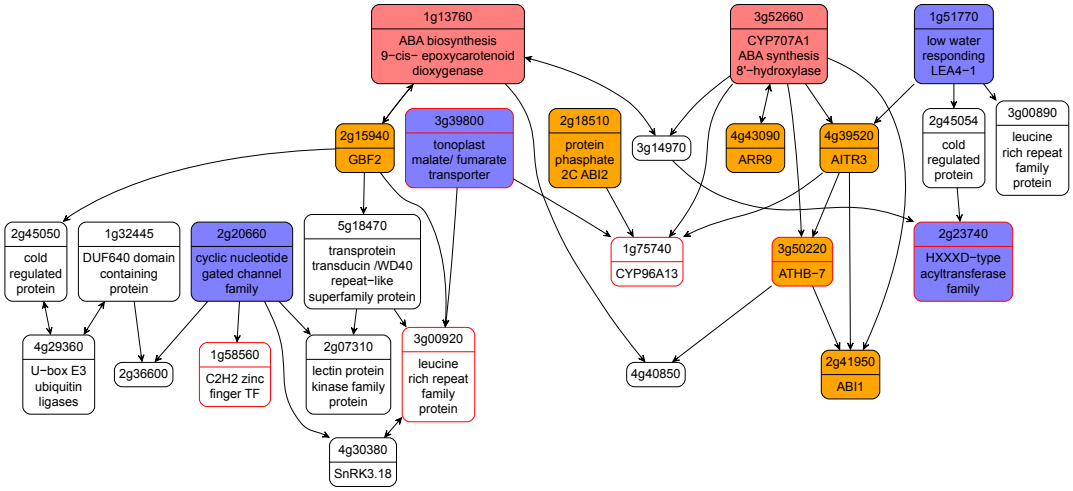
Name (# of included modules)	**Common GO**	**Correlated Trait (p<1e-03)**	**DGE Enrichment (p<1e-03; percentage when >75%)**
Clique 1 (4)	circadian rhythm		Bd3-1c6: E
Clique 2 (1)		glucose (1.1e-05)	E (97%),GxE
Clique 3 (2)		Bd21d13: Glucose (9.4e-08) Bd3-1d8: Glucose (1.6e-06)	Bd21d13: E (100%); Bd3-1d8: 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

327

328 *Inferring regulatory interactions within modules*

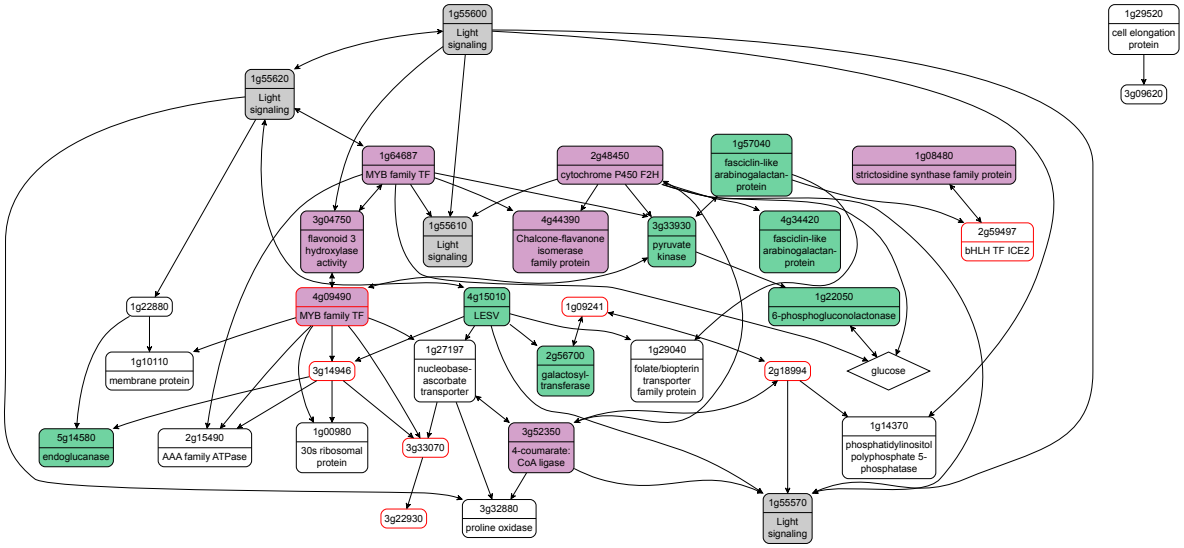
329 The preceding analyses demonstrate that the gene compositions of some co-
330 expression modules are conserved across genotypes and treatments while others vary
331 as a function of genotype, treatment, or GxE. Because co-expression may arise from
332 regulatory interactions, we next assess a possible cause of module-level variation:
333 variation in regulatory interactions among genes in modules. We use graphical models
334 to infer the conditional independence of genes to estimate causal relationships among
335 them, allowing us to infer the directionality of gene regulation. Here we work with the
336 simplifying assumption that the regulator(s) of a focal gene is in the same module as
337 that gene. We focus on three exemplar modules. Since the regulation is based on
338 estimates of transcript abundance, gene "interaction" might represent either direct
339 transcription factor-target regulation or indirect regulation that might arise from system-
340 level phenomena such as metabolism or the cell cycle [59].

341 **A**



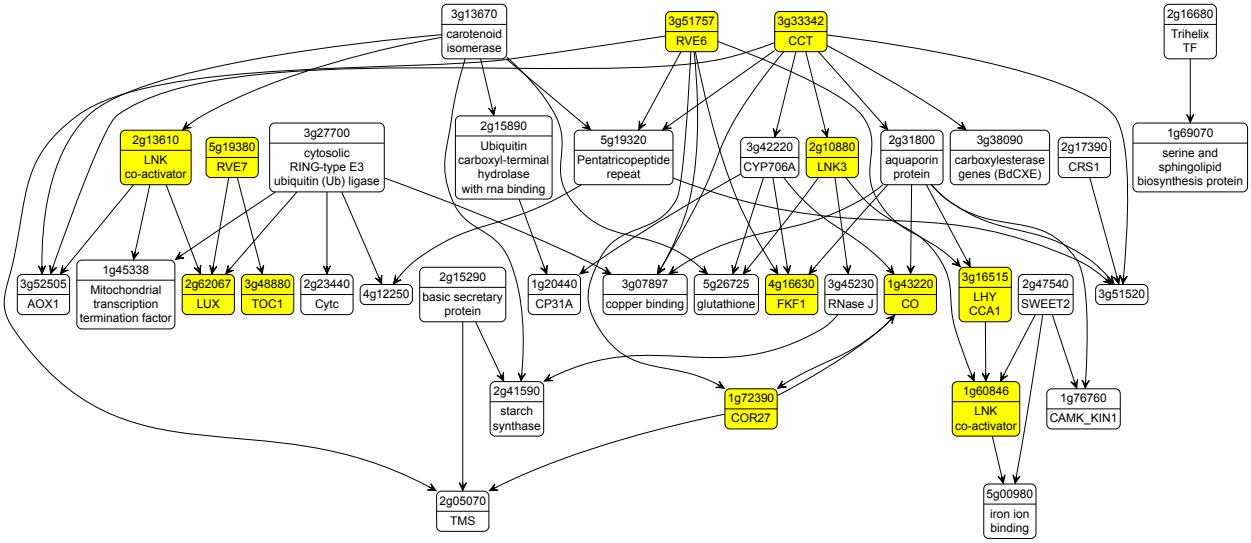
342

343 **B**



344

345 **C**



346

347 Figure 4. Inferred directed interactions within modules. **A)** Module 12 in Bd21d, **B)**
348 Module 14 in Bd3-1d and **C)** Module 10 in Bd21d inferred using the UTI-GSP method.
349 Nodes are genes and edges are inferred causal interactions. Edges with two arrows
350 indicate either the direction of interaction is uncertain or that the interaction is inferred to
351 be in both directions. Second row of each node shows annotation of orthologs in
352 Arabidopsis or rice. Glucose is included as a node in module Bd3-1d12 as an example
353 of a module with a significant correlation between leaf glucose content and the module
354 eigengene (Table 1). Genes exhibiting a significant pattern of GxE in our DGE analyses
355 are highlighted with red outline.
356
357 Bd21d12 is the sole module in Clique 2, which we above infer to be a module
358 whose coherence is GxE-dependent. Many of the genes in Bd21d12 are annotated with
359 functions in ABA biosynthesis and signaling and two of these genes are predicted to act
360 upstream many other genes in the module (Figure 4A red fill): Bradi1g13760, an
361 ortholog of AtNCED2, a 9-cis-epoxycarotenoid dioxygenase, and Bradi3g52660,
362 orthologous to AtCYP707A1, with inferred ABA 8'-hydroxylase activity. The module also
363 contains genes whose orthologs are ABA-induced or ABA-dependent regulators (Figure
364 4A orange fill), including Bradi4g39520 (orthologous to a transcription repressor,
365 AtAITR3), Bradi3g50220 (orthologous to ABA-responsive ATHB-7), Bradi2g15940
366 (orthologous to AtGBF2, an ABA regulated bZIP transcription factor), and Bradi2g41950
367 and Bradi2g18510 (orthologous to AtHAB2 and AtHAB1, respectively). Other genes are
368 associated with drought response in Arabidopsis (Figure 4A blue fill), including stomatal
369 movement (Bradi2g20660, Bradi3g39800), wax biosynthesis (Bradi2g23740), and a late
370 embryo abundant (LEA) protein (Bradi1g51770). The gene with strongest GxE effect in
371 our dataset, Bradi3g00920 (Figure 2C), has the highest eigenvalue in this module.
372 Module 12 in Bd3-1d is correlated with leaf glucose content and is a member of
373 Clique 8-8 (Figure 4B). Bradi1g55600 is inferred to be upstream of many other genes in
374 Bd3-1d12 (Figure 4B grey fill); its Arabidopsis ortholog is an early light-induced protein
375 (ELIP2) and regulates chlorophyll synthesis[60]. Paralogs of Bradi1g55600 are also
376 present in this module. Downstream in the inferred network there are genes whose
377 Arabidopsis orthologs encode enzymes in primary metabolism (Figure 4B green fill),
378 including Bradi3g33930 (pyruvate kinase), Bradi1g22050 (6-phosphogluconolactonase
379 in the oxidative phosphorylation pathway), Bradi5g14580 (an endoglucanase),
380 Bradi2g56700 (a galactosyltransferase), and Bradi1g18650 (a 4-alpha-

381 glucanotransferase in starch degradation). Above, we reported that leaf glucose content
382 is correlated with this module; when we include glucose as a node in the module, we
383 find that it is predicted to directly interact with Bradi1g22050, the 6-
384 phosphogluconolactonase. Module Bd3-1d12 also contains Bradi4g15010, whose
385 Arabidopsis ortholog is predicted to have starch synthesizing and degrading activity
386 (LESV). Additional genes in this module have predicted functions in secondary
387 metabolism (Figure 4B purple fill), including orthologs of an Arabidopsis MYB family
388 transcription factor involved in flavonol biosynthesis (Bradi1g64687), flavonoid 3'-
389 hydroxylase (Bradi3g04750), flavanone 2-hydroxylase (Bradi2g48450) and a 4-
390 coumarate: CoA ligase in the phenylpropanoid pathway (Bradi3g52350). Notably,
391 several genes with unknown functions are connected in this module and exhibit GxE in
392 their individual transcript abundances.
393 Finally, we report an example of a clique that is conserved in all four library
394 types. We investigated its constituent module, Bd21d10, as a representative of the
395 clique and find that it is enriched in genes relate to circadian processes (Figure 4C,
396 yellow fill). There are five predicted MYB TFs in Bd21d10, including Bradi3g51757
397 (orthologous to AtRVE6), Bradi2g62067 (orthologous to AtBOA and OsLUX),
398 Bradi5g19380 and Bradi1g29680 (orthologous to AtRVE2/7), Bradi3g16515
399 (orthologous to AtLHY1, and OsCCA1). Additional genes in this module are orthologous
400 to CONSTANS (CO) and related proteins in Arabidopsis. Genes in circadian-controlled
401 pathways show conserved oscillation over 24 hour intervals[61], suggesting that
402 stochastic clock variance among our sampled individuals may have provided the signal
403 for reconstructing these conserved interactions in our data.
404
405 *Genes predicted to interact exhibit various patterns of differential gene expression*
406 Several of the reconstructed co-expression modules are enriched with genes that
407 exhibit differential gene expression (Table 1). We next ask whether pairs of genes
408 predicted to interact in our directed regulatory networks (Figure 4) share a common
409 pattern of differential gene expression. For example, two genes sharing a common edge
410 may differ in their response to genotype (e.g., one gene shows no DGE and the second
411 gene shows a genotypic effect on gene expression), environment, GxE, or no

