Genetic analysis of drought stress response in *Arabidopsis thaliana* and *Brassica rapa*

Mohamed El-Soda

Thesis committee

Promotor
Prof. Dr M. Koornneef
Personal chair at the Laboratory of Genetics
Wageningen University

Co- promotor
Dr M.G.M. Aarts
Associate Professor, Laboratory of Genetics
Wageningen University

Other members
Prof. Dr F.A. van Eeuwijk, Wageningen University
Prof. Dr H.J. Bouwmeester, Wageningen University
Dr X. Yin, Wageningen University
Dr C.G. van der Linden, Wageningen University

This research was conducted under the auspices of the Graduate School of Experimental Plant Sciences

Genetic analysis of drought stress response in *Arabidopsis thaliana* and *Brassica rapa*

Mohamed El-Soda

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr M.J. Kropff,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Monday 28 October 2013
at 4 p.m. in the Aula.

Mohamed El-Soda

Genetic analysis of drought stress response in *Arabidopsis thaliana* and *Brassica rapa*

125 pages.

PhD thesis, Wageningen University, Wageningen, NL (2013)
With references, with summary in English
ISBN: 978-94-6173-744-1

Contents

Chapter 1: Introduction

Part of this introduction has been submitted as an invited review to Trends in Plant Sciences entitled "Genotype by environment interaction and QTL mapping in plants. Lessons from Arabidopsis thaliana", by Mohamed El-Soda, Marcos Malosetti, Bas J. Zwaan, Maarten Koornneef and Mark G.M. Aarts.

Plant response to drought

Drought has been, and will continue to be, a strong and important selective regime in plant populations (Juenger, 2013). As plants cannot migrate when challenged by fluctuations in environmental conditions, such as drought, they need to cope and adapt to the new fluctuating environment. Therefore, plants have evolved three adaptive strategies to colonize terrestrial ecosystems with limited water availability and to optimise their growth under such conditions. The first strategy is drought escape by early flowering to avoid that major periods of growth and reproduction overlap with periods of water-deficit. The second strategy is drought avoidance via adaptive responses that maintain plant water status even in the face of a drying environment. This can happen by increasing water uptake via increasing the root to shoot biomass ratio, and/or by reducing water loss via transpiration by closing stomata. Stomatal closure results in a concomitant decrease in CO_2 availability in leaves and hence a decrease in availability of assimilates. Also, stomatal closure forces the Rubisco enzyme to lose its favoured ^{12}C isotope and to decrease its discrimination against the ^{13}C isotope, causing an increase in the ^{13}C isotope. The carbon isotope ratio, $\delta\ ^{13}C$, is a useful and simple measure of water use efficiency (WUE) which is defined as carbon gain at the cost of water loss. Increasing WUE is a drought tolerance mechanism that protect plant cells and tissues from water deficit in a way that allows for recovery (Tuberosa, 2012; Assmann, 2013; Juenger, 2013; Tardieu, 2013). Another measure of WUE that is readily quantified is the transpiration efficiency (TE), measured as grams of biomass produced per gram of water used. In general, WUE is higher in C_4 photosynthesis species, such as maize, sorghum, than in C_3 photosynthesis species, but $\delta\ ^{13}C$ can be measured only in C_3 (Assmann, 2013; Tardieu, 2013). One of the few successful applications of $\delta\ ^{13}C$ in breeding programs was

achieved in wheat by selecting genotypes with low $\delta^{13}C$ that showed high TE and increasing aerial biomass under drought conditions (Rebetzke et al., 2002).

Another drought tolerance mechanism is osmotic adjustment (OA), a metabolic process by which plants protect their cellular function, and entails an increase in intercellular solutes, such as proline, glycine betaine, sugar, alcohols and other compounds, in response to water stress (Verslues and Juenger, 2011; Tuberosa, 2012). However, the trade-off between metabolic requirements for OA and the potential benefits for crops, varies on a case-by-case basis as a function of the crop and the dynamics and severity of the drought episodes (Tuberosa, 2012). OA and all other mechanisms that protect plant cells from the harmful effects of dehydration are regulated by genes that are modulated by the plant hormone abscisic acid (ABA). ABA is a fundamental component of the mechanisms allowing plants to avoid or tolerate low water availability. For example, the ABA concentration in leaf tissues is one of the main factors affecting transpiration through stomatal closure and water uptake through the root to shoot biomass ratio (Tuberosa, 2012). In addition, ABA regulates the accumulation of the amino acid proline in plants experiencing abiotic stresses (Sharma et al., 2011). Proline, antioxidants such as ascorbic acid, anthocyanins and tocophorol, are known to be scavengers for Reactive Oxygen Species (ROS) that together with enzymes such as superoxide dismutase and catalases that degrade ROS, reduce the damage of ROS to the structure of stressed plant cells (Mittler, 2002; Cruz de Carvalho, 2008; Bhatt et al., 2011; Setter, 2012; Bhargava and Sawant, 2013; Liang et al., 2013). In general, drought stress shares common aspects with other stresses such as low temperature and salt stress (Zhu, 2002).

Plasticity and genotype by environment interaction

The ability of a genotype to produce distinct phenotypes in different environments, often as a way to adapt to the environment, is known as phenotypic plasticity. This is in general trait and developmental stage specific (Scheiner, 1993) and depends very often on the genotype. Phenotypic plasticity can be described in so-called reaction norms, depicting the phenotypic values of different genotypes over several environments (Figure 1). In the absence of an interaction between genotype and environment, the reaction norms will run parallel to each other, also called “non-cross over interaction” (DeWitt et al., 1998; Sultan, 2000; Pigliucci, 2005; Van Kleunen and Fischer, 2005; Mitchell-Olds and Schmitt, 2006; Ellers and Stuefer, 2010; Nicotra et

al., 2010; Herman and Sultan, 2011; Assmann, 2013; Juenger, 2013). When phenotypic plasticity differs between genotypes, making the reaction norms to potentially cross, this is considered as genotype by environment interaction (GxE). In practice this means that a genotype performing superior in one environment may perform less well in another environment (Kang and Donald, 1997; Weigel, 2012; Assmann, 2013; Juenger, 2013).

