Plant, Cell & Environment

Plant, Cell and Environment (2015) **38**, 585–599 doi: 10.1111/pce.12418

Original Article

Quantitative trait loci and candidate genes underlying genotype by environment interaction in the response of *Arabidopsis thaliana* to drought

Mohamed El-Soda[1]*, Willem Kruijer[2], Marcos Malosetti[2], Maarten Koornneef[1,3] & Mark G. M. Aarts[1]

[1]Laboratory of Genetics, [2]Biometris – Applied Statistics, Department of Plant Science, Wageningen University, Wageningen 6708PB, The Netherlands and [3]Department of Plant Breeding and Genetics, Max Planck Institute for Plant Breeding Research, Cologne 50829, Germany

ABSTRACT

Drought stress was imposed on two sets of *Arabidopsis thaliana* genotypes grown in sand under short-day conditions and analysed for several shoot and root growth traits. The response to drought was assessed for quantitative trait locus (QTL) mapping in a genetically diverse set of Arabidopsis accessions using genome-wide association (GWA) mapping, and conventional linkage analysis of a recombinant inbred line (RIL) population. Results showed significant genotype by environment interaction (G×E) for all traits in response to different watering regimes. For the RIL population, the observed G×E was reflected in 17 QTL by environment interactions (Q×E), while 17 additional QTLs were mapped not showing Q×E. GWA mapping identified 58 single nucleotide polymorphism (SNPs) associated with loci displaying Q×E and an additional 16 SNPs associated with loci not showing Q×E. Many candidate genes potentially underlying these loci were suggested. The genes for *RPS3C* and *YLS7* were found to contain conserved amino acid differences when comparing Arabidopsis accessions with strongly contrasting drought response phenotypes, further supporting their candidacy. One of these candidate genes co-located with a QTL mapped in the RIL population.

Key-words: genome-wide association mapping; G×E; QTL mapping; Q×E.

INTRODUCTION

Abiotic stresses, including drought, negatively affect plant growth and limit crop productivity. To cope with the negative effects of drought, plants have evolved one or more adaptive strategies. This adaptation is expected to reveal a phenotypic plasticity response, in which significant phenotypic changes occur because of drought. If the drought stress has a differential effect on the phenotype of different genotypes, meaning there is genetic variation for the phenotypic plasticity, this will be expressed as genotype by environment interaction (G×E; Assmann 2013). G×E can be genetically dissected into its underlying quantitative trait loci (QTL), illustrating their interactions with the environment (QTL by environment interactions; Q×E), by incorporating the environmental factors in the genetic analysis (Assmann 2013; Juenger 2013; El-Soda *et al*. 2014). The key development in genetically dissecting such complex responses is the analysis of trait-marker associations via conventional linkage mapping and genome-wide association (GWA) mapping (Weigel 2012). While a GWA study (GWAS) allows high accuracy mapping of trait associated loci, when compared with conventional linkage analysis, it is often not powerful enough to detect the effect of rare alleles in the association panel, even if they have large phenotypic effects (Eichler *et al*. 2010; Gibson 2012). Therefore, using both GWAS and traditional linkage mapping to analyse a trait is an attractive combination of methods to have high accuracy mapping while accounting for false positives and avoiding false negatives (Nordborg & Weigel 2008; Atwell *et al*. 2010; Bergelson & Roux 2010; Brachi *et al*. 2010; Sterken *et al*. 2012; Weigel 2012).

Considering G×E in QTL mapping increases the statistical power to detect more QTLs with better explained variance, compared with methods in which GxE is not considered (Joosen *et al*. 2012; El-Soda *et al*. 2014). Based on the phenotypic effect of a QTL in all tested environments, it can be classified as a main effect QTL, with a comparable effect on the phenotype regardless of the environment, or as an environment-specific QTL (indicating Q×E), affecting the phenotype in one environment, but not, or to a different degree, in another (El-Soda *et al*. 2014). In case a QTL effect in one environment is undetectable, it is referred to as a conditionally neutral QTL. When a trait is controlled by several of such QTLs, it will be difficult to use the QTL mapping data obtained in one condition to predict the QTL effects in another without knowledge of G×E (Kamoshita *et al*. 2002; Malosetti *et al*. 2004, 2013; Tuberosa & Salvi 2006; Holland 2007; Tuberosa 2012; Zhao & Xu 2012). The commonly used approach for GWAS in plants is a univariate analysis for each trait and treatment. This will identify QTLs with main effects, but does not account for G×E. Therefore, a multi-trait mixed model (MTMM) approach was recently proposed for multi-trait or multi-environment association

Correspondence: M. G. M. Aarts. e-mail: mark.aarts@wur.nl

*Present address: Department of Genetics, Faculty of Agriculture, Cairo University, Cairo, Egypt.

mapping (Korte *et al*. 2012). In this approach, a marker can have different effects in different environments, therefore explaining at least part of G×E in terms of Q×E.

Arabidopsis has previously been used for the identification of drought tolerance QTLs (McKay *et al*. 2003, 2008; Tisné *et al*. 2010). Those studies addressed Arabidopsis response to drought using short-day conditions to avoid the complication of induction of flowering, which is enhanced by growing in long days and which may considerably change plant physiology (Andres & Coupland 2012). However, naturally growing Arabidopsis plants experience spells of water shortage in the spring (short days) and summer (long days) months, but the response of flowering time (FT) QTL to drought stress, in the form of Q×E, has not been addressed so far.

We assessed the drought response of Arabidopsis using a GWA mapping panel consisting of 350 genetically diverse accessions collected from all over its natural distribution range (Li *et al*. 2010). To reflect the whole plant performance under drought stress, rosette and root traits were measured. The MTMM approach we used identified several main effect QTLs as well as Q×E, including candidate genes. In addition, we assessed drought response in a recombinant inbred lines (RIL) population based on a cross between Sha and Col-0 (Simon *et al*. 2008), two accessions with contrasting phenotypes for root length response to drought, to confirm some of the QTLs identified in the GWA mapping. Both populations were grown under short days, but we also grew the RIL population under long days, to assess the effect of FT on drought response.

MATERIALS AND METHODS

Mapping population and experimental set-up

A GWA mapping population consisting of 360 diverse accessions selected from a global collection of Arabidopsis accessions, genotyped with over 200 000 bi-allelic (Col or non-Col) non-singleton single nucleotide polymorphism (SNPs; Li *et al*. 2010; http://www.naturalvariation.org/hapmap) was used for GWA QTL analysis. Out of this set, lines CS76117, CS76118, CS76130, CS76178, CS76204, CS76237, CS76241, CS76246, CS76271 and CS76294 were not grown. A population of 164 F6 lines from a Sha × Col RILs population, genotyped with 86 SNP markers (Simon *et al*. 2008), was used for conventional QTL analysis. The experiments were performed in a completely randomized block design with three replicate blocks. All experiments were conducted in the summer/autumn of 2011 in a temperature and day length-controlled greenhouse with an average temperature of 21.3 and 17.5 °C, and an average relative humidity of 69 and 74%, during day and night, respectively.

Both populations were grown under short-day (SD) conditions (10 h light), and harvested before flowering. Plants were grown for 34 d in 4 × 4 × 7 cm (length × width × height) plastic pots containing silver sand covered with a thin layer of sieved peat. The RIL population was also grown under long-day (LD) conditions (20 h light), at which nearly all lines flowered at the end of the experiment. To ensure drying rates comparable with the SD experiment, plants in the LD experiment were grown for 27 d on a 1:1 mix of sand and peat in 4 × 4 × 5 cm plastic pots. An additional SD experiment was carried out on the RIL population grown under well-watered conditions, in silver sand covered with thin layer of sieved soil, to determine FT and rosette fresh weight (RosFW) on the same day when the first flower appeared.

At the start of each experiment, all pots were watered with nutrient solution (contained 1 mM N, 1.1 mM P and 5.9 mM K), at 100% soil water holding capacity (SWHC). Per pot, two seeds were sown and 3 d after germination, seedlings were thinned to one per pot. All pots were watered every 2–3 d when SWHC dropped below 80%. To induce drought, watering was stopped 12 d after germination in the SD experiment and 9 d after germination in the LD experiment. In the control treatment, pots were watered every 2 or 3 d, when SWHC was less than 80%, until the end of the experiment.

At the end of the experiments, rosettes were harvested and RosFW was measured. Rosettes were oven-dried at 65 °C for 3 d for rosette dry weight (RosDW) measurements. Water content (WC) was calculated as WC = (RosFW – RosDW)/RosDW. Only for the SD experiment, roots were washed carefully, placed in a plastic tray filled with water and scanned with a flatbed scanner. Total root system length (RL) was measured from the scans using WinRhizo (Regent Instruments Inc., Quebec, Canada) and thereafter root DW (RDW) was determined. Finally, the ratio between RL and RosDW (RL/RosDW) was calculated.

Statistical and quantitative trait loci analysis

Statistical analysis was performed using GenStat for Windows, 15th Edition (VSN International Ltd., Hemel Hempstead, UK). For each experiment, analysis of variance (ANOVA) was used to test the significance of differences between treatments and lines and for the G×E. Broad-sense heritability was estimated as the ratio between the genetic variance Vg, and the total phenotypic variance $Vt = Vg + Ve$. Vg and the residual variance (Ve) are estimated by, respectively, $(MS(\text{genotype}) - MS(\text{residual}))/r$ and $MS(\text{residual})$. The terms $MS(\text{genotype})$ and $MS(\text{residual})$ are the mean sums of squares for RILs or accessions and residual error in an ANOVA, and r is the number of replicates.

Statistical models for linkage and GWA QTL detection

A general multi-environment mixed model approach for GWA and traditional linkage QTL detection was used (see van Eeuwijk *et al*. 2010 and Malosetti *et al*. 2013 for a detailed discussion of models). Here, we first give a general formulation of the models and testing procedures, and then we describe the specific features for GWAS and for linkage QTL analysis, respectively.

The general mixed model

The phenotypic multi-environment data consisted of single entries per genotype and per environment (so no replicates within environments). With the observations of the n_g

genotypes in n_e environments collected in a vector $y = [y_1, y_2, \ldots, y_{n_e}]'$, being y_1 all the observations in environment 1, and so on, the basic mixed model assuming fixed environments and random genotypic effects can be written as

$$y = X\beta + Zu. \quad (1)$$

The design matrices X and Z associate trait values with fixed and random effects, β is the vector of fixed effects including intercept and environmental main effects, and u is the vector of random genetic by environment effects, which are assumed to follow a multivariate normal distribution $N(0, \Sigma)$, Σ being a variance–covariance matrix reflecting genetic correlations between observations.