412 difference. (See Supplementary File 1 for illustrative examples.) We summarize these
413 different classes of shared or divergent DGE in Supplementary Table 1 and
414 Supplementary Figure 3. For example, of 41 total edges in module Bd21d12 (which we
415 only observe in Bd21 under drought), there are five edges showing G changes, 11
416 edges showing E changes, 12 edges showing both G and E changes, 11 edges
417 showing GxE changes, and two edges linking nodes with no difference in DGE. Of 62
418 total edges in Bd21c5 (unique to Bd21), there are 23 edges showing no DGE changes,
419 25 edges showing G changes, nine edges showing E changes, eight edges showing
420 both G and E changes, and no edges showing GxE changes.
421
422 *Genetic and environmental effects on the interactions among genes*
423 An edge between nodes in our inferred gene regulation network represents a
424 hypothesis about the regulatory interaction between two genes. The high biological
425 replication of our RNASequencing across four library types enables us to fit linear
426 models to explore the effects of genotype, environment, and their interaction on the
427 expression relationship between pairs of genes. There are four patterns of interest in
428 our dataset.
429 First, the regulatory interaction between two genes may be present under one or
430 some combinations of genotype and treatment but absent in others; this pattern would
431 be diagnosed by identifying a significant change of the slope of the regression
432 relationship between two genes. Environmentally dependent regulatory interaction is
433 one possible cause of such a pattern (Figure 5B), and we see many clear examples of
434 environmental regulation. In module Bd21d12, described above, the regression of
435 Bradi4g40850 (unknown function) on upstream genes shows a significant E effect of
436 slope from gene Bradi3g50220 (orthologous to the ABA-dependent ATHB-7
437 transcription factor; Figure 5B). Similarly, in drought specific module Bd21d13, the
438 regression of Bradi1g37410 (which encodes a dehydrin family protein) on upstream
439 genes suggests that Bradi1g44480 (predicted dehydrogenase E1 component domain
440 containing protein) regulates Bradi1g37410 only in drought (Supplementary Figure 4A).
441 Across modules, we see clustering of such context-dependent edges, such as module

442 Bd21d13 in Clique 3 – reconstructed only under drought conditions – comprising 32
443 edges of which 14 (44%) are environmentally dependent (Supplementary Table 1).
444 A regulatory interaction between two genes also may be observed in only one of
445 the two genotypes (Figure 5D). For example, in Bd21 specific module Bd21c5, the
446 regression of Bradi1g15840 (putative allene oxide cyclase in jasmonic acid synthesis)
447 on upstream genes suggest Bradi3g44990 (putative cell-wall invertase) regulates
448 Bradi1g15840 only in Bd21 (Figure 5D). Consistent with the observation that the genes
449 in Bd21c5 are only co-expressed in Bd21, we find that Bd21c5 has 12 edges showing a
450 significant effect of genotype among its among 65 edges (18%; Supplementary Table
451 1).
452 We might also observe that a regression is only significant in a single library type,
453 indicative of GxE for a given pairwise regulatory interaction (Figure 5A). In module
454 Bd21d12, discussed above and enriched for ABA-responsive genes, the regression of
455 Bradi3g00920 (Leucine Rich Repeat family protein) on upstream genes suggests that
456 Bradi2g15940 (GBF2) interacts with Bradi3g00920 with a significant GxE effect of the
457 slope (Figure 5A).
458 In the second general pattern, we observe a change in the slope of the
459 regression of the target gene on its putative regulator without a complete loss of the
460 correlation; this might be caused if the binding interaction between a transcription factor
461 and a regulatory motif is stronger in one genotype than a second genotype (Figure 5E).
462 As one example, Bradi4g20520 (unknown function) is predicted to regulate
463 Bradi3g19730 (jacalin-like lectin domain containing protein) in all library types but with a
464 greater slope in the Bd21 libraries than in Bd3-1 libraries.
465 Third, the baseline expression of genes might be different across library types
466 (Figure 5G), possibly due to the presence of cofactors or the general resource and
467 energy availability of the cell; this type of regulation may be detected as a difference in
468 the regression line's intercept among library types. In module Bd21c10, the
469 Bradi3g45230 (RNA-metabolising metallo-beta-lactamase) regression on Bradi2g10880
470 (LNK3) shows is consistent with the hypothesis that their interaction is similar in all
471 library types, but that the baseline expression in Bd3-1 is constitutively higher than in
472 Bd21. This pattern is observed as a higher intercept in Bd3-1 libraries which leads to

473 genetically dependent expression of the downstream gene. Such intercept differences
474 as a function of environment, genotype, or GxE are widely found among the pairwise
475 interactions identified in our data (Supplementary Table 1).
476 The final pattern is no difference in the pairwise regulation between genes
477 among library types but an apparent transmission of a regulatory response between the
478 genes. This pattern is manifested as neither a significant slope change nor an intercept
479 effect in the regression, and a common differential expression between putative
480 regulator and its target gene (e.g. each gene exhibits a significant effect of GxE on its
481 own transcript abundance). One such case is Bradi5g18277 (predicted serine-type
482 endopeptidase) and Bradi1g53680 (predicted zinc transporter 3 precursor) in module
483 Bd31c16 (Figure 5F). In module Bd31d12, GxE genes are found more likely to be
484 regulated by genes exhibiting GxE than by non-GxE genes in the module (t-test, p-value
485 = 0.013; see Supplementary Figure 4B for an example).
486 Finally, we use our regression approach to study the transcriptional control of leaf
487 glucose content, which shows a significant GxE effect in our metabolic assays (Figure
488 1H) and is highly correlated with four module eigengenes (module Bd21d12, Bd21d13,
489 Bd3-1d8, Bd3-1d12; Table 1). Nine genes are inferred to regulate glucose when it is
490 included as a node in these modules. Among these, we find that Bradi4g39520 only
491 exhibits an interaction with glucose under drought stress (Figure 5C); the glucose
492 content is same between the genotypes and is not dependent on the expression level of
493 Bradi4g39520 under control conditions. Bradi4g39520 is an ortholog of AITR3, an ABA-
494 induced transcription repressor that acts as feedback regulator in ABA signaling in
495 Arabidopsis.
496

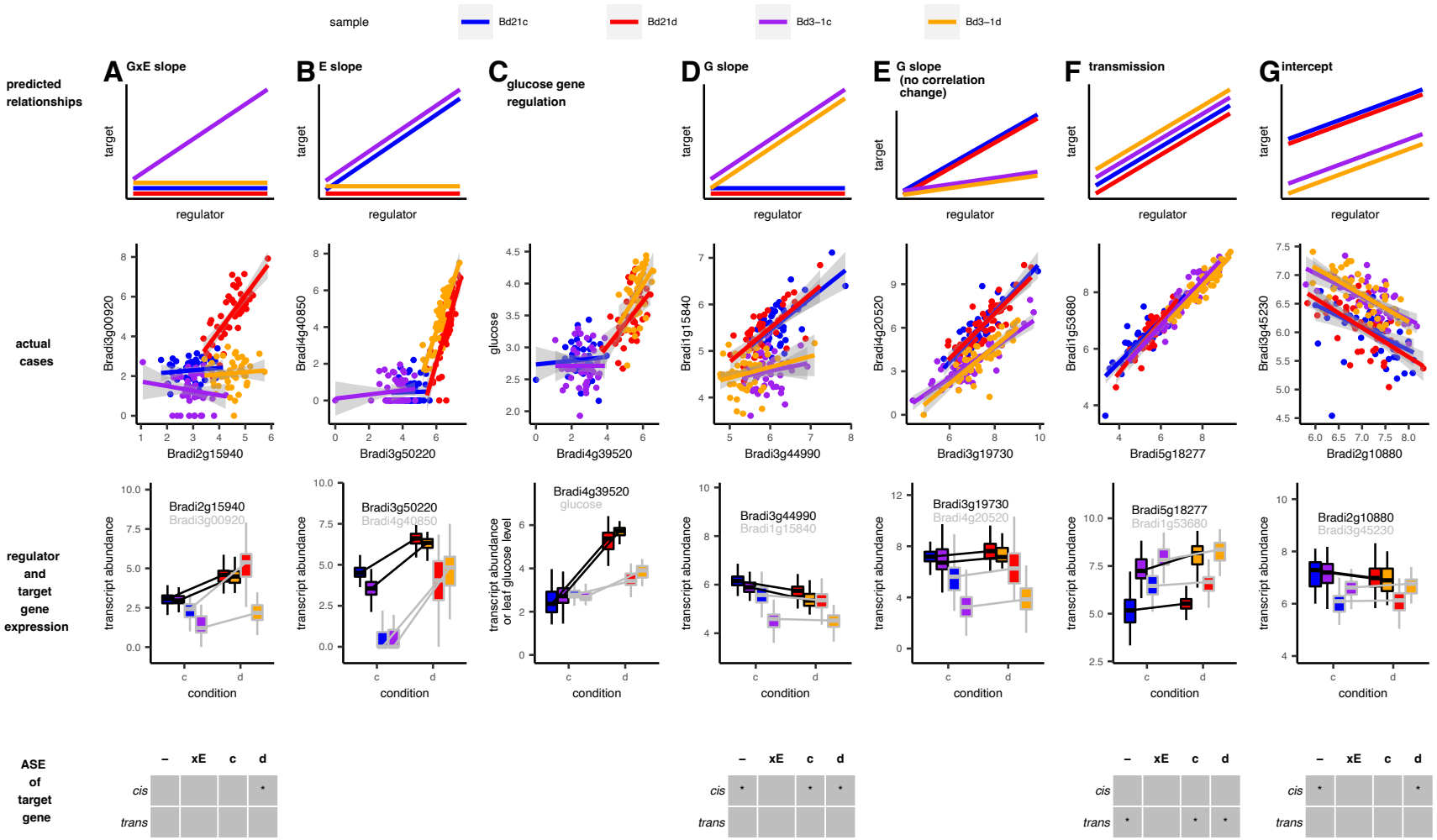


Figure 5. Predicted and actual examples of regulations between pairs of genes selected. **A-B, D-G)** show regulation slope changes, transmission, and regulation intercept changes. These are illustrated with linear regression between regulator and target, the expression level of each gene, and allele 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

503 content on upstream identified gene Bradi4g39520 as a putative regulator of leaf glucose content. Transcript abundances
504 are log2 (norm_count+1).

505 *Identifying allele-specific expression in target genes*
506 We next used allele specific expression assays to infer whether the observed genetic or
507 GxE-associated changes in regulatory interactions arise from genetic factors acting in
508 *cis* or in *trans*. In the above example of genotype-dependent interaction between
509 Bradi3g44990 and Bradi1g15840 (Figure 5D), Bradi1g15840 is inferred to be *cis*-
510 regulated (Supplementary Figure 5B). In the example of putative GxE regulation of
511 Bradi3g00920 by Bradi2g15940 (Figure 5A), we identified Bradi3g00920 is regulated by
512 a *cis*- acting factor under drought stress (Supplementary Figure 5A). In the example of
513 an apparent transmission of expression GxE from Bradi5g18277 to Bradi1g53680
514 (Figure 5F), we observed that transcript variation of Bradi1g53680 is controlled by a
515 *trans*- acting variant (Supplementary Figure 5C). We observe that the genotype-
516 dependent regression intercept of Bradi3g45230 is associated with a cis-variant (Figure
517 5G and Supplementary Figure 5D), consistent with a model in which Bradi2g10880
518 regulates Bradi3g45230 differently in two genotypes due to a variable enhancer or other
519 *cis*- regulatory element.
520
521 **DISCUSSION**
522 Gene expression has long been a focus of both evolutionary and functional genetic
523 studies. Here, we aimed to integrate these two perspectives by studying the
524 environmental and genetic control of regulatory interactions between genes using highly
525 replicated RNASequencing libraries of *Brachypodium distachyon* accessions exposed
526 to moderate soil drying and control conditions.
527 We extend past work showing that the expression of individual genes can vary as
528 a function of genotype and environmental treatment to demonstrate variation in
529 interactions between genes. One example highlighted throughout the analyses in this
530 study is Bradi3g00920, which shows a pattern of GxE in its transcript abundance. By
531 developing a directed graphical network for the module in which the gene is located, we
532 infer that Bradi3g00920 is relatively downstream of several other genes putatively
533 regulating it. Multiple regression of its gene expression suggests that Bradi3g00920 is
534 regulated by Bradi2g19540 only under drought in Bd21. We infer that this pattern
535 suggests a possible cause of the observed GxE in gene expression for Bradi3g00920.

536 Further, allele-specific expression reveals that this gene is associated with a *cis*-
537 regulated expression pattern.
538 We next consider how a differentially expressed gene may transmit its
539 expression pattern through its interactions with other genes. Here, we refer to the
540 correlated pattern of differential expression as “transmission.” As shown in the example
541 in Figure 6F, an inferred target gene (Bradi1g53680) shows the same dependency on a
542 putative regulator (Bradi5g18277) across all library types. Indeed, the target gene is
543 found to be *trans*- regulated, suggesting that the genetic variant(s) causing the
544 differential expression is in the direct regulator or upstream of it. An implication of this
545 finding is that many *trans*- acting loci in QTL studies might include such gene
546 expression transmission and not a protein coding change in the *trans*-acting regulator.
547 The transmission of the expression response is also a direct example showing certain
548 genetic variants can affect more than one gene, and that the expression of individual
549 genes should not be considered independent variables in transcriptome studies.
550 A second key implication of our work is that variation in gene-gene interactions
551 observed in modules can lead to a module not being reconstructed in certain
552 combinations of genotype and environment. For example, 44% of the inferred edges in
553 module Bd21d13 exhibit environmentally dependent regulation and, perhaps as a result,
554 this module is found only under drought stress. In contrast, in modules that have fewer
555 pairwise gene-gene regulatory variations, or when gene expression states are
556 transmitted faithfully from upstream genes in the network, these modules may be more
557 likely observed under all sampled conditions.
558 Genetic variants affecting expression, whether constitutive differences between
559 varieties (in our study, those with an only significant G effect) or GxE, may be acted on
560 by selection if they ultimately have an effect on phenotype. Past work addressed the
561 possibility of spatial separation across a network between a genetic variant and its
562 effect[35,62], or between the perception of an environmental stimulus and the downstream
563 transcriptional response[63]. The network context of a genetic variant is important for
564 evolutionary studies because the number and functions of the genes which interact with
565 a variable gene influences the pleiotropy of that genetic variant[64]. Recent work
566 considered the network topology or pleiotropy of gene expression in the context of the

567 evolution of plasticity[48,65]; however, most studies do not account for the prospect of
568 pleiotropic effects of these genetic variants and, as a result, cannot account for variation
569 in the network topology itself. We specifically consider such regulatory interaction
570 variation here. Our large RNASequencing dataset also allowed us to estimate gene co-
571 expression independently in each combination of genotype and environmental
572 treatment, thus relaxing the assumption of an invariant network topology across
573 conditions used previously[48,65,66]. We note that other approaches to account for network
574 structure, such as mapping covariance or principle components as traits in QTL analysis
575 [3,11,67] may detect similar regulatory structure.
576 Changes in gene-gene interactions are found to be clustered in some modules,
577 suggesting a spatial pattern of regulatory variation in networks and supporting the
578 hypothesis that modularity allows for reduced pleiotropic effects of mutations[64]. A
579 related hypothesis is that certain regulatory pathways are under weaker purifying
580 selection, thus leading to accumulation of genetic variation in a subset of modules.
581 Consistent with this latter hypothesis, we observe that variable interactions are more
582 common in modules not associated with essential biological pathways (e.g., not
583 enriched for GO terms such as translation and ATP synthesis processes). The small
584 overlap between variable modules and the chain structure in Figure 4 is representative
585 of these patterns. Previously, we observed a similar pattern in which individual genes
586 exhibiting expression GxE were localized at the periphery of a gene co-expression
587 network under drought[48]. The pattern we observe here is similar, but now at the level of
588 gene-gene regulatory interactions.
589 While we have highlighted a number of cases where the interactions among
590 genes vary, we do observe several modules that, while conserved among library types,
591 are enriched for genes that respond to stress. Clique 8-6, for example (Figure 4), is
592 comprised of one module from each of the four library types but is significantly enriched
593 for genes that individually show a significant response to soil drying stress in our study.
594 These gene lists are also enriched for GO terms associated with the light reactions of
595 photosynthesis. Thus, while the gene interaction network of the pathway forms a system
596 whose interactions are constant across various conditions and genotypes, this system is
597 nonetheless environmentally responsive. The mechanism of this response -- whether

598 via active regulation or more systems-level metabolic differences -- cannot be
599 determined with this type of data but we highlight this difference from the
600 environmentally dependent gene-gene interactions that we observe in, e.g., Clique 3
601 (Figure 4).
602 Our study also revealed a functional connection between cellular and whole-plant
603 traits to specific gene regulatory features. First, modules associated with traits or
604 enriched with GO terms represent hypotheses about the transcriptional control of key
605 drought responses. However, even in modules representing housekeeping pathways,
606 fewer than 50% of genes have annotations; our network-based approach therefore
607 represents a novel means to discover gene function. Second, certain traits, like leaf
608 glucose content, might be controlled by several pathways, as we found glucose to be
609 associated with four modules (i.e., the module overlapping with three modules in
610 glycogen biosynthesis process clique, two drought specific modules, and the module
611 enriched with ABA processes; Table 1). We also show how these associations may
612 change in different environments as some of the glucose-associated modules only exist
613 under drought stress. Collectively, these results suggest that our approach represents
614 an unbiased and systematic way of developing hypotheses about the genes and
615 pathways underlying biological processes.
616 Collectively, our work highlights how natural selection may modify gene
617 regulation networks to have new functions or environmental responses and may also
618 shed light on how we can modify traits through breeding. First, a gene's expression can
619 be modified at various steps of its regulation, e.g., at its direct regulator or further
620 upstream. Conversely, modification of a gene's expression to be responsive to an
621 environmental factor will likely require accounting for gene-gene interactions; our study
622 shows one way to identify such targets.
623 Second, regulatory variation or environmentally responsive features seem to
624 cluster in specific modules. Variants whose effects are restricted to such modules may
625 have fewer pleiotropic effects and thereby be more likely to be retained in populations.
626 Third, it might be possible to change plasticity to certain stressors without having
627 pleiotropic effects in other environments; modules with many GxE edges but few G
628 edges might be fruitful targets to such avoid trade-offs. On the other hand, targeting

629 modifications to modules with E edges but few GxE edges may prevent rank-changing
630 interactions when breeding for a range of environments[19,20]. While our results suggest
631 how natural or artificial selection may shape, or be shaped by, physiological and
632 regulatory responses to the environment, how these factors directly contribute to
633 drought tolerance and lifetime fitness cannot be determined in the chamber growth
634 experiments like those performed here. Future studies need to build the connection
635 from the current transient gene expression and regulation characterization to fitness in
636 the natural environment.
637
638 **METHODS**
639 **Experimental growth conditions and drought treatments**
640 The experiment protocol was adapted from a study by Des Marais et al.[68] Two
641 accessions of *Brachypodium distachyon* (Bd21, Bd3-1) seeds were placed at 4 °C for
642 seven days to synchronize seedling germination. After that, the seeds were planted in
643 Deepot D16H pots (Stuewe&Sons) filled with Profile porous ceramic rooting media
644 (Profile Products, Buffalo Grove, IL). The dry weight of each pot (DW) and its field
645 capacity for water (FC) were recorded. FC was determined by saturating a pot with
646 water and allowing it to drain under gravity overnight and calculating the weight
647 difference from DW; this was used as the basis for the subsequent controlled soil dry
648 down. Plants were grown in growth chambers with 50% humidity, 500 μmol photons m^{-2}
649 s^{-1} light intensity, and 25°C day/20°C night (12h/12h) and bottom watered with tap water
650 (pH=5-6) every other day throughout the experiment, supplemented with DYNA-GRO
651 GROW Liquid Plant Food 7-9-5 (Dyna-Gro, Richmond, CA, USA) over the first 32 days.
652 On day 33 after sowing, control plants were watered to 85% FC every night while dry-
653 down plants were allowed to decrease their soil water content to 50% FC over 7 days
654 (5% decrease each day). The plants were divided into two groups with Group 1 used to
655 characterize plant responses along the dry-down process, and Group 2 used for gene
656 expression and metabolite content.
657 On days 0 (85%FC), 2 (75%FC), 4 (65% FC), 5 (60% FC), 6 (55% FC), and 7
658 (50% FC) nine individuals of each genotype from each treatment (control and drying)
659 were used for gas exchange measurements at 3–5 hours after lights-on, non-

660 photochemical quenching (NPQ) at 5-6 hours after lights-on, and destructive sampling
661 for leaf hydraulic potential, leaf metabolites, and total plant biomass at 6-8 hours after
662 lights-on. Our preliminary analyses determined that gas exchange and leaf water
663 potentials of droughted plants began to deviate from water controls on days 5 and 6 of
664 the dry-down; accordingly, we selected day 6 for more detailed analysis. 48-58
665 replicates of leaf tissue in each genotype and treatment were measured for gas
666 exchange and photosynthesis at 1.5-5 hours after lights-on on day 6 and harvested for
667 RNA extraction and metabolites assay between 6-7.5 hours after lights-on (equivalent to
668 mid-day). Sampling of individuals was randomized to avoid confounding time-of-day
669 harvest with genotype or treatment.
670
671 **Physiological and metabolic measurements**
672 Net CO_2 assimilation rate (A), and stomatal conductance (g_{sw}) of the youngest fully
673 expanded leaf of a selected tiller were determined using a Li-6800 (LI-COR
674 Biosciences, Lincoln, NE). The LiCor chamber was set to a flow speed of 200 μmol s^{-1},
675 humidity at 50%, CO_2 concentration at 400 μmol mol^{-1}, fan speed of 10,000 rpm, and
676 photosynthetic photon flux density (PPFD) at 1000 μmol (m^{-2} s^{-1}). Leaves were left for 5
677 min to stabilize in the chamber before each reading. Photochemical quenching (qP),
678 non-photo-chemical quenching (qN), and Nonphotochemical chlorophyll fluorescence
679 quenching (NPQ), PSII maximum efficiency of light adapted leaves (fv'/fm') were
680 measured on leaves at noon; Parameters follow Li-6800 protocol of Determining PSII
681 efficiency. Leaf chlorophyll content was estimated by taking the average of two leaves
682 measured with MultispeQ V2.0 prior to lights-on in the growth chamber.
683 The leaf water potential was measured by the Scholander Pressure Chamber
684 Model 670 (PMS Instrument Company, Albany, OR) on the youngest fully expanded
685 leaf of sampled plants.
686 Above-ground tissue was excised at soil level with a razor blade. Roots were
687 cleaned with tap water to gently remove growth media and dried with tissue paper.
688 Fresh tissue biomass was recorded before samples were dried for 48 h at 80°C and
689 then weighed for dry biomass.

690 Tissue for RNA and metabolites was flash frozen in liquid nitrogen, stored at -
691 70 °C and freeze-dried overnight prior to grinding with bead-beating. 20mg dry weight is
692 utilized for the following assays. Metabolites were extracted using sequential ethanol
693 extractions following established protocols[69,70]. Protein, starch, and high degree of
694 polymerization (HDP) fructan, total free amino acids, ethanol-soluble carbohydrates
695 (glucose, fructose, sucrose, and low degree of polymerization (LDP) fructan), and free
696 nitrate were quantified following Burnett et al[69]. Metabolite data are expressed here per
697 g dry weight, or as glucose equivalents per dry weight for carbohydrates. Technical and
698 analytical replicates were run for all assays.
699
700 **Statistical analyses**
701 For the traits collected during dry-down (Supplementary Data Set 9), most traits were
702 fitted into linear mixed models using lm and lmer [71] in R[72]. Glucose, fructose were fitted
703 with generalized linear mixed models with Gamma distribution in glm and glmer [71], with
704 family=Gamma(link= “inverse”). Accession (two levels), Treatment (two levels), Harvest
705 day (six levels), and their interactions, were modeled as fixed factors and freeze dry
706 session (four levels) and assay batches (five levels) for metabolites were modeled as
707 random factors. Model selection was first performed on the fixed-effect parts of each
708 model, based on maximum likelihood Akaike information criterion (AIC) using the
709 dredge function in the MuMin package [73]. Random-effects factors of each model are
710 based on maximum likelihood of models using the anova function. Models are shown in
711 Supplementary Table 2. Throughout the dry-down process, the pairwise significance
712 test between treatments for each accession on each harvesting day (Supplementary
713 Figure 1 and 2) was estimated with a t-test (or Z-test for glm, glmer) adjusted by Tukey
714 using emmeans function [74] in R based on the model selected. Outliers were removed
715 before model fitting. The results are visualized using ggplot2 and ggpubr[75,76].
716 Similarly, leaf trait data collected on Day 6 (Supplementary Data Set 10) were fit
717 with (generalized) linear mixed models using lm, lmer, glm, or glmer, as appropriate,
718 with fixed factors of Accession, Treatment, and their interaction, and random factors of
719 growth batch, freeze dry sessions, and assay batch for metabolites (Figure 1). The p-

720 values shown are based on likelihood ratio test (Chisq test) of dropping each term using
721 anova function.
722
723 **Transcriptome sampling and processing**
724 Leaf tissue flash frozen on liquid nitrogen was ground using an MP FastPrep-24™
725 Classic bead beating grinder with CoolPrep™ 24 x 2mL Adapter (MP Biomedicals)
726 loaded with dry ice. Approximately 100mg of tissue was used for RNA extraction using
727 Spectrum Plant Total RNA Kit (Sigma-Aldrich, USA) according to the manufacturer
728 instructions, followed by DNase treatment (On-Column DNase I Digestion Set, Sigma-
729 Aldrich, USA). The quality and quantity of the RNA samples were determined both by
730 absorbance measurements (Nanodrop 1000, NanoDrop Technologies) and with a Tape
731 Station (4200 TapeStation System, Agilent). 25ng of RNA were processed for High-
732 Throughput 3' Digital Gene Expression [77,78] by the MIT BioMicro Center, pooled and
733 then sequenced across three flowcells on an Illumina Novaseq (NovaSeq6000
734 Sequencing System, Illumina) generating, on average, 3.9±1.3 million raw reads per
735 sample.
736 Raw sequence reads were trimmed by Trimmomatic [79] and Cutadapt [80] to
737 remove adapters and poly(A) tails. All reads were mapped to the *B. distachyon* Bd21
738 primary transcriptome sequence
739 (Bdistachyon_556_v3.2.transcript_primaryTranscriptOnly, downloaded from
740 Phytozome) and the annotated sequence of the whole Bd21 plastome, which is
741 available at Genbank (LT558597 accession). Read alignments from Bowtie2 [81] were
742 corrected by RNA-Seq by Expectation-Maximization (RSEM) [82] to account for non-
743 uniquely aligned reads. Bowtie2 alignment is done with seedlength -L 25 and mismatch
744 allowance -n 2 which allows mismatches of Bd3-1 read alignments. Alignment with --
745 estimate-rspd specifies the reads position distributions that are highly 5' or 3'- biased [82].
746 Reads with the same unique molecular identifier and aligned to the same genes were
747 removed to reduce PCR biases during library generation [77,78] prior to adjustment in
748 RSEM. RNASeq reads from Bd21 and Bd3-1 mapped to the Bd21 references at
749 77.0±1.7% and 78.1±1.8 %, respectively, suggesting no genotypic effect on mapping.

750 Reads which did not map to the genome during this first round were mapped onto
751 ribosomal RNA, after which the total mapping rates increased to ~96%.
752
753 **Differential gene expression (DGE)**
754 RNASeq reads were normalized using Relative Log Expression normalization
755 implemented by DESeq2 [83]; these normalized data with $\log_2$ (Count+1) transformation
756 are used for downstream analyses. The grouping is visualized with PCA using
757 factoextra package in R.[84] In DGE step, genes with >0 counts in more than 10 samples
758 are kept. DESeq2 fit the transcript count of each gene into the generalized linear model
759 Transcript Count ~ Genotype + Treatment (+ Genotype * Treatment) or Transcript
760 Count ~ Treatment, stratified by genotype or Transcript Count ~ genotype, stratified by
761 treatment. Each gene is fitted into a negative binomial distribution. Results in Figure 3
762 were visualized with ggplot in R. [75,76]
763
764 **Identifying patterns of variation in clusters of gene-gene interactions**
765 For this and following analyses, we removed transcripts exhibiting low variance (<0.1)
766 among replicates in each library type; in total, 16,207 genes are included. Co-
767 expression modules in each library type (Bd21 Control, Bd21 Drought, Bd3-1 Control,
768 Bd3-1 Drought) were generated using Weighted Gene Co-expression Network Analysis
769 (WGCNA)[57]. For WGCNA, the starting correlation matrix is unsigned, as both negative
770 and positive correlations are of interest in this study. Weight power=5 is added to the
771 correlation matrix to achieve a scale-free adjacency matrix. The topology similarity
772 matrices in the Bd21c, Bd3-1d, and Bd3-1c data sets are scaled to have a similar
773 distribution to Bd21d, which we chose arbitrarily as the reference library type.
774 Dendrograms were generated from topology similarity matrices, and modules were
775 generated by cutting at two parameter settings: (1) moderate size modules, used for
776 most analysis below, high dissimilarity (height= NULL(0.99) with a dynamic hybrid
777 algorithm (minClusterSize =30, deepsplit=2, pamStage = FALSE, mergeheight=0.25);
778 (2) conservative (smaller) modules, used for causal inference analyses discussed
779 below, cutHeight = 0.984, minClusterSize = 15, deepSplit = 3, pamStage = FALSE,
780 mergeheight=0.25. The two different settings are used to better understand the effect of

781 threshold choice within each library type. In brief, our WGCN analyses recovered 15
782 and 24 modules in Bd21c in the two settings, respectively; 14 and 18 modules in Bd21d
783 in the two settings, respectively; 14 and 16 modules in Bd3-1c in two settings,
784 respectively; and 15 and 20 modules in Bd3-1d in the two settings, respectively. Gene
785 membership is summarized in Supplementary Data Set 9.

786 We next addressed metrics of module preservation across our four library types,
787 Bd21c, Bd21d, Bd3-1c, and Bd3-1d. First, we assessed the degree to which the lists of
788 genes in a given module are conserved among our two genotypes and two
789 environmental conditions using a test for membership overlap. We required that module
790 overlap must be significant in all two parameter settings (moderate and conservative,
791 described above, as shown in Figure 4, three modules with unstable overlapping are not
792 shown). To examine the overlap in gene module membership among library types we
793 used the SuperExact test [58] and a p-value cut-off of with $< 10^{-12}$ with the alternative
794 hypothesis that the size intersect between two groups or among more groups is larger
795 than expected by chance (Supplementary Data Set 5). The cut off is chosen to ensure a
796 p-value <0.05 after applying Bonferroni correction of multiple tests. Hypergraphs
797 representing module overlap networks -- here, termed “cliques” for subnetworks -- were
798 visualized using the R package igraph [85,86].

799 Second, we tested for enrichment of transcripts exhibiting significant DGE in our
800 co-expression modules using Fisher’s Exact Tests [58]. Significant enrichments (p-value
801 <0.001 uncorrected by test number) are reported in Supplementary Data Set 6 and
802 summarized in Table 1. Here the enrichment is reported for modules estimated from our
803 conservative parameter settings to be consistent with following network inference
804 analysis.

805 Third, biological roles of each module were inferred using GO term enrichment,
806 by inspecting gene annotations, and by testing for module correlation with the sampled
807 physiological and metabolic traits. Module GO enrichment is conducted on using the
808 PANTHER classification[87] and implemented in rbioapi[88], with Fisher’s exact statistical
809 overrepresentation tests of GO biological process. Only the significantly enriched GO
810 terms at FDR = 0.1 and the lowest-rank GO categories are reported. Here the
811 enrichment is reported if positive in modules of any of the parameter settings

812 (Supplementary Data Set 7). For common GO term comparison among modules in an
813 overlapping clique, all GO categories are considered (summarized in Table 1). Common
814 genes and annotations in overlapping cliques are list in Supplementary Data Set 3-4.
815 Gene annotations are from MapMan (https://mapman.gabipd.org/mapmanstore) based
816 on the Phytozome database. Correlations between module eigengenes (i.e., their first
817 principal component) and traits was performed in each library type. The correlation p-
818 values are reported in Supplementary Data Set 11 and summarized in Table 1 based on
819 moderate criteria modules, uncorrected for the number of tests.
820 Finally, we tested the hypothesis that genes within modules share common DNA
821 motifs corresponding to known transcription factor binding sites. All possible 6-mer
822 sequences were tested for enrichment in the -500 to +200 upstream window regions[89]
823 relative to each gene's predicted transcriptional start site in the Bd21 reference genome,
824 in the module using AME in MEME suite[90]. The AME parameters were: --verbose 1 --
825 scoring avg --method fisher --hit-lo-fraction 0.25 --evalue-report-threshold 10.0, with the
826 control set based on all the genes. The Arabidopsis DAPSeq database [91] was used as a
827 reference to interpret enriched motif sequence identified. Here the enrichment is
828 reported for moderate criteria modules.
829
830 **Causal network inference in modules**
831 Methods such as WGCNA, based on correlations in transcript abundance, cannot
832 provide information about the direction of regulatory interactions. Our high replication
833 provides the power to model genes as random variables and infer conditional
834 independence of gene expression to represent possible causality. Here, we use the
835 Unknown-Targeted Interventional Greedy Sparest Permutation (UT-IGSP) algorithm [92]
836 to infer regulatory causality, representing a group of genes as a directed acyclic graph
837 (DAG). We use all genes in a given WGCNA module – estimated from one library type -
838 - as priors to reduce latent variables in the model, relying on the simplifying assumption
839 that there are no edges among genes between modules. We include both samples that
840 the module is inferred from (here, representing "observational" data) as well as samples
841 from all other library types (here, representing "interventional data"), representing the
842 assumption that most of the gene regulation structure is the same between library

843 types. In these analyses, we use the conservative module criteria (most restrictive
844 inclusion in the WGCNA analyses) and limit the analysis to modules with number of
845 genes fewer than 70% of samples in the specific library type. When testing for modules
846 that are conserved within a treatment or within a genotype, modules with as many as 70
847 genes can be used because more samples are available for inference. We use stability
848 selection [93] as implemented in Belyaeva et al [94] to minimize the effect of
849 hyperparameters. Edges with a selection probability higher than upper 0.95 (or 0.97 for
850 bigger conserved modules to keep sparse edges, and 0.95 for modules with glucose as
851 a node) quantile are kept. Genes not found to be connected to other genes are not
852 shown. DAGs were exported from R using the rgraphviz and visualized using Graphviz
853 Visual Editor [95].
854
855 **Identifying patterns of variation in gene-gene interactions**
856 We extended our analysis of gene expression responses using a regression approach
857 that explicitly models the magnitude and direction of putative interaction between genes
858 based on their transcript abundances. For a putative regulator and target pair identified
859 in our DAGs, we regress the target gene transcript abundance on the regulator
860 transcript abundance, and test whether the slope and intercept coefficients of
861 regression changes as a function of genotype or environmental treatment. When there
862 are multiple inferred regulators for a target gene, a multivariate linear regression is
863 applied.
864 For the hypothesis that gene x_i $(i = 1, \ldots, n)$ regulates gene y, to infer the
865 intercept or slope differences between genotypes, treatment, or their interactions, we
866 modeled their ($\log_2$ -transformed) gene expression x_i and y as:

867 $$y = \gamma_0 + \gamma_g g + \gamma_e e + \gamma_{ge} g * e + \sum_{i=1}^{n} \left(\beta_{0,i} + \beta_{g,i} g + \beta_{e,i} e + \beta_{ge,i} g * e\right) * x_i$$

868 Here g, e are used to differentiate the dataset used to train the parameters (i.e., g, e are
869 "0,0", "0,1", "1,0", "1,1" for Bd21 control, Bd21 drought, Bd3-1 control and Bd3-1 drought
870 respectively). Thus, consistent with the DGE test using DESeq2, above, a significant p-
871 value of γ_g, β_g suggest a genotypic effect on regression intercept and slopes,
872 γ_e, β_e suggest an environmental effect on intercept and slopes, and γ_{ge}, β_{ge} suggest a

873 GxE effect of intercept and slopes. Only genes with regression R^2 greater than >0.5 are
874 considered, as estimated with the Car package in R. [96] P-value of <0.05 is reported to
875 be significant without multiple test correction. Note that if the transcript abundances are
876 standardized before fitting our regression models, the slope differentiation test should
877 closely match the results of CILP tests introduce by Lea et al[67] which we include as
878 Supplementary Figure 6.
879
880 **Allele specific expression**
881 We used allele-specific expression to assess whether GxE in transcript abundances
882 were caused by *cis*- or *trans*-acting variants. Bd21 served as a pollen donor on
883 emasculated Bd3-1 flowers, resulting in ten F1 hybrid seeds. To confirm successful
884 outcrossing in this largely selfing species, we extracted DNA from leaves (DNeasy Plant
885 Pro Kit; Qiagen, MD, US) and tested for a known length polymorphism on a 3% agarose
886 gel. We performed a dry-down experiment on these F1s and extracted leaf RNA, as
887 above. 100ng of RNA for each of 10 samples along with 12 samples (three replicated of
888 each genotype and condition) from the initial experiment were aliquoted for sequencing.
889 For DGE, and to identify SNPs between alleles, we generated libraries using NEB Ultra
890 II Directional RNA with Poly(A) Selection (New England Biolabs, Ipswich, MA). Libraries
891 were sequenced on an Illumina Novaseq generating 34.1 ± 13.7 million raw reads per
892 sample.
893 To generate a reference diploid transcriptome, RNAseq reads of each Bd3-1
894 sample were aligned to the Bd21 reference genome (Bdistachyon_556_v3.2) based on
895 a RNAseq short variant discovery pipeline [97], with STAR alignment [98] of two pass and
896 alignSJoverhangMin=30 and alignSJDBoverhangMin=10, variants calling GATK [99] and
897 the flag of dont-use-soft-clipped-bases. The variants called in each sample were then
898 filtered to find SNPs and in/dels with homozygous alternative genotype mapping. To
899 reduce false discovery only variants identified in more than half of the replicates of each
900 genotype were kept. The synthesized Bd21 or Bd3-1 vcf together with Bd21 genome
901 and annotation gff was then used to generate *de novo* Bd21 or Bd3-1 genome using
902 g2gtools [100], and a diploid genome with emase [101]. The reads from all the samples were
903 mapped to this diploid transcriptome using bowtie[102] and a count matrix was estimated

904 from alignments using emase (an alternate pipeline employing bowtie2 followed by
905 RSEM provides similar results; not shown) to account for the multi-mapping particularly
906 accounting for allelic expression. Isoforms are not differentiated in this study.
907 Genes with different counts between two alleles are considered to have
908 sequence variants but sparse differences in a gene might not accurately differentiate
909 two alleles. Accordingly, we included 3963 genes that have more than >1 count of the
910 allele in one accession and <20% total count of the allele in the opposing accession in
911 all samples, modified from Gould et al [103]. The RNASeq count matrix was normalized
912 using RLE in DESeq2, as above. All genes (>0 in more than 10 samples, including
913 allele-indifferentiable genes) are included in the normalization step with allele-
914 distinguishable genes shown as allelic expression and other genes as total expression
915 to ensure the normalization factors are calculated based on most genes for the
916 comparable of genes with allelic variants. Possible read-mapping bias was assessed as
917 in Lovell et al. [104] shown in Supplementary Figure 7. PCA of samples are shown in
918 Supplementary Figure 8.
919 We employed a pipeline based on DESeq2 as introduced by Lovell et al [104].
920 Briefly, the transcript abundance of a target gene is fit into full model
921 $log_2 q_{ij} = \beta_0 + \beta_T T_i + \beta_A A_i + \beta_G G_i + \beta_{AG} A_i * G_i + \beta_{TG} T_i * G_i + \beta_{AT} T_i * A_i + \beta_{AGT} A_i * T_i * G_i$
922 where q_{ij} is gene expression of sample i for gene j, T_i is treatment (control: 0, drought:
923 1), A_i is allele type (Bd21: 0, Bd3-1: 1), G_i is generation (F1: 0, Parents: 1). Accordingly,
924 β_A tests for *cis*-acting regulation; β_{AG} tests for *trans*-acting; β_{AT} tests for *cis*-by-
925 environment-acting; and β_{AGT} tests for *trans*-by-environment-acting. Here, the approach
926 was also modified to perform tests in each environmental condition (as a "half model").
927 To be specific, in half model, each test of cis/trans is done separately in each condition
928 because of our limited power and possible low expression of the gene in samples of
929 some conditions; cis/trans by environment interaction cannot be directly tested in this
930 analysis. P-values <0.1 is reported to be significant without multiple test correction.
931 Results are summarized in Supplementary Data Set 12.
932
933 **DATA AVAILABILITY**

934 The datasets and computer code produced in this study are available in the following
935 databases:

936 - RNA-Seq data: Gene Expression Omnibusm GSE267880
937 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE267880)
938 - RNA-Seq data for allele specific expression: Gene Expression Omnibusm
939 GSE267878 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE267878)
940 - Analysis scripts: GitHub
941 (https://github.com/jyjiey/GxE_in_gene_regulation_network).

942

943 **SUPPLEMENTARY DATA**
944 **Supplementary Data Set 1.** Gene DGE list
945 **Supplementary Data Set 2.** GxE DGE gene GO enrichment
946 **Supplementary Data Set 3.** Module common genes
947 **Supplementary Data Set 4.** Module common gene annotation sorted
948 **Supplementary Data Set 5.** Module overlap p-value and intersect gene number
949 **Supplementary Data Set 6.** Module G,E,GxE DGE gene enrichment p-value
950 **Supplementary Data Set 7.** Module GO enrichment
951 **Supplementary Data Set 8.** Module traits correlation p-value
952 **Supplementary Data Set 9.** Traits data along dry-down
953 **Supplementary Data Set 10.** Traits data on day6 matching RNA samples
954 **Supplementary Data Set 11.** Gene module membership
955 **Supplementary Data Set 12.** ASE type
956 **Supplementary Figure 1.** Physiological traits during dry-down
957 **Supplementary Figure 2.** Other physiological traits during dry-down in addition to
958 Supplementary Figure 1
959 **Supplementary Figure 3.** Inferred directed interactions within selected modules with
960 gene DGE and edge DGE changes labeled
961 **Supplementary Figure 4.** Examples of regulations between pairs of genes selected in
962 addition to Figure 5
963 **Supplementary Figure 5.** ASE of the genes in Figure 5

964 **Supplementary Figure 6.** Heatmap comparing CILP and correlation variation from
965 regression method
966 **Supplementary Figure 7.** Mean reads of Bd21 allelic count vs. mean reads of Bd3-1
967 allelic count across all samples
968 **Supplementary Figure 8.** PCA of allelic gene expression
969 **Supplementary Table 1.** Percentage of E, G, G+E and GxE DGE change and slope
970 change, and intercept changes among edges identified in small modules with inferred
971 directed interactions
972 **Supplementary Table 2.** Models selected for physiological traits during dry-down
973 reported in Supplementary Figure 1 and 2
974 **Supplementary File 1.** Direct test of DGE change between two genes
975

976 **ACKNOWLEDGEMENTS**

977 This work was supported by a grant from the MIT Jameel Water and Food Systems Lab
978 (JWAFS) and NSF CAREER Award (IOS 2239070) to DLD. JY was supported, in part,
979 by a graduate research fellowship from MIT JWAFS. ACW & AR were supported
980 through the United States Department of Energy contract no. DE-SC0012704 to
981 Brookhaven National Laboratory. We acknowledge the MIT SuperCloud and Lincoln
982 Laboratory Supercomputing Center for providing (HPC, database, consultation)
983 resources that have contributed to the research results reported within this paper. We
984 also acknowledge the computational support of Massachusetts Green High-
985 Performance Computer the Commonwealth Computational Cloud for Data Driven
986 Biology (C3DDB).
987

988 **AUTHOR CONTRIBUTIONS**

989 DLD and JY designed the research; JY, ACB, AR and DLD performed research; JY and
990 DLD design the analytic and computational tools, analyzed data. DLD and JY wrote the
991 paper with contributions from AR and ACB.
992

993 **CONFLICT OF INTEREST**

994 The authors declare that they have no conflict of interest.

995

996 **REFERENCE**

997 1. Hill, M. S., Vande Zande, P. & Wittkopp, P. J. Molecular and evolutionary processes
998 generating variation in gene expression. *Nature Reviews Genetics* 1–13 (2020)
999 doi:10.1038/s41576-020-00304-w.
1000 2. Wittkopp, P. J. & Kalay, G. Cis-regulatory elements: molecular mechanisms and
1001 evolutionary processes underlying divergence. *Nat Rev Genet* **13**, 59–69 (2012).
1002 3. Veltsos, P. & Kelly, J. K. The quantitative genetics of gene expression in Mimulus
1003 guttatus. *PLOS Genetics* **20**, e1011072 (2024).
1004 4. Huang, W., Carbone, M. A., Lyman, R. F., Anholt, R. R. H. & Mackay, T. F. C.
1005 Genotype by environment interaction for gene expression in Drosophila
1006 melanogaster. *Nat Commun* **11**, 5451 (2020).
1007 5. National Academies of Sciences, E. *Next Steps for Functional Genomics:*
1008 *Proceedings of a Workshop*. (2020). doi:10.17226/25780.
1009 6. Rockman, M. V. Reverse engineering the genotype–phenotype map with natural
1010 genetic variation. *Nature* **456**, 738–744 (2008).
1011 7. Des Marais, D. L., Hernandez, K. M. & Juenger, T. E. Genotype-by-Environment
1012 Interaction and Plasticity: Exploring Genomic Responses of Plants to the Abiotic
1013 Environment. *Annual Review of Ecology, Evolution, and Systematics* **44**, 5–29
1014 (2013).
1015 8. Lowry, D. B. *et al.* Expression Quantitative Trait Locus Mapping across Water
1016 Availability Environments Reveals Contrasting Associations with Genomic Features
1017 in Arabidopsis. *The Plant Cell* **25**, 3266–3279 (2013).
1018 9. Lovell, J. T. *et al.* The genomic landscape of molecular responses to natural drought
1019 stress in Panicum hallii. *Nat Commun* **9**, 5213 (2018).
1020 10. Josephs, E. B. *et al.* The Evolutionary Forces Shaping Cis- and Trans-Regulation of
1021 Gene Expression within a Population of Outcrossing Plants. *Molecular Biology and*
1022 *Evolution* **37**, 2386–2393 (2020).
1023 11. Josephs, E. B., Lee, Y. W., Stinchcombe, J. R. & Wright, S. I. Association mapping
1024 reveals the role of purifying selection in the maintenance of genomic variation in

1025 gene expression. *Proceedings of the National Academy of Sciences* **112**, 15390–
1026 15395 (2015).
1027 12. Varoquaux, N. *et al.* Transcriptomic analysis of field-droughted sorghum from
1028 seedling to maturity reveals biotic and metabolic responses. *Proceedings of the*
1029 *National Academy of Sciences* **116**, 27124–27132 (2019).
1030 13. Winkelmüller, T. M. *et al.* Gene expression evolution in pattern-triggered immunity
1031 within Arabidopsis thaliana and across Brassicaceae species. *The Plant Cell* (2021)
1032 doi:10.1093/plcell/koab073.
1033 14. Wu, T.-Y. *et al.* Evolutionarily conserved hierarchical gene regulatory networks for
1034 plant salt stress response. *Nat. Plants* **7**, 787–799 (2021).
1035 15. Reynoso, M. A. *et al.* Evolutionary flexibility in flooding response circuitry in
1036 angiosperms. *Science* **365**, 1291–1295 (2019).
1037 16. Bradshaw, A. D. Evolutionary Significance of Phenotypic Plasticity in Plants. in
1038 *Advances in Genetics* (eds. Caspari, E. W. & Thoday, J. M.) vol. 13 115–155
1039 (Academic Press, 1965).
1040 17. Smith, H. Signal perception, differential expression within multigene families and the
1041 molecular basis of phenotypic plasticity. *Plant, Cell & Environment* **13**, 585–594
1042 (1990).
1043 18. Fox, R. J., Donelson, J. M., Schunter, C., Ravasi, T. & Gaitán-Espitia, J. D. Beyond
1044 buying time: the role of plasticity in phenotypic adaptation to rapid environmental
1045 change. *Philosophical Transactions of the Royal Society B: Biological Sciences* **374**,
1046 20180174 (2019).
1047 19. Xiong, W. *et al.* Increased ranking change in wheat breeding under climate change.
1048 *Nat. Plants* **7**, 1207–1212 (2021).
1049 20. Choquette, N. E. *et al.* Environment-specific selection alters flowering-time plasticity
1050 and results in pervasive pleiotropic responses in maize. *New Phytologist* **238**, 737–
1051 749 (2023).
1052 21. Levis, N. A. & Pfennig, D. W. Evaluating 'Plasticity-First' Evolution in Nature: Key
1053 Criteria and Empirical Approaches. *Trends Ecol Evol* **31**, 563–574 (2016).

1054 22. Rispail, N. *et al.* Multi-Environmental Trials Reveal Genetic Plasticity of Oat
1055 Agronomic Traits Associated With Climate Variable Changes. *Front. Plant Sci.* **9**,
1056 (2018).
1057 23. Des Marais, D. L. & Juenger, T. E. Pleiotropy, plasticity, and the evolution of plant
1058 abiotic stress tolerance. *Annals of the New York Academy of Sciences* **1206**, 56–79
1059 (2010).
1060 24. Remold, S. K. & Lenski, R. E. Pervasive joint influence of epistasis and plasticity on
1061 mutational effects in Escherichia coli. *Nat Genet* **36**, 423–426 (2004).
1062 25. Wei, S. *et al.* A transcriptional regulator that boosts grain yields and shortens the
1063 growth duration of rice. *Science* **377**, eabi8455 (2022).
1064 26. Da Ros, L. *et al.* Multi-omics atlas of combinatorial abiotic stress responses in
1065 wheat. *The Plant Journal* **116**, 1118–1135 (2023).
1066 27. Shu, J. *et al.* Transcriptomic, proteomic, metabolomic, and functional genomic
1067 approaches of Brassica napus L. during salt stress. *PLOS ONE* **17**, e0262587
1068 (2022).
1069 28. Xu, K. *et al.* Proteomic analysis reveals the molecular mechanism underlying the
1070 cold acclimation and freezing tolerance of wheat (Triticum aestivum L.). *Plant Sci*
1071 **318**, 111242 (2022).
1072 29. Amiour, N. *et al.* The use of metabolomics integrated with transcriptomic and
1073 proteomic studies for identifying key steps involved in the control of nitrogen
1074 metabolism in crops such as maize. *Journal of Experimental Botany* **63**, 5017–5033
1075 (2012).
1076 30. Zhang, H., Zhu, J., Gong, Z. & Zhu, J.-K. Abiotic stress responses in plants. *Nat Rev*
1077 *Genet* **23**, 104–119 (2022).
1078 31. Passioura, J. B. Translational research in agriculture. Can we do it better? *Crop*
1079 *Pasture Sci.* **71**, 517–528 (2020).
1080 32. Nicotra, A. B. *et al.* Plant phenotypic plasticity in a changing climate. *Trends Plant*
1081 *Sci* **15**, 684–692 (2010).
1082 33. Des Marais, D. L. *et al.* Physiological Genomics of Response to Soil Drying in
1083 Diverse Arabidopsis Accessions[W][OA]. *Plant Cell* **24**, 893–914 (2012).

1084 34. Shin, J. H. *et al.* Transcriptomic changes due to water deficit define a general
1085 soybean response and accession-specific pathways for drought avoidance. *BMC*
1086 *Plant Biol* **15**, 26 (2015).
1087 35. Liu, X., Li, Y. I. & Pritchard, J. K. Trans Effects on Gene Expression Can Drive
1088 Omnigenic Inheritance. *Cell* **177**, 1022-1034.e6 (2019).
1089 36. Control of Transcription Factor Activity. in *Plant Genes, Genomes and Genetics*
1090 149–160 (John Wiley & Sons, Ltd, 2015). doi:10.1002/9781118539385.ch10.
1091 37. Transcription Factors Interpret cis-Regulatory Information. in *Plant Genes, Genomes*
1092 *and Genetics* 133–148 (John Wiley & Sons, Ltd, 2015).
1093 doi:10.1002/9781118539385.ch9.
1094 38. Making mRNAs – Control of Transcription by RNA Polymerase II. in *Plant Genes,*
1095 *Genomes and Genetics* 121–131 (John Wiley & Sons, Ltd, 2015).
1096 doi:10.1002/9781118539385.ch8.
1097 39. Marand, A. P., Eveland, A. L., Kaufmann, K. & Springer, N. M. cis-Regulatory
1098 Elements in Plant Development, Adaptation, and Evolution. *Annu Rev Plant Biol* **74**,
1099 111–137 (2023).
1100 40. Jones, D. M. & Vandepoele, K. Identification and evolution of gene regulatory
1101 networks: insights from comparative studies in plants. *Current Opinion in Plant*
1102 *Biology* **54**, 42–48 (2020).
1103 41. Alon, U. Network motifs: theory and experimental approaches. *Nat Rev Genet* **8**,
1104 450–461 (2007).
1105 42. Lachowiec, J., Queitsch, C. & Kliebenstein, D. J. Molecular mechanisms governing
1106 differential robustness of development and environmental responses in plants. *Ann*
1107 *Bot* **117**, 795–809 (2016).
1108 43. Thompson, D., Regev, A. & Roy, S. Comparative analysis of gene regulatory
1109 networks: from network reconstruction to evolution. *Annu Rev Cell Dev Biol* **31**,
1110 399–428 (2015).
1111 44. Sun, Y. & Dinneny, J. R. Q&A: How do gene regulatory networks control
1112 environmental responses in plants? *BMC Biology* **16**, 38 (2018).
1113 45. Wagner, G. P., Pavlicev, M. & Cheverud, J. M. The road to modularity. *Nature*
1114 *Reviews Genetics* **8**, 921–931 (2007).

1115 46. Grishkevich, V. *et al.* A genomic bias for genotype–environment interactions in C.
1116 elegans. *Molecular Systems Biology* **8**, 587 (2012).
1117 47. Grishkevich, V. & Yanai, I. The genomic determinants of genotype × environment
1118 interactions in gene expression. *Trends Genet* **29**, 479–487 (2013).
1119 48. Des Marais, D. L., Guerrero, R. F., Lasky, J. R. & Scarpino, S. V. Topological
1120 features of a gene co-expression network predict patterns of natural diversity in
1121 environmental response. *Proceedings of the Royal Society B: Biological Sciences*
1122 **284**, 20170914 (2017).
1123 49. Luscombe, N. M. *et al.* Genomic analysis of regulatory network dynamics reveals
1124 large topological changes. *Nature* **431**, 308–312 (2004).
1125 50. Scholthof, K.-B. G., Irigoyen, S., Catalan, P. & Mandadi, K. K. Brachypodium: A
1126 Monocot Grass Model Genus for Plant Biology. *The Plant Cell* **30**, 1673–1694
1127 (2018).
1128 51. Draper, J. *et al.* Brachypodium distachyon. A New Model System for Functional
1129 Genomics in Grasses. *Plant Physiol* **127**, 1539–1555 (2001).
1130 52. Des Marais, D. L., Lasky, J. R., Verslues, P. E., Chang, T. Z. & Juenger, T. E.
1131 Interactive effects of water limitation and elevated temperature on the physiology,
1132 development and fitness of diverse accessions of Brachypodium distachyon. *New*
1133 *Phytologist* **214**, 132–144 (2017).
1134 53. López-Alvarez, D. *et al.* Environmental niche variation and evolutionary
1135 diversification of the Brachypodium distachyon grass complex species in their native
1136 circum-Mediterranean range. *American Journal of Botany* **102**, 1073–1088 (2015).
1137 54. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic Gene
1138 Expression in a Single Cell. *Science* **297**, 1183–1186 (2002).
1139 55. Cortijo, S., Bhattarai, M., Locke, J. C. W. & Ahnert, S. E. Co-expression Networks
1140 From Gene Expression Variability Between Genetically Identical Seedlings Can
1141 Reveal Novel Regulatory Relationships. *Front. Plant Sci.* **11**, (2020).
1142 56. Livingston, D. P., Hincha, D. K. & Heyer, A. G. Fructan and its relationship to abiotic
1143 stress tolerance in plants. *Cell. Mol. Life Sci.* **66**, 2007–2023 (2009).
1144 57. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation
1145 network analysis. *BMC Bioinformatics* **9**, 559 (2008).

1146 58. Wang, M., Zhao, Y. & Zhang, B. Efficient Test and Visualization of Multi-Set
1147 Intersections. *Sci Rep* **5**, 16923 (2015).
1148 59. Xiong, Y. *et al.* Glucose–TOR signalling reprograms the transcriptome and activates
1149 meristems. *Nature* **496**, 181–186 (2013).
1150 60. Tzvetkova-Chevolleau, T. *et al.* The light stress-induced protein ELIP2 is a regulator
1151 of chlorophyll synthesis in Arabidopsis thaliana. *The Plant Journal* **50**, 795–809
1152 (2007).
1153 61. Global Profiling of Rice and Poplar Transcriptomes Highlights Key Conserved
1154 Circadian-Controlled Pathways and cis-Regulatory Modules. *PLOS ONE* **6**, e16907
1155 (2011).
1156 62. Marchand, G. *et al.* Bridging physiological and evolutionary time-scales in a gene
1157 regulatory network. *New Phytol* **203**, 685–696 (2014).
1158 63. Chevin, L.-M., Leung, C., Le Rouzic, A. & Uller, T. Using phenotypic plasticity to
1159 understand the structure and evolution of the genotype-phenotype map. *Genetica*
1160 **150**, 209–221 (2022).
1161 64. Wagner, G. P. & Zhang, J. The pleiotropic structure of the genotype–phenotype
1162 map: the evolvability of complex organisms. *Nature Reviews Genetics* **12**, 204–213
1163 (2011).
1164 65. Laitinen, R. A. E. & Nikoloski, Z. Genetic basis of plasticity in plants. *J Exp Bot* **70**,
1165 739–745 (2019).
1166 66. Hämälä, T., Gorton, A. J., Moeller, D. A. & Tiffin, P. Pleiotropy facilitates local
1167 adaptation to distant optima in common ragweed (Ambrosia artemisiifolia). *PLOS*
1168 *Genetics* **16**, e1008707 (2020).
1169 67. Lea, A. *et al.* Genetic and environmental perturbations lead to regulatory
1170 decoherence. *eLife* **8**, e40538 (2019).
1171 68. Des Marais, D. L., Razzaque, S., Hernandez, K. M., Garvin, D. F. & Juenger, T. E.
1172 Quantitative trait loci associated with natural diversity in water-use efficiency and
1173 response to soil drying in Brachypodium distachyon. *Plant Science* **251**, 2–11
1174 (2016).

1175 69. Burnett, A. C., Rogers, A., Rees, M. & Osborne, C. P. Carbon source-sink limitations
1176 differ between two species with contrasting growth strategies. *Plant Cell Environ.* **39**,
1177 2460–2472 (2016).
1178 70. Ainsworth, E. A. *et al.* Does elevated atmospheric [CO2] alter diurnal C uptake and
1179 the balance of C and N metabolites in growing and fully expanded soybean leaves?
1180 *Journal of Experimental Botany* **58**, 579–591 (2007).
1181 71. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models
1182 Using lme4. *Journal of Statistical Software* **67**, 1–48 (2015).
1183 72. R Core Team (2022). R: A Language and Environment for Statistical Computing. R
1184 Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
1185 73. Bartoń, K. MuMIn: Multi-Model Inference. (2023).
1186 74. Lenth, R. V. *et al.* emmeans: Estimated Marginal Means, aka Least-Squares Means.
1187 (2024).
1188 75. ggplot2 Based Publication Ready Plots. https://rpkgs.datanovia.com/ggpubr/.
1189 76. ggplot2: Elegant Graphics for Data Analysis | SpringerLink.
1190 https://link.springer.com/book/10.1007/978-3-319-24277-4.
1191 77. Xiong, Y. *et al.* A Comparison of mRNA Sequencing with Random Primed and 3′-
1192 Directed Libraries. *Scientific Reports* **7**, 14626 (2017).
1193 78. Soumillon, M., Cacchiarelli, D., Semrau, S., Oudenaarden, A. van & Mikkelsen, T. S.
1194 Characterization of directed differentiation by high-throughput single-cell RNA-Seq.
1195 *bioRxiv* 003236 (2014) doi:10.1101/003236.
1196 79. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for
1197 Illumina Sequence Data. Bioinformatics, btu170.
1198 80. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing
1199 reads. *EMBnet.journal* **17**, 10–12 (2011).
1200 81. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat*
1201 *Methods* **9**, 357–359 (2012).
1202 82. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data
1203 with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
1204 83. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and
1205 dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

1206 84. Kassambara, A. & Mundt, F. factoextra: Extract and Visualize the Results of
1207 Multivariate Data Analyses. (2020).
1208 85. Csárdi, G. & Nepusz, T. The igraph software package for complex network research.
1209 in (2006).
1210 86. Csárdi, G. *et al.* igraph: Network Analysis and Visualization. (2023).
1211 87. Thomas, P. D. *et al.* PANTHER: A Library of Protein Families and Subfamilies
1212 Indexed by Function. *Genome Res.* **13**, 2129–2141 (2003).
1213 88. Rezwani, M., Pourfathollah, A. A. & Noorbakhsh, F. rbioapi: user-friendly R interface
1214 to biologic web services' API. *Bioinformatics* **38**, 2952–2953 (2022).
1215 89. Ksouri, N. *et al.* Tuning promoter boundaries improves regulatory motif discovery in
1216 nonmodel plants: the peach example. *Plant Physiology* **185**, 1242–1258 (2021).
1217 90. McLeay, R. C. & Bailey, T. L. Motif Enrichment Analysis: a unified framework and an
1218 evaluation on ChIP data. *BMC Bioinformatics* **11**, 165 (2010).
1219 91. O'Malley, R. C. *et al.* Cistrome and Epicistrome Features Shape the Regulatory
1220 DNA Landscape. *Cell* **165**, 1280–1292 (2016).
1221 92. Squires, C., Wang, Y. & Uhler, C. Permutation-Based Causal Structure Learning
1222 with Unknown Intervention Targets. Preprint at
1223 https://doi.org/10.48550/arXiv.1910.09007 (2020).
1224 93. Meinshausen, N. & Bühlmann, P. Stability Selection. *Journal of the Royal Statistical*
1225 *Society Series B: Statistical Methodology* **72**, 417–473 (2010).
1226 94. Belyaeva, A. *et al.* Causal network models of SARS-CoV-2 expression and aging to
1227 identify candidates for drug repurposing. *Nat Commun* **12**, 1024 (2021).
1228 95. Rgraphviz. *Bioconductor* http://bioconductor.org/packages/Rgraphviz/.
1229 96. Fox, J. *et al.* car: Companion to Applied Regression. (2023).
1230 97. RNAseq short variant discovery (SNPs + Indels). *GATK*
1231 https://gatk.broadinstitute.org/hc/en-us/articles/360035531192-RNAseq-short-
1232 variant-discovery-SNPs-Indels- (2023).
1233 98. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15
1234 (2013).
1235 99. Auwera, G. A. V. der & O'Connor, B. D. *Genomics in the Cloud: Using Docker,*
1236 *GATK, and WDL in Terra*. (O'Reilly Media, Inc., 2020).

100. g2gtools documentation. https://churchill-lab.github.io/g2gtools/.

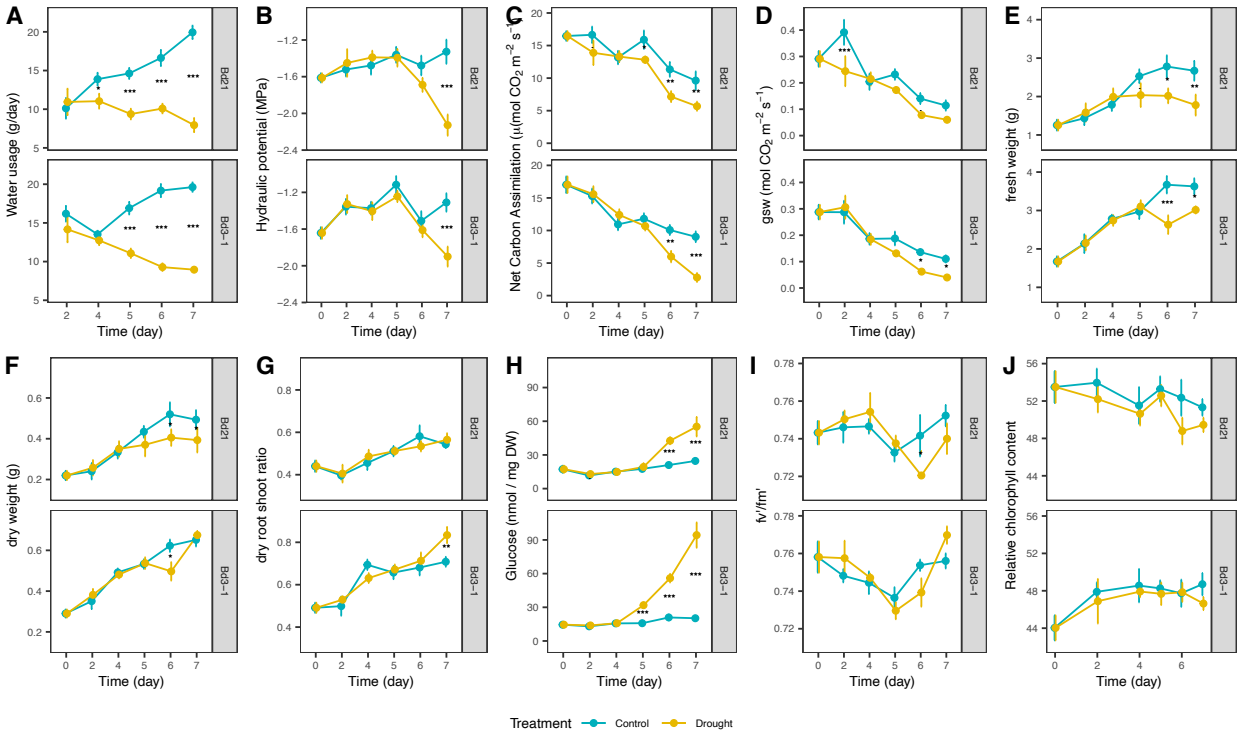
101. Raghupathy, N. *et al.* Hierarchical analysis of RNA-seq reads improves the accuracy of allele-specific expression. *Bioinformatics* **34**, 2177–2184 (2018).

102. 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).

103. 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).

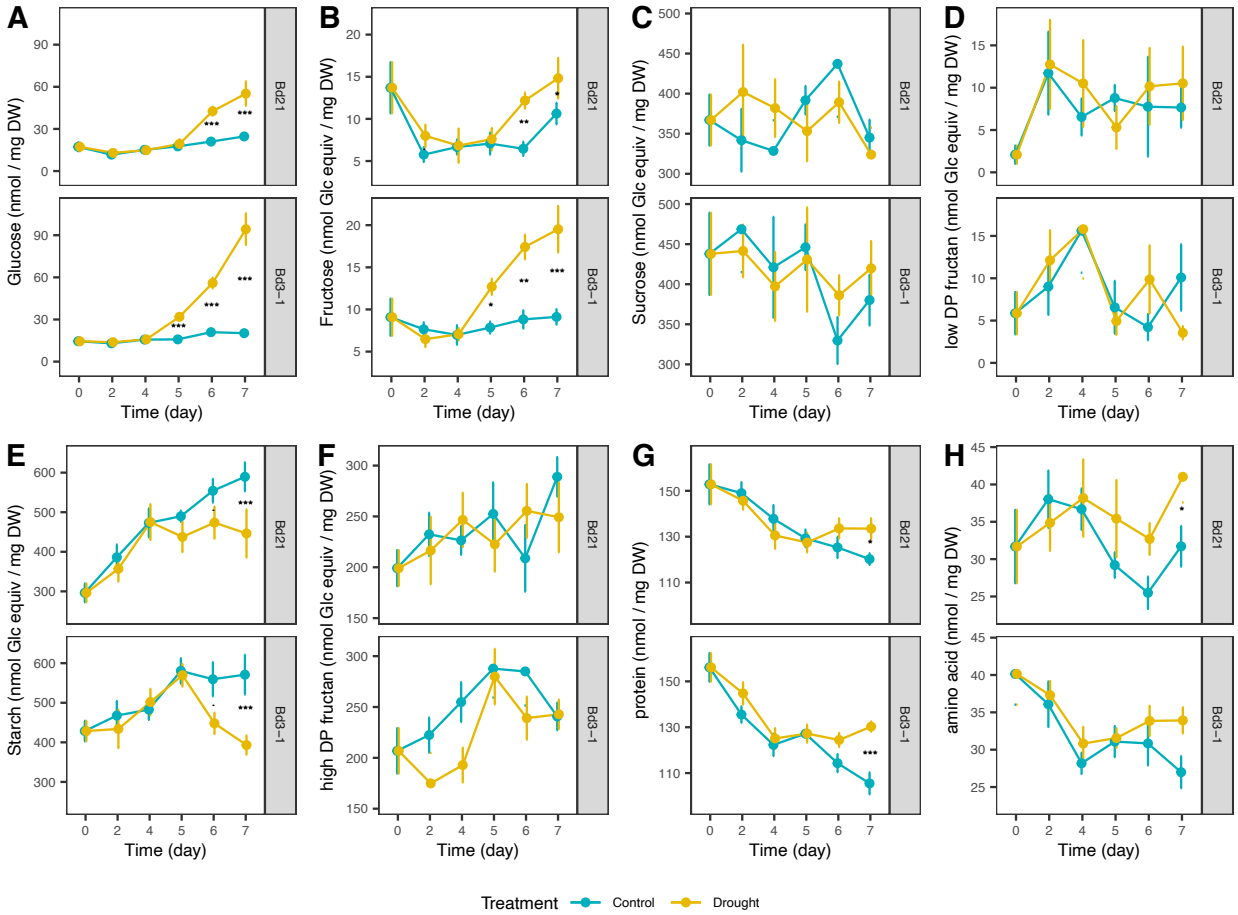
104. Lovell, J. T. *et al.* Drought responsive gene expression regulatory divergence between upland and lowland ecotypes of a perennial C4 grass. *Genome Res* **26**, 510–518 (2016).

SUPPLEMENTARY INFORMATION



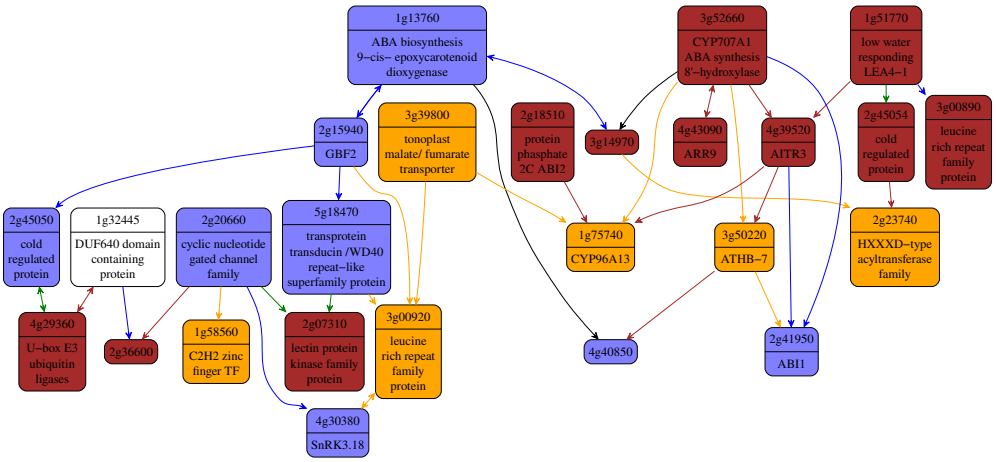
Supplementary Figure 1. Comparison of physiological and metabolic traits between control and dry-down in two accessions from Day 0 to Day 7. (Generalized) Linear

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.

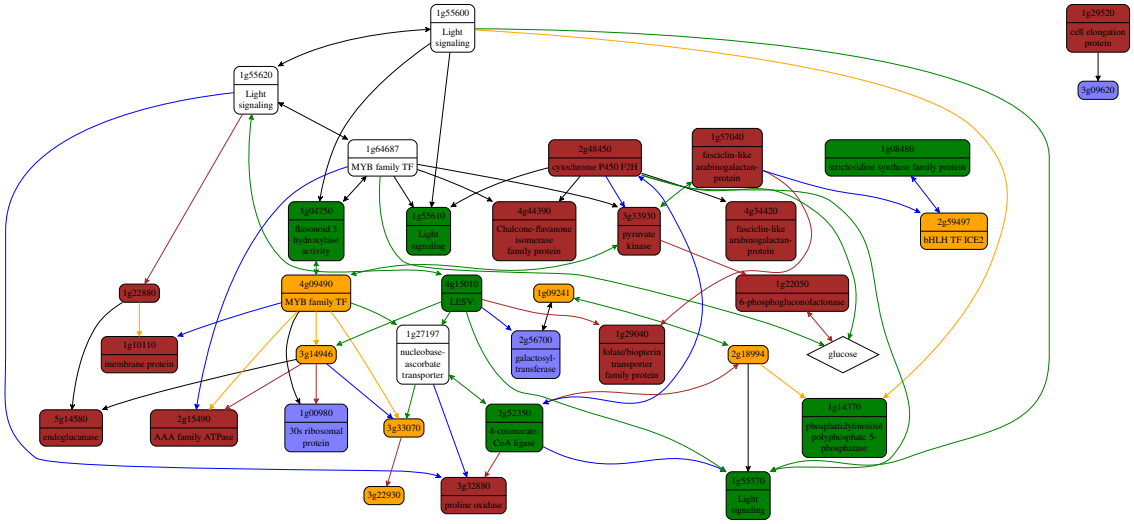


Supplementary Figure 2. Other physiological traits during dry-down in addition to Supplementary Figure 1. Comparison of physiological and metabolic traits between control and dry-down in two accessions from Day 0 to Day 7. (Generalized) Linear models were determined based on AIC score. Random-effects 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.

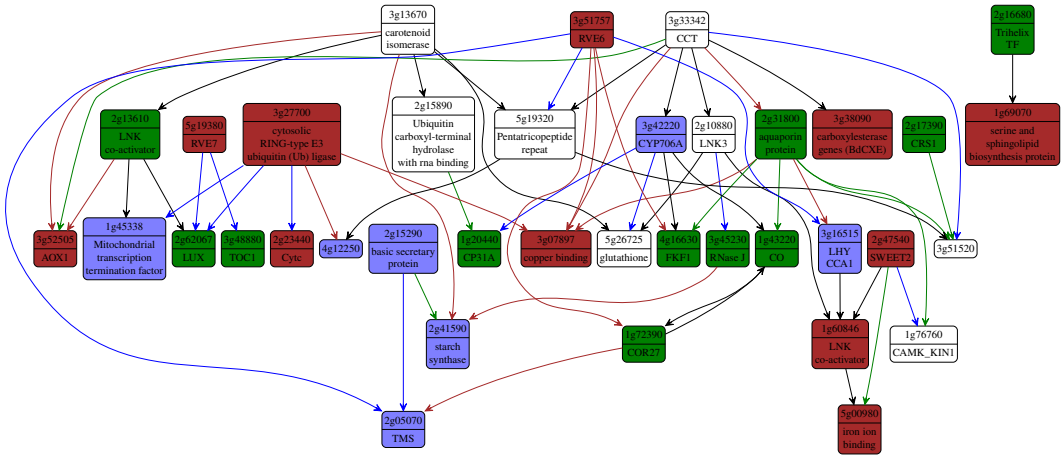
A


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

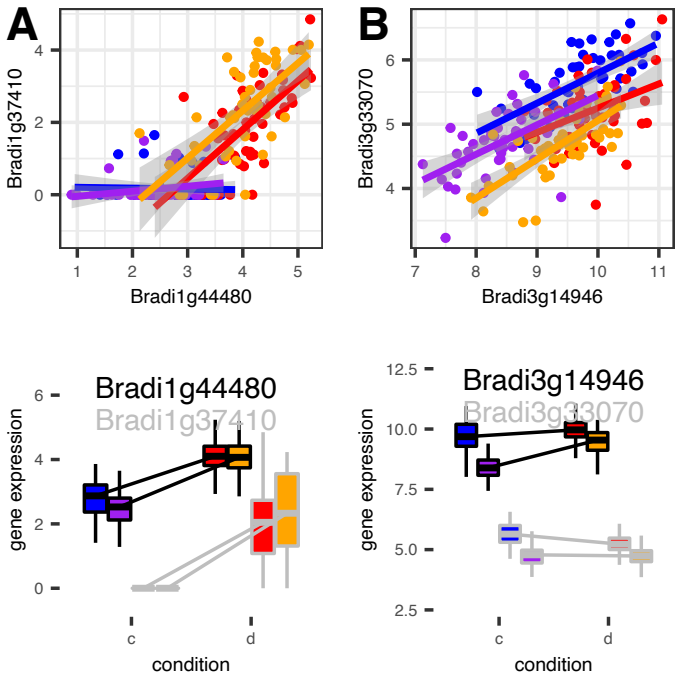
B


1g55600
Light signaling
1g29520
cell elongation protein
1g55620
Light signaling
3g09620
1g64687
MYB family TF
2g48450
cytochrome P450 F2H
1g57040
fasciclin-like arabinogalactan-protein
1g08480
strictosidine synthase family protein
3g04750
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
3g15610
LESV
1g09241
1g22050
6-phosphogluconolactonase
1g10110
membrane protein
1g27197
nucleobase-ascorbate transporter
2g56700
galactosyl-transferase
1g29040
folate/biopterin transporter family protein
2g18994
glucose
3g14946
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
1g35570
Light signaling

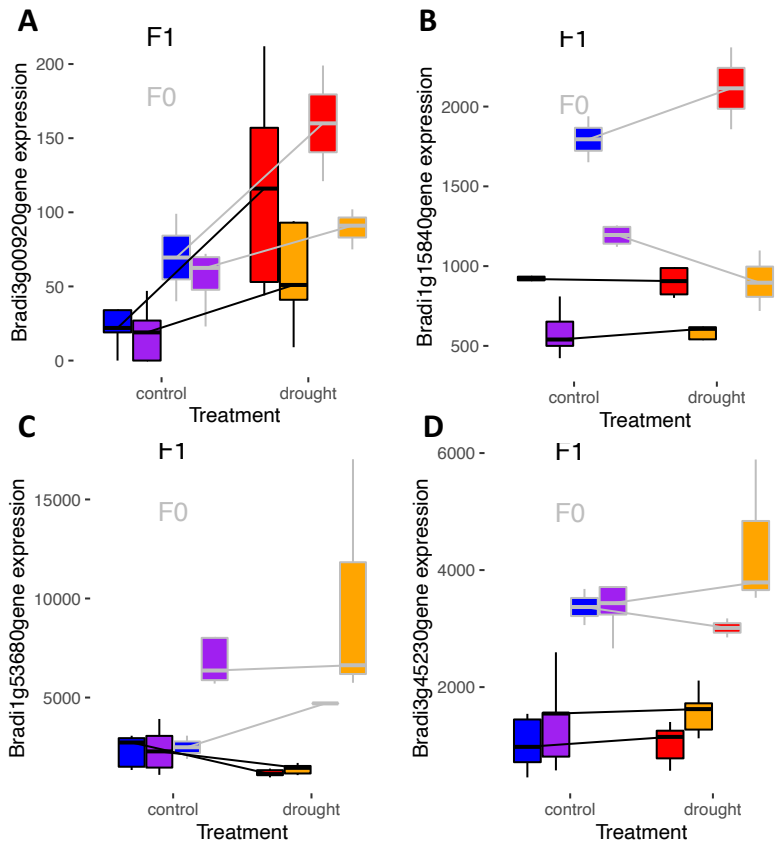
C


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
5g48880
TOC1
2g23440
Cytc
4g12250
2g15290
basic secretary protein
1g20440
CP31A
3g07897
copper binding
5g26725
glutathione
4g16630
FKF1
3g45230
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

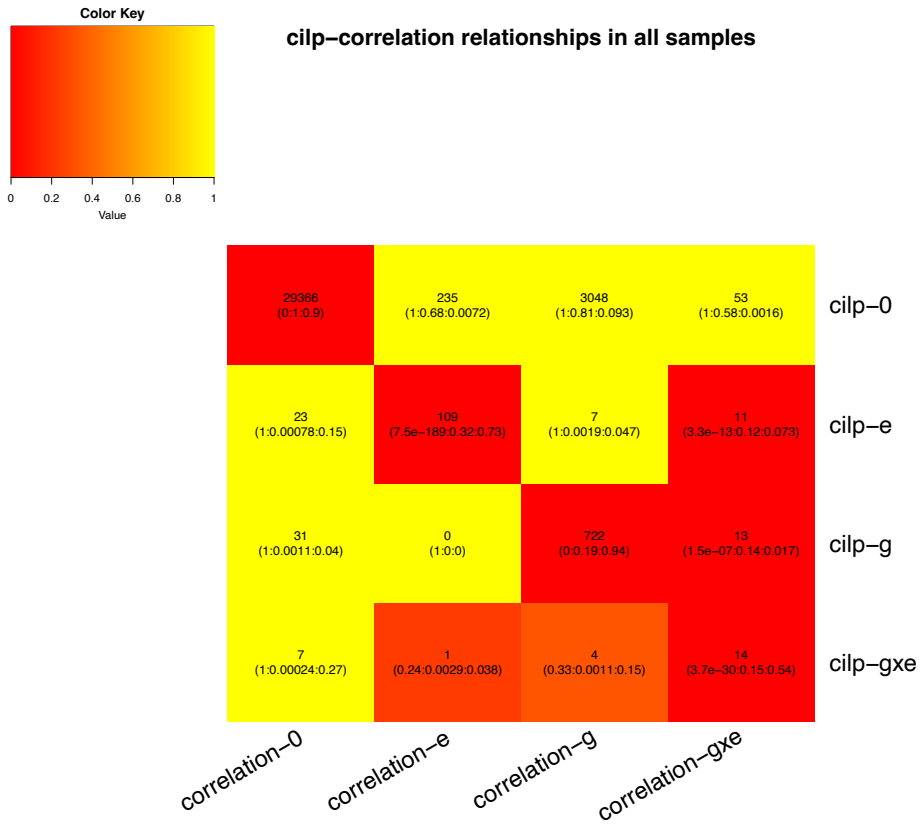
Supplementary Figure 3. Inferred directed interactions within selected modules with gene DGE and edge DGE changes, corresponding to **A)** Module 12 in Bd21d, **B)** Module 14 in Bd3-1d and **C)** Module 10 in Bd21d shown in Figure 5, with node colors are based on gene DGE and edge colors are based on edge DGE changes (blue: E, green: G, brown: G+E, orange: GxE, black/white: none).



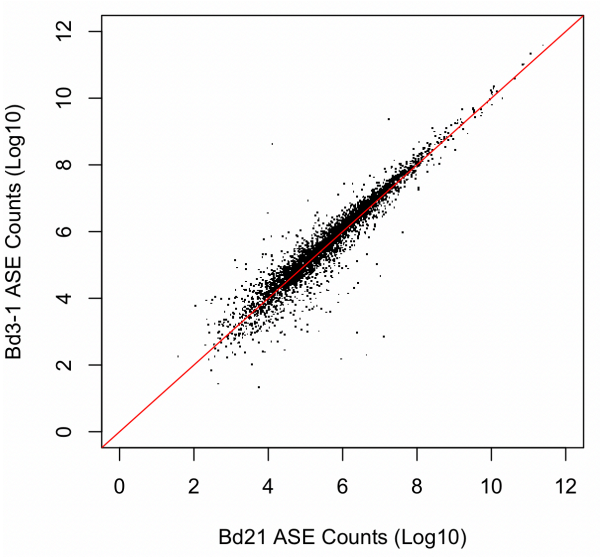
Supplementary Figure 4. Examples of regulations between pairs of genes selected in addition to Fig.6 **A)** shows regulation E slope changes, and **B)** shows transmission of GxE DGE. They are illustrated with linear regression between regulator and target, their expression levels based on log2(norm_count+1). Bd21c, Bd21d, Bd3-1c, Bd3-1d are colored with blue, red, purple, orange, respectively.



Supplementary Figure 5. 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.



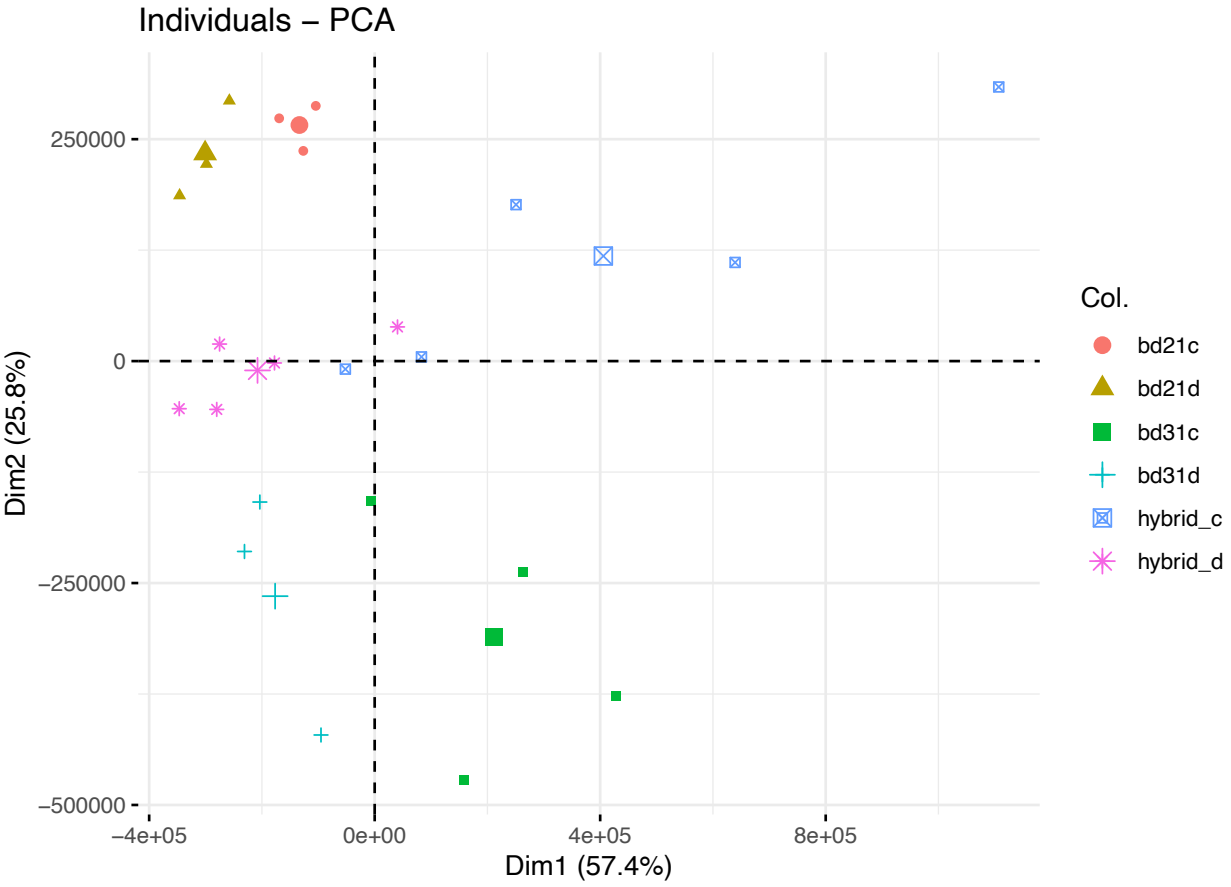
Supplementary Figure 6. Heatmap comparing CILP and correlation variation from regression method based on gene pairs with correlation >0.75 in any of the library types. The number represent the # of the cases found in the intersect. The three number in the parenthesis are p-value of fisher's exact test, 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.



1302

1303 Supplementary Figure 7. mean reads of Bd21 allelic count vs. mean reads of Bd3-1
1304 allelic count across all samples of allele specific genes suggest no alignment bias to
1305 Bd21 or Bd3-1

1306



1307

1308 Supplementary Figure 8. PCA of allelic gene counts colored based on samples
1309 genotype and treatment (after normalization)
1310
1311 Supplementary Table 1. Percentage of E, G, G+E and GxE DGE change, and slope
1312 change identified in edges or intercept change in genes in small module (<40 genes)
1313 inferred gene regulation network (e.g., Figure 4).

		edge DGE change				edge rewiring %					gene intercept change %				
module no.	total edges	G	E	G+E	GxE	G	E	G+E	GxE	total genes	E	G	GxE	G+E	Non
bd21c_13	67	25	4	10	6	10%	12%	3%	9%	33	3%	24%	24%	6%	39%
bd21c_14	52	19	8	5	0	8%	17%	0%	6%	32	6%	25%	13%	9%	28%
bd21c_17	37	16	4	6	0	5%	5%	3%	5%	27	11%	11%	15%	19%	33%
bd21c_21	30	8	4	12	0	10%	20%	3%	7%	19	5%	5%	21%	21%	37%
bd21c_22 (part of 9)	22	5	3	3	0	5%	23%	9%	5%	18	6%	22%	17%	0%	28%
bd21c_23 (part of 10)	25	5	3	10	0	4%	8%	4%	0%	18	6%	6%	17%	22%	33%
bd21d_12	41	5	11	12	11	34%	15%	5%	12%	29	3%	17%	17%	7%	31%
bd21d_13	32	2	11	7	5	44%	3%	6%	6%	25	0%	4%	32%	8%	16%
bd21d_16 (part of 5)	23	6	9	6	1	17%	4%	4%	17%	20	5%	15%	5%	25%	30%
bd31d_12	68	15	16	11	6	10%	9%	0%	1%	37	8%	8%	27%	24%	11%
bd31d_14	51	17	3	8	0	10%	4%	4%	4%	28	0%	25%	14%	21%	25%
bd31d_18 (part of 21)	32	4	10	10	2	13%	19%	9%	9%	19	11%	5%	16%	21%	37%
bd31d_19	108	44	15	18	4	7%	24%	5%	6%	33	6%	12%	15%	21%	39%
bd31c_13	79	8	4	0	0	8%	5%	3%	13%	31	19%	10%	10%	10%	48%
bd21c_5 (intersect with Bd21d_2)	65	25	9	8	0	5%	18%	2%	9%	36	14%	8%	8%	17%	39%
bd31c_10	55	11	19	10	0	7%	9%	2%	11%	33	12%	15%	21%	27%	21%
bd31c_16	27	13	1	3	0	15%	19%	0%	15%	15	7%	20%	7%	27%	33%

1314

1315 Supplementary Table 2. Models used for physiological traits during dry-down reported in
1316 both Supplementary Figure 1 and 2. (Generalized) Linear models were determined
1317 based on AIC score. Random-effects components of the models were based on
1318 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

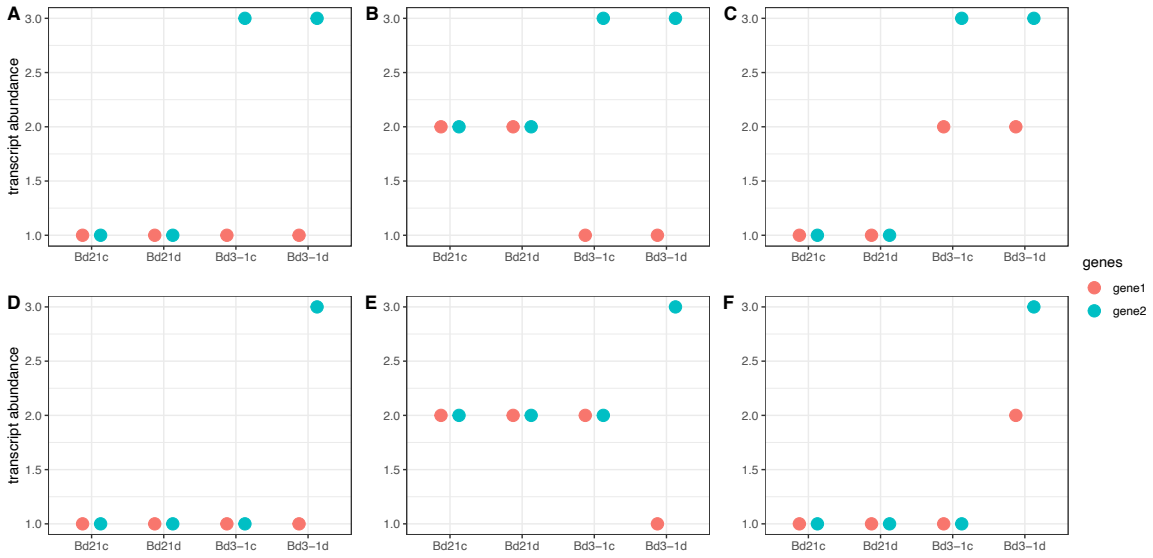
Supplementary File 1 Direct test of DGE change between two genes

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 2. This method avoids the indirect comparison by DGEs of two genes, but generally consistent with it (see Supplementary Figure 3).



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.