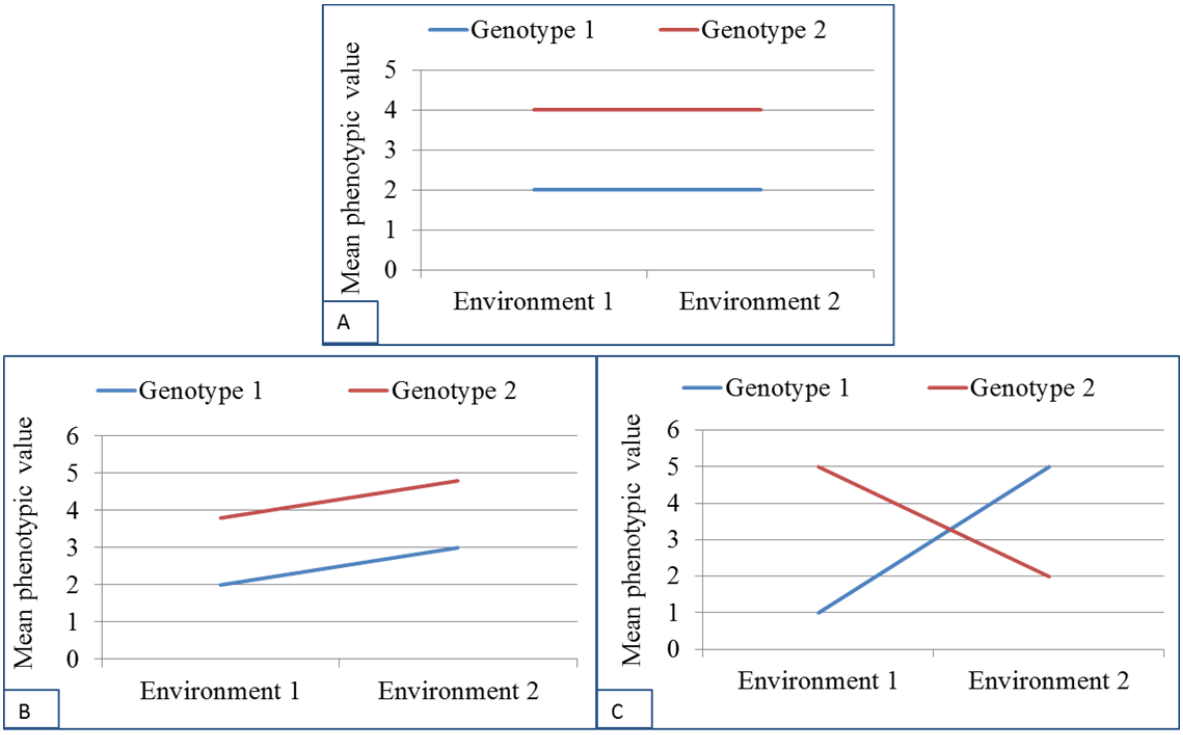


Figure 1: Reaction norms of two genotypes illustrating three cases; (A) no plasticity or GxE. (B) phenotypic plasticity, the trait value changes across environments but the rank order of genotypes across environments remains unchanged (non-crossed parallel reaction norms). (C) GxE causes significant rank changes among genotypes evaluated in different environments (crossed reaction norms).

GxE is not limited to the interaction between a single plant or genotype, and its surrounding environmental conditions, but may extend to the interaction between individual genotypes, or between plant genotypes and pathogen genotypes in the same environment, known as genotype by genotype by environment interaction (GxGxE) (Agrawal, 2001; Mutic and Wolf, 2007; Jorgensen, 2012). This means an expansion of the description of the plant phenotype (P) as the combined effect of genetic composition (G) and heterogonous environmental factors (E), with a

GxE component, leading to P = G + E + GxE (Mackay, 2001; Bernardo, 2008; Visscher et al., 2008). As a consequence, if there is no GxE, a variety trial in a breeding program conducted at only one location will provide universal results (Kang and Donald, 1997). Often this is not the case, as there is GxE, meaning that discarding genotypes evaluated in only one environment in early stages of breeding programmes, bears the risk to discard genotypes with the potential to perform well in other environments, or genotypes that perform relatively well in many environments. Thus some potentially useful genetic variation and favourable alleles can be lost due to limited testing (Ceccarelli et al., 1994; Kang and Donald, 1997). This is why evaluating genotypes in multiple environments will help to identify superior and stable genotypes across different environments, providing more robust breeding results, and offering insights into the genetic basis of G×E (Bergelson and Roux, 2010; van Eeuwijk et al., 2010). The latter is of great interest for crop breeders to realise how much of the selection progress achieved in one environment can be carried over to other environments (Kang and Donald, 1997; Van Kleunen and Fischer, 2005; Nicotra et al., 2010).

The caution to be careful with extrapolation of the results of plants grown in unnatural, controlled, environments to plants grown in the field, because of the expected large GxE (Tuberosa, 2012), is supported by several results from *Arabidopsis thaliana* (Arabidopsis). For instance, in several studies very poor correlation was reported between variation for flowering time (FT) scored in field experiments and FT variation observed under greenhouse conditions (Brachi et al., 2010; Hancock et al., 2011; Méndez-Vigo et al., 2013). The majority of the candidate genes found to control FT in the field, were involved in the regulation of the circadian clock and were not associated with the genes controlling flowering time under greenhouse conditions. This means that under natural conditions (in the field), FT is modulated by more complex environmental cues that are most likely absent under controlled conditions. Another example reported significant differences in leaf sizes, shapes and pigment composition when comparing field and climate chamber grown Arabidopsis raised under different light conditions (Mishra et al., 2012). Similar major differences were found in adjusting the functions of individual proteins involved in the photosynthetic apparatus when shifting Arabidopsis from climate chamber to the field (Jänkänpää et al., 2012). So, to fully understand GxE and its role in shaping adaptive variation under natural conditions, and to extrapolate such knowledge to plant breeding and ecological studies, it is vital to also consider plants grown in the field or similar

conditions rather than focus solely on plants grown under climate chamber conditions. Alternatively, controlled climate chamber conditions reflecting natural conditions should be found that allow extrapolation of the results to field grown plants, as was done by Li et al (Li et al., 2010).

Fitness trade-offs, antagonistic pleiotropy and GxE

Evolutionary biologists have long recognized that traits improving fitness in local environments can have neutral effects or even be deleterious in non-local environments, thus explaining apparent fitness trade-offs between environments. One of the mechanisms underlying trade-offs is antagonistic pleiotropy, a single allele yielding higher fitness in one environment causes a fitness decrease in another environment. Potentially this is crucial in determining a species' range (Moran, 1992; DeWitt et al., 1998; Agrawal, 2001; Elena and Sanjuán, 2003; Hall et al., 2010; Anderson et al., 2011; Fournier-Level et al., 2011; Leinonen et al., 2013). To date, only few examples of antagonistic pleiotropy in plants have been reported, mainly for Arabidopsis (Scarcelli et al., 2007; Todesco et al., 2010; Fournier-Level et al., 2011), but also for example for *Arabidopsis lyrata* (Leinonen et al., 2013) or *Mimulus guttatus* (Hall et al., 2010). The reason for this is probably the high statistical power required to demonstrate antagonistic pleiotropy, as the fitness advantage of the local allele can only be considered if it contributes to statistically significant phenotypic differences when comparing two contrasting environments (Anderson et al., 2011).