For QTL detection, Eqn 1 is extended with marker information to fit environment-specific QTL effects:

$$y = X\beta + X^*\alpha + Zu, \quad (2)$$

where $X^* = I_{n_e} \otimes M$ is the extra design matrix, I_{n_e} is the identity matrix of length n_e, and M is a column vector of length n_g with genotypic scores derived from molecular marker information. In its simplest form, and with two alleles per marker (A and a), the additive genetic score consists of the number of A alleles in each genotype. The vector α contains the environment-specific QTL additive effects. Note that in our case, with $n_e = 2$, the effects are $\alpha = [\alpha_1, \alpha_2]'$ and correspond to the QTL effect under control and under drought conditions, respectively. Inference about QTL was made by the following hypotheses testing procedure:

1. Perform a global test for the presence of a QTL at the specific position by assessing the $H0 : \alpha_i = 0 \forall i$. The hypothesis testing involved a multiple testing correction (see details in the GWAS and linkage analysis section). Rejection of the null hypothesis indicates a QTL is present with an effect in at least one of the environments.
2. When hypothesis 1 was rejected, a more specific test for a differential QTL effect between environments (i.e. QTL by environment interaction) was performed. We did so by decomposing the QTL effect into a main effect α^* (consistent QTL effect across conditions), and interaction effects $\alpha_i^* (i = 1 \ldots n_e)$. Note that the genetic score design matrix becomes $X^* = [1_{n_e} \otimes M, I_{n_e} \otimes M]$, and $\alpha = [\alpha^*, \alpha_1^*, \ldots, \alpha_{n_e}^*]$, with the constraint $\sum \alpha_i^* = 0$, 1_{n_e} being the vector of ones of length n_e. The test for a significant QTL by environment interaction was assessed by the $H_0: \alpha_i^* = 0 \; \forall \; i$. This test was performed at a type I error rate of 0.05.
3. Depending on the results of the hypothesis 2, the QTL effects were reported as environment-specific QTL effects $\left(\alpha_i = a^* + \alpha_i^*\right)$ when QTL by environment interaction was significant, or as main QTL effects (α^*) when no significant QTL by environment interaction was found.

The GWA model

The major feature in the GWA mixed model is that of an appropriate set-up of the variance–covariance model as discussed by Korte *et al*. (2012) to reflect the genetic correlation introduced by the genetic relatedness between individuals in the population. The covariance structure given by Σ should accommodate genetic correlations imposed by the kinship information in addition to that caused by the repeated measures across conditions: $\Sigma = \begin{bmatrix} \sigma_1^2 K & \sigma_{12} K \\ \sigma_{12} K & \sigma_2^2 K \end{bmatrix} + \begin{bmatrix} \sigma_{e1}^2 I_{n_g} & 0 \\ 0 & \sigma_{e2}^2 I_{n_g} \end{bmatrix}$, with K an identity-by-state matrix obtained from SNP information. Marker-trait association was assessed with type I error rate of 10^{-4} (for visualization purposes, associated P-values were displayed in a minus log10 scale: $-\log10(P) = 4$). A minor allele frequency of 0.05 was used to discard markers with very rare alleles. Korte *et al*. (2012) showed that it is straightforward to extend this model to situations where some observations are missing and only available for one of the environments. This extension is however not implemented in the original software; therefore, we added this feature to the R-code, which is available on request.

QTL×E linkage QTL mapping

The major feature in the linkage analysis is that of the construction of genetic scores not only at marker positions, but also in between markers. Using linkage information and hidden Markov chain methods (Jiang & Zeng 1997) conditional identity-by-descend (IBD) probabilities can be estimated at any position on the genome, that is, at and in between markers (for details of a RIL population, see Boer *et al*. 2007). The conditional IBD probabilities were used to form the design matrix $X^* = I_{n_e} \otimes M$, the elements of M being the difference of the probabilities on the two homozygous genotypes. The probabilities were calculated from marker information, as implemented in GenStat 15. With respect to the variance–covariance model still covariances between observations in the two treatments should be accounted for, so in this case $\Sigma = \begin{bmatrix} \sigma_1^2 I_{n_g} & \sigma_{12} I_{n_g} \\ \sigma_{12} I_{n_g} & \sigma_2^2 I_{n_g} \end{bmatrix}$.

The QTL search (i.e. IBD estimation) was done at a step size of 10 cM. We allowed a number of cofactors in the model to control for genetic background noise (Jansen & Stam 1994; Zeng 1994), allowing a minimum cofactor proximity of 50 cM. For automatic selection of QTLs, a minimum separation of selected QTLs of 25 cM was used. A multiple testing threshold value of $-\log10 = 2.8$ was calculated based on the approach implemented in GenStat (Li & Ji 2005), with 0.05 set as the genome-wide type I error level. The allelic effect of each QTL in each environment, the effect of Q × E, and the explained phenotypic variance of each QTL per environment were determined by fitting a final multi-QTL model after running a backward selection using all candidate QTLs from the last composite interval mapping round (two rounds of CIM performed).

The identification of candidate genes by the analysis of sequence polymorphisms

To provide experimental support for candidate genes, the publicly available sequences of accessions of the GWA

mapping population with strong contrasting phenotypes were investigated. The 350 accessions were classified based on the drought: control RosDW and RL ratios. For each classification, two contrasting groups were made, each composed of 10 accessions, one with the smallest (drought sensitive) and another with the largest (drought tolerant) RosDW or RL ratios. Both sets were compared regarding the predicted amino acid sequences of candidate genes, using the Arabidopsis 1001 genomes browser http://signal.salk.edu/atg1001/3.0/gebrowser.php.

RESULTS

Comparing two mapping populations grown in well-watered and drought treatments in SD conditions

Both GWA and RIL mapping populations were phenotyped for plant performance traits such as RosFW, RosDW, RL, RDW and RL/RosDW. When the trait averages of the accessions in the GWA population were considered, the values of RosFW and RosDW were higher in the control treatment than in the drought treatment, while those for RL/RosDW were lower (Fig. 1a). ANOVA showed a significant difference between control and drought treatments for all traits (Supporting Information Table S1). A correlation analysis showed a positive correlation between all measured traits in both treatments, except for RL/RosDW, which showed a negative correlation with RosDW in the drought treatment (Supporting Information Table S2).

Frequency distributions of the measured traits for the RIL population showed transgression beyond both parental lines for all traits except for RosFW, where the transgression was only in one direction (Fig. 1b,c). Like for the GWA population, RosFW and RosDW were higher in the control than in the drought treatment, whereas RL/RosDW was lower (Supporting Information Table S3). When the phenotypic variations for traits observed in both the GWA and RIL populations are compared (Fig. 1), they are in the same range, except for RosFW, which is somewhat higher in the RIL population. A correlation analysis for the RIL population (Supporting Information Table S4), comparing the same traits in control and drought treatments, showed that all traits were positively correlated. Considering correlations between different traits in either control or drought treatments revealed that RL was positively correlated with RosFW, RosDW and RDW. A remarkable difference between Col and Sha was that while RosDW of both did not differ much either under control or drought treatments, there was considerable difference for RL and RDW (Fig. 1), which can have interesting consequences for drought tolerance. Sha appears to invest in maintaining its root system under drought, showing little reduction in RL and RDW compared with the control treatment, while Col appears to invest less in its roots system under drought, showing strong reduction in RL and RDW.

GWA mapping and G×E

For the five examined traits, we identified 74 SNPs with $-\log10(P) > 4$, of which 58 SNPs showed a significant Q×E effect while 16 SNPs showed a main effect (Fig. 2). These 74 SNPs correspond to 69 loci, each containing several genes when considering the region in linkage disequilibrium with the significant SNPs (Supporting Information Table S5). For further analysis, we only considered the loci showing a significant Q×E effect and within those, the genes that have functions related to abiotic stress response, based on their biological function as described in The Arabidopsis Information Resource (TAIR; www.arabidopsis.org).

A SNP in the *TARGET OF RAPAMYCIN* (*TOR*; At1g50030) gene (Dobrenel *et al*. 2011; Ren *et al*. 2012; Caldana *et al*. 2013) was associated with RosDW. For this SNP, the non-Col allele increased RosDW in both treatments, but the increase was three times higher in the drought than in the control environments. A SNP associated with RDW was mapped to the *SNF1-RELATED PROTEIN KINASE 2.2* (*SnRK2.2*; At3g50500) gene (Fujii *et al*. 2007, 2011; Kulik *et al*. 2011). The Col allele of this SNP had a strong effect on RDW in the drought treatment, but a moderate effect, in opposite direction, in the control treatment. An additional association with RDW was found for a SNP mapped in the *SHORT ROOT HAIR 1* (*SRH1*; At4g34580) gene (Huang *et al*. 2013), again with a stronger effect in the drought treatment and an opposite effect in the control treatment. Similar opposite effects are seen for several of the associated SNPs (Supporting Information Table S5).

Mapping QTLs in the RIL population

Like for the GWA population, QTLs were mapped for the measured traits in the Sha × Col RIL population, and classified as main effect QTLs or QTLs with significant Q×E based on the absolute effect of each allele on the trait value in every treatment (Table 1). In total, 27 QTLs were mapped for the measured traits (Fig. 3). Out of the 15 QTLs mapped for shoot traits, the large majority, 13 QTLs, showed significant Q×E related to the drought treatment (Table 1). Of the 12 QTLs mapped for root traits, only two showed significant Q×E related to drought (Table 2). Some of the QTLs identified in the RIL population co-localized with significant SNPs identified in the GWAS. For example, the *RDW4* QTL co-localized with a SNP associated with the *SHORT ROOT HAIR 1* (*SRH1*; At4g34580) gene. The *FW1*, *FW4* and *DW2* QTLs co-located with SNPs associated with, respectively, *TARGET OF RAPAMYCIN* (*TOR*; At1g50030), *CYTOCHROME P450 83A1* (*CYP83A1*; At4g13770) and the sulphate transporter gene *SULTR4;1* (At5g13550).

The effects of FT on the response to drought in LD conditions

Since Arabidopsis may experience drought both in early spring as in summer, thus with different photoperiods, we also grew the RIL population under control and drought

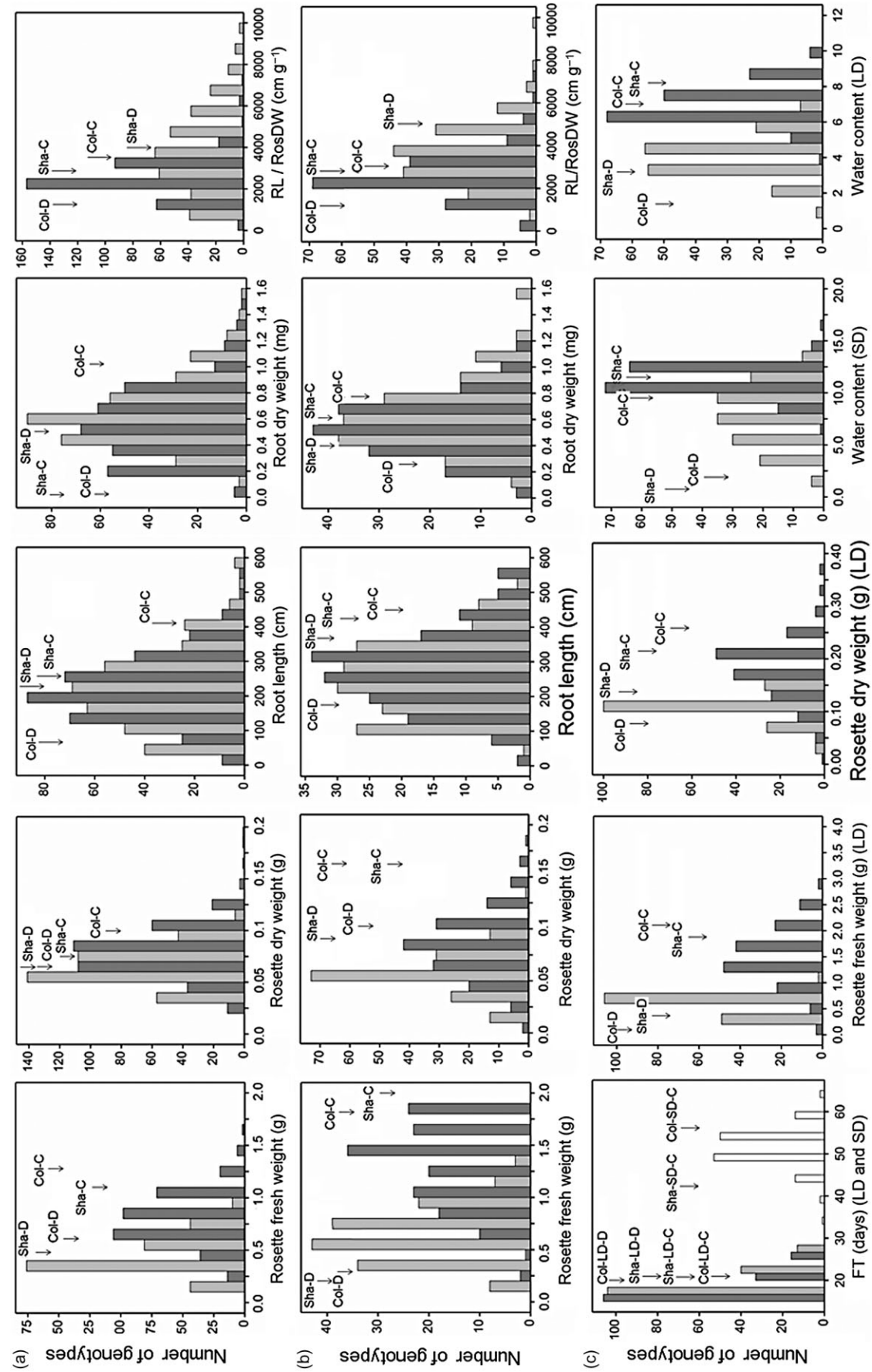

Figure 1. Frequency distributions of the non-normalized values of traits phenotyped for the GWA population of Arabidopsis accessions grown in short days (SD) (a); and the Sha × Col Arabidopsis RIL population grown in SD (b); as well as of additional traits phenotyped for the Sha × Col RIL population grown either in SD or long days (LD) (c). Sha-C and Col-C refer to the values of the parental lines Sha and Col grown in control treatments and Sha-D and Col-D to the values of the parental lines grown in drought treatments, all indicated with arrows. Data for control treatments are presented with dark grey bars and data from drought treatments with light grey bars. The white bars indicate flowering time in the SD flowering experiment (only control treatment). The vertical axes indicate the numbers of genotypes per trait value class and the horizontal axes indicate the different trait value classes. RL/RosDW is the ratio between root length (RL) and rosette dry weight (RosDW); water content is calculated as (fresh weight – dry weight)/dry weight.

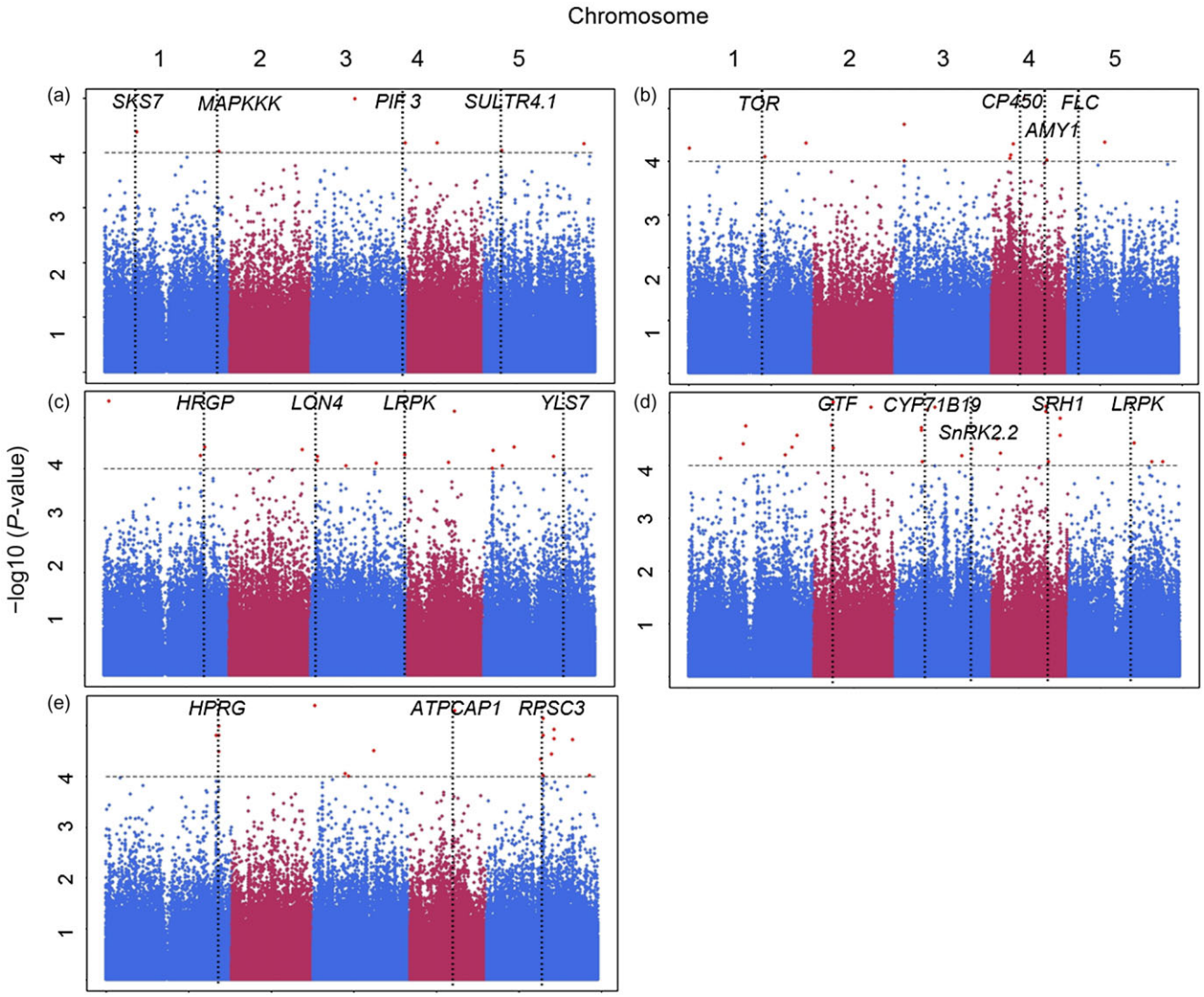

Figure 2. GWA mapping results of the MTMM approach showing the –log10(*P*) values for SNPs associated with rosette fresh weight (a), rosette dry weight (b), root length (c), root dry weight (d) and the ratio between root length and rosette dry weight (RL/RosDW) (e). In each panel, the SNPs corresponding to the five Arabidopsis chromosomes are indicated in alternating blue/purple colours, with the horizontal axes indicating genome sequence positions. Vertical axes indicate –log10(*P*) values. The –log10(*P*) significance threshold of 4 is indicated with a horizontal dashed line. Vertical dashed lines indicate significant SNPs in genes with annotated functions, as reported in Supporting Information Table S5. SNPs in *Yellow Leaf Senecence7* and *Flowering Locus C*, which were two loci within the significance range of relevant QTLs identified in the RIL population, are also indicated.

treatments using LD conditions. Next to FT, RosFW, RosDW and WC were determined (Fig. 1c; Supporting Information Table S3). The drought treatment had no significant effect on FT; however, two FT QTL, *QFT3* and *QFT4*, showed significant Q×E in response to drought (Table 1). QTLs mapped for RosFW, RosDW and WC in the LD experiment co-located with QTLs mapped in SD experiment; however, significant Q×E was observed for the majority of QTLs because of day length when comparing the two control and the two drought treatments. The *FW3*, *FW4* and *FW6* QTLs showed Q×E because of day length in the control treatments but not in the drought treatments, while the *FW5*, *DW2* and *WC5* QTLs showed Q×E because of day length only in the drought treatments and not in the control treatments.

The RIL population was also grown under well-watered SD conditions and left to flower to obtain SD FT data. This time, only FT and RosFW were scored (Supporting Information Table S3). FT in SD is positively correlated with FT in LD. Under SD conditions, the *QFT2* and *QFT6* loci explained most of the phenotypic variance, with *QFT2* showing Q×E, affecting FT more under SD than under LD conditions. For *QFT6*, the major FT QTL in LD, explaining around 30% of the phenotypic variance, the Col allele contributed to increased FT. *QFT3* was only mapped under control conditions, with the Col allele reducing FT. *QFT4* was mapped in all conditions with the Col allele contributing to early flowering in SD and the Sha allele contributing to early flowering in LD. *QFT4* co-localized with QTL for

Table 1. QTLs detected for shoot traits of the Sha × Col RIL population grown under short-day (SD) and long-day (LD) conditions in control and drought treatments using the MTMM approach

	QTL					SD (control)		SD (drought)		LD (control)		LD (drought)		SD flowering (control)	
Trait	Name	Chr.	Position in cM	–log10(*P*)	Q×E	effect	R^2	Effect	R^2	Effect	R^2	Effect	R^2	Effect	R^2
Rosette FW	*FW1*	1	51.2	4.0	ns	–0.036	0.7	–0.036	2.0	–0.036	0.5	–0.036	7.2	–0.036	0.1
	FW2	2	30.7	3.8	s	–0.113	6.7	–	–	–0.169	10.1	–0.023	3.6	–0.103	1.0
	FW3	4	0.0	3.1	s	–0.052	1.4	–0.036	2.0	–0.155	8.4	–0.026	3.7	–0.212	4.1
	FW4	4	30.8	2.2	s	–0.055	1.6	–0.013	0.3	–0.092	3.0	–0.013	1.0	–0.326	9.7
	FW5	5	30.5	4.4	s	–0.118	7.3	–0.080	9.6	–0.103	3.5	0.020	2.3	–	–
	FW6	5	87.3	8.6	s	–0.063	2.1	0.021	0.7	0.159	8.9	0.024	4.0	0.389	13.8
Rosette DW	*DW1*	2	30.7	3.0	s	–0.008	7	–0.003	1.9	–0.017	8.6	–0.005	5.0	–	–
	DW2	5	11.8	3.6	s	–0.007	4.5	–0.002	1.0	–0.009	2.3	–0.008	10.7	–	–
	DW3	5	87.3	5.1	s	–0.003	0.9	–0.001	0.3	0.013	5.5	0.004	2.8	–	–
Water content	*WC1*	1	65.7	5.0	ns	–0.235	2.9	–0.235	0.6	–0.235	4.7	–0.235	4.0	–	–
	WC2	2	12.2	1.8	s	–0.205	2.2	–0.180	0.4	–0.115	1.1	0.200	2.9	–	–
	WC3	3	63.4	2.7	s	–0.178	1.6	–0.391	1.8	0.148	1.9	–0.248	4.4	–	–
	WC4	4	35.5	6.3	s	–0.189	1.8	–0.073	0.1	–0.446	16.9	–0.056	0.2	–	–
	WC5	5	30.5	8.7	s	–0.262	3.5	–0.904	9.4	–0.255	5.5	0.445	14.2	–	–
	WC6	5	80.9	4.1	s	–0.229	2.7	0.598	4.1	0.288	7.0	0.056	0.2	–	–
Flowering time	*QFT1*	1	65.7	4.6	ns	–	–	–	–	–0.829	7.4	–0.829	8.0	–0.829	3.1
	QFT2	1	91.5	6.6	s	–	–	–	–	0.407	1.8	0.346	1.4	1.428	9.2
	QFT3	2	25.0	3.4	s	–	–	–	–	0.241	0.6	–	–	1.104	5.5
	QFT4	3	0.0	3.0	s	–	–	–	–	–0.199	0.4	–0.089	0.1	0.756	2.6
	QFT5	4	47.9	4.9	ns	–	–	–	–	–0.689	6.4	–0.689	7.1	–	–
	QFT6	5	9.0	15.3	ns	–	–	–	–	1.586	27.0	1.586	29.1	1.586	11.3
	QFT7	5	87.3	5.6	ns	–	–	–	–	0.985	10.4	0.985	11.2	0.985	4.4

Chromosome numbers (Chr.) and chromosomal positions (in centimorgans, cM) are indicated. The abbreviated trait name column ('name') lists the identified QTL names arranged according to chromosome number and position. '–log10(*P*)' indicates the QTL significance level. 's' and 'ns' refer to significant and non-significant QTL by environment interaction (Q×E) related to the watering condition. Positive 'effect' values indicate that the Col allele contributes to an increase in the trait value, while negative values indicate that the Sha allele contributes to a trait value increase. Units for effect are the same as the trait units. R^2 indicates the percentage of total phenotypic variance, which is explained by each QTL.

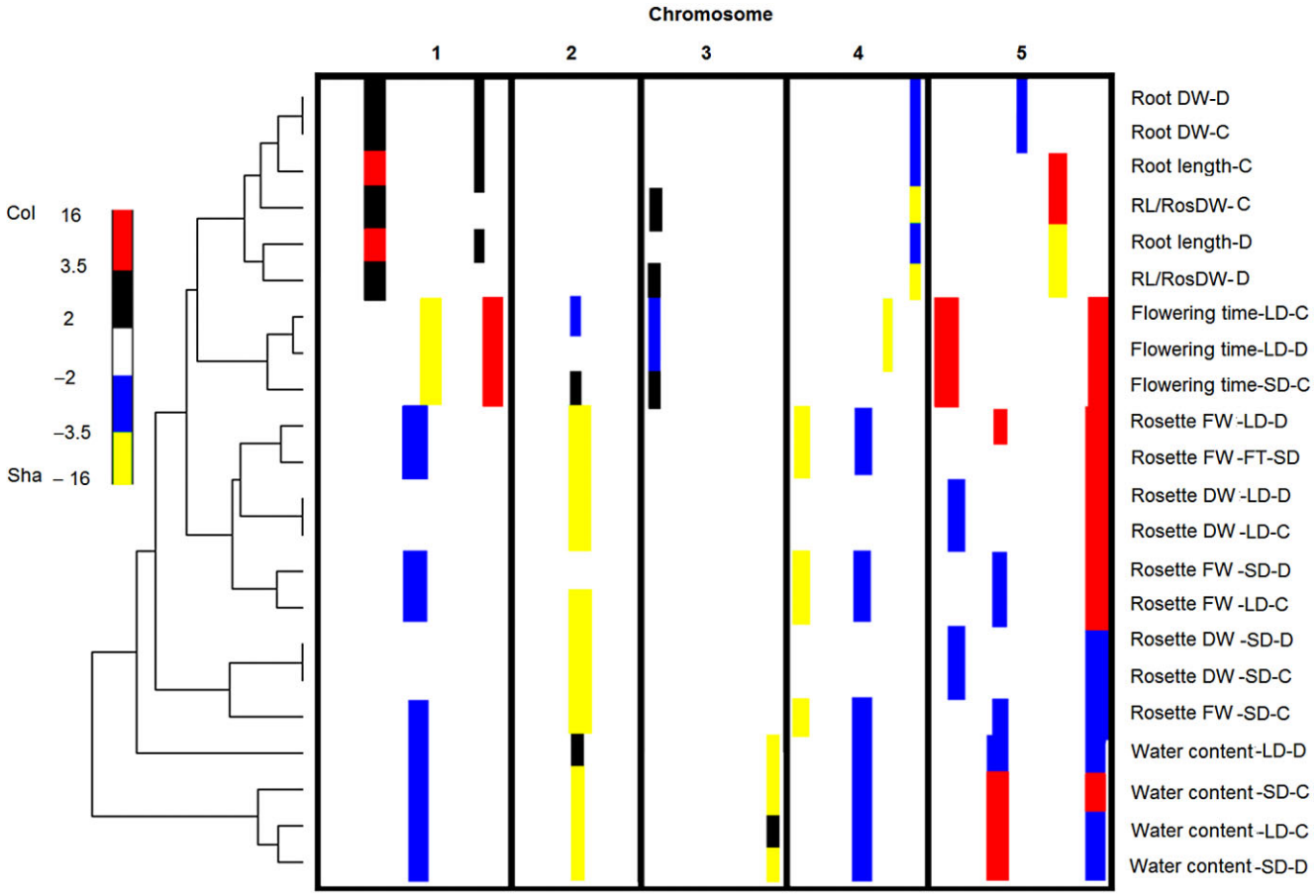

Figure 3. A clustered heat map showing the QTL significance ranges based on the –log10(*P*) profiles of the measured traits in the Sha × Col RIL population. The five columns indicate the five chromosomes, scaled in centimorgans, ascending from the left to right. Rows indicate individual trait profiles. A colour scale is used to indicate the QTL significance level corresponding to the –log10(*P*) score, with a –log10(*P*) score of 2 used as significance threshold. Red and black indicate a positive effect on the trait by the Col allele, blue and yellow indicate a negative effect on the trait by the Col allele (meaning a positive effect on the trait by the Sha allele). The width of a bar indicates the significance interval of the QTL. Hierarchical clustering, shown on the left, reflects the correlation between traits based on the QTL profiles. LD and SD refer to long-day and short-day conditions and C and D refer to control and drought treatments, respectively. Rosette FW-F-SD means fresh weight at flowering time under short days, RL/RosDW means the ratio between root length and rosette dry weight (RosDW).

RL/RosDW2, both with the Col allele contributing to higher trait values. *QFT5*, of which the Sha allele reduced FT, was only mapped under LD conditions. Sha provided the allele increasing FT as well as the alleles contributing to higher values for the co-locating *FW1* and *WC1* QTLs. *QFT7* co-located with the *FW6*, *DW3* and *WC6* QTLs, all with the Col allele contributing to higher trait values.

Validation of root length QTLs

As there was a striking difference between Col and Sha concerning root length under drought stress conditions (Fig. 1), we were interested to confirm the detected RL QTLs. To first investigate the allelic contributions, the RILs were classified into 16 genotypic groups based on all possible allelic combinations at the four detected RL QTL in the SD drought treatment (Fig. 4a). For *RL1* and *RL2*, the Col alleles increased RL, while for *RL3* and *RL4*, the Col alleles decreased RL. Only for *RL4*, Q×E was observed (Table 2). ANOVA indicated that under drought, there was significant epistasis between *RL2* and *RL4* ($P = 0.04$), with the Col allele at *RL2* contributing to longer roots regardless of the nature of the allele at *RL4*, while the Sha allele at *RL2* contributed to shorter roots only if there was a Col allele at *RL4*.

The effects of the *RL* loci on other root traits did not show the same strong contrasts as for RL. Out of all 16 possible allelic combinations, the longest and shortest roots under drought are, respectively, caused by the AABB and BBAA genotypes (Fig. 4e). To confirm this in an independent experiment, and to determine any consequences for related traits, seven RILs with the AABB genotype and six RILs with the BBAA genotype were regrown under drought conditions. RL, RosFW, RosDW and WC were determined for these plants (Fig. 4f–i). ANOVA confirmed the significant difference

Table 2. QTLs detected for root traits of the Sha × Col RIL population grown under short-day (SD) conditions in control and drought treatments using the MTMM approach

Trait	QTL					SD (control)		SD (drought)	
Unit	Name	Chr.	cM	–log10(*P*)	Q×E	Effect	R^2	Effect	R^2
Root length (cm)	*RL1*	1	31.9	3.8	ns	25.34	4.9	25.34	5.6
	RL2	1	83.1	2.2	ns	18.17	2.5	18.17	2.9
	RL3	4	68.3	2.7	ns	–21.28	3.5	–21.28	4.0
	RL4	5	63.6	3.6	s	12.50	1.2	–27.74	6.8
Root dry weight (g)	*RDW1*	1	31.9	3.1	ns	0.008	4.9	0.008	3.2
	RDW2	1	83.1	2.1	ns	0.002	3.1	0.002	2.0
	RDW3	4	68.3	2.9	ns	–0.006	4.7	–0.006	3.1
	RDW4	5	43.1	2.2	ns	–0.004	3.2	–0.004	2.1
RL/RosDW (cm g^{-1})	*RLROSDW1*	1	31.9	3.4	ns	223.5	4.6	223.5	2.8
	RLROSDW2	3	3.8	2.2	ns	174.3	2.8	174.3	1.7
	RLROSDW3	4	68.3	4.4	ns	–263.6	6.4	–263.6	3.9
	RLROSDW4	5	63.6	3.6	s	237.0	5.2	–278.7	4.3

Chromosome numbers (Chr.) and chromosomal positions (in centimorgans, cM) are indicated. The abbreviated trait name column ('name') lists the identified QTL names arranged according to chromosome number and position. RL/RosDW is the ratio between root length. –log10(*P*) indicates the QTL significance level. 's' and 'ns' refer to significant and non-significant QTL by environment interaction (Q×E). Positive 'effect' values indicate that the Col allele contributes to an increase in the trait value, while negative values indicate that the Sha allele contributes to a trait value increase. Units for effect are the same as the trait units. R^2 indicates the percentage of total phenotypic variance, which is explained by each QTL.

for RL, RosFW and RosDW between AABB and BBAA, but not for WC (data not shown).

Analysis of candidate genes identified after GWA and RIL mapping

Candidate genes were identified on the basis of SNPs found to be significantly associated with one of the investigated traits of the GWA population. Priority was given to genes that were reported to be associated with abiotic stress (Supporting Information Table S5) and genes mapped to QTL confidence intervals as identified in the RIL population. To further provide experimental support for potential candidate genes, one set of accessions with highly contrasting phenotypes was composed based on the ratio between RosDW under drought and RosDW under control conditions (Supporting Information Table S6), thus distinguishing drought-tolerant (with a high ratio) and drought-sensitive accessions (with a low ratio). The selected accessions with contrasting phenotypes did not exhibit any distinguishing geographic distribution. Most of the drought-tolerant accessions had shorter total root system lengths in drought than in control treatments, while most drought-sensitive accessions showed much less variation in total root system length when comparing both treatments (RL ratio often >1). Another 20 accessions were selected based on the ratio between RL under drought and RL under control conditions (Supporting Information Table S7). Again, the selected accessions did not exhibit any marked geographic distribution. In general, accessions with longer roots in the drought environment exhibited higher RDW than accessions with shorter roots.

Subsequently, the Arabidopsis 1001 genomes browser (signal.salk.edu/atg1001/3.0/gebrowser.php; Weigel & Mott 2009) was used to compare the predicted amino acid sequences of all genes reported in Supporting Information Table S5 for the accessions with contrasting phenotypes, meaning the genes closest to the significantly associated SNP and the genes in linkage disequilibrium with such SNP. Most of the examined genes did not show any clear amino acid sequence differences between the extreme groups. However, two genes did, the *RIBOSOMAL PROTEIN S3C* (*RPS3C*; At5g35530; Supporting Information Fig. S1) and the *YELLOW LEAF SPECIFIC 7* (*YLS7*) gene (At5g51640; Yoshida *et al.* 2001), also known as the *TRICHOME BIREFRINGENCE-LIKE 17 (TBL17*) gene (Supporting Information Fig. S2). The *RPS3C* gene was associated with RL/RosDW and showed a significant Q×E effect. The GWA accessions for which genome sequence information was available were classified into five haplotypes for this gene and compared with each other with respect to average RL/RosDW values in control and drought treatments (Supporting Information Fig. S1b). The haplotype group 'Wei-0', which included Col, showed a significantly different RL/RosDW under drought when compared with the haplotype group including Sha. To verify that the observed variation was not just due to a generally variable region of the genome, we also compared the coding regions of 25 genes upstream and 25 genes downstream of the *RPS3C* gene, but no common differences were observed between both groups (data not shown). A haplotype analysis of *RPS3C* showed that accessions with the 'Sha' haplotype had a higher RL/RosDW ratio in the drought treatment than accessions with the 'Col' haplotype, in line with the QTL results found for the RIL population. Thus, the unique variation in the *RPS3C* gene, corresponding to the observed SNP association with RosDW, strongly supports its candidacy as the causal gene underlying the trait variation.

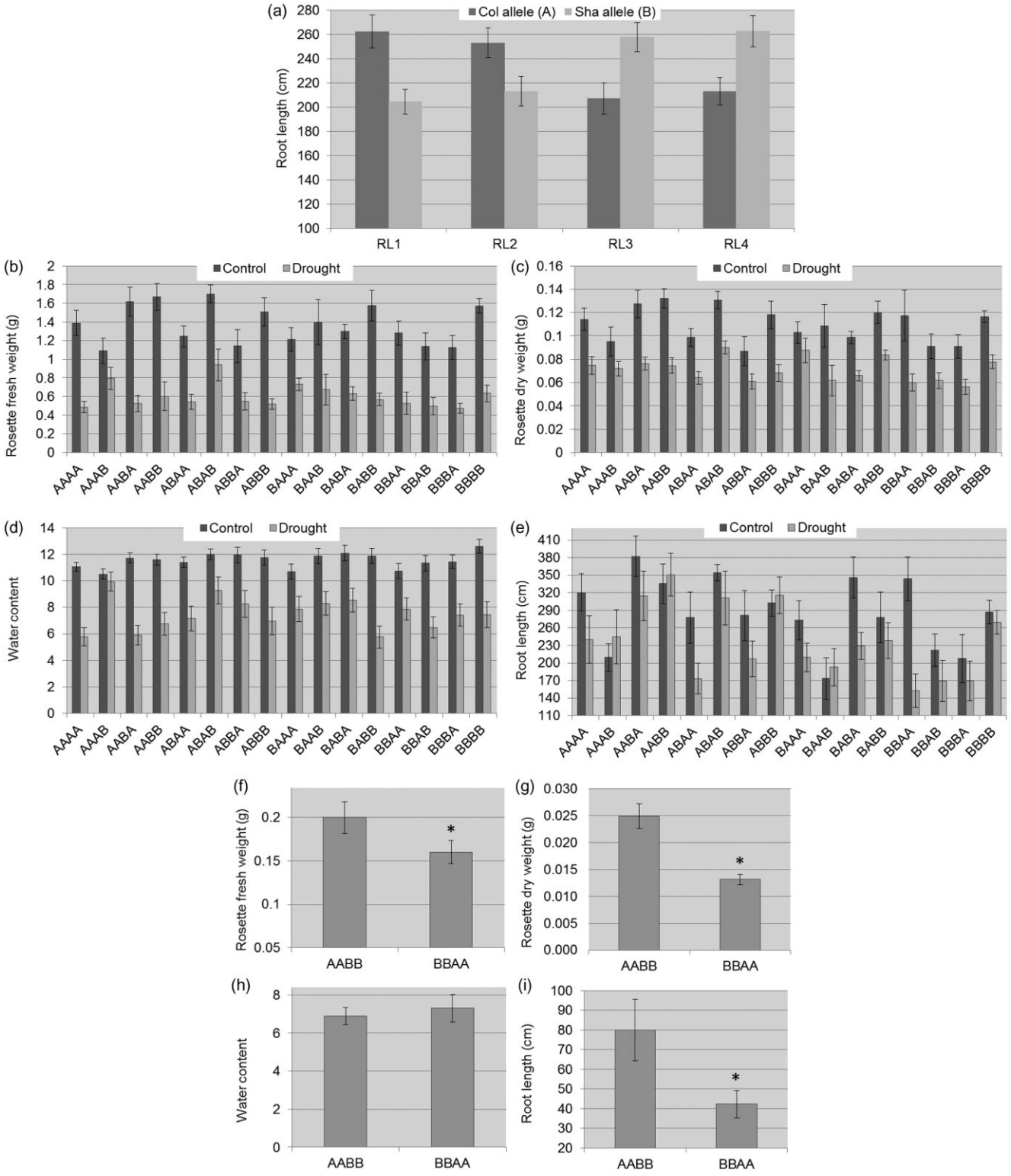

Figure 4. Analysis of the phenotypic effect of variation at the *RL1*, *RL2*, *RL3* and *RL4* root length QTLs detected in the Sha × Col RIL population (Table 2), on rosette fresh weight, rosette dry weight, water content and root length. Error bars refer to standard errors of the means. (a) Root lengths for RILs classified according to their genotypes for each of the *RL* QTLs. (b–e) All RILs were classified based on the four possible allelic combinations at all four *RL* QTL and rosette fresh weight (b), rosette dry weight (c), water content (d) and root length (e) were determined. (f–i) Results of a QTL validation experiment in which 13 RILs, seven with the AABB genotype and six with the BBAA genotype, were regrown in a drought treatment, with three replications each and phenotyped for rosette fresh weight (f), rosette dry weight (g), water content (h) and root length (i). A significant difference of the BBAA genotype from the AABB genotype is indicated with * ($P < 0.05$).

The SNP mapped to the *YLS7* gene was associated with RL with $-\log10(P) = 3.7$. This SNP was in the confidence interval of *RL4*, and for both *RL4* and the SNP associated with *YLS7*. the Col allele contributed to increased RL in the control treatment and to decreased RL in the drought treatment. When the two sets of 10 accessions with contrasting RL ratio phenotypes (Supporting Information Table S7) were examined for their *YLS7* predicted amino acid sequence variations, this showed common differences between both sets (Supporting Information Fig. S2a). Comparing control and drought root length phenotypes of four haplotype groups for *YLS7* showed that the 'Sha' haplotype group was clearly distinct from the other three haplotype groups, showing a longer root system in drought than in the control treatment, with the largest difference with the 'Col' haplotype group (Supporting Information Fig. S2b). When examining the sequences of 25 genes upstream and 25 genes downstream of *YLS7*, a similar distinction between accessions was found for the neighbouring gene, At5g51630, annotated to encode a TIR-NBS-LRR class of disease resistance protein (Supporting Information Fig. S3). The haplotype analysis for this gene showed that the RL under drought of the haplotype group 'Sha' is not different from that of the 'Col' haplotype group, and 'Sha' also shows no significant difference between RL in the drought or control treatments (Supporting Information Fig. S3b). Thus, these results suggest that At5g51630 is less likely than *YLS7* to be the candidate gene underlying *RL4*.

DISCUSSION

To assess the extent of natural variation in rosette and root morphological responses to drought, we studied the response to drought in a GWA population and the Sha × Col RIL population, grown on sand under greenhouse conditions. A general mechanism observed here to cope with drought was by increasing the root to shoot biomass ratio under drought, as reflected in the RL/RosDW trait. When comparing Col and Sha, the two parental lines of the RIL population, Col had longer roots, and consequently higher RL/RosDW, than Sha in the control environment, while Sha had longer roots, and consequently higher RL/RosDW, in the drought environment. This is similar to what was reported for their response to potassium starvation (Kellermeier *et al*. 2013). In general, longer roots enable plants to take up more water, illustrated here by the positive correlation observed between RL and WC. However, WC was negatively correlated with RosDW in the drought treatment in the RIL population, which can be explained by closing stomata under drought and as a result reducing transpiration and nutrient uptake. It seems that Sha and Col represent accessions with extreme contrasting phenotypes as in the HapMap population; RosDW and WC were positively correlated.

The positive correlation observed between some traits under control and drought stress suggests commonalities in their genetic regulation, which was confirmed by finding co-location of significant SNPs for correlated traits. However, also, contrasting correlations between traits and between treatments were observed, which suggest condition-specific genetic determination of at least some of the traits (QTL×E). In contrast to many plant GWAS, which used univariate analysis, we applied a bivariate analysis, the MTMM approach (Korte *et al*. 2012), which accounts for context-dependent QTL effects. Statistical models that explicitly account for G×E will help to discover novel genes that act synergistically with environment (Thomas 2010), potentially leading to the identification of superior and stable genotypes across different environments if applied in crop breeding (Filiault & Maloof 2012).

We compared plants grown under SD conditions with LD-grown plants. In general, the LD-grown plants exhibited higher RosDW than the SD-grown plants, but in terms of identified QTLs, there was hardly any difference (Table 1), suggesting that loci involved in drought tolerance will do so irrespective of day length. In a study where different traits are analysed, co-location of QTL for different traits can be found. This may be due to pleiotropy, but such is difficult to distinguish from close linkage of different loci without further analysis. Although FT itself was not notably affected by drought, there were several growth-related QTLs that co-located with FT loci. Such is known to occur frequently in Arabidopsis (as reviewed previously (Alonso-Blanco *et al*. 2009)). *QFT3* was mapped only in control environments, with significantly different effects between LD and SD flowering. It co-located with *WC2*, which in SD conditions is also only found in the control treatment, suggesting pleiotropy. The *QFT3* mapping interval comprises the *EARLY FLOWERING 3* (*ELF3*) gene, which was mapped earlier in the Bay × Sha RIL population (Jiménez-Gómez *et al*. 2010). In contrast, the *QFT5* locus was only found in LD. It did not co-locate with other growth trait QTLs, but co-located with another LD-FT QTL, previously mapped in the Nd × Col population (Werner *et al*. 2005). The *QFT5* confidence interval comprises the *TWIN SISTER OF FT (TSF)* gene, a gene known to affect FT (Brachi *et al*. 2010). *QFT3* and *QFT5* are typical examples of loci showing conditional neutrality, having a phenotypic effect in some environments, but not all (Anderson *et al*. 2011; El-Soda *et al*. 2014). The *QFT4* locus co-located with the *RLROSDW2* loci to the top of chromosome 3. A FT QTL was previously mapped in the same population (Simon *et al*. 2008), as well as in the Landsberg *erecta* (L*er*) × Sha population (El-Lithy *et al*. 2004). *QFT4* showed Q×E in response to drought and day length, but as this was not seen for the *RLROSDW2* locus, pleiotropy is a less likely reason for the co-location.

In addition, some of the identified root trait QTLs co-located with root trait QTLs mapped in previous studies. For example, Galpaz & Reymond (2010) used the same RIL population to map six QTLs for RL under control and salt stress conditions, of which two co-located with the *RL1* and *RL2* QTLs we found and a third QTL co-located with the *RLROSDW2* QTL. *RLROSDW2* also co-located with QTLs for similar traits previously mapped in the hydroponically grown Bay-0 × Sha RIL population (Bouteillé *et al*. 2012). For all loci mapping in this region, the Sha allele contributed to a decrease in the trait values.

While the great advantage of GWAS over other mapping studies is the high mapping resolution that can be achieved once a suitable, well-genotyped, mapping population has been constructed, there are also a few disadvantages. One is that because of the large numbers of markers, it is difficult to maintain statistical power while controlling for false positives. Thus, it is not easy to choose an appropriate significance threshold. To exclude false positives, a conservative Bonferroni correction should be applied, which would mean a significance threshold of $-\log10(P) = 6.5$. This would result in no significant SNPs with our dataset. Since previous approaches using a threshold of $-\log10(P) = 4$ gave good enrichment for *a priori* candidates (Atwell *et al*. 2010; Li *et al*. 2010), we used the same threshold. Unfortunately, so far, very few of the candidate genes coming from GWAS in Arabidopsis have been confirmed by subsequent molecular genetic analysis, questioning if this threshold is indeed sufficiently selective. One example of such confirmation is the recently identified variation in a new F-box gene, *Kurz-und-Klein* (*KUK*). GWA identified an associated SNP almost reaching the significance threshold after Bonferroni correction, which was confirmed at the gene sequence level to be associated with variation in root development (Meijon *et al*. 2014).

Another disadvantage of GWAS is a poor ability in detecting rare alleles even if they have large phenotypic effects. A good example of this was recently explained for the Arabidopsis *GA5* locus (Barboza *et al*. 2013). Combining traditional linkage mapping with GWA is attractive to reduce the rate of false positives and to detect false negatives (Nordborg & Weigel 2008; Atwell *et al*. 2010; Bergelson & Roux 2010; Brachi *et al*. 2010; Sterken *et al*. 2012; Weigel 2012). That is why we searched the confidence intervals of QTLs mapped in the RIL population, for SNPs associated with the same traits. Two of such associations were found, for RosDW, associated with *FLC*, and for RL, associated with *YLS7*. This may seem few; however, one has to consider that Col and Sha will reflect only a limited fraction of the genetic variation present in the GWA panel. The *FLC* gene is mainly known for controlling FT, but it is also known for its pleiotropic effects on other traits, for example, rosette size in Arabidopsis (Boss *et al*. 2004), and leaf size and biomass in tobacco (Salehi *et al*. 2005). Although it was mapped with $-\log10(P) = 3.6$, we find this sufficient reason to consider *FLC* a good candidate gene underlying the variation for RosDW, and corresponding to the *DW2* and *QFT6* QTLs mapped in the RIL population.

A similar case was found for the *YLS7* gene, also known as *TBL17*, which is considered as the causal candidate gene underlying the *RL4* QTL. For both *RL4* as well as the associated *YLS7* SNP, the RL promoting allele was from the Col haplotype in the control treatment and from the Sha haplotype in the drought treatment. The *YLS7* gene was first identified in a screen for senescence-induced genes (Yoshida *et al*. 2001), but no function was suggested for it. A recent study of Ponkan mandarin (*Citrus reticulata*) showed that the *CrYLS7* gene was transcriptionally down-regulated when fruits were stored in the cold (Zhu *et al*. 2011). *YLS7* belongs to the *TBL* (*TRICHOME BIREFRINGENCE-LIKE*) family of 46 genes in Arabidopsis, which are characterized by encoding a Domain of Unknown Function 231 (DUF231; Moreno *et al*. 2012). Very few of these genes have been functionally analysed. The *ESKIMO1* gene, also known as *TBL29*, is acting as a negative regulator of freezing tolerance (Xin & Browse 1998; Xin *et al*. 2007) and is involved in tolerance to salt and drought (Lugan *et al*. 2009; Lefebvre *et al*. 2011). The *PMR5* or *TBL44* gene induces resistance to powdery mildew and enriches the cell wall with pectin (Kreimer *et al*. 2012). Mutants of the *TRICHOME BIREFRINGENCE* (*TBR*) and co-expressed *TBL3* genes, also show altered cell wall pectin compositions, as well as altered stem, rosette and root growth phenotypes (Moreno *et al*. 2012). Altogether, these examples illustrate that several members of the *TBL* gene family have a role in pectin composition of the cell wall, which easily envisions a role for *YLS7/TBL17* alleles to affect root growth. However, final proof supporting the candidacy of *YLS7* as the gene underlying the *RL4* QTL will need to come from confirming the differential functions of the Col and Sha *YLS7* alleles on RL. Such will also exclude the possibility that not *YLS7*, but the neighbouring TIR-NBS-LRR disease resistance-like gene (At5g51630) is the causal gene. Our results showed that this gene is less likely to affect root growth when compared with *YLS7* and SNPs in this gene are not in LD with the SNP in *YLS7*.

Like the *YLS7* gene, the *RPS3C* gene (Barakat *et al*. 2001) associated with RL/RosDW, was identified to have different alleles distinguishing accessions with contrasting phenotypes for RosDW. This gene is transcriptionally induced by salt stress in Arabidopsis roots. The RPS3C protein binds phosphatidic acid (McLoughlin *et al*. 2013). Phosphatidic acid is a signalling lipid that rapidly accumulates in roots in response to a wide array of abiotic stress stimuli including drought (McLoughlin & Testerink 2013). It is the only gene in a large genomic region comprising the 25 genes upstream and downstream of *RPS3C* to show sequence differences distinguishing accessions with contrasting RL/RosDW ratios (Supporting Information Fig. S1). This included the other gene in the region for which an associated SNP was found, the AT5G35580 gene, encoding a disease resistance-related protein kinase gene (Supporting Information Table S5).

Next to these genes, several genes that were corresponding to associated SNPs are implicated in drought stress response and are worthwhile following up. A SNP mapped to the *SNF1-RELATED PROTEIN KINASE 2.2* (*SnRK2.2*) gene (Fujii *et al*. 2007) was associated with RDW. The double-mutant *snrk2.2/snrk2.3* was showing abscisic acid (ABA)-insensitive phenotypes in seed germination and root growth (Fujii *et al*. 2007) and the Arabidopsis triple-mutant *snrk2.2/snrk2.3/snrk2.6* was found to be nearly completely ABA insensitive, exhibiting greatly reduced tolerance to drought (Fujii & Zhu 2009; Kulik *et al*. 2011). Thus, this gene is a key regulator of ABA signalling and an interesting candidate for a drought responsive QTL.

Another association was found between RosDW and a SNP in the *TARGET OF RAPAMYCIN* (*TOR*) gene,

encoding a growth regulator involved in sensing nutrient availability (Dobrenel *et al.* 2011; Liao *et al.* 2011; Ren *et al.* 2012; Caldana *et al.* 2013). Overexpression of *TOR* increased shoot biomass and resistance to stresses (Dobrenel *et al.* 2011), while the inhibition of *TOR* resulted in reduced root and leaf growth, leading to poor nutrient uptake and poor light energy utilization (Ren *et al.* 2012). Thus, different alleles of this gene may well contribute to the variation in RosDW we found for drought-sensitive and tolerant accessions.

Finally, a SNP for the *SHORT ROOT HAIR1* (*SHR1*) or *CAN OF WORMS1* (*COW1*) gene was associated with RDW. This gene encodes a phosphatidylinositol transfer protein needed for root hair elongation (Bohme *et al.* 2004). Although this is not directly involved in root growth, impaired root hair elongation is likely to affect other growth traits, especially under drought stress.

In conclusion, this study investigated the differential genotypic response of Arabidopsis to drought stress. We showed promising associations between SNPs in relevant genes and drought tolerance related rosette and root traits. However, additional confirmation will need to be obtained from the phenotypic analysis of null mutants and reciprocal allele transformations, as well as from studying differential gene expression and co-expression networks to validate their candidacy as genes underlying the observed phenotypic variation. If so, this will also reveal more on the functions they have in the response of Arabidopsis to drought, in which the knowledge can potentially be used to improve drought stress tolerance in crops.

ACKNOWLEDGMENTS

We acknowledge the Dutch Experimental Plant Science graduate school for financially supporting this research. We thank Charles N. Moreira, Pingping Huang, Christina May, Corrie Hanhart, Nihal Erol-Öztolan and Ana Carolina A.L. Campos for their help in sowing and harvesting the plants used in the described experiments and Pingping Huang for critically reading the manuscript prior to submission.

REFERENCES

Alonso-Blanco C., Aarts M.G.M., Bentsink L., Keurentjes J.J.B., Reymond M., Vreugdenhil D. & Koornneef M. (2009) What has natural variation taught us about plant development, physiology, and adaptation? *The Plant Cell* **21,** 1877–1896.

Anderson J.T., Willis J.H. & Mitchell-Olds T. (2011) Evolutionary genetics of plant adaptation. *Trends Genetics* **27,** 258–266.

Andres F. & Coupland G. (2012) The genetic basis of flowering responses to seasonal cues. *Nature Review Genetics* **13,** 627–639.

Assmann S.M. (2013) Natural variation in abiotic stress and climate change responses in Arabidopsis: implications for twenty-first-century agriculture. *International Journal of Plant Sciences* **174,** 3–26.

Atwell S., Huang Y.S., Vilhjalmsson B.J., Willems G., Horton M., Li Y., Meng D., *et al.* (2010) Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465,** 627–631.

Barakat A., Szick-Miranda K., Chang I.F., Guyot R., Blanc G., Cooke R., Bailey-Serres J. (2001) The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. *Plant Physiology* **127,** 398–415.

Barboza L., Effgen S., Alonso-Blanco C., Kooke R., Keurentjes J.J.B., Koornneef M. & Alcázar R. (2013) *Arabidopsis* semidwarfs evolved from independent mutations in *GA20ox1*, ortholog to green revolution dwarf alleles in rice and barley. *Proceedings of the National Academy of Sciences* **110,** 15818–15823.

Bergelson J. & Roux F. (2010) Towards identifying genes underlying ecologically relevant traits in *Arabidopsis thaliana*. *Nature Review Genetics* **11,** 867–879.

Boer M.P., Wright D., Feng L., Podlich D.W., Luo L., Cooper M. & van Eeuwijk F.A. (2007) A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in Maize. *Genetics* **177,** 1801–1813.

Bohme K., Li Y., Charlot F., Grierson C., Marrocco K., Okada K., Nogue F. (2004) The Arabidopsis COW1 gene encodes a phosphatidylinositol transfer protein essential for root hair tip growth. *The Plant Journal* **40,** 686–698.

Boss P.K., Bastow R.M., Mylne J.S. & Dean C. (2004) Multiple pathways in the decision to flower: Enabling, promoting, and resetting. *Plant Cell* **16,** S18–S31.

Bouteillé M., Rolland G., Balsera C., Loudet O. & Muller B. (2012) Disentangling the intertwined genetic bases of root and shoot growth in Arabidopsis. *PLoS ONE* **7,** e32319.

Brachi B., Faure N., Horton M., Flahauw E., Vazquez A., Nordborg M., Roux F. (2010) Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genetics* **6,** e1000940.

Caldana C., Li Y., Leisse A., Zhang Y., Bartholomaeus L., Fernie A.R., Giavalisco P. (2013) Systemic analysis of inducible target of rapamycin mutants reveal a general metabolic switch controlling growth in *Arabidopsis thaliana*. *The Plant Journal* **73,** 897–909.

Dobrenel T., Marchive C., Sormani R., Moreau M., Mozzo M., Montane M.H., Meyer C. (2011) Regulation of plant growth and metabolism by the TOR kinase. *Biochemical Society Transactions* **39,** 477–481.

van Eeuwijk F.A., Bink M.C.M., Chenu K. & Chapman S.C. (2010) Detection and use of QTL for complex traits in multiple environments. *Current Opinion Plant Biology* **13,** 193–205.

Eichler E.E., Flint J., Gibson G., Kong A., Leal S.M., Moore J.H. & Nadeau J.H. (2010) Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Review Genetics* **11,** 446–450.

El-Lithy M.E., Clerkx E.J., Ruys G.J., Koornneef M. & Vreugdenhil D. (2004) Quantitative trait locus analysis of growth-related traits in a new Arabidopsis recombinant inbred population. *Plant Physiology* **135,** 444–458.

El-Soda M., Malosetti M., Zwaan B.J., Koornneef M. & Aarts M.G.M. (2014) Genotype × environment interaction QTL mapping in plants: lessons from Arabidopsis. *Trends in Plant Science* **9,** 390–398.

Filiault D.L. & Maloof J.N. (2012) A genome-wide association study identifies variants underlying the *Arabidopsis thaliana* shade avoidance response. *PLoS Genetics* **8,** e1002589.

Fujii H. & Zhu J.K. (2009) Arabidopsis mutant deficient in 3 abscisic acid-activated protein kinases reveals critical roles in growth, reproduction, and stress. *Proceedings of the National Academy of Sciences of the United States of America* **106,** 8380–8385.

Fujii H., Verslues P.E. & Zhu J.K. (2007) Identification of two protein kinases required for abscisic acid regulation of seed germination, root growth, and gene expression in Arabidopsis. *The Plant Cell* **19,** 485–494.

Fujii H., Verslues P.E. & Zhu J.K. (2011) Arabidopsis decuple mutant reveals the importance of SnRK2 kinases in osmotic stress responses *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* **108,** 1717–1722.

Galpaz N. & Reymond M. (2010) Natural variation in *Arabidopsis thaliana* revealed a genetic network controlling germination under salt stress. *PLoS ONE* **5,** e15198.

Gibson G. (2012) Rare and common variants: twenty arguments. *Nature Review Genetics* **13,** 135–145.

Holland J.B. (2007) Genetic architecture of complex traits in plants. *Current Opinion Plant Biology* **10,** 156–161.

Huang J., Kim C.M., Xuan Y.H., Park S.J., Piao H.L., Je B.I., Han C.D. (2013) OsSNDP1, a Sec14-nodulin domain-containing protein, plays a critical role in root hair elongation in rice. *Plant Molecular Biology* **82,** 39–50.

Jansen R.C. & Stam P. (1994) High resolution of quantitative traits into multiple loci via interval mapping. *Genetics* **136,** 1447–1455.

Jiang C. & Zeng Z.B. (1997) Mapping quantitative trait loci with dominant and missing markers in various crosses from two inbred lines. *Genetica* **101,** 47–58.

Jiménez-Gómez J.M., Wallace A.D. & Maloof J.N. (2010) Network analysis identifies *ELF3* as a QTL for the shade avoidance response in Arabidopsis. *PLoS Genetics* **6,** e1001100.

Joosen R.V.L., Arends D., Willems L.A.J., Ligterink W., Jansen R.C. & Hilhorst H.W.M. (2012) Visualizing the genetic landscape of Arabidopsis seed performance. *Plant Physiology* **158,** 570–589.

Juenger T.E. (2013) Natural variation and genetic constraints on drought tolerance. *Current Opinion Plant Biology* **16,** 274–281.

Kamoshita A., Zhang J., Siopongco J., Sarkarung S., Nguyen H.T. & Wade L.J. (2002) Effects of phenotyping environment on identification of quantitative trait loci for rice root morphology under anaerobic conditions. *Crop Science* **42,** 255–265.

Kellermeier F., Chardon F. & Amtmann A. (2013) Natural variation of Arabidopsis root architecture reveals complementing adaptive strategies to potassium starvation. *Plant Physiology* **161,** 1421–1432.

Korte A., Vilhjalmsson B.J., Segura V., Platt A., Long Q. & Nordborg M. (2012) A mixed-model approach for genome-wide association studies of correlated traits in structured populations. *Nature Genetics* **44,** 1066–1071.

Kreimer A., Litvin O., Hao K., Molony C., Pe'er D. & Pe'er I. (2012) Inference of modules associated to eQTLs. *Nucleic Acids Research* **40,** e98.

Kulik A., Wawer I., Krzywinska E., Bucholc M. & Dobrowolska G. (2011) SnRK2 protein kinases – key regulators of plant response to abiotic stresses. *Omics: A Journal of Integrative Biology* **15,** 859–872.

Lefebvre V., Fortabat M.N., Ducamp A., North H.M., Maia-Grondard A., Trouverie J., Durand-Tardif M. (2011) ESKIMO1 disruption in Arabidopsis alters vascular tissue and impairs water transport. *PLoS ONE* **6,** e16645.

Li J. & Ji L. (2005) Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* **95,** 221–227.

Li Y., Huang Y., Bergelson J., Nordborg M. & Borevitz J.O. (2010) Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. *Proceedings of the National Academy of Sciences of the United States of America* **107,** 21199–21204.

Liao Y.Y., Buckhout T.J. & Schmidt W. (2011) Phosphate deficiency-induced cell wall remodeling: linking gene networks with polysaccharide meshworks. *Plant Signal & Behavior* **6,** 700–702.

Lugan R., Niogret M.F., Kervazo L., Larher F.R., Kopka J. & Bouchereau A. (2009) Metabolome and water status phenotyping of Arabidopsis under abiotic stress cues reveals new insight into ESK1 function. *Plant, Cell & Environment* **32,** 95–108.

McKay J.K., Richards J.H. & Mitchell-Olds T. (2003) Genetics of drought adaptation in *Arabidopsis thaliana*: I. Pleiotropy contributes to genetic correlations among ecological traits. *Molecular Ecology* **12,** 1137–1151.

McKay J.K., Richards J.H., Nemali K.S., Sen S., Mitchell-Olds T., Boles S., Juenger T.E. (2008) Genetics of drought adaptation in *Arabidopsis thaliana* II. QTL analysis of a new mapping population, KAS-1 × TSU-1. *Evolution* **62,** 3014–3026.

McLoughlin F. & Testerink C. (2013) Phosphatidic acid, a versatile water-stress signal in roots. *Frontiers in Plant Science* **4,** 525.

McLoughlin F., Arisz S.A., Dekker H.L., Kramer G., de Koster C.G., Haring M.A., Testerink C. (2013) Identification of novel candidate phosphatidic acid-binding proteins involved in the salt-stress response of *Arabidopsis thaliana* roots. *Biochemical Journal* **450,** 573–581.

Malosetti M., Voltas J., Romagosa I., Ullrich S.E. & van Eeuwijk F.A. (2004) Mixed models including environmental covariables for studying QTL by environment interaction. *Euphytica* **137,** 139–145.

Malosetti M., Ribaut J.M. & van Eeuwijk F.A. (2013) The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Frontiers in Physiology* **4,** 44.

Meijon M., Satbhai S.B., Tsuchimatsu T. & Busch W. (2014) Genome-wide association study using cellular traits identifies a new regulator of root development in Arabidopsis. *Nature Genetics* **46,** 77–81.

Moreno V., Sole X., Cordero D., Crous-Bou M., Berenguer A., Sanz-Pamplona R., Castells A. (2012) Colorectal cancer eQTLs as susceptibility loci candidates. *European Journal of Cancer* **48,** S282–S282.

Nordborg M. & Weigel D. (2008) Next-generation genetics in plants. *Nature* **456,** 720–723.

Ren M., Venglat P., Qiu S., Feng L., Cao Y., Wang E., Datla R. (2012) Target of rapamycin signaling regulates metabolism, growth, and life span in Arabidopsis. *The Plant Cell* **24,** 4850–4874.

Salehi H., Ransom C.B., Oraby H.F., Seddighi Z. & Sticklen M.B. (2005) Delay in flowering and increase in biomass of transgenic tobacco expressing the Arabidopsis floral repressor gene *Flowering Locus C*. *Journal of Plant Physiology* **162,** 711–717.

Simon M., Loudet O., Durand S., Bérard A., Brunel D., Sennesal F.-X., Camilleri C. (2008) Quantitative trait loci mapping in five new large recombinant inbred line populations of *Arabidopsis thaliana* genotyped with consensus single-nucleotide polymorphism markers. *Genetics* **178,** 2253–2264.

Sterken R., Kiekens R., Boruc J., Zhang F., Vercauteren A., Vercauteren I., *et al.* (2012) Combined linkage and association mapping reveals CYCD5;1 as a quantitative trait gene for endoreduplication in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* **109,** 4678–4683.

Thomas D. (2010) Gene–environment-wide association studies: emerging approaches. *Nature Review Genetics* **11,** 259–272.

Tisné S., Schmalenbach I., Reymond M., Dauzat M., Pervent M., Vile D. & Granier C. (2010) Keep on growing under drought: genetic and developmental bases of the response of rosette area using a recombinant inbred line population. *Plant, Cell & Environment* **33,** 1875–1887.

Tuberosa R. (2012) Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology* **3,** 347.

Tuberosa R. & Salvi S. (2006) Genomics-based approaches to improve drought tolerance of crops. *Trends Plant Science* **11,** 405–412.

Weigel D. (2012) Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiology* **158,** 2–22.

Weigel D. & Mott R. (2009) The 1001 genomes project for *Arabidopsis thaliana*. *Genome Biology* **10,** 107.

Werner J.D., Borevitz J.O., Warthmann N., Trainer G.T., Ecker J.R., Chory J. & Weigel D. (2005) Quantitative trait locus mapping and DNA array hybridization identify an FLM deletion as a cause for natural flowering-time variation. *Proceedings of the National Academy of Sciences of the United States of America* **102,** 2460–2465.

Xin Z. & Browse J. (1998) Eskimo1 mutants of Arabidopsis are constitutively freezing-tolerant. *Proceedings of the National Academy of Sciences of the United States of America* **95,** 7799–7804.

Xin Z., Mandaokar A., Chen J., Last R.L. & Browse J. (2007) Arabidopsis ESK1 encodes a novel regulator of freezing tolerance. *The Plant Journal* **49,** 786–799.

Yoshida S., Ito M., Nishida I. & Watanabe A. (2001) Isolation and RNA gel blot analysis of genes that could serve as potential molecular markers for leaf senescence in *Arabidopsis thaliana*. *Plant and Cell Physiology* **42,** 170–178.

Zeng Z.B. (1994) Precision mapping of quantitative trait loci. *Genetics* **136,** 1457–1468.

Zhao F. & Xu S. (2012) Genotype by environment interaction of quantitative traits: a case study in Barley. *G3* **2,** 779–788.

Zhu A., Li W., Ye J., Sun X., Ding Y., Cheng Y. & Deng X. (2011) Microarray expression profiling of postharvest Ponkan mandarin (*Citrus reticulata*) fruit under cold storage reveals regulatory gene candidates and implications on soluble sugars metabolism. *Journal of Integrative Plant Biology* **53,** 358–374.

Received 23 December 2013; received in revised form 7 July 2014; accepted for publication 14 July 2014

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1. (a) Comparison of predicted amino acid sequences (AA) for the *RPS3C* gene (At5g35530) identified in the GWA mapping population, based on the image provided by the 1001 genome browser (http://signal.salk.edu/atg1001/3.0/gebrowser.php). The dashed line indicates the position of the SNP associated to the mapped trait, RL/RosDW. The comparison is shown for the 10 accessions with the highest and lowest RosDW ratios as shown in Supporting Information Table S6. Based on the predicted AA sequences of this gene, 160 re-sequenced accessions of the GWA population were classified to five haplotypes of which RL/RosDW

ratios as determined in the control and drought treatments are shown (b). The number of accessions in each haplotype is indicated (*n*). Significant differences between the phenotypic values are indicated with a, b and c.

Figure S2. (a) Comparison of predicted amino acid sequences (AA) for At5g51640 (*YLS7*), a gene identified in the GWA mapping population, based on the image provided by the 1001 genome browser (http://signal.salk.edu/atg1001/3.0/gebrowser.php). The position of the SNP associated with RL was close to *YLS7* and co-locating with *RL4* as identified in the Sha × Col RIL population. The comparison is shown for the 10 accessions with the highest and lowest RL ratios as shown in Supporting Information Table S7. Based on the predicted AA sequences of this gene, 160 re-sequenced accessions of the GWA population were classified to four haplotypes of which RL values as determined in the control and drought treatments are shown (b). The number of accessions in each haplotype is indicated (*n*). Significant differences between the phenotypic values are indicated with a and b.

Figure S3. (a) Comparison of predicted amino acid sequences (AA) for At5g51630, a putative disease resistance gene neighbouring *YLS7* (Supporting Information Fig. S2), between the accessions with the highest and lowest RL ratios as shown in Supporting Information Table S7, based on the image provided by the 1001 genome browser (http://signal.salk.edu/atg1001/3.0/gebrowser.php). Based on the predicted AA sequences of this gene, 160 re-sequenced accessions of the GWA population were classified to four haplotypes of which RL values as determined in the control and drought treatments are shown (b). The number of accessions in each haplotype is indicated (*n*). Significant differences between the phenotypic values are indicated with a and b.

Table S1. ANOVA table showing the population mean (with standard deviation, SD), maximum and minimum values for three replications of rosette fresh weight and dry weight and two replications of root length and root dry weight of Arabidopsis accessions of the GWA population grown in a well-watered control treatment (control) and a drought treatment (drought).

Table S2. Pearson correlations for the indicated traits phenotyped in the GWA population grown in a control (C) and drought (D) treatment.

Table S3. Phentoypic analysis of indicated traits determined for the parental lines and the Sha × Col RIL population grown in a control (C) and drought (D) treatment in long and short-day conditions.

Table S4. Pearson correlations for the analysed traits of the Sha × Col RIL population grown in control (C) and drought (D) treatments under long-day (LD) and short-day (SD) conditions.

Table S5. List of candidate genes mapped for shoot and root traits, as identified in the analysis of the Arabidopsis GWA mapping population.

Table S6. The 10 most drought-sensitive and drought-tolerant accessions identified in the GWA mapping population, ranked according to the lowest, respectively, highest drought to control rosette dry weight ratios (RosDW).

Table S7. The 10 most drought-sensitive and drought-tolerant accessions identified in the GWA mapping population, ranked according to the lowest, respectively, highest drought to control root length ratios (RL).