Integrating transcriptomic network reconstruction and QTL analyses reveals mechanistic connections between genomic architecture and *Brassica rapa* development

Robert L. Baker[*†]
Wen Fung Leong[‡]
Marcus T. Brock[§]
Matthew J. Rubin[§]
R. J. Cody Markelz[**]
Stephen Welch[‡]
Julin N. Maloof[**]
Cynthia Weinig[§]

[*]Corresponding author

[†] Department of Biology, Miami University, Oxford OH 45056

[‡] Department of Agronomy, Kansas State University, Manhattan KS, 66506

[§] Department of Botany, University of Wyoming, Laramie, WY 82071

[**] Department of Plant Biology, University of California Davis, Davis CA 95616

Running Title:

Integrating FVT, QTLs, and eQTLs

Keywords:

Brassica rapa, Function-Valued Trait, QTL mapping, Mutual Rank, Weighted Gene Co-expression Network Analysis, eQTL Mapping, developmental genetics

Corresponding Author:

Mailing address:

Robert L. Baker

167G Pearson Hall

Department of Biology

Miami University

700 E. High St.

Oxford, OH 45056

Telephone:

1.513.529.3175

Email:

robert.baker@miamioh.edu

ABSTRACT

Plant developmental dynamics can be heritable, genetically correlated with fitness and yield, and undergo selection. Therefore, characterizing the mechanistic connections between the genetic architecture governing plant development and the resulting ontogenetic dynamics of plants in field settings is critically important for agricultural production and evolutionary ecology. We use a hierarchical Bayesian Function-Valued Trait (FVT) approach to estimate *Brassica rapa* growth curves throughout ontogeny, across two treatments and in two growing seasons. We find that the shape of growth curves is relatively plastic across environments compared to final height, and that there are trade-offs between growth rate and duration. We determined that combining FVT Quantitative Trait Loci (QTL) and genes/eigengene expression identified via transcriptomic co-expression network reconstructions best characterized phenotypic variation. Further, targeted eQTL analyses identified regulatory hotspots that colocalized with FVT QTL and co-expression network identified genes and mechanistically link FVT QTL with structural trait variation throughout development in agroecologically relevant field settings.

INTRODUCTION

Plant developmental genetics are correlated with fitness and yield (Baker *et al.* 2015; Kulbaba *et al.* 2017). Therefore, characterizing the mechanistic connections between the genetic architecture governing plant development and the resulting ontogenetic dynamics of plants in field settings is critically important to improving agricultural production and understanding evolutionary performance. Forward genetic approaches such as quantitative trait mapping are an attractive method of characterizing genetic architecture because they do not require *a priori* information such as candidate loci and can be used to describe pleiotropic and epistatic loci as well as polygenic traits (Prioul *et al.* 1997; Mackay 2013; Csilléry *et al.* 2018). Transcriptomic co-expression analyses and expression QTL (eQTL) have also been used to identify the underlying genetic architecture responsible for phenotypic variation (e.g. Nozue *et al.* 2018). Recently, combining information from genomic association studies and transcriptomic expression analyses has been used to pinpoint candidate genes (Hitzemann *et*

al. 2003; Li *et al.* 2018; Luo *et al.* 2018; Schaefer *et al.* 2018). However, co-expression network analyses can also provide insight into the mechanistic connections between QTL genotypes and phenotypes. Here, we ask whether QTL, co-expression analyses, or a combination thereof best predict phenotypic variation. In combination with a targeted eQTL analyses in agroecologically relevant field settings, we characterize the mechanistic connections between the genomic architecture, transcriptomic expression networks, and phenotypic variation throughout plant development.

Development rarely occurs in discrete steps, yet developmental data are typically collected at multiple distinct but inter-dependent time points. Function-Valued Trait (FVT) modeling is one method of estimating the underlying continuous nature of development and avoiding complicated repeated measures analyses, which often compromise statistical power in downstream analyses (Wu *et al.* 1999; Griswold *et al.* 2008). One approach to FVT modeling involves fitting mathematical functions to discrete data to estimate continuous curves that represent the change of a trait or character as a function, typically of time (Kingsolver *et al.* 2001; Wu and Lin 2006; Stinchcombe and Kirkpatrick 2012). Although there are multiple approaches to modeling continuous growth, one particular advantage of FVT modeling is that parameters describing developmental growth curves can be extracted from the FVT models and used as biologically interpretable and inter-relatable traits such as the relationship between growth rates, durations, inflection points, and final sizes. This 'parameters as data' approach enables a broad array of analyses at both genetic and phenotypic levels (Hernandez 2015; Kulbaba *et al.* 2017). In the current study, we employ a Bayesian hierarchical approach to FVT modeling that leverages global information from the entire dataset as well as each genotype to estimate replicate-level parameters describing growth curves that underlie the developmental dynamics of plant height.

One inherent but seldom addressed complication in studying developmental genetics is that development of a given trait rarely occurs independently of organism-level attributes. For instance, in plants carbon availability can severely limit and alter development, even in determinate structures such as leaves (Schneidereit *et al.* 2005; Raines and Paul 2006). Further, including physiological parameters in plant breeding models is predicted to accelerate and

improve yield gains (Hammer *et al.* 2005). One solution is using a hierarchical Bayesian approach to FVT modeling that incorporates genotype-specific values for physiological conditions such as carbon availability (for instance, estimated using A_{max}) to statistically factor out variation caused by resource availability. Accounting for carbon availability in FVT parameter estimation can increase estimates of heritability and improve QTL mapping results (Baker 2018a, b).

QTL mapping provides a well-tested method of uncovering the genetic architecture of Function-Valued Traits (FVT). FVT variation may arise from structural or regulatory genes that differ among sampled genotypes. Examining gene expression can therefore provide insight into the mechanistic connections between genomic architecture and developmental dynamics of phenotypes (Schmid *et al.* 2005; Li *et al.* 2010; Jiang *et al.* 2015; Zhu *et al.* 2016). We use Mutual Rank (MR) and Weighted Gene Co-expression Network Analyses (WGCNA) to identify expression networks associated with FVT trait variation. These networks are then used to focus our analysis to specific expression traits for eQTL mapping (Munkvold *et al.* 2013; Ponsuksili *et al.* 2015). Interestingly, the genomic architecture of eQTL appears to depart from that of other phenotypic QTL such as FVT QTL in two important respects: first, gene expression traits tend to have only one or a few eQTL whereas morphological phenotypic traits are often highly polygenic (Gibson and Weir 2005). Second, eQTL from multiple expression traits in diverse taxa from yeast to *Brassica* can be highly colocalized into eQTL "hotspots". These hotspots may indicate a regulatory gene or switch that has a disproportionate impact on downstream gene expression (Schadt *et al.* 2003; West *et al.* 2007; Hammond *et al.* 2011). In contrast, QTL for morphological traits may colocalize, but typically they do not do so to the same extent (Schadt *et al.* 2003; Tian *et al.* 2016). Whether general eQTL trends hold for targeted expression traits in agroecologically relevant field settings remains unknown. Further, to the best of our knowledge eQTL mapping has not been used to examine the mechanistic basis of developmental morphology captured via function-valued trait modeling.

Here, we estimate continuous developmental growth curves of plant height, a trait that when selected upon can lead to more effective increases in yield than directly selecting on yield itself (Law *et al.* 1978), in a set of *Brassica rapa* Recombinant Inbred Lines (RILs) while

mathematically factoring out the effects of carbon availability. We examine the patterns of genetic correlations among parameters describing change in height over time such as growth duration and final plant size, and we ask whether these developmental parameters correlate with yields. Using QTL mapping, we outline the genetic architecture of plant height development. Next, we use MR and WGCNA to identify genes and gene network module eigengenes whose expression patterns correlate with FVT parameters. We compare the predictive capacity of QTL and co-expression approaches in two ways: first, we test the relative effectiveness of QTL vs. MR genes vs. WGCNA module eigengenes (and combinations thereof) in explaining genetic variation of developmental traits. Second, we test whether QTL for FVT traits are enriched for genes identified via co-expression approaches. To explore the mechanistic basis of FVT QTL, we perform eQTL mapping on our MR genes and WGCNA module eigengenes. For eQTL and FVT QTL that colocalize, we explore the relative proportion *cis*- vs. *trans*-eQTL and their effect sizes. We ask whether eQTL colocalize to regulatory hotspots and if so how these compare to FVT QTL. Our eQTL analysis offers an additional line of inference for candidate gene identification as well as a potential mechanistic explanation for the regulation of yield-related FVT QTL.

MATERIALS AND METHODS

Species description

Brassica rapa (Brasssicaceae) is an herbaceous crop species first domesticated in Eurasia. This study was conducted on Recombinant Inbred Lines (RILs) derived from crossing R500, a yellow sarson oil seed variety, with IMB211, which is a rapid cycling line derived from the Wisconsin Fast Plant line (WFP). All RILs are expected to be >99% homozygous (Kokichi and Shyam 1984; Brock and Weinig 2007; Iniguez-Luy *et al.* 2009; Markelz *et al.* 2017). In comparison with IMB211, R500 flowers later, attains a larger size and greater biomass, and allocates more resources to seed production. This experiment includes 120 RILs as well as R500 and representative IMB211 genotypes.

Experimental Design and Data Collection

In 2011, and 2012, the IMB211 × R500 RILs were germinated in the University of Wyoming greenhouse in fertilized field soil, and transplanted into the field at two planting densities, as previously described (Baker *et al.* 2015). Briefly, crowded (CR) plants consisted of 5 plants of the same genotype per 4" peat pot with the central plant designated as a focal individual. The uncrowded (UN) treatment consisted of a single plant per pot. When the cotyledons were expanded, plants were transplanted to the field into randomly located blocks that consisted of either UN or CR plants. Each block contained a full RIL set (and representatives of the RIL parental genotypes), and RIL locations were randomized within blocks with 25cm between each focal plant. For phenotypic data collection 6 UN blocks were transplanted into the field in 2011 and in 2012 8 CR and 8 UN blocks were transplanted. In 2011, an additional 5 UN blocks were transplanted into the field for RNAseq. Plants were watered daily to field capacity and treated with pesticides as needed following Baker *et al.* (2015). Each year, we collected data on the timing of germination, bolting, and flowering by surveying plants 5-7×/week. We recorded temperature data every 5s in the greenhouse and field using a series of Onset® Hobo data loggers (Bourne, MA, USA) and a Campbell Scientific (Logan, UT, USA) CR23X data logger equipped with a Vaisala (Helsinki, Finland) HMP-50 sensor. Temperature data were used to produce hourly and daily means, as well as hourly and daily minimums and maximums, for Degree Day (DD) calculations, which used a *B. rapa*-specific base value of 0.96°C (Vigil *et al.* 1997).

Morphological data. Plant height was recorded for all plants starting at leaf emergence. In 2011, height was measured 6 times during the growing season, and these measurements captured final heights. In 2012, height was measured 2-3 times per week until senescence. Perhaps because of the increased precision in 2012 trait estimates, RNAseq data corresponds more closely to 2012 plant-level phenotypic data compared to 2011, and we focus on 2012 plant-level phenotypic data. Full analyses of FVT traits and QTL including 2011 data can be found in the supplemental materials. Flowering phenology and performance were estimated based on 2012 fruit and seed numbers, as described in Baker *et al.* (2015).

208 **Function-Valued Trait (FVT) modeling and data analysis.** Height data were visually inspected
209 for erroneous data points on a replicate level following Baker et al (2015). FVT modeling for
210 trait estimation used Bayesian approaches that fit logistic growth curves to longitudinal height
211 data (Eqn 1; adapted from Baker *et al.* 2018a). Height for each individual replicate plant is
212 represented by a minimum of 5 and maximum of 13 sequential measurements. Briefly, we
213 utilized a three-level hierarchical Bayesian model that retains the measurement data
214 structure to account for information across all plants and genetic lines within the
215 population, including replicate plants within each line.

216 $$\frac{d}{dt}H = rH\left(\frac{H_{Hmax}-H}{H_{Hmax}}\right) \qquad \text{(Eq. 1)}$$

217 Replicate-level parameters were extracted from the fitted logistic growth curves and treated as
218 trait data (Jaffrézic and Pletcher 2000; Kingsolver *et al.* 2001, Wu and Lin 2006; Stinchcombe *et*
219 *al.* 2010; Baker *et al.* 2018a). These parameters include the growth rate (*r*, cm/DD), and an
220 estimate of the maximum height based on the asymptote of the logistic growth curve (*Hmax*, in
221 cm). Additional parameters were algebraically extracted from the growth curve and include the
222 duration of growth (*d*, in DD) and the inflection point of the growth curve in Degree Days (*iD*, in
223 DD). The parameter *d* was defined as the time in DD when 95% of the final size (*Hmax*) was
224 achieved. The parameter *iD* reflects the transition from exponentially accelerating to
225 decelerating growth rates.

226 The hierarchical Bayesian model was implemented using PyMC, a Bayesian Statistical
227 Modeling Python module. The model parameters were estimated via MCMC using the
228 Metropolis-Hastings algorithm (Chib and Greenberg 1995; Patil *et al.* 2010). The MCMC
229 estimations were performed using a single chain to sample 500,000 iterations, which includes
230 the first discarded 440,000 burn-in iterations; the remaining 60,000 iterations were retained. By
231 thinning to 1 iteration in 20, the retained iterations were reduced to 3,000 samples for every
232 FVT parameter from which the posterior distributions were tabulated. All parameters' trace and
233 auto-correlation plots were examined to ensure that the MCMC chain had adequate mixing and
234 had reached convergence. All observed data for each genotype were plotted with two 95%
235 credible interval envelopes. The inner, yellow envelope represents the credible intervals for the

model based on the observed data, and the green envelope is the 95% credible interval where future observations from the same environment are expected (Fig. 1; Kruschke 2014; Baker *et al.* 2018b).

Phenotypic plasticity: To detect environmental factors that might affect the correspondence between genotype and phenotype, we analyzed replicate level phenotypic datasets from 2012. We tested for the main effects of genotype and treatment and all possible interactions using the *lme4* and *pbkrtest* packages in the R statistical environment (Halekoh and Højsgaard 2014; R Core Team 2016; Bates *et al.* 2018). In these tests, all effects were considered random and block was nested within the treatment effect. Significant main effects of environment (treatment) were considered evidence of phenotypic plasticity, and interactions of treatment × genotype was considered evidence for genetic variation in phenotypic plasticity.

***Best Linear Unbiased Predictions (BLUPs)*:** BLUPs were calculated independently for UN and CR treatments in R using the *lmer* function in the *lme4* package while controlling for block effects (Bates *et al.* 2018; Kuznetsova *et al.* 2018). Broad sense heritability (H^2) was calculated as the genotypic variance divided by the sum of genotypic, block, and residual variances.

***Genetic Correlations*:** We assessed the genetic correlations among height FVT and previously published phenology and fitness traits (Baker *et al.* 2015) across environments using Pearson's correlations of trait BLUPs. Bonferroni corrections for multiple testing were applied to all genetic correlations.

***QTL mapping*:** QTL analyses were performed in R/qtl (Broman *et al.* 2003) based on a map with 1451 SNPs having an average distance of 0.7 cM between informative markers (Markelz *et al.* 2017). The *scanone* function was used to perform interval mapping (1cM resolution with estimated genotyping errors of 0.001 using Haley Knott regression) to identify additive QTL. QTL model space was searched using an iterative process (*fitqtl*, *refineqtl*, and *addqtl)* to identify additional QTL while taking into account the effects of QTL identified by *scanone* and

addqtl. All significance thresholds (0.95) were obtained using 10,000 *scanone* permutations (Broman *et al.* 2003; Broman and Sen 2009). QTL and their 1.5LOD confidence intervals are displayed using MapChart2.0 (Voorrips 2002). Percent variance explained (PVE) is calculated as PVE=100 × (1 - 10^(-2 LOD/ n)). We compared QTL peaks to the *B. rapa* genome (Version 1.5; Cheng *et al.* 2011) to identify positional candidate genes underlying each QTL. A similar approach was used for mapping eigengene QTL (see below). However, the R/qtl implementation of composite interval mapping (Broman and Sen 2009) was used.

RNAseq. We used the RNA sequencing data previously reported in Markelz et al (2017). Briefly, in 2011 five UN blocks of plants designated for destructive sampling were transplanted into the field and allowed to establish for three weeks. Apical meristem tissue, consisting of the upper 1cm of the bolting inflorescence, was collected from three individual replicate plants per RIL and immediately flash frozen on liquid nitrogen as described in Markelz et al (2017). RNA library preparation and sequencing were performed as previously described (Kumar *et al.* 2012; Markelz *et al.* 2017). Reads were mapped to the *B. rapa* CDS reference described in Devisetty *et al.* (2014) using BWA (Li and Durbin 2009), with an average of 6.52 Million mapped reads per replicate. Read counts were imported to R (R Core Team 2016) and filtered to retain genes where more than 2 counts per million were observed in at least 44 RILs. Libraries were normalized using the trimmed mean of M-values (TMM) method (Robinson and Oshlack 2010) and a variance stabilizing transformation was done using voom (Law *et al.* 2014).

Genetic network reconstruction. To reconstruct gene co-expression networks, the fitted gene expression values for each RIL from the limma-voom fit (expression ~ RIL) were used and filtered to keep the top 10,000 genes most variable between RILs.

For each sample type, two network reconstruction methods were used. First, mutual correlation rank (MR) networks (Obayashi and Kinoshita 2009) were constructed. Pairwise MRs were calculated between each of the 10,000 genes and also between each gene and the BLUP parameter estimates from the 2011 and 2012 FVT models. A series of increasingly large growth-related networks were defined using genes directly connected to the FVT parameters with MR

thresholds of ≤ 10, 20, 30, and 50. Multiple different phenotypes were used to jointly seed each network, therefore networks may contain more nodes (and more genes) than the thresholds suggest. However, because some gene expression levels are uniquely correlated with specific phenotypes while others may be correlated with multiple phenotypes, the number of nodes is less than the product of the threshold value and number of phenotypes used to seed the network. Permutation analysis was used to test the network size expected by random chance at each threshold; 95 or more of 100 permutation networks had zero edges connecting FVT BLUPs and gene expression, showing that our MR networks are recovering statistically significant connections. We used the *blastn* algorithm (Altschul *et al.* 1990) with the discontiguous megablast option and an E-value cutoff of 0.001 to compare *B. rapa* genes to *Arabidopsis thaliana* genes (TAIR10 annotation; ftp://ftp.arabidopsis.org/home/tair/Sequences/blast_datasets/TAIR10_blastsets/TAIR10_cds_20101214_updated).

Second, we constructed networks using a Weighted Gene Correlation Network Analysis (WGCNA; Zhang and Horvath 2005; Langfelder and Horvath 2008). For these networks a soft threshold power of 3 was used, corresponding to the lowest power that had a correlation coefficient > 0.9 with a scale-free network topology. We used the "signed hybrid" network, which only connects genes with positive correlation coefficients. This network consisted of 50 modules with a median of 91 genes per module. The eigengene expression value of each module was determined using WGCNA functions. The Pearson correlation between each module's eigengene expression value and each FVT BLUP was calculated to identify modules potentially related to FVTs. Modules were considered significantly associated with a FVT BLUP if the multiple-testing corrected p-value (method = "holm" in R function *p.adjust*) for the correlation test was less than 0.05. Gene Ontology (GO) category enrichment was performed on each significant module; we only examined the Biological Process (BP) and Cellular Compartment (CC) categories. Categories were considered significantly enriched if the false discovery rate adjusted p-value was < 0.05.

Comparing approaches for genetic architecture. We compared the effectiveness of QTL, MR, and WGCNA approaches for predicting phenotypic variation in *r* and *Hmax* through a series of multivariate linear regression models (*lm* function in R). We extracted the effect size and direction for each QTL using the *effectplot* function in r/qtl (Broman and Sen 2009). In all cases, the trait BLUPs were the dependent variable, and all allele-specific effect sizes, gene expression, and eigengene expression values were independent variables. For each trait we generated three types of additive models: 1) models with one type of independent variable (genotypic information based on alleles harbored at each QTL including allele-specific effect sizes and direction or genotype specific gene expression values for MR genes or genotype specific eigengene expression values), 2) models with two types of independent variables (QTL and MR gene expression, QTL and eigengene expression, or MR gene expression and eigengene expression), and 3) full models with all three data types as independent variables. For each trait we included only significant QTL, genes from the MR30 network, and eigengenes that were significantly correlated with the trait of interest. Each model was subjected to a backwards model reduction routine where non-significant terms were iteratively removed until all terms in the model had significant effects on the dependent variable ($p<0.10$). We used AIC scores to compare final models.

Relationships between co-expression and FVT QTL. We performed Fisher's exact test to determine whether the FVT QTL regions were enriched for genes and/or eigengenes identified via MR and WGCNA network analyses. Enrichment of FVT QTL for MR-identified genes was interpreted as evidence that the MR-identified genes are candidate causal genes for the FVT trait of interest.

eQTL Analyses. To explore the regulatory mechanisms of MR-identified genes and WGCNA-identified eigengenes, as well as their potential connection to FVT QTL, we performed eQTL analyses. Our network analyses effectively allowed us to reduce the number of expression traits mapped from 10,000 to less than 75. Therefore, we used composite interval mapping (Zeng 1993), which is usually considered too computationally intensive for eQTL studies. Permutation

testing (Doerge and Churchill 1996) was used to establish a p < 0.05% significance threshold for each gene. The *bayesint* function in r/qtl was used to define 99% confidence intervals for each eQTL. For some eQTL with very high LOD scores the resulting confidence interval was a single basepair (clearly unrealistic given the limitations imposed by the number of recombination events in a mapping population). For such eQTL we used a window of +/- 20kb around the identified base pair as the eQTL interval. We defined *cis*-eQTL as eQTL that include the physical gene generating the mRNA transcript and *trans*-eQTL as any eQTL that does not include the physical location of the gene. For MR-identified genes, *cis*-eQTL are interpreted as evidence of variation in *cis* regulatory elements such as promoters whereas *trans*-eQTL are interpreted as evidence for *trans*-acting regulatory proteins such as transcription factors, other signaling proteins, or small RNAs that modulate gene expression. Because eigengenes represent the composite expression of a median of 90 genes, one cannot assign *cis*- vs. *trans*-eQTL identity for these traits (although the majority of their action is expected to be in trans). MR gene or eigengene eQTL that colocalize with FVT QTL may explain the underlying basis for the FVT QTL, and such colocalizing eQTL represent candidate causal genes for the FVT eQTL locus. An alternative explanation is that eQTL that co-localize with FVT QTL are in linkage disequilibrium with the FVT QTL candidate. eQTL that do not co-localize with FVT QTL may still be affecting plant development, but at a level not directly detectable in the FVT QTL mapping.

Data availability: The linkage map used for QTL and eQTL analyses is available in Markelz et al (2017). Replicate level FVT parameters are presented in S1; RIL-specific gene expression values will be made available in supplemental materials (via FigShare) upon acceptance of the manuscript and are available to the editor and reviewers upon request.

RESULTS

FVT Modeling: For all FVT modeling, the data were sufficient to support all aspects of the growth curves modeled, and the models fit the data well (Fig. 1 for example model fits). Plots for all FVT models can be found in S2.

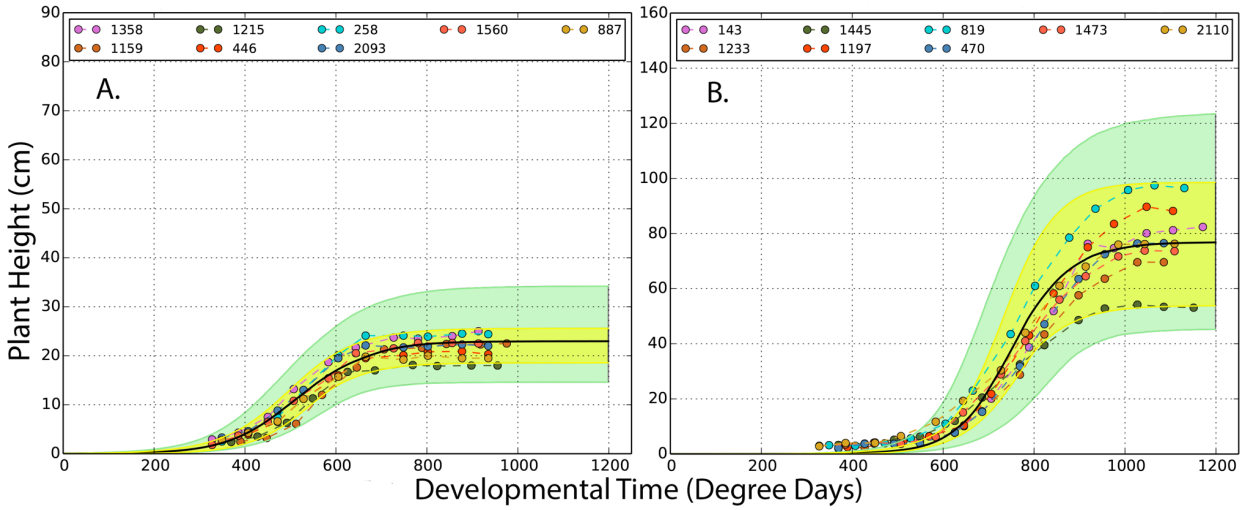


Fig. 1. Representative genotypes (A, IMB211; B, R500) of Bayesian FVT trait estimation approaches for uncrowded plants from the 2012 season. Within each panel, dots represent observed data. Colors indicate replicates within each genotype, and indicate that each replicate was measured multiple times throughout the growing season. The black line is the Bayesian estimate of logistic growth curve that best represents each genotype. The yellow envelope is a 95% credible envelope for the observed data; the green envelope is a 95% credible envelope for where new data is predicted to occur for a specific genotype and environment combination.

Phenotypic Plasticity and Heritability: To assess the effects of the environment on plastic growth responses, we analyzed raw replicate level data. Although there were main effects of Block (nested within treatment) and genotype (RIL ID) for all traits, there were no significant main effects treatment (Table 1). However, there was genetic variation for a plastic response to crowding for all traits except *iD* (inflection time, in Degree Days; treatment-by-genotype interaction; Table 1).

In general, heritabilities were higher for plants grown in the UN relative to CR treatments for all traits. This may reflect the relatively stochastic nature of the crowding response: in some cases in the CR treatment the focal plant may have outcompeted its neighbors whereas in others it may have been outcompeted.

Table 1. Phenotypic plasticity and heritabilities of FVT parameters. Block is nested within the Treatment effect. Treat corresponds to the crowded and uncrowded treatments in 2012 and Genotype indicates RIL id. Significant effects are emphasized by bold text.

		Random effects – Chi Square value (degrees of freedom)				Heritabilities (%)	
Trait	Model t-value (df)	Block (Treat)	Treat	Geno-type	Treat* Geno-type	UN 2012	CR 2012
R	**16.62 (1.08) ***	**80.2 (2) *****	7.28e-12 (1) NS	**136 (1) *****	**211 (1) *****	74.5	76.0
D	**43.32 (1.57) ****	**58.5 (2) *****	3.64e-12 (1) NS	**294 (1) *****	**4.88 (1) ***	79.5	79.3
iD	**37.16 (1.65) ****	**98.2 (2) *****	1.42e-10 (1) NS	**369 (1) *****	0.34 (1) NS	86.8	83.7
Hmax	**8.70 (1.83) ***	**116 (2) *****	0.0 (1) NS	**226.4 (1) *****	**42.3 (1) *****	81.2	68.1

Signif. codes: $p < 0.001$ '***'; $p < 0.01$ '**'; $p < 0.05$ '*'; $p < 0.1$ '.'; $p > 0.1$ 'NS'

***Genetic Correlations*:** To explore the genetic relationships among the height FVT parameters and previously published estimates of plant phenology and fitness, we conducted a correlation analysis on BLUPs of each trait. In general, the pattern of genetic correlations within years and treatments was similar. UN*r* from 2012 was correlated with all traits except *Hmax* (Fig 2). In contrast, CR*r* in 2012 was negatively correlated with other all other 2012 CR FVT traits, with all CR phenology traits (except the bolting-to-flowering interval) and CR fitness traits (S3). UN*r* in 2012 was negatively correlated with UN*d* and *iD* but not *Hmax*. UN*r* 2012 was also negatively correlated with phenology and fitness. These patterns of genetic correlations are largely consistent across years and treatments; a representative subset are presented in Fig 2.

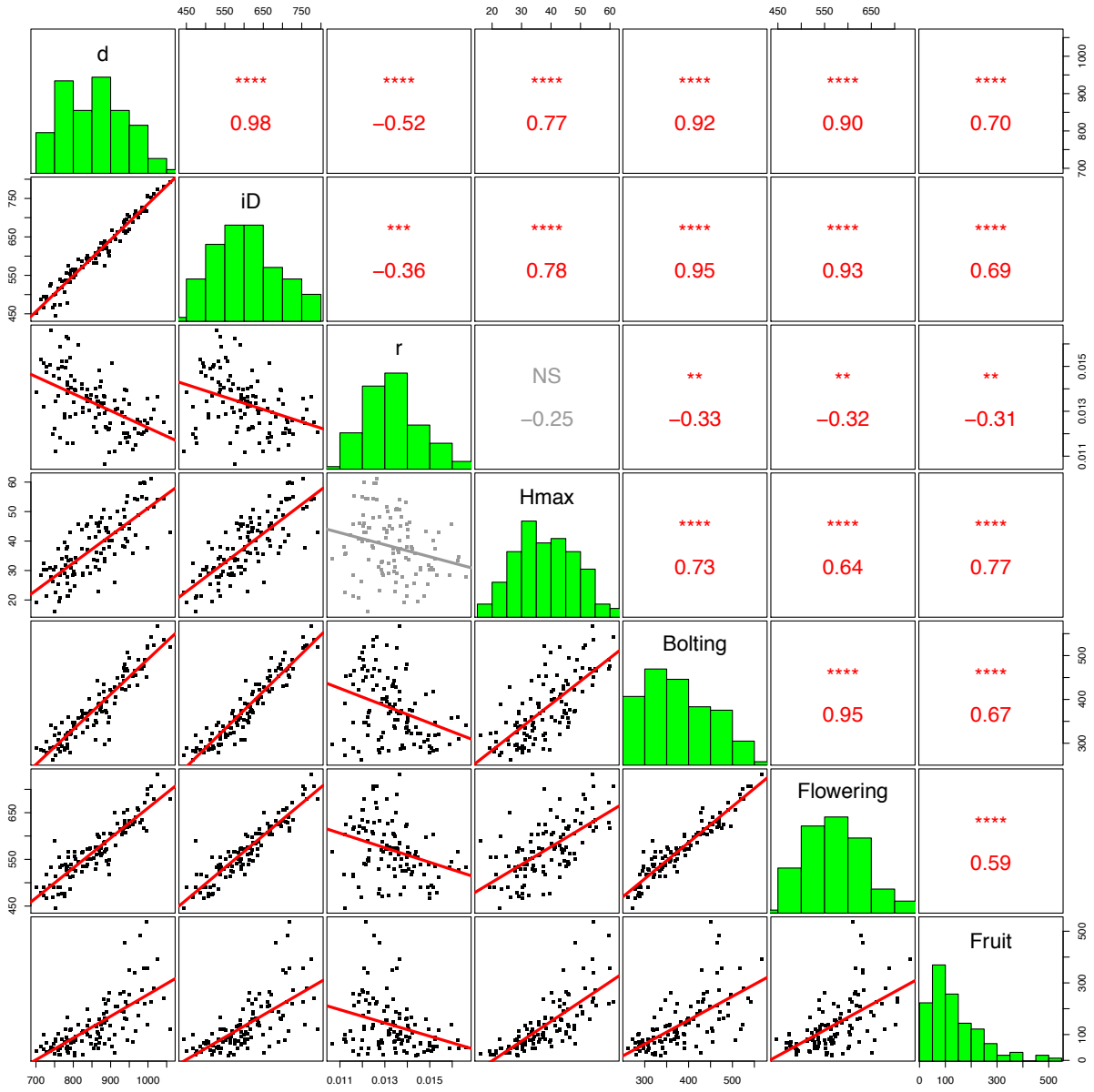


Fig 2. Genetic correlations among UN 2012 FVT height, phenology, and fitness traits. Each point is a genotypic mean (BLUP). Bonferroni corrections for multiple tests (n=7) have been applied. Non-significant correlations are in gray. All time is expressed in Degree Days. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001, NS p ≥ 0.05.

QTL mapping: To further explore the genetic architecture of the height FVT parameters, we conducted QTL mapping analyses of the height FVT traits. In total we mapped 32 individual QTL from 2012 (2011 FVT QTL are presented in S4); however, an alternative interpretation is that we mapped as few as 9 highly pleiotropic QTL. QTL were observed throughout the genome,

except on chromosomes 2, 4, and 8. Most QTL localized to chromosome 3, 9 and 10. Across all traits, each QTL explained 29% of trait variation on average. The minimum explained variance was 9.5% and the maximum was 73% of variance (Fig 3 & S4).

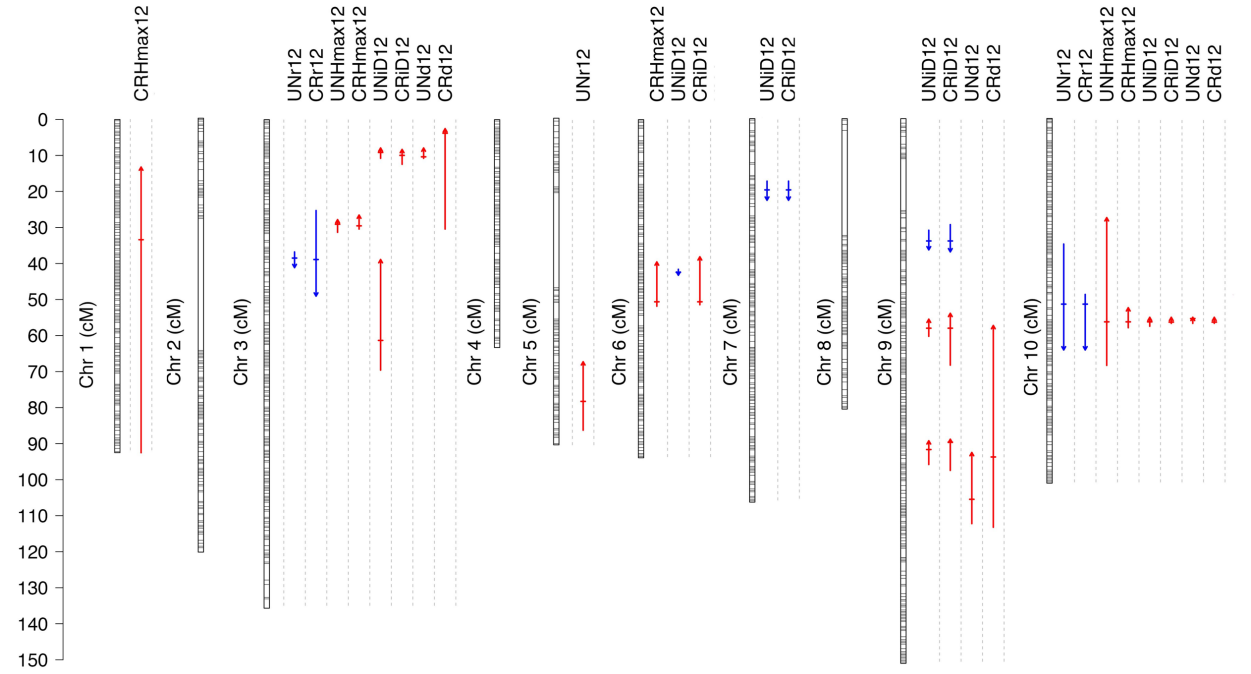


Fig 3. A map of all QTL identified in 2012. Horizontal lines on chromosomes indicate the position of RNAseq markers used to genetic map construction. Each QTL is indicated with a vertical arrow under the trait name. Horizontal hatches indicate QTL position, the arrow length indicates 1.5 LOD support limits. Arrow heads and color (up, red = positive; down, blue = negative) indicate QTL direction relative to the R500 parent. Exact locations, markers, and LOD scores for all QTL can be found in S4.

Genes under FVT QTL: To determine positional candidates within mapped FVT QTL, we compared our FVT QTL to the *B. rapa* genome and identified genes underlying the QTL. We restricted our search to QTL with LOD > 9 (Table 2). All 9 of these QTL were on either chromosome 3 or 10. Because several of the QTL co-localized (had overlapping 1.5 LOD confidence intervals), we often found the same genes under multiple QTL. After removing duplicate entries, we found 490 unique genes underlying the 9 QTL investigated (S5).

RNAseq. We used RNA sequencing (RNAseq) to understand the genetic mechanisms underlying FVT QTL and as an alternative approach for examining the genetic architecture of our FVT traits without *a priori* knowledge. 21,147 genes of 28,668 genes with detectable expression in UN treatment were differentially expressed among RILs (FDR < 0.01). The 10,000 genes with the most variable expression among RILs were used for downstream network analysis.

Mutual Rank Network Analysis:

To find gene co-expression networks relevant to the FVT model parameters, we built Mutual Rank (MR) networks nucleated on each FVT model parameter and performed permutation analyses to determine the statistical significance of our networks. Ninety-five or more of 100 permutations had zero connections between FVT parameters and gene expression. Therefore, our MR networks are enriched for *bona fide* connections at a variety of MR threshold cutoffs (The MR30 network is shown in Fig 4; larger networks become difficult to visualize and are presented in S6). Complete gene membership for all MR-thresholds annotated with the best hit obtained by *blastn* against the predicted *A. thaliana* proteome are presented in supplemental materials S7.

We used Fisher's exact test to determine whether FVT QTL were enriched for MR-identified genes. We found no evidence for enrichment for MR10 networks (p=1.0) but significant evidence for enrichment for MR20, MR30, and MR50 networks (p<5E-09; Table 4). In theory, MR10 networks should contain only those genes whose expression values are most highly correlated with FVT phenotypes. The non-significant results for MR10 may be caused by low power due to the single gene identified.

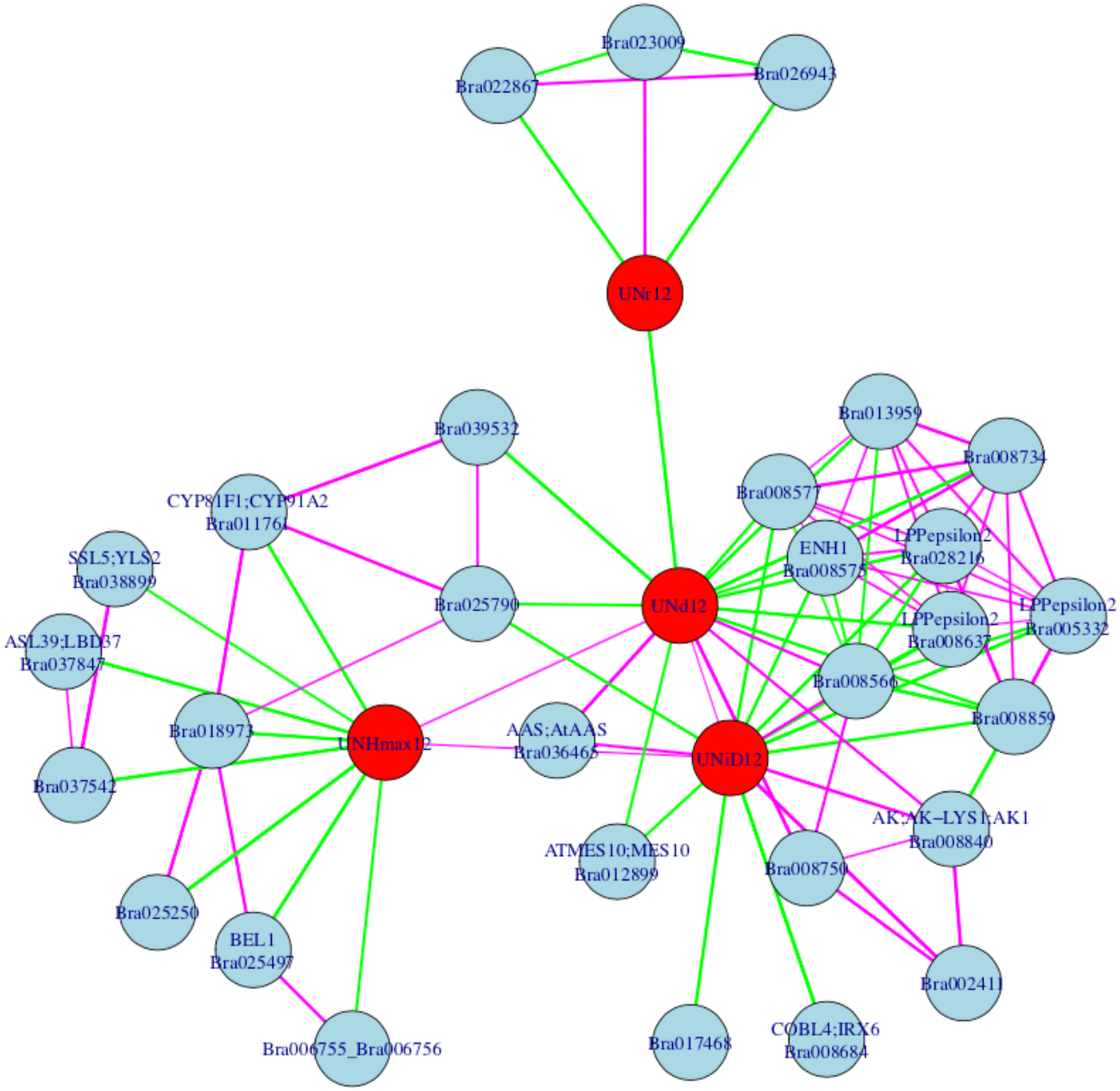


Fig 4: A scale-free diagram of the Mutual Rank network nucleated around FVT traits from 2012 with a cutoff of 30. Network nodes consist of either FVT traits or co-expressed genes. FVT traits are shown in red circles and genes are indicated in blue circles. Network edges indicate significant correlations. Purple lines indicate positive correlation values while green lines indicate negative correlation values and line thickness corresponds to strength of the correlation. UN, uncrowded; *r*, growth rate; *d*, duration of growth; *iD*, time in degree days when the growth curve reached its inflection point; *Hmax*, estimated maximum height based on FVT modeling. Additional network cutoffs, 2011, and 2012 crowded networks are in S6; gene names and annotations are in S7.

Table 4. Fishers exact tests for enrichment of FVT QTL for MR-identified genes.

		QTL		
		Yes	No	p-value
MR10	Yes	0	2	1.0
	No	5,816	37,645	(NS)
MR20	Yes	16	0	6.91e-13
	No	6,800	37,647	***
MR30	Yes	25	4	4.98e-09
	No	5,791	37,643	***
MR50	Yes	46	10	9.93e-21
	No	5,770	37,637	***

p> 0.05, NS; p<0.0001, ****

Weighted Gene Co-expression Network Analysis (WGCNA):

In a second approach to identifying gene expression networks related to estimates of FVT trait parameters, we used a Weighted Gene Co-expression Network Analysis (WGCNA) to identify eigengene modules. Modules of interest were identified as those showing a significant correlation between eigengene expression values and FVT model parameters across the RILs (Figure 5). Gene Ontology (GO) enrichment analysis was performed to examine the potential function of correlated module (S8); below we discuss correlations with modules that had at least one GO term enriched. There are positive correlations between 2012 BLUPs for maximum height (*Hmax*), growth *duration* (*d*), and the time that the growth curve reached its inflection point (*iD*) and the *cyan* module (related to protein translation), the *midnight blue* module (related to wounding/herbivore defense responses as well as some abiotic stress responses), and the *blue* module (enriched for genes related to cell division and development). This suggests that plants that have a longer duration of growth and reach a higher maximum height are producing more protein, undergoing more rounds of cell division, and have increased defense signaling. These three parameters also showed negative correlation with the brown module (enriched for actin cytoskeleton and protein dephosphorylation terms). *Hmax* is negatively correlated with *yellow* (enriched for terms related to photosynthesis). This

correlation could be caused by a difference in cellular maturation rates: plants with more rapid cellular differentiation would be expected to show an upregulation of chloroplast genes and reduced growth due to earlier differentiation and consequently relative lack of cell elongation.

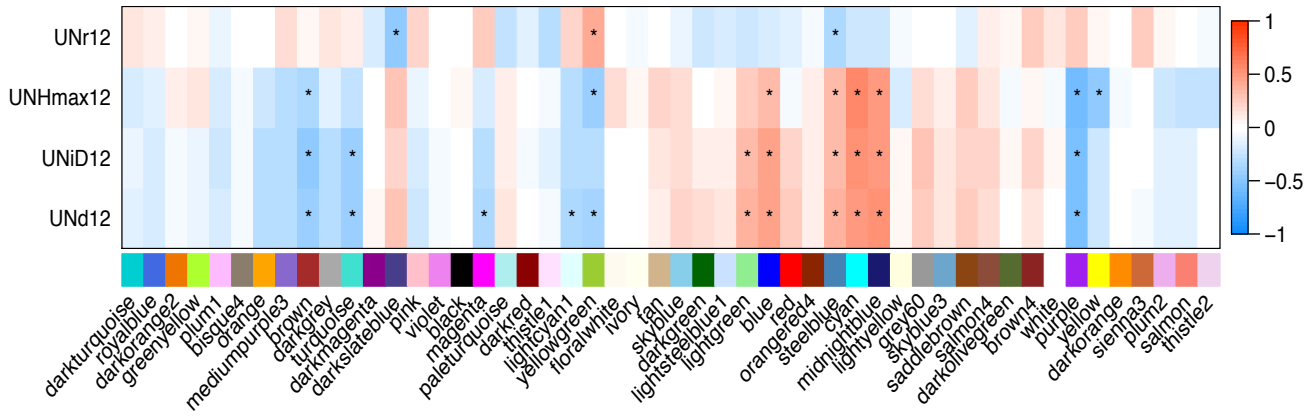


Fig 5. Correlations among WGCNA identified eigengenes and UN 2012 FVT traits. Significant correlations are denoted with an asterisk. *r*, growth rate; *d*, duration of growth; *iD*, time in degree days when the growth curve reached its inflection point; *Hmax*, estimated maximum height based on FVT modeling.

Comparisons of QTL and network modeling for phenotypic prediction: To compare the effectiveness of various approaches and combinations of these approaches in explaining the variation in FVT trait estimates, we compared a series of additive linear models based on QTL, MR genes, or WGCNA eigengenes both singly and in combination. For UN*r* (in 2012), models containing only QTL outperformed models containing either MR30 identified gene expression or WGCNA-identified eigengene expression (Table 5). For two-data type models, models with only QTL outperformed those containing multiple data types. For *Hmax*, MR gene expression outperformed both QTL and WGCNA-identified eigengene expression as well as combinations of two data types. For both traits, the full model (with all three data types for *r*, but which reduced to WGCNA and MR gene expression values for *Hmax*) were the best models for explaining phenotypic variation (*r*: $F_{(5,110)}$=25.31, p<0.0001; *Hmax*: $F_{(9,106)}$=33.16, p<0.0001). Similarly, the best two-data type models were a significantly better fit to the data than the best single-data type models (*r*: $F_{(5,114)}$=40.182, p<0.0001; *Hmax*: F(4,113)=80.398, p<0.0001). For all comparisons, the significantly better model according to ANOVA also had lower AIC scores

(Table 5). Taken together, these results indicate that although each approach has significant predictive capacity, combining multiple approaches improves estimation of trait variation.

Table 5. Comparison of additive linear models using genetic and transcriptomic data to explain 2012 uncrowded phenotypic data.

Trait	Best single-data type model	AIC	Next best AIC (next best model)	Formula[§]	Best model F-value (DF), significance and adjusted R^2
r	QTL	-1305.97	-1256.43 (WGCNA)	y ~ rQTL2 + rQTL2 +r QTL3	F(3, 113)= 30.9 *** R^2=0.4361
Hmax	MR	735.5348	783.6546 (WGCNA)	y ~ Bra03899 + Bra011761 + Bra006755_Bra006756 + Bra036465 + Bra008859 + Bra037542	F(6,109)=45.48 *** R^2=0.6989
Best 2-data type model					
r	QTL + WGCNA (reduces to just QTL)	-1305.97	--1297.869 (MR+WGCNA)	y ~ rQTL1 + rQTL2 + rQTL3	F(3,113)=30.9 *** R^2=0.4361
Hmax	MR + WGCNA (reduces to just MR)	734.2895	752.3889 (QTL+MR; reduces to just MR*)[†]	y ~ Bra011761 + Bra006755_Bra006756 + Bra13959 + Bra08840 + Bra008859 + Bra037542 _ Bra002411	F(7,108)=40.16 *** R^2=0.7045
Best overall model					
r	Full model (QTL+ MR+ WGCNA)	-1308.602	-1305.97 (QTL + WGCNA)	y ~ rQTL2 + yellowgreen + Bra006755_Bra06756 + Bra025790 + Bra028216	F(5, 110)=25.31 *** R^2=0.5138
Hmax	Full model (reduces to MR + WGCNA)	731.63	-734.2895 (MR + WGCMA; reduces to just MR*)	y ~ yellow + Bra011761 + Bra006755_Bra006756 + Bra008575 + Bra008577 + Bra008840 + Bra008859 + Bra037542 + Bra002411	F(9,106)=33.16 *** R^2=0.7157

*** $p < 0.0001$

[*] This model reduced to include just MR gene expression values but is different from the best Hmax single-data type model that also includes just MR gene expression values.

[§] *r*QTL 1-3 have markers at A03x 6417941, A05x23393567, and A10x11427369, respectively,

eQTL analyses and colocalization of eQTL with of FVT QTL

Because including MR and WGCNA results both improved upon linear models for FVT traits that contained just QTL (Table 5) and because all models that included MR and WGCNA

gene/eigengene expression values were significant and predicted FVT trait variation, we used eQTL analyses to assess the mechanistic relationship between gene/eigengene expression and FVT QTL. For the 29 MR30-identified genes, we found significant eQTL on all chromosomes except 5 and 8. In congruence with FVT QTL mapping results, there were eQTL with particularly high LOD scores on chromosomes 3 and 10 (LOD >75; Figure 6). There was significant overlap among 2012 FVT-QTL confidence intervals and MR-eQTL confidence intervals based on permutation tests (n=1000, p=0.003).

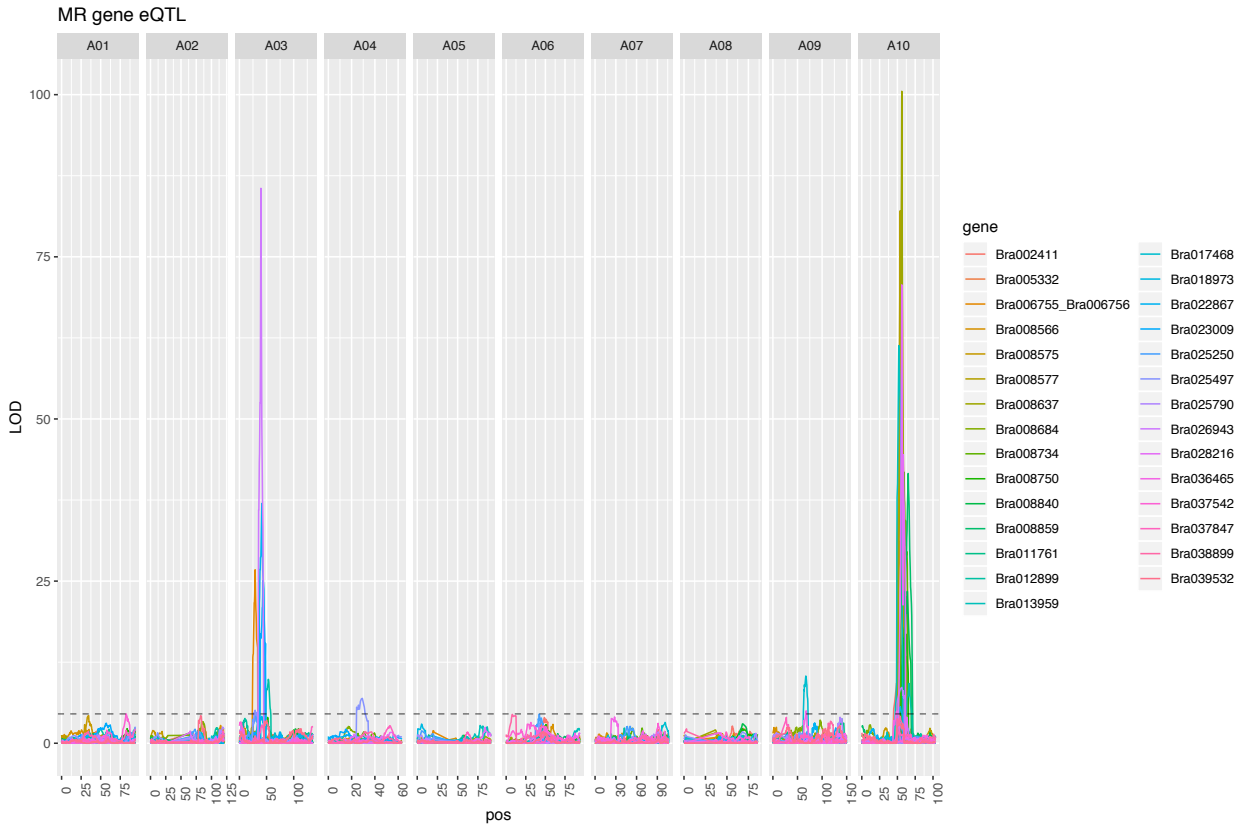


Fig 6. Expression trait QTL (eQTL) identified using Composite Interval Mapping (CIM) for MR30-identified genes where MR networks were nucleated around UN FVT traits. Note the eQTL hotspots on chromosomes 3 and 10.

Of the 57 MR50 genes, 42 genes had a total of 47 eQTL that overlapped with FVT QTL with LOD scores ranging from 100.5-4.6. Six of the 42 MR50 genes with eQTL that colocalized with FVT QTL had *cis*-eQTL, and of those six, three were in networks with cutoffs of MR30 or below (Table 6). The co-occurrence of these loci as MR-identified *cis*-eQTL and FVT QTL

545 indicates that they are strong candidate genes for regulating the FVT traits. For any given FVT
546 trait, none of the MR genes with *cis*-eQTL also had *trans*-eQTL that colocalized with other FVT
547 QTL. Of the 36 MR genes with *trans*-eQTL that colocalized with FVT QTL, 33 had a single *trans*-
548 eQTL that colocalized with FVT QTL. Three genes (Bra012899, Bra014655, and Bra029573) had
549 *trans*-eQTL that colocalized with two or more distinct FVT QTL (Table 7).
550
551 Table 6. MR-identified genes with *cis*-eQTL that co-localize with UN 2012 FVT QTL. Note that because FVT QTL
552 overlap a single MR *cis*-eQTL may colocalize with FVT QTL for multiple traits.

MR gene	MR Network	Chromo-some	FVT trait	eQTL LOD range	AGI	*A. thaliana* symbol
Bra008840	20	10	*r, Hmax*	22.526-23.419	AT5G13280	*AK;AK-LYS1;AK1*
Bra008859	20	10	*r, Hmax*	41.593-41.593	AT5G13070	*NA*
Bra008750	30	10	*r, iD, Hmax*	15.375-17.671	AT5G14600	*NA*
Bra008711	50	10	*r, Hmax*	24.585-26.594	AT5G15250	*ATFTSH6;FTSH6*
Bra008931	50	10	*r, Hmax*	8.515-10.223	AT5G11880	*NA*
Bra029100	50	3	*r*	28.288-30.055	AT5G53045	*NA*

553
554 Table 7. MR-identified genes with multiple *trans*-eQTL that co-localize with FVT QTL.

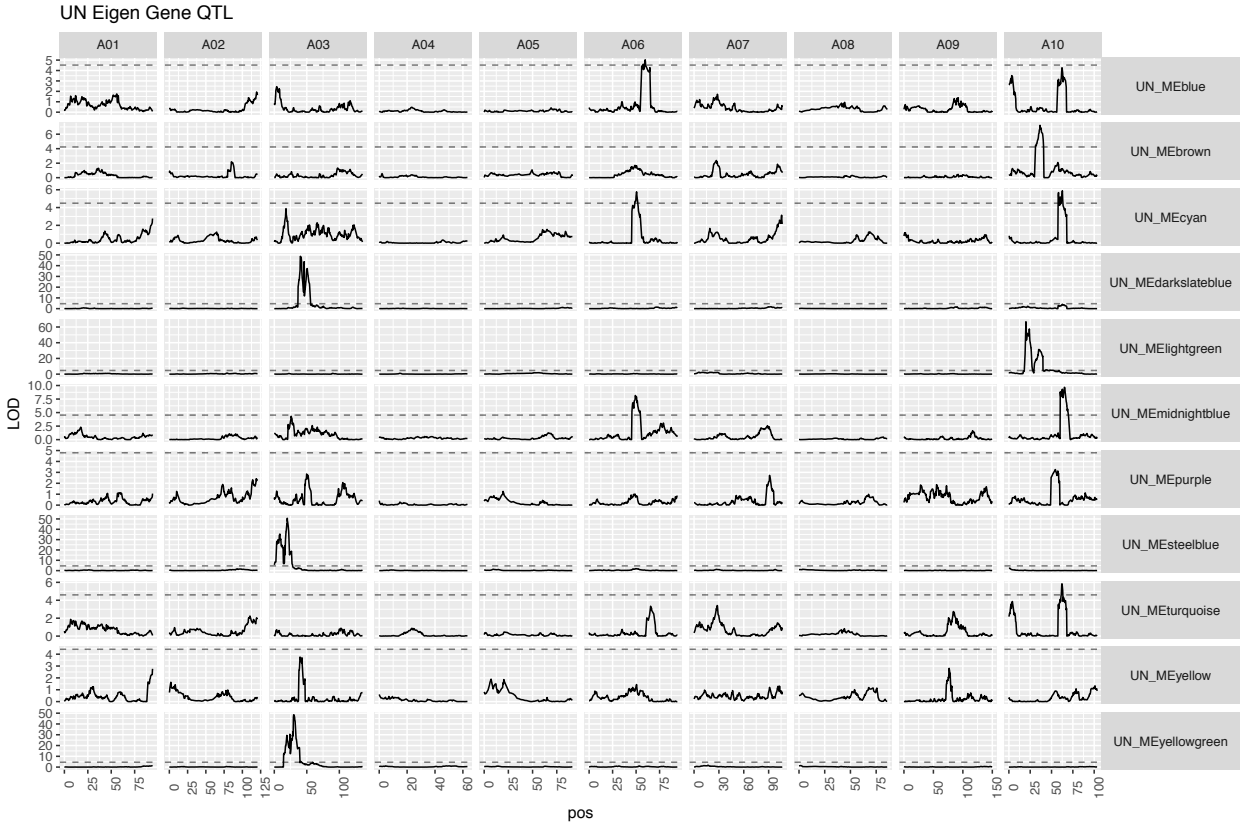
MR gene	MR Network	Chromosome	FVT Trait	eQTL LOD range
Bra012899	10	3	*iD*	8.315-9.854
		10	*Hmax*	6.970-9.214
Bra014655	50	3	*r, iD, d, Hmax*	2.857-4.930
		6	*iD*	6.526-8.333
		10	*r, Hmax*	3.290-5.176
Bra029573	50	3	*Hmax*	5.096-7.279
		6	*iD*	3.052-5.367
		10	*Hmax*	2.906-4.812

555
556 Next we performed eQTL analyses (Figure 7) for the 11 WGCNA-identified eigengene
557 modules based on UN 2012 FVT (see Figure 5). Chromosome 3 harbored strong eQTL for
558 “darkslateblue”, “steelblue”, and “yellowgreen” (all with no go enrichment; nge). Chromosome
559 6 had QTL for “blue” (cell division), “cyan” (translation), and “midnightblue”
560 (herbivore/wounding). Chromosome 10 had Eigengene eQTL in two locations, one for “brown”
561 (actin cytoskeleton) and “lightgreen” (nge), the other for “cyan” (translation), “midnightblue”

562 (herbivore/wounding), “turquoise” (nge), and a suggestive peak for “blue” (cell division). Five of
563 the eleven eigengenes had eQTL also colocalized with FVT QTL, indicating a potential causative
564 connection between eigengenes and FVT for *r*, *iD*, and *Hmax* (Table 8). However, each
565 eigengene had only one eQTL that colocalized with an FVT QTL.

566 The second chromosome 10 location (“cyan”, “midnightblue”, and “turquoise”) overlaps
567 with the FVT QTL9 and the Eigengenes has significant correlations with *d* and *iD* FVTs indicating
568 a possible causative connection. We then performed permutation tests and determined that
569 FVT-QTL were enriched for WGCNA-eQTL (n=1000, p=0.005).

570



571

572 Fig 7. Expression trait QTL analysis (eQTL) for WGCNA-identified eigengenes that significantly correlate with UN
573 FVT traits.

574

575 Table 8. Eigengene eQTL and FVT QTL colocalization.

Trait (eigengene)	Chromo-some	FVT trait	LOD range
brown	10	*r, Hmax*	6.052-7.223

cyan	10	*iD, r, Hmax*	4.900-5.880
darkslateblue	3	*r, iD*	47.828-47.828
midnightblue	10	*r, Hmax,*	8.826-9.674
turquoise	10	*r, Hmax*	3.967-5.840
yellowgreen	3	*Hmax*	48.424-48.550

DISCUSSION

Plant height is often correlated with fitness and yield. Height is a complex and dynamic trait that changes over the course of development, and variation in plant height is necessarily generated through variation in developmental dynamics. However, similar heights can be achieved through multiple different growth curves. Quantifying the underlying genetic architecture and mechanistic basis of growth dynamics may result in improved estimations of final plant height, fitness, and yield. Here, we use Bayesian hierarchical modeling to estimate Function-Valued Trait (FVT) parameters describing continuous plant growth and explore their correlations with phenology and fitness. We test whether QTL mapping, genes identified through Mutual Rank (MR) co-expression, or eigengenes identified through Weighted Gene Network Co-expression Analysis (WGCNA) co-expression, or combining these information types best explain genetic variation in agroecologically relevant FVT traits in the field. Further, we employ eQTL analyses to explore the molecular genetic regulatory mechanisms that mechanistically connect FVT QTL with phenotypic variation.

Although development typically occurs in a continuous fashion, most studies quantifying development necessarily collect data at discrete timepoints. We take a “parameters as data” approach to FVT modeling to estimate the continuous nature of plant development (Hernandez 2015; Kulbaba *et al.* 2017). Much as floral development or leaf development has well defined core molecular genetic pathways that govern organ formation, elaboration, or elongation (reviewed in Bowman *et al.* 2012), there is likely a core genetic architecture that contributes to plant height. However, exogenous and endogenous factors can influence the outputs of these developmental programs. For instance, crowding may trigger a shade avoidance response and lead to rapid increases in height (e.g. Schmitt *et al.* 2003). Similarly, plant carbon status can affect the developmental morphology and final size of organs such as leaves (Schneidereit *et al.*

2005; Raines and Paul 2006; Baker *et al.* 2018a). We took two approaches to examining the core developmental genetics of plant height. First, we grew plants across multiple growing seasons and in crowded and uncrowded conditions. Second, we included a genotype-specific co-factor in our FVT models that accounts for variation in photosynthetic rates (approximated through *Amax*), thereby statistically factoring out variation due to carbon availability and allowing us to more directly interrogate the developmental genetic architecture and molecular mechanisms contributing to plant height (Baker *et al.* 2018a; b). In our study, all FVT traits had relatively high broad sense heritabilities (>70%), and all had significant main effects of genotype. Interestingly, although there were no significant main effects of treatment (i.e. population means did not differ), all FVT trait estimates (except *iD*) exhibited genetic variation for carbon-independent phenotypic plasticity via a treatment-by-genotype interaction, likely because of rank-order differences across treatments at the genotypic level (Table 1).

Morphological phenotypes, such as components of yield and height, can be highly integrated throughout development (reviewed in Klingenberg 2014). Final height is often used as a proxy for yield or fitness, yet plant growth dynamics throughout ontogeny may also be correlated with aspects of yield such as fruit and seed set (Yin *et al.* 2011; Tanger *et al.* 2017). In our experimental set of *Brassica rapa* Recombinant Inbred Lines (RILs), plant developmental dynamics including duration of growth (*d*), the inflection point in the growth curve that represents the change from exponentially accelerating to decelerating growth (*iD*), and estimates of final plant height (*Hmax*) were all significantly and positively genetically correlated (Fig 2). Interestingly, growth rates (*r*) were negatively correlated with *d* and *iD*, but were not correlated with *Hmax*, indicating that while there is a trade-off between growth rates and durations, duration of growth may be more important for final plant height than growth rate. All of our estimates of plant growth and final size were significantly genetically correlated with both phenology and yield traits. The significant correlations of *r* with yields indicates that developmental dynamics of a given trait can be related to crop yields and plant fitness through mechanisms that may be at least partially independent of final size. Because final size is positively correlated with yields while growth rates are negatively correlated with yields,

630 selection for maximum yields at early harvest dates may come at the expense of late harvest
631 yields and vice versa.

632 To examine the genetic architecture underlying the FVT estimates of growth rates,
633 durations, and final sizes, we used standard QTL mapping procedures, which revealed a number
634 of QTL. Of particular note, when QTL for *r* colocalized with *d*, the QTL were of opposite sign,
635 confirming our negative genetic correlations between growth rates and durations, and
636 indicating potentially pleiotropic loci contributing to both traits. On average, FVT QTL explained
637 24% of trait variation and the number of genes under each QTL ranged in to the hundreds. In
638 part to narrow down the list of candidate genes and in part to understand the mechanistic
639 regulation of FVT via QTL, we took two additional transcriptomic co-expression approaches to
640 exploring the genetic architecture of FVT traits: First, we seeded a Mutual Rank (MR) co-
641 expression network with FVT traits and asked which gene expression values correlated with
642 variation in FVT traits. Second, we constructed 50 eigengenes based on a Weighted Gene Co-
643 expression Network Analysis (WGCNA) and asked which eigengenes were correlated with
644 individual FVT trait. We found that FVT QTL were significantly enriched for MR genes, indicating
645 that these two approaches were identifying some common drivers of FVT traits. To compare
646 the effectiveness of all three approaches, we asked whether QTL, MR genes, or eigengenes best
647 explained variance in FVT traits. Although QTL outperformed both co-expression network
648 modeling approaches for *r*, combining data from multiple approaches yielded improvements in
649 our models, indicating that even though QTL, MR genes, and eigengenes often physically co-
650 localize within the genome, they are not synonymous with one another (Table 5).

651 To better understand the potential function of genes related to growth WGCNA and MR
652 networks, we used gene annotations and homology to *A. thaliana*. Although about half of the
653 eigengenes that correlated with FVT BLUPs had no gene ontology enrichment, three eigengenes
654 with eQTL on chromosome 10 were enriched for actin/cytoskeleton, herbivore/wounding and
655 cell division, respectively. The MR30 genes include a homolog of the homeodomain gene *BEL1*
656 (*NACA3* (Reiser *et al.* 1995) which is negatively correlated with *Hmax*); *BEL1* homologs have
657 been implicated in regulation of the shoot apical meristem (Rutjens *et al.* 2009) and thus could
658 be related to plant growth. An additional gene was identified with homology to the COBRA

family gene *COBL4/IRX6* (negatively correlated with *iD*), involved in secondary cell wall biosynthesis. The MR30 network also contains a number of genes involved in metabolic homeostasis. Four of these genes are localized to the plastid and negatively correlated with *d* and *iD*, including three orthologs of the *plastidic lipid phosphate phosphatase epsilon 2* gene (*LPPε2*), which is potentially involved in synthesis of diacylglycerol, a precursor to essential photosynthetic membrane components (Nakamura *et al.* 2007). Another plastid-localized MR30 network gene is *ENHANCER OF SOS3-1 (ENH1)*; *ENH1* functions to mitigate the effects of reactive oxygen species (Zhu *et al.* 2007). Thus, plants with longer growing periods appear to put less resources into photosynthesis. The MR30 network also includes a homolog of the *A. thaliana LATERAL ORGAN BOUNDARY DOMAIN37 (LBD37)* gene, an important regulator of nitrogen response in both *A. thaliana* and *Oryza sativa* (Rubin *et al.* 2009; Albinsky *et al.* 2010). *LDB37* is negatively correlated with *Hmax.* Two genes involved in amino acid synthesis or homeostasis are present in the MR30 network and show positive correlations with *d* and *iD*: a homolog of *ASPARTATE KINASE1 (AK1),* required for regulation of aspartate, lysine, and methionine (Clark and Lu 2015), and *AROMATIC ALDEHYDE SYNTHASE (AAS)*, which converts phenylalanine into phenylacetaldehyde (Gutensohn *et al.* 2011). Overall the MR30 network results point to a close connection between metabolic regulation and growth.

Transcriptomic data allowed us to further explore the regulatory control of the FVT using eQTL mapping of WGCNA eigengenes and MR genes. eQTL mapping treats gene expression levels as quantitative traits. When combined with QTL studies of morphological phenotypes, the ultimate goal of eQTL mapping is to identify the molecular genetic changes in gene expression that lead to structural phenotypic variation, thus providing mechanistic explanations for the associations between genotype and phenotype (Schadt *et al.* 2008). In humans, such studies demonstrate that eQTL can be used in a cell-type specific fashion to annotate GWAS associations (Brown *et al.* 2013). In our study, 42 MR genes had eQTL that colocalized with FVT QTL and 6 of the 11 WGCNA eigengenes that correlated with FVT also had eQTL that colocalized with FVT QTL. These data demonstrate that the relationship between genomic loci (FVT QTL) and phenotypic variation in FVT traits is likely mediated by gene

687 expression, specifically the expression of the genes and eigengenes we identified via MR and
688 WGCNA.

689 Our eQTL results qualitatively departed from common morphological trait QTL analyses
690 in two ways. First, MR-identified gene expression traits mapped to all chromosomes except
691 chromosome 2, but two locations had multiple eQTL with very high LOD scores (>75): the top of
692 chromosome 3 and the middle of chromosome 10. Virtually all genes had eQTL that mapped to
693 one of these two locations, a common result potentially indicating an eQTL 'hotspot'. A
694 previous study of the effects of soil phosphorous using the same *B. rapa* RILs also identified
695 eQTL hotspots (Hammond *et al.* 2011), but on different chromosomes. The colocalization of
696 eQTL hotspots and FVT QTL may indicate novel regions involved in pleiotropic co-regulation of
697 several downstream genes in the regulatory network contributing to change in plant height
698 (Gibson and Weir 2005).

699 Although the presence of eQTL hotspots indicates pleiotropic gene regulation, our eQTL
700 analyses also qualitatively departed from the FVT QTL analysis in that most of the gene
701 expression traits we mapped were not polygenic. Of the 42 MR gene expression traits mapped,
702 only three had eQTL that colocalized with more than one FVT QTL. eQTL studies commonly find
703 a relative paucity of polygenic regulation compared to structural QTL studies, and our results
704 support the general consensus that expression traits and structural phenotypes have distinctly
705 different genetic architectures (but see West *et al.* 2007 for a counter-example). However, most
706 eQTL are of relatively large effect, meaning that many small effect eQTL could remain
707 undetected and contribute to polygenic regulation of gene expression traits (Gibson and Weir
708 2005), and these eQTL may or may not occur in regulatory hotspots.

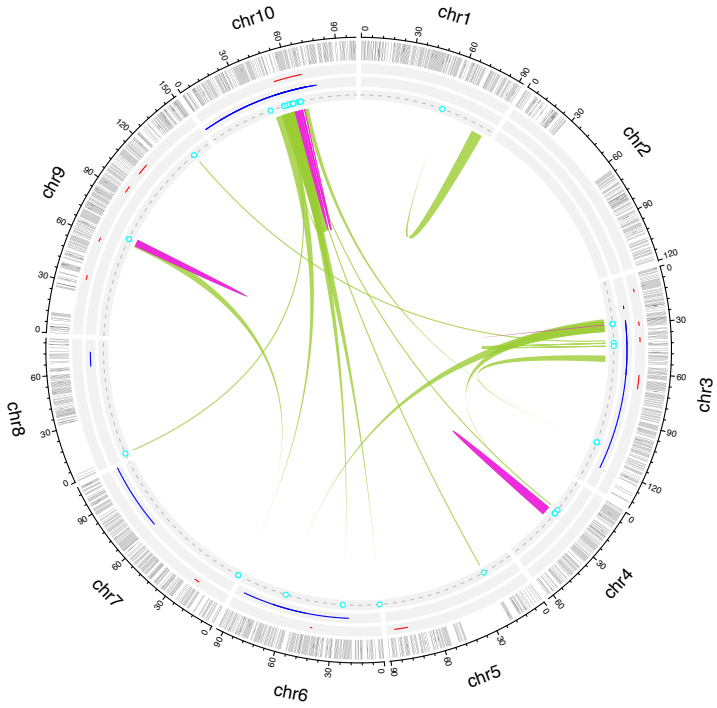


Figure 8. Function-Valued Trait QTL (2012 uncrowded data), Weighted Gene Co-expression Network Analysis (WGCNA) identified eigengene eQTL, and genes identified via Mutual Rank (MR) co-expression occur at regulatory hotspots on chromosomes 10 and 3, indicating that these MR genes are candidate master regulators that integrate information to generate developmental trait variation. MR gene *cis*-eQTL (pink links) on chr10 and 3 lend further credence to this relationship. MR genes with *trans*-eQTL (green links) that map to these hotspots are putative upstream genes feeding in to the FVT regulatory network. By integrating information from multiple analyses. From exterior to center: chromosomes in black, linkage map in gray, FVT QTL in red, eigengene eQTL in blue, MR genes in cyan, MR *trans*-eQTL in light green and MR *cis*-eQTL in pink.

To further understand the regulation of expression traits and FVT QTL, we divided MR eQTL into two classes: putative *cis*- and *trans*-eQTL where *cis*-eQTL likely correspond to *cis*-regulatory elements influencing gene expression (Doss *et al.* 2005). In contrast, *trans*-eQTL do not contain the gene whose expression pattern is mapped and likely correspond to *trans*-acting factors such as transcription factors that influence the MR gene expression (Hansen *et al.* 2008). In our study, of the 42 MR genes with eQTL that colocalized with FVT QTL, only five were in *cis* and the remaining 37 were in *trans*, which is only slightly higher than the proportion of *trans*-eQTL identified in an intraspecific maize cross (Swanson-Wagner *et al.* 2009). Because the *B. rapa* RILs are also generated from an intraspecific cross, our results are consistent with theoretical and experimental work suggesting that *trans* gene regulation should be more

729 prevalent than *cis* regulation at the intraspecific level (Wittkopp *et al.* 2008; Goncalves *et al.*
730 2012, but see O'Quin *et al.* 2012 for an exception). Although unlikely given the genetic
731 architecture of our eQTL, biases towards *trans* regulation may also stem from highly pleiotropic
732 genes (reviewed in Signor and Nuzhdin 2018). Other authors have offered an alternative
733 interpretation: in *A. thaliana* the proportion of *cis*- to *trans*-eQTL appears to scale with
734 statistical power and the ability to detect small effect eQTL. *Trans*-eQTL are typically assumed
735 to be of small effect and so increasing sample size, replicate number, or density of markers on
736 the genetic map should in theory increase the proportion of *trans*-eQTL detected (Hansen *et al.*
737 2008). The fact that we detected so many *trans*-eQTL may indicate that our study system has
738 ample power to detect small effect *trans*-eQTL (our percent variance explained was 10%).
739 Interestingly, a subset of the *trans*-eQTL we identified (located in eQTL hotspots) had
740 exceptionally high LOD scores (75-100) that were twice as large as the largest *cis*-eQTL LOD
741 score. Clearly, not all *trans*-eQTL have small effect sizes.

742 Our study demonstrates the importance of examining not just final plant height, but the
743 developmental dynamics that contribute to height growth curves in agroecologically relevant
744 field settings. We fit function-valued trait models to our data and, while statistically factoring
745 out aspects of physiology such as carbon assimilation rates, demonstrate that parameters
746 describing continuous developmental growth curves are correlated with plant fitness and yield.
747 The shape of these growth curves (as described by *r,* d, and iD) is phenotypically plastic, while
748 estimates of final height (*Hmax*) are relatively robust across environments. However, changes in
749 the sign of bivariate correlations indicate a trade-off between yields at given final size vs. yields
750 at early developmental times. We map FVT QTL to multiple chromosomes and utilize a guided
751 eQTL mapping approach to investigate the regulatory mechanisms connecting genotype to FVT
752 phenotype. Specifically, we use WGCNA to identify eigengenes for actin/cytoskeleton and cell
753 division processes whose expression values that correlate with FVT traits. FVT trait seeded MR
754 co-expression networks had an overall association with metabolic regulation and growth
755 processes. We demonstrate that combining multiple approaches yields the best explanation of
756 phenotypic variance. We identify more *trans*- than *cis*-eQTL and these *trans*-eQTL are highly
757 colocalized at regulatory hotspots, likely including transcription factors that influence

758 downstream gene regulation. Because our *cis*- and trans-eQTL hotspots colocalize with FVT
759 QTL, these expression traits are likely components of the molecular regulatory mechanisms
760 mediating the generation of FVT phenotypic variation from genomic variation (Fig 8).
761
762 **ACKNOWLEDGEMENTS**
763 University of Wyoming undergraduates E. Gimpel, J. Whipps, K. Anderson, M. Pratt, J. Beckius,
764 C. Blemenshine, S. Cheeney, M. Yorgason, W. Gardner, C. Planche, C. Gifford, L. Lucas, K. Riggs,
765 D. Lorimer, D. Nykodym, and L. Steinken assisted with data collection and entry. C. Seals and R.
766 Pendleton facilitated plant growth. This work is supported by National Science Foundation
767 Grants IOS-1306574 to RLB and IOS-0923752 to CW, SW, and JM.
768
769 **WORKS CITED**
770 Albinsky D., Kusano M., Higuchi M., Hayashi N., Kobayashi M., *et al.*, 2010 Metabolomic
771 Screening Applied to Rice FOX *Arabidopsis* Lines Leads to the Identification of a Gene-
772 Changing Nitrogen Metabolism. Mol. Plant 3: 125–142.
773 Altschul S. F., Gish W., Miller W., Myers E. W., Lipman D. J., 1990 Basic local alignment search
774 tool. J. Mol. Biol. 215: 403–410.
775 Baker R. L., Leong W. F., Brock M. T., Markelz R. J. C., Covington M. F., *et al.*, 2015 Modeling
776 development and quantitative trait mapping reveal independent genetic modules for leaf
777 size and shape. New Phytol. 208: 257–268.
778 Baker R. L., Leong W. F., An N., Brock M. T., Rubin M. J., *et al.*, 2018a Bayesian estimation and
779 use of high-throughput remote sensing indices for quantitative genetic analyses of leaf
780 growth. Theor. Appl. Genet. 131: 283–298.
781 Baker R. L., Leong W. F., Welch S., Weinig C., 2018b Mapping and Predicting Non-linear *Brassica*
782 *rapa* Growth Phenotypes Based on Bayesian and Frequentist Complex Trait Estimation. G3
783 Genes|Genomes|Genetics 8:1247-1258.
784 Bates D., Maechler M., Bolker B., Walker S., Christensen R. H. B. C., *et al.*, 2018 lme4: Linear
785 Mixed-Effects Models using "Eigen" and S4.
786 Bowman J. L., Smyth D. R., Meyerowitz E. M., 2012 The ABC model of flower development:
787 Then and now. Development 139: 4095 LP-4098.
788 Brock M. T., Weinig C., 2007 Plasticity and Environment-Specific Covariances: An investigation
789 of floral-vegetative and within flower correlations. Evolution (N. Y). 61: 2913–2924.
790 Broman K. W., Wu H., Sen S., Churchill G. A., 2003 R/qtl: QTL mapping in experimental crosses.
791 Bioinformatics 19: 899–890.
792 Broman K. W., Sen S., 2009 *A guide to QTL Mapping with R/qtl*. Springer, New York.
793 Brown C. D., Mangravite L. M., Engelhardt B. E., 2013 Integrative Modeling of eQTLs and Cis-
794 Regulatory Elements Suggests Mechanisms Underlying Cell Type Specificity of eQTLs (G
795 Gibson, Ed.). PLoS Genet. 9: e1003649.

796 Cheng F., Liu S., Wu J., Fang L., Sun S., *et al.*, 2011 BRAD, the genetics and genomics database
797 for *Brassica* plants. BMC Plant Biol. 11: 136.
798 Chib S., Greenberg E., 1995 Understanding the Metropolis-Hastings Algorithm. Am. Stat. 49:
799 327–335.
800 Clark T. J., Lu Y., 2015 Analysis of Loss-of-Function Mutants in Aspartate Kinase and Homoserine
801 Dehydrogenase Genes Points to Complexity in the Regulation of Aspartate-Derived Amino
802 Acid Contents. Plant Physiol. 168: 1512–1526.
803 Csilléry K., Rodríguez-Verdugo A., Rellstab C., Guillaume F., 2018 Detecting the genomic signal
804 of polygenic adaptation and the role of epistasis in evolution. Mol. Ecol. 27: 606–612.
805 Devisetty U. K., Covington M. F., Tat A. V, Lekkala S., Maloof J. N., 2014 Polymorphism
806 Identification and Improved Genome Annotation of *Brassica rapa* Through Deep RNA
807 Sequencing. G3 Genes|Genomes|Genetics 4: 2065–2078.
808 Doerge R. W., Churchill G. A., 1996 Permutation Tests for Multiple Loci Affecting a Quantitative
809 Character. Genetics 142: 285 LP-294.
810 Doss S., Schadt E. E., Drake T. A., Lusis A. J., 2005 Cis-acting expression quantitative trait loci in
811 mice. Genome Res. 15: 681–691.
812 Gibson G., Weir B., 2005 The quantitative genetics of transcription. Trends Genet. 21: 616–623.
813 Goncalves A., Leigh-Brown S., Thybert D., Stefflova K., Turro E., *et al.*, 2012 Extensive
814 compensatory cis-trans regulation in the evolution of mouse gene expression. Genome
815 Res. 22: 2376–2384.
816 Griswold C. K., Gomulkiewicz R., Heckman N., 2008 Hypothesis testing in comparative and
817 experimental studies of Function-Valued Traits. Evolution (N. Y). 62: 1229–1242.
818 Gutensohn M., Klempien A., Kaminaga Y., Nagegowda D. A., Negre-Zakharov F., *et al.*, 2011 Role
819 of aromatic aldehyde synthase in wounding/herbivory response and flower scent
820 production in different *Arabidopsis* ecotypes. Plant J. 66: 591–602.
821 Halekoh U., Højsgaard S., 2014 A Kenward-Roger Approximation and Parametric Bootsrap
822 Methods for Tests in Linear Mixed Models - The R Package pbkrtest. J. Stat. Software2 59:
823 1–30.
824 Hammer G. L., Chapman S., Oosterom E. van, Podlich D. W., 2005 Trait physiology and crop
825 modelling as a framework to link phenotypic complexity to underlying genetic systems.
826 Aust. J. Agric. Res. 56: 947–960.
827 Hammond J. P., Mayes S., Bowen H. C., Graham N. S., Hayden R. M., *et al.*, 2011 Regulatory
828 Hotspots Are Associated with Plant Gene Expression under Varying Soil Phosphorus Supply
829 in *Brassica rapa*. Plant Physiol. 156: 1230 LP-1241.
830 Hansen B. G., Halkier B. A., Kliebenstein D. J., 2008 Identifying the molecular basis of QTLs:
831 eQTLs add a new dimension. Trends Plant Sci. 13: 72–77.
832 Hernandez K. M., 2015 Understanding the genetic architecture of complex traits using the
833 function-valued approach. New Phytol. 208: 1–3.
834 Hitzemann R., Malmanger B., Reed C., Lawler M., Hitzemann B., *et al.*, 2003 A strategy for the
835 integration of QTL, gene expression, and sequence analyses. Mamm. Genome 14: 733–
836 747.
837 Iniguez-Luy F. L., Lukens L., Farnham M. W., Amasino R. M., Osborn T. C., 2009 Development of
838 public immortal mapping populations, molecular markers and linkage maps for rapid
839 cycling *Brassica rapa* and *B. oleracea*. Theor. Appl. Genet. 120: 31–43.

840 Jaffrézic F., Pletcher S. D., 2000 Statistical models for estimating the genetic basis of repeated
841 measures and other function-valued traits. Genetics 156: 913–922.
842 Jiang L., Clavijo J. A., Sun L., Zhu X., Bhakta M. S., *et al.*, 2015 Plastic expression of heterochrony
843 quantitative trait loci (hQTLs) for leaf growth in the common bean (*Phaseolus vulgaris*).
844 New Phytol. 207: 872–882.
845 Kingsolver J. G., Gomulkiewicz R., Carter P. A., 2001 Variation, selection and evolution of
846 function-valued traits. Genetica 112: 87–104.
847 Klingenberg C. P., 2014 Studying morphological integration and modularity at multiple levels:
848 concepts and analysis. Philos. Trans. R. Soc. B Biol. Sci. 369: 20130249.
849 Kokichi H., Shyam P., 1984 Ethnobotany and Evolutionary Origin of Indian Oleiferous Brassicae.
850 Indian J. Genet. Plant Breed. 44: 102–112.
851 Kruschke J. K., 2014 *Doing Bayesian Data Analysis: A tutorial with R, BUGS, and Stan*. Academic
852 Press.
853 Kulbaba M. W., Clocher I. C., Harder L. D., 2017 Inflorescence characteristics as function-valued
854 traits: Analysis of heritability and selection on architectural effects. J. Syst. Evol. 55: 559–
855 565.
856 Kumar R., Ichihashi Y., Kimura S., Chitwood D. H., Headland L. R., *et al.*, 2012 A High-Throughput
857 Method for Illumina RNA-Seq Library Preparation. Front. Plant Sci. 3: 202.
858 Kuznetsova A., Brockhoof P. B., Christensen R. H. B. C., 2018 lmerTest: Tests in Linear Mixed
859 Effects Models.
860 Langfelder P., Horvath S., 2008 WGCNA: an R package for weighted correlation network
861 analysis. BMC Bioinformatics 9: 559.
862 Law C. N., Snape J. W., Worland A. J., 1978 The genetical relationship between height and yield
863 in wheat. Heredity (Edinb). 40: 133.
864 Law C. W., Chen Y., Shi W., Smyth G. K., 2014 voom: precision weights unlock linear model
865 analysis tools for RNA-seq read counts. Genome Biol. 15: R29.
866 Li H., Durbin R., 2009 Fast and accurate short read alignment with Burrows-Wheeler transform.
867 Bioinformatics 25: 1754–1760.
868 Li P., Ponnala L., Gandotra N., Wang L., Si Y., *et al.*, 2010 The developmental dynamics of the
869 maize leaf transcriptome. Nat. Genet. 42: 1060.
870 Li R., Jeong K., Davis J. T., Kim S., Lee S., *et al.*, 2018 Integrated QTL and eQTL Mapping Provides
871 Insights and Candidate Genes for Fatty Acid Composition, Flowering Time, and Growth
872 Traits in a F2 Population of a Novel Synthetic Allopolyploid *Brassica napus*. Front. Plant Sci.
873 9: 1632.
874 Luo J., Xu P., Cao P., Wan H., Lv X., *et al.*, 2018 Integrating Genetic and Gene Co-expression
875 Analysis Identifies Gene Networks Involved in Alcohol and Stress Responses . Front. Mol.
876 Neurosci. 11: 102.
877 Mackay T. F. C., 2013 Epistasis and quantitative traits: using model organisms to study gene–
878 gene interactions. Nat. Rev. Genet. 15: 22.
879 Markelz R. J. C., Covington M. F., Brock M. T., Devisetty U. K., Kliebenstein D. J., *et al.*, 2017
880 Using RNA-seq for Genomic Scaffold Placement, Correcting Assemblies, and Genetic Map
881 Creation in a Common *Brassica rapa* Mapping Population. G3 Genes|Genomes|Genetics.
882 Munkvold J. D., Laudencia-Chingcuanco D., Sorrells M. E., 2013 Systems Genetics of
883 Environmental Response in the Mature Wheat Embryo. Genetics 194: 265–277.

884 Nakamura Y., Tsuchiya M., Ohta H., 2007 Plastidic Phosphatidic Acid Phosphatases Identified in
885 a Distinct Subfamily of Lipid Phosphate Phosphatases with Prokaryotic Origin. J. Biol.
886 Chem. 282: 29013–29021.
887 Nozue K., Devisetty U. K., Lekkala S., Mueller-Moule P., Bak A., *et al.*, 2018 Network analysis
888 reveals a role for salicylic acid pathway components in shade avoidance. Plant Physiol.
889 O'Quin K. E., Schulte J. E., Patel Z., Kahn N., Naseer Z., *et al.*, 2012 Evolution of cichlid vision via
890 trans-regulatory divergence. BMC Evol. Biol. 12: 251.
891 Obayashi T., Kinoshita K., 2009 Rank of Correlation Coefficient as a Comparable Measure for
892 Biological Significance of Gene Coexpression. DNA Res. 16: 249–260.
893 Patil A., Huard D., Fonnesbeck C., 2010 PyMC: Bayesian Stochastic Modelling in Python. J. Stat.
894 Software, Artic. 35: 1–81.
895 Ponsuksili S., Siengdee P., Du Y., Trakooljul N., Murani E., *et al.*, 2015 Identification of Common
896 Regulators of Genes in Co-Expression Networks Affecting Muscle and Meat Properties.
897 PLoS One 10: e0123678.
898 Prioul J.-L., Quarrie S., Causse M., Vienne D. de, 1997 Dissecting complex physiological functions
899 through the use of molecular quantitative genetics. J. Exp. Bot. 48: 1151–1163.
900 R Core Team, 2016 R: A language and environment for statistical computing. R Foundation for
901 Statistical Computing.
902 Raines C. A., Paul M. J., 2006 Products of leaf primary carbon metabolism modulate the
903 developmental programme determining plant morphology. J. Exp. Bot. 57: 1857–1862.
904 Reiser L., Modrusan Z., Margossian L., Samach A., Ohad N., *et al.*, 1995 The *BELL1* gene encodes
905 a homeodomain protein involved in pattern formation in the *Arabidopsis* ovule
906 primordium. Cell 83: 735–742.
907 Robinson M. D., Oshlack A., 2010 A scaling normalization method for differential expression
908 analysis of RNA-seq data. Genome Biol. 11: R25.
909 Rubin G., Tohge T., Matsuda F., Saito K., Scheible W.-R., 2009 Members of the *LBD* Family of
910 Transcription Factors Repress Anthocyanin Synthesis and Affect Additional Nitrogen
911 Responses in *Arabidopsis*. Plant Cell 21: 3567–3584.
912 Rutjens B., Bao D., Eck-Stouten E. Van, Brand M., Smeekens S., *et al.*, 2009 Shoot apical
913 meristem function in *Arabidopsis* requires the combined activities of three *BEL1*-like
914 homeodomain proteins. Plant J. 58: 641–654.
915 Schadt E. E., Monks S. A., Drake T. A., Lusis A. J., Che N., *et al.*, 2003 Genetics of gene expression
916 surveyed in maize, mouse and man. Nature 422: 297.
917 Schadt E. E., Molony C., Chudin E., Hao K., Yang X., *et al.*, 2008 Mapping the Genetic
918 Architecture of Gene Expression in Human Liver (G Abecassis, Ed.). PLoS Biol. 6: e107.
919 Schaefer R., Michno J.-M., Jeffers J., Hoekenga O. A., Dilkes B. P., *et al.*, 2018 Integrating co-
920 expression networks with GWAS to prioritize causal genes in maize. Plant Cell.
921 Schmid M., Davison T. S., Henz S. R., Pape U. J., Demar M., *et al.*, 2005 A gene expression map
922 of *Arabidopsis thaliana* development. Nat. Genet. 37: 501.
923 Schmitt J., Stinchcombe J. R., Heschel M. S., Huber H., 2003 The Adaptive Evolution of Plasticity:
924 Phytochrome-Mediated Shade Avoidance Responses. Integr. Comp. Biol. 43: 459–469.
925 Schneidereit J., Häusler R. E., Fiene G., Kaiser W. M., Weber A. P. M., 2005 Antisense repression
926 reveals a crucial role of the plastidic 2-oxoglutarate/malate translocator DiT1 at the
927 interface between carbon and nitrogen metabolism. Plant J. 45: 206–224.

928 Signor S. A., Nuzhdin S. V, 2018 The Evolution of Gene Expression in cis and trans. Trends Genet.
929 34: 532–544.
930 Stinchcombe J. R., Izem R., Shane H. M., McGoey B. V., Schmitt J., 2010 Across-Environment
931 genetic correlations and the frequency of selective environments shape the evolutionary
932 dynamics of growth rate in *Impatiens capensis*. Evolution (N. Y). 64: 2887–2903.
933 Stinchcombe J. R., Kirkpatrick M., 2012 Genetics and evolution of function-valued traits:
934 understanding environmentally responsive phenotypes. Trends Ecol. Evol. 27: 637–647.
935 Swanson-Wagner R. A., DeCook R., Jia Y., Bancroft T., Ji T., *et al.*, 2009 Paternal Dominance of
936 Trans-eQTL Influences Gene Expression Patterns in Maize Hybrids. Science (80-.). 326:
937 1118 LP-1120.
938 Tanger P., Klassen S., Mojica J. P., Lovell J. T., Moyers B. T., *et al.*, 2017 Field-based high
939 throughput phenotyping rapidly identifies genomic regions controlling yield components in
940 rice. Sci. Rep. 7: 42839.
941 Tian J., Keller M. P., Broman A. T., Kendziorski C., Yandell B. S., *et al.*, 2016 The Dissection of
942 Expression Quantitative Trait Locus Hotspots. Genetics 202: 1563 LP-1574.
943 Vigil M. F., Anderson R. L., Beard W. E., 1997 Base Temperature and Growing-Degree-Hour
944 Requirements for the Emergence of Canola. Crop Sci. 37: 844–849.
945 Voorrips R. E., 2002 MapChart: Software for the Graphical Presentation of Linkage Maps and
946 QTLs. J. Hered. 93: 77–78.
947 West M. A. L., Kim K., Kliebenstein D. J., Leeuwen H. van, Michelmore R. W., *et al.*, 2007 Global
948 eQTL Mapping Reveals the Complex Genetic Architecture of Transcript-Level Variation in
949 *Arabidopsis*. Genetics 175: 1441–1450.
950 Wittkopp P. J., Haerum B. K., Clark A. G., 2008 Regulatory changes underlying expression
951 differences within and between *Drosophila* species. Nat. Genet. 40: 346.
952 Wu W. R., Li W. M., Tang D. Z., Lu H. R., Worland A. J., 1999 Time-related mapping of
953 quantitative trait loci underlying tiller number in rice. Genetics 151: 297–303.
954 Wu R., Lin M., 2006 Functional mapping — how to map and study the genetic architecture of
955 dynamic complex traits. Nat. Rev. Genet. 7: 229.
956 Yin X., McClure M. A., Jaja N., Tyler D. D., Hayes R. M., 2011 In-Season Prediction of Corn Yield
957 Using Plant Height under Major Production Systems. Agron. J. 103: 923–929.
958 Zeng Z. B., 1993 Theoretical basis for separation of multiple linked gene effects in mapping
959 quantitative trait loci. Proc. Natl. Acad. Sci. U. S. A. 90: 10972–10976.
960 Zhang B., Horvath S., 2005 A General Framework for Weighted Gene Co-Expression Network
961 Analysis. Stat. Appl. Genet. Mol. Biol. 4.
962 Zhu J., Fu X., Koo Y. D., Zhu J.-K., Jenney F. E., *et al.*, 2007 An Enhancer Mutant of *Arabidopsis*
963 Salt Overly Sensitive 3 Mediates both Ion Homeostasis and the Oxidative Stress Response.
964 Mol. Cell. Biol. 27: 5214–5224.
965 Zhu Z., Zhang F., Hu H., Bakshi A., Robinson M. R., *et al.*, 2016 Integration of summary data from
966 GWAS and eQTL studies predicts complex trait gene targets. Nat. Genet. 48: 481.
967