Studying Arabidopsis fitness response to early flowering showed that the epistasis between *FRI* and *FLC* has an effect on plant survival in different environments. For example, fall-germinating Arabidopsis accessions with functional *FRI* alleles showed higher winter survival in one *FLC* allelic background, whereas, spring-germinating accessions with deletion of the *FRI* allele had greater seed production in another *FLC* allelic background (Korves et al., 2007). A recent genetic analysis in *A. thaliana* reported many genetic markers associated with conditional neutrality and weak antagonistic pleiotropic effects controlled by allelic variation at two genes, *CHROMATIN REMODELING 8* (*CHR8)*, involved in DNA repair after viral infection, and *SENESCENCE-ASSOCIATED GENE 21* (*SAG21*), involved in water stress tolerance. *CHR8* was associated with survival in Germany although it was established in the northern distribution range of the species, while *SAG21* was associated with survival in Finland although it was established in a more

southern distribution range (Fournier-Level et al., 2011). An example that is related to trade-offs between growth performance and resistance to biotic stresses was found in the effect of an allele of the *ACCELERATED CELL DEATH 6* (*ACD6*) gene that was found in about 20% of all *A. thaliana* plants sampled throughout the world. Although this allele slows down plant growth, it also has antagonistic pleiotropic effects (Figure 2-II) by increasing plant resistant to pathogens (Todesco et al., 2010).

Mapping populations

A common way to do the genetic analysis of a trait is to perform a so-called Quantitative Trait Loci (QTL) mapping (Koornneef et al., 2004; Alonso-Blanco et al., 2009), a process of constructing linkage maps and conducting QTL analysis to identify genomic regions associated with traits (Collard et al., 2005). Conducting QTL studies requires segregating populations, often derived from the cross of two parental lines, such as a population of F2 or F3 lines, back-cross (BC) lines, doubled haploid (DH) lines, or Recombinant Inbred Lines (RILs). The advantage of a BC, DH or RIL population is that if a sufficient number of generations is used, individual lines are nearly homozygous and they can be easily propagated by selfing. BC and RIL populations have the additional advantage that when earlier generations are kept, they still contain sufficient residual heterozygosity to allow the confirmation of QTL via heterogeneous inbred families (HIF), in which only a small portion of the genome segregates for the two parental alleles (Tuinstra et al., 1997; Joosen et al., 2012). To overcome the limited number of recombinations found in the RIL progeny of a biparental cross, advanced intercross RIL populations (AIC-RIL) (Balasubramanian et al., 2009) have been developed, in which the number of recombinations is increased by intermating F2 and later generations before inbred lines are derived. To even further increase the number of alleles and number of recombinations, the multiple parents populations such as the Multiple Advanced Generation Inter-Cross (MAGIC) population (Kover et al., 2009) and the Arabidopsis multiparent RIL (AMPRIL) population (Huang et al., 2011) were introduced. Both MAGIC and AMPRIL populations, allow the use of both linkage and association methodologies without the difficulties of highly structured populations (Cavanagh et al., 2008). Mapping using MAGIC or AMPRIL populations is more complex than with RIL populations, but with a sufficiently high density of intermediate frequency markers, it is possible to infer the most likely local founder genotype. In addition, the accuracy of QTL mapping increases with using the

MAGIC and AMPRIL populations. but not all possible QTL that can be found in pairwise crosses between some of the parents are detected (Weigel, 2012).

An alternative approach to the QTL analysis of the progeny of selected crosses to identify genes underlying natural variation is to use genome wide association studies (GWAS). The advantage of GWAS over selected cross progeny analysis is that genotypes from naturally evolved and adapted populations can be used, which will make elegant use of historical recombinations accumulated over thousands of generations in random mating populations, as part of the evolutionary events that have shaped natural genetic variation. Since large numbers of diverse genotypes should be used to obtain sufficient statistical power, the genetic resolution to identify single nucleotide polymorphisms (SNPs) associated with QTL for the examined trait is generally high, potentially sufficiently high to narrow down the associated region to one or few genes, without the need for additional fine-mapping (Zhu et al., 2008; Mackay et al., 2009; Rafalski, 2010; Weigel, 2012). In *A. thaliana*, community efforts have been made to generate appropriate GWAS panels, including the required high density SNP genotyping (Atwell et al., 2010; Li et al., 2010). In *A. thaliana* the decay of linkage disequilibrium is less than 5 to 10 kb, often comprising only one or two genes, which means that a very dense haplotype map created with around 250,000 SNPs should be sufficient for proper GWAS (Li et al., 2010; Weigel, 2012).

From GxE to mapping QTL by environment interaction

In order to predict the genotypic response to selection, the broad sense heritability should be estimated, which is the proportion of total phenotypic variance (V_p) for a trait, among individuals in a population that is due to heritable genetic variance (V_g). This should include an estimate for the environmental variance, V_e, as $V_p = V_g + V_e$ (Holland et al., 2003; Visscher et al., 2008; Chevin et al., 2010; Eichler et al., 2010). Although heritability components, genetic variance (V_g) and phenotypic variance (V_p), are significantly associated, the sum of V_g of individual QTL rarely explains the whole heritability of a trait, which results in questioning the missing heritability. This missing heritability can be attributed to many reasons, such as variation resulting from epistasis, epigenetic variants, rare variants associated with association mapping studies, many small effect QTL contributing to phenotypic differences (Manolio et al., 2009; Eichler et al., 2010; Brachi et al., 2011), and GxE (Thomas, 2010).

To dissect GxE into its individual genetic components, the genetic complexity of phenotypic responses to the environment and their allelic composition underlying QTL should be considered. As a consequence, GxE will be corresponding to, often several, QTL by environment interactions (QxE), which together reflect the sensitivity of the QTL to the environment (Mackay, 2001; Doerge, 2002; Koornneef et al., 2004; Malosetti et al., 2004; Boer et al., 2007; Holland, 2007; Bernardo, 2008; Mackay et al., 2009; Weigel, 2012; Assmann, 2013). Based on the effect of each QTL in all tested environments, QTL can be classified into: (i) constitutive QTL, also called **e**nvironmentally stable QTL, with a small QxE effect but a large main effect in different environments, which are the main target for breeding programs, as the effect of a QTL is similar in all environments (Figure 3 - QTL1 and 2). (ii) Environment specific QTL, with the QTL showing an effect in one environment but no effect in another (Kamoshita et al., 2002; Malosetti et al., 2004; Tuberosa and Salvi, 2006; Holland, 2007; Tuberosa, 2012; Zhao and Xu, 2012) (Figure 3 – QTL3). Such effect is commonly introduced as conditional neutrality which obstructs the transferability of QTL mapping results (Mackay, 2001). (iii) A QTL has significantly different effects in all environments (Figure 3 – QTL4), which is considered as QxE. (iv) Finally, the QTL can have opposite effects in different environments (Figure 3 – QTL5) which was introduced earlier as antagonistic pleiotropic effect.

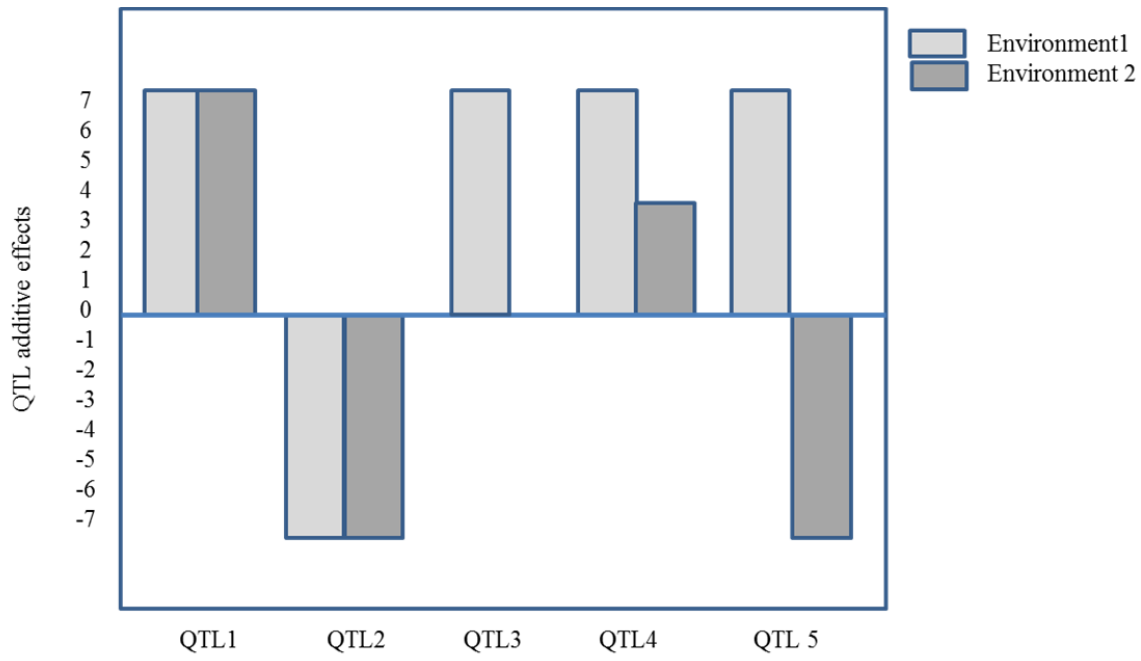


Figure 3: Environment-specific QTL effects. QTL1 and QTL2 are main effect QTL with a positive and negative additive effect respectively. QTL3 is detected under one environment only. QTL4 represents a QTL that has a strong effect in one environment and a weaker effect in another. QTL5 represents a QTL that has opposite additive effects under both environments.

Statistical models for detecting QxE

From a statistical point of view, detecting GxE and QxE consists in finding significant deviations from the additive, major effect QTL, model when modelling the response of at least two genotypes in at least two environments. In the analysis, a trait measured in two environments typically is regarded as two different traits implying that the physiological mechanisms underlying the same plant trait might be different across environments, and consequently, the loci underlying that differential performance are also different (Kang and Donald, 1997).

Different approaches have been used to detect QxE. (i) A standard two-way ANOVA (Carena et al., 2009). (ii) Multivariate composite interval mapping (CIM) as an extension of the univariate mixture model interval mapping approach (Jiang and Zeng, 1995). This provides an exact test for QxE, taking into account the correlation structure between environments. The advantage of this method is that the full variance-covariance (VCOV) matrix of the genetic residual effect can be incorporated into the model. The disadvantage is that with an increasing number of environments the number of parameters to estimate increases dramatically (Chen et al., 2010). Another disadvantage is that the models to use are rather limited, as extensions of the model to account for other factors in the experiment or co-variables are not easy to implement. Therefore, (iii) a mixed-model approach was proposed to model a highly structured covariance matrix (Piepho, 2005), which showed that this approach outperforms all the methods described above (Zhao and Xu, 2012). (iv) Within the least squares framework, a multivariate regression approach was described (Knott and Haley, 2000), allowing to test whether a QTL affects more than one trait (environment), and also testing for linkage between QTL. (v) A mixed-model strategy where a suitable variance-covariance model is selected first for the random genetic effects in the GxE data (Boer et al., 2007). In this strategy, the model selected for the random GxE effects is used for testing environment-specific QTL effects along the genome, in an iterative forward selection approach with cofactors. Their model was able to detect QTL with consistent effects across environments (main effects), and QTL whose effects changed from environment to environment

(QTL with significant QxE). In a final stage, environment-specific QTL effects were related to specific environmental/climatic covariates to develop an eco-physiological QTL model. Finally (vi) a new Bayesian shrinkage method has been used for estimating all main effects and QxE interactions simultaneously in a single model (Chen et al., 2010).

QxE analysis: a case study in Arabidopsis

To visualise the advantage of considering GxE in a multi environments analysis (MEA) and not in a single environment analysis (SEA), one previously published trait will be discussed here. It concerns the "area under the germination curve until 100 h" (AUC), a measure that combines seed germination speed with final germination, tested in a number of germination environments for Arabidopsis (Joosen et al., 2012). This trait was previously analysed using the mixed model approach, as implemented in the QTL library in GenStat (VSNi International), for MEA, while for SEA a protocol was designed in the R/qtl package, using mixture models. Using two different statistical methods makes comparisons less straightforward, which is why we performed SEA using the same statistical framework as used for MEA, that is, the mixed model approach as implemented in the QTL library in GenStat 15 (VSNi International), similar to what was used by Joosen et al. (2012) for MEA. Thereafter, the SEA and MEA analyses for AUC were compared in terms of the number of detected QTL and the explained variance (R^2). In total 57 vs. 80 QTL were detected in SEA and MEA respectively in all environments. For example, SEA couldn't detect any significant QTL in an environment called Cold_WS_AR, while MEA detected three significant QTL. Detecting significant QTL in one environment but not others can be due to type 2 error (false negatives): a QTL is present but just below the operational significance threshold in that environment (MacMillan et al., 2006; Bernardo, 2008) which explains the power gained of considering GxE in QTL mapping (Wei et al., 2010). Although the results (Table 1) showed that three environments highlighted in gray, showed less QTL in MEA than in SEA, one of these environments, CD_NS_AR, showed higher total R^2 in MEA than SEA and Heat_NS_AR showed almost the same R^2, which supports the suggestion (Thomas, 2010) that accounting for GxE increases the explained heritable variance of a trait.

This comparison shows that MEA is a more powerful QTL mapping approach than SEA, because it integrates all the information simultaneously and allows the use of a more realistic model in terms of variances-co-variances leading to a more appropriate or powerful statistical test of hypotheses.

Table 1: Re-analysed data from Joosen et al., (2012) for the "Area Under the germination Curve until 100 h" (AUC) in seed germination experiments performed with the *A. thaliana* Bay x Sha RIL population using multi environment analysis (MEA) and single environment analysis (SEA). R^2 = explained variance, NS = without cold stratification, WS = with cold stratification, ABA = germination with ABA, AR = germination of after-ripened seeds with cold stratification, AR_NS = germination of after-ripened seeds without cold stratification, CD_NS_AR = germination of after-ripened seeds after controlled deterioration, Cold_NS_Fresh = germination of freshly harvested seeds at 10^0 C, Heat_NS_Fresh = germination of freshly harvested seeds at 30^0 C, Manitol_NS = germination on mannitol.

	MEA		SEA	
Environment	No. of QTL	R^2	No. of QTL	R^2
ABA_NS	7	0.484	3	0.369
ABA_WS	5	0.538	5	0.534
AR	5	0.594	3	0.579
AR_NS	5	0.280	2	0.322
CD_NS_AR	3	0.257	5	0.193
CD_WS_AR	1	0.127	2	0.157
Cold_NS_AR	4	0.260	3	0.195
Cold_NS_Fresh	6	0.380	3	0.298
Cold_WS_AR	3	0.107	-	0.009
Heat_NS_AR	5	0.562	6	0.599
Heat_NS_Fresh	5	0.252	2	0.171
Heat_WS_AR	6	0.417	4	0.380
Mannitol_NS	2	0.263	2	0.275
Mannitol_WS	3	0.167	3	0.259
NaCl_NS	4	0.453	3	0.504
NaCl_WS	6	0.451	5	0.574
Stratification_AR	4	0.263	3	0.244
Stratification_Fresh	6	0.615	3	0.554

Cytoplasmic contribution to GxE

Not only the nuclear genotype, but also the cytoplasmic genotype can affect the evolution of adaptation to the environment (Campbell and Waser, 2001). For instance, the RIL progenies of

the two reciprocal crosses between *A. thaliana* accessions Kas-1 and Tsu-1 showed significant differences in water use efficiency when examined under drought conditions (McKay et al., 2008), suggesting the cause was not due to nuclear but cytoplasmic differences or to differences in cytoplasmic-nuclear interactions. Another study (El-Lithy et al., 2005) showed that a mutation in a chloroplast gene had antagonistic pleiotropic effect by lowering photosynthesis efficiency but increasing atrazine resistant in the parental accession Ely and the F2BC4 backcross progeny carrying the Ely cytoplasm, but not in the other parent, L*er*, indicating a cytoplasmic origin of the trait, agreeing with the observed maternal inheritance. Although in general, little attention has been paid in the past to the contribution of the cytoplasmic genome to GxE, it may certainly be worthwhile to examine this, e.g. by using reciprocal crosses. It at least will help to detect cytoplasm × QTL interactions and may contribute to identify useful chloroplast or mitochondrial variation contributing to environmental adaptation (McKay et al., 2008).

From GxE to gene by environment interaction

So far, GWAS has not often been used to detect GxE, mainly because of lack of statistical power and the difficulty to detect rare alleles with GWAS, which appear to be frequently contributing to strong GxE (Eichler et al., 2010; Thomas, 2010; Gibson, 2012). Still, the wide representation of genotypes potentially covering a large fraction of the total genetic spectrum of a species, can be a very useful resource to especially unravel GxE. Recent GWAS that successfully accounted for GxE did so by examining the trait under different conditions using the same population of genotyped accessions and analysing the response to each environment separately (Ehrenreich et al., 2009; Fournier-Level et al., 2011; Hancock et al., 2011; Filiault and Maloof, 2012). This in turn will not allow testing the effect of the same SNP in all tested environments and will lower the power to discover novel alleles/genes that act synergistically to environmental adaptation and plasticity (Thomas, 2010; Korte et al., 2012). Therefore, a recent multi-trait mixed model approach was proposed for multi-trait or multi-environment association mapping (Korte et al., 2012). In this approach a marker can have different effects in different environments, hence QxE, which will explain part of the GxE.

While GWAS suffer from a high frequency of false positives due to population structure, some statistical approaches were introduced to overcome this problem (Yu et al., 2006; Zhao et al., 2007; Atwell et al., 2010; Brachi et al., 2010). However, correcting for population structure will

increase the frequency of false negatives, meaning that when corrected for statistically, causative genetic markers may not be detected when they are strongly associated with population structure (Brachi et al., 2010). Therefore, combining QTL mapping with GWAS can be an alternative for reducing the rate of false positives, for detecting false negatives (Nordborg and Weigel, 2008; Atwell et al., 2010; Bergelson and Roux, 2010; Brachi et al., 2010; Sterken et al., 2012; Weigel, 2012) and obviously it still can account for GxE. This can be achieved by selecting RIL populations generated from extreme parental lines for the studied trait. To better understand the genomic architecture of adaptive quantitative traits, GWAS can be combined with transcriptional networks to rapidly identify and validate a large number of novel genes (Chan et al., 2011; Yano et al., 2013). Characterizing QTL phenotypic effects and examining if and how the expression of the causal genes differs across environments is an important first step in explaining at the molecular level how GxE determines phenotypes. As a next step, and to further understand the molecular basis of plasticity and GxE, a more detailed study is necessary, involving analysis of differential protein levels and function, and detailed analysis of phenotypic differences caused by different alleles, over many more environments.

Arabidopsis thaliana and *Brassica rapa*

In my thesis I will mainly deal with the plant model species, *A. thaliana*, and a related crop species, *Brassica rapa*. The *A. thaliana* genome sequence became available more than a decade ago (Arabidopsis-Genome-Initiative, 2000). This enabled the initiation of the genome re-sequencing project for 1001 Arabidopsis accessions (Weigel and Mott, 2009) where first the sequencing of 80 accessions was accomplished (Cao et al., 2011) and now the sequence information for over 450 accessions is already available (Assmann, 2013). In addition, the existing natural variation among this large number of accessions collected from all over the world, facilitates understanding the genetic mechanisms governing traits of interest (Fournier-Level et al., 2011; Hancock et al., 2011; Weigel, 2012; Assmann, 2013) and has inspired the creation of over 60 RIL populations available in stock centres (Weigel, 2012). All information available for Arabidopsis can be easily used to improve other important crop species such as *Brassica rapa* via comparative alignment (Schranz et al., 2006) which means that the gene order is very similar over large regions of the genome (genome synteny). This makes *B. rapa* an ideal crop model for genetic studies, together with having the smallest genome in the *Brassica* genus,

the availability of its whole genome sequence (Wang et al., 2011) and the availability of tools for genome analysis provided in the *Brassica* database (BRAD) (Cheng et al., 2011). All of this facilitates fine mapping and cloning of desired QTL and identifying potential candidate genes to improve *B. rapa* marker-assisted breeding programs. Equally attractive is that such facilities are not only useful for breeding *B. rapa*, but also for breeding other closely related Brassica species like *B. oleracea* or the allotetraploid species *B. napus* and *B. juncea* that also carry the *B. rapa* genome *(Cheng et al., 2011; Li et al., 2013)*. Both, diploid and allotetraploid brassica plants are members of the *Brassicaceae* family. This family
is an important source of edible roots, stems, leaves, buds and inflorescences, as well as of edible or industrial oils, condiments and forage which is widely grown in Europe, Western Russia, Central Asia, and the Near East (Quijada et al., 2007).

Scope, objectives and outline of the thesis

This thesis focused on phenotyping shoot and root traits of *A. thaliana* and *B. rapa* grown on sand and in greenhouses, to further understand how plants can adapt to natural drought stress. The main objectives were to (i) establish a well characterised *B. rapa* RIL population; (ii) genetically dissect and identify QTL, and preferentially even the causal candidate genes, underlying the plant morphological or physiological response to drought; and (iii) to understand how plants differentially interact with their control and drought environments.
In order to achieve these objectives and to use the advantages of the plant model Arabidopsis, an already existing Arabidopsis RIL population was selected based on the differential root drought response of the two parental lines, Sha and Col, to be evaluated under control and drought **(chapter 2)**. A large population of Arabidopsis accessions collected from all over its natural distribution area, genotyped with > 200,000 SNPs, was used in a GWAS approach to identify candidate genes that are associated with pant response to drought **(chapter 3)**. In order to apply the same methodology in crop breeding, I first contributed to the genetic mapping of a new *B. rapa* RIL population, consisting of 160 lines and genotyped with 270 different markers **(chapter 4)**. Thereafter, the morphological and physiological responses of this population to drought was evaluated, which showed significant GxE and QxE **(chapter 5)**. Finally the overall results and the implications of the studies carried out in this thesis were discussed **(chapter 6)**.

References

Agrawal AA. 2001. Phenotypic plasticity in the interactions and evolution of species. *Science* **294**, 321-326.

Alonso-Blanco C, Aarts MGM, Bentsink L, Keurentjes JJB, 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 JT, Willis JH, Mitchell-Olds T. 2011. Evolutionary genetics of plant adaptation. *Trends in Genetics* **27**, 258-266.

Arabidopsis-Genome-Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.

Assmann SM. 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 YS, Vilhjalmsson BJ, *et al.* 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627-631.

Balasubramanian S, Schwartz C, Singh A, *et al.* 2009. QTL mapping in new Arabidopsis thaliana advanced intercross-recombinant inbred lines. *PLoS ONE* **4**, e4318.

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

Bernardo R. 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years *Crop Science* **48**, 1649-1664.

Bhargava S, Sawant K. 2013. Drought stress adaptation: metabolic adjustment and regulation of gene expression. *Plant Breeding* **132**, 21-32.

Bhatt D, Negi M, Sharma P, Saxena S, Dobriyal A, Arora S. 2011. Responses to drought induced oxidative stress in five finger millet varieties differing in their geographical distribution. *Physiology and Molecular Biology of Plants* **17**, 347-353.

Boer MP, Wright D, Feng L, Podlich DW, Luo L, Cooper M, van Eeuwijk FA. 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.

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

Brachi B, Morris GP, Borevitz JO. 2011. Genome-wide association studies in plants: the missing heritability is in the field. *Genome Biology* **12**, 232. doi:210.1186/gb-2011-1112-1110-1232.

Campbell DR, Waser NM. 2001. Genotype-by-environment interaction and the fitness of plant hybrids in the wild. *Evolution* **55**, 669-676.

Cao J, Schneeberger K, Ossowski S, *et al.* 2011. Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. *Nature Genetics* **43**, 956-U960.

Carena MJ, Romagosa I, Eeuwijk FAv, Thomas WTB. 2009. Statistical analyses of genotype by environment data. *Cereals*, Vol. 3: Springer US, 291-331.

Cavanagh C, Morell M, Mackay I, Powell W. 2008. From mutations to MAGIC: resources for gene discovery, validation and delivery in crop plants. *Current Opinion Plant Biology* **11**, 215-221.

Ceccarelli S, Erskine W, Hamblin J, Grando S. 1994. Genotype by environment interaction and international breeding programmes. *Experimental Agriculture* **30**, 177-187.

Chan EKF, Rowe HC, Corwin JA, Joseph B, Kliebenstein DJ. 2011. Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in *Arabidopsis thaliana*. *PLoS Biology* **9**, e1001125.

Chen X, Zhao F, Xu S. 2010. Mapping environment-specific quantitative trait loci. *Genetics* **186**, 1053-1066.

Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, Li P, Hua W, Wang X. 2011. BRAD, the genetics and genomics database for Brassica plants. *BMC Plant Biology* **11**, 136.

Chevin L-M, Lande R, Mace GM. 2010. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biology* **8**, e1000357.

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* **142**, 169-196.

Cruz de Carvalho MH. 2008. Drought stress and reactive oxygen species: production, scavenging and signaling. *Plant Signaling and Behaviour* **3**, 156-165.

DeWitt TJ, Sih A, Wilson DS. 1998. Costs and limits of phenotypic plasticity. *Trends in Ecology & Evolution* **13**, 77-81.
Doerge RW. 2002. Mapping and analysis of quantitative trait loci in experimental populations. *Nature Review Genetics* **3**, 43-52.
Ehrenreich IM, Hanzawa Y, Chou L, Roe JL, Kover PX, Purugganan MD. 2009. Candidate gene association mapping of Arabidopsis flowering time. *Genetics* **183**, 325-335.
Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH. 2010. Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Reviews Genetics* **11**, 446-450.
El-Lithy ME, Rodrigues GC, van Rensen JJS, Snel JFH, Dassen HJHA, Koornneef M, Jansen MaK, Aarts MGM, Vreugdenhil D. 2005. Altered photosynthetic performance of a natural Arabidopsis accession is associated with atrazine resistance. *J Exp Bot* **56**, 1625-1634.
Elena SF, Sanjuán R. 2003. Climb every mountain? *Science* **302**, 2074-2075.
Ellers J, Stuefer J. 2010. Frontiers in phenotypic plasticity research: new questions about mechanisms, induced responses and ecological impacts. *Evolutionary Ecology* **24**, 523-526.
Filiault DL, Maloof JN. 2012. A genome-wide association study identifies variants underlying the *Arabidopsis thaliana* shade avoidance response. *PLoS Genetics* **8**, e1002589.
Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM. 2011. A map of local adaptation in *Arabidopsis thaliana*. *Science* **334**, 86-89.
Gibson G. 2012. Rare and common variants: twenty arguments. *Nature Reviews Genetics* **13**, 135-145.
Hall MC, Lowry DB, Willis JH. 2010. Is local adaptation in Mimulus guttatus caused by trade-offs at individual loci? *Molecular Ecology* **19**, 2739-2753.
Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG, Toomajian C, Roux F, Bergelson J. 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. *Science* **334**, 83-86.
Herman JJ, Sultan SE. 2011. Adaptive transgenerational plasticity in plants: case studies, mechanisms, and implications for natural populations. *Frontiers in Plant Science* **2. doi: 10.3389/fpls.2011.00102**.
Holland JB. 2007. Genetic architecture of complex traits in plants. *Curr Opin Plant Biol* **10**, 156-161.
Holland JB, Nyquist WE, Cervantes-Martínez CT. 2003. Estimating and interpreting heritability for plant breeding: an update. *Plant Breeding Reviews*: John Wiley & Sons, Inc., 9-112.
Huang XQ, Paulo MJ, Boer M, Effgen S, Keizer P, Koornneef M, van Eeuwijk FA. 2011. Analysis of natural allelic variation in Arabidopsis using a multiparent recombinant inbred line population. *Proc Natl Acad Sci U S A* **108**, 4488-4493.
Jänkänpää HJ, Mishra Y, SchrÖDer WP, Jansson S. 2012. Metabolic profiling reveals metabolic shifts in Arabidopsis plants grown under different light conditions. *Plant, Cell and Environment* **35**, 1824-1836.
Jiang C, Zeng ZB. 1995. Multiple trait analysis of genetic mapping for quantitative trait loci. *Genetics* **140**, 1111-1127.
Joosen RVL, Arends D, Willems LAJ, Ligterink W, Jansen RC, Hilhorst HWM. 2012. Visualizing the genetic landscape of Arabidopsis seed performance. *Plant Physiol* **158**, 570-589.
Jorgensen TH. 2012. The effect of environmental heterogeneity on RPW8-mediated resistance to powdery mildews in *Arabidopsis thaliana*. *Annals of Botany* **109**, 833-842.
Juenger TE. 2013. Natural variation and genetic constraints on drought tolerance. *Curr Opin Plant Biol* **16**, 274-281.
Kamoshita A, Zhang J, Siopongco J, Sarkarung S, Nguyen HT, Wade LJ. 2002. Effects of phenotyping environment on identification of quantitative trait loci for rice root morphology under anaerobic conditions. *Crop Science* **42**, 255-265.
Kang MS, Donald LS. 1997. Using genotype-by-environment interaction for crop cultivar development. *Advances in Agronomy*, Vol. Volume 62: Academic Press, 199-252.
Knott SA, Haley CS. 2000. Multitrait least squares for quantitative trait loci detection. *Genetics* **156**, 899-911.
Koornneef M, Alonso-Blanco C, Vreugdenhil D. 2004. Naturally occuring genetic variation in *Arabidopsis thaliana Annual Review of Plant Biology* **55**, 141-172.
Korte A, Vilhjalmsson BJ, 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.
Korves TM, Schmid KJ, Caicedo AL, Mays C, Stinchcombe JR, Purugganan MD, Schmitt J. 2007. Fitness effects associated with the major flowering time gene FRIGIDA in *Arabidopsis thaliana* in the field. *The American Naturalist* **169**, E141-E157.

Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R. 2009. A Multiparent Advanced Generation Inter-Cross to fine-map quantitative traits in *Arabidopsis thaliana*. *PLoS Genetics* **5**, e1000551.
Leinonen PH, Remington DL, Leppálá J, Savolainen O. 2013. Genetic basis of local adaptation and flowering time variation in *Arabidopsis lyrata*. *Molecular Ecology* **22**, 709-723.
Li X, Ramchiary N, Dhandapani V, Choi SR, Hur Y, Nou I-S, Yoon MK, Lim YP. 2013. Quantitative trait loci mapping in *Brassica rapa* revealed the structural and functional conservation of genetic loci governing morphological and yield component traits in the A, B, and C subgenomes of Brassica species. *DNA Research* **20**, 1-16.
Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO. 2010. Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. *Proce Nat Acad Sci USA* **107**, 21199-21204.
Liang X, Zhang L, Natarajan SK, Becker DF. 2013. Proline mechanisms of stress survival. *Antioxidants and Redox Signaling* **doi:10.1089/ars.2012.5074**.
Mackay TFC. 2001. The genetic architecture of quantitative traits *Annual Review of Genetics* **35**, 303-339.
Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics* **10**, 565-577.
MacMillan K, Emrich K, Piepho HP, Mullins C, Price A. 2006. Assessing the importance of genotype × environment interaction for root traits in rice using a mapping population II: conventional QTL analysis. *Theoretical and Applied Genetics* **113**, 953-964.
Malosetti M, Voltas J, Romagosa I, Ullrich SE, van Eeuwijk FA. 2004. Mixed models including environmental covariables for studying QTL by environment interaction. *Euphytica* **137**, 139-145.
Manolio TA, Collins FS, Cox NJ, *et al.* 2009. Finding the missing heritability of complex diseases. *Nature* **461**, 747-753.
McKay JK, Richards JH, Nemali KS, Sen S, Mitchell-Olds T, Boles S, Stahl EA, Wayne T, Juenger TE. 2008. Genetics of drought adaptation in *Arabidopsis thaliana* II. QTL analysis of a new mapping population, KAS-1 × TSU-1. *Evolution* **62**, 3014-3026.
Méndez-Vigo B, Gomaa NH, Alonso-Blanco C, Xavier Picó F. 2013. Among- and within-population variation in flowering time of Iberian *Arabidopsis thaliana* estimated in field and glasshouse conditions. *New Phytologist* **197**, 1332-1343.
Mishra Y, Jankanpaa HJ, Kiss AZ, Funk C, Schroder WP, Jansson S. 2012. Arabidopsis plants grown in the field and climate chambers significantly differ in leaf morphology and photosystem components. *BMC Plant Biollogy* **12**, 6. doi:10.1186/1471-2229-1112-1186.
Mitchell-Olds T, Schmitt J. 2006. Genetic mechanisms and evolutionary significance of natural variation in Arabidopsis. *Nature* **441**, 947-952.
Mittler R. 2002. Oxidative stress, antioxidants and stress tolerance. *Trends in Plant Science* **7**, 405-410.
Moran NA. 1992. The evolutionary maintenance of alternative phenotypes. *The American Naturalist* **139**, 971-989.
Mutic JJ, Wolf JB. 2007. Indirect genetic effects from ecological interactions in *Arabidopsis thaliana*. *Molecular Ecology* **16**, 2371-2381.
Nicotra AB, Atkin OK, Bonser SP, *et al.* 2010. Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**, 684-692.
Nordborg M, Weigel D. 2008. Next-generation genetics in plants. *Nature* **456**, 720-723.
Piepho HP. 2005. Statistical tests for QTL and QTL-by-environment effects in segregating populations derived from line crosses. *Theoretical and Applied Genetics* **110**, 561-566.
Pigliucci M. 2005. Evolution of phenotypic plasticity: where are we going now? *Trends in Ecology & Evolution* **20**, 481-486.
Rafalski JA. 2010. Association genetics in crop improvement. *Current Opinion Plant Biology* **13**, 174-180.
Rebetzke GJ, Condon AG, Richards RA, Farquhar GD. 2002. Selection for reduced carbon isotope discrimination increases aerial biomass and grain yield of rainfed bread wheat. *Crop Science* **42**, 739-745.
Scarcelli N, Cheverud JM, Schaal BA, Kover PX. 2007. Antagonistic pleiotropic effects reduce the potential adaptive value of the FRIGIDA locus. *Proce Nat Acad Sci USA* **104**, 16986-16991.
Scheiner SM. 1993. Genetics and evolution of phenotypic plasticity. *Annual Review of Ecology and Systematics* **24**, 35-68.
Schranz ME, Lysak MA, Mitchell-Olds T. 2006. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends in Plant Science* **11**, 535-542.
Setter TL. 2012. Analysis of constituents for phenotyping drought tolerance in crop improvement. *Frontiers in Physiology* **3**, 180.

Sharma S, Villamor JG, Verslues PE. 2011. Essential role of tissue-specific proline synthesis and catabolism in growth and redox balance at low water potential. *Plant Physiol* **157**, 292-304.
Sterken R, Kiekens R, Boruc J, *et al.* 2012. Combined linkage and association mapping reveals CYCD5;1 as a quantitative trait gene for endoreduplication in Arabidopsis. *Proc Natl Acad Sci U S A* **109**, 4678-4683.
Sultan SE. 2000. Phenotypic plasticity for plant development, function and life history. *Trends in Plant Science* **5**, 537-542.
Tardieu F. 2013. Plant response to environmental conditions: assessing potential production, water demand and negative effects of water deficit. *Frontiers in Physiology* **4:17. 10.3389/fphys.2013.00017**.
Thomas D. 2010. Gene--environment-wide association studies: emerging approaches. *Nature Review Genetics* **11**, 259-272.
Todesco M, Balasubramanian S, Hu TT, *et al.* 2010. Natural allelic variation underlying a major fitness trade-off in *Arabidopsis thaliana*. *Nature* **465**, 632-636.
Tuberosa R. 2012. Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology* **3:347. doi:10.3389/fphys.2012.00347**.
Tuberosa R, Salvi S. 2006. Genomics-based approaches to improve drought tolerance of crops. *Trends in Plant Science* **11**, 405-412.
Tuinstra MR, Ejeta G, Goldsbrough PB. 1997. Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. *Theoretical and Applied Genetics* **95**, 1005-1011.
van Eeuwijk FA, Bink MCM, Chenu K, Chapman SC. 2010. Detection and use of QTL for complex traits in multiple environments. *Current Opinion Plant Biology* **13**, 193-205.
Van Kleunen M, Fischer M. 2005. Constraints on the evolution of adaptive phenotypic plasticity in plants. *New Phytologist* **166**, 49-60.
Verslues PE, Juenger TE. 2011. Drought, metabolites, and Arabidopsis natural variation: a promising combination for understanding adaptation to water-limited environments. *Current Opinion Plant Biology* **14**, 240-245.
Visscher PM, Hill WG, Wray NR. 2008. Heritability in the genomics era - concepts and misconceptions. *Nature Reviews Genetics* **9**, 255-266.
Wang X, Wang H, Wang J, *et al.* 2011. The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics* **43**, 1035-1039.
Wei X, Jackson PA, Hermann S, Kilian A, Heller-Uszynska K, Deomano E. 2010. Simultaneously accounting for population structure, genotype by environment interaction, and spatial variation in marker–trait associations in sugarcane. *Genome* **53**, 973-981.
Weigel D. 2012. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiol* **158**, 2-22.
Weigel D, Mott R. 2009. The 1001 genomes project for *Arabidopsis thaliana*. *Genome Biology* **10**, 107.
Yano R, Takebayashi Y, Nambara E, Kamiya Y, Seo M. 2013. Combining association mapping and transcriptomics identify HD2B histone deacetylase as a genetic factor associated with seed dormancy in *Arabidopsis thaliana*. *The Plant Journal* **74**, 815-828.
Yu J, Pressoir G, Briggs WH, *et al.* 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* **38**, 203-208.
Zhao F, Xu S. 2012. Genotype by environment interaction of quantitative traits: a case study in Barley. *G3: Genes, Genomes, Genetics* **2**, 779-788.
Zhao K, Aranzana MJ, Kim S, *et al.* 2007. An Arabidopsis example of association mapping in structured samples. *PLoS Genetics* **3**, e4.
Zhu C, Gore M, Buckler ES, Yu J. 2008. Status and prospects of association mapping in plants. *The Plant Genome* **1**, 5-20.
Zhu JK. 2002. Salt and drought stress signal transduction in plants. *Annual Review Plant Biology* **53**, 247-273.

Chapter 2: Genotype by environment interaction and plasticity for drought response under short and long photoperiods in *Arabidopsis thaliana*

Mohamed El-Soda[1,2] , Marcos Malosetti[3], Maarten Koornneef[1,4], Mark G.M. Aarts[1]

[1]Laboratory of Genetics, Wageningen University, Wageningen, the Netherlands; [2]Department of Genetics, Faculty of Agriculture, Cairo University, Egypt; [3]Biometris - Applied Statistics, Department of Plant Science, Wageningen University, Wageningen, the Netherlands, [4]Max Planck Institute for Plant Breeding Research, Cologne, Germany

Abstract

In the present study we investigated the differential genotypic response to drought stress in long day (LD) and short day (SD) photoperiods in *Arabidopsis thaliana* (Arabidopsis) via quantitative trait loci (QTL) mapping using the Sha x Col recombinant inbred line (RIL) population. Our results showed significant genotype by environment interaction (GxE) for all traits in response to drought in both photoperiods, which was reflected in 27 significant QTL by environment interactions (QxE). In addition, 21 QTL were mapped with a main effect on drought response. The drought treatment did not affect flowering time under LD conditions, however, significant GxE was observed, which was explained by a QTL with significant QxE on the top of chromosome 3. Two flowering time QTL showed conditional neutrality when comparing day length conditions. In addition, we mapped the ability of Arabidopsis to adapt to changing environments by mapping its plasticity in different environments. Two plasticity specific QTL were identified, in addition to loci that co-located with main effect QTL.

Key words: Genotype by environment interaction, RIL population, flowering, drought

Introduction

Plant growth is greatly affected by environmental abiotic stresses of which drought remains the major abiotic factor limiting crop productivity worldwide. Drought is likely to become more threatening with the predicted global temperature increase and the global human population growth (Smith and De Smet, 2012; Assmann, 2013). Plants have evolved three adaptive strategies to colonize terrestrial ecosystems with limited water availability and optimise their

growth under such conditions. These strategies are: (i) drought escape by early flowering; (ii) drought tolerance via increasing water use efficiency (WUE); and (iii) drought avoidance by reducing water loss via transpiration and increasing water uptake by increasing root dry weight to shoot dry weight ratio (Tuberosa, 2012; Assmann, 2013; Juenger, 2013; Tardieu, 2013). An accurate estimate of drought adaptation *per se* implies the absence of other biotic or abiotic stress factors that might influence plant growth and function (Tuberosa, 2012). Recent studies revealed that plant response to drought cannot be predicted from its phenotype in well-watered conditions (Bouchabke et al., 2008), nor does it reflect its phenotype when drought stress is combined with one or two more additional stresses (Mittler, 2006; Vile et al., 2012). In addition to several drought response studies that compared plants exposed to drought with plants growing under well-watered conditions (e.g. (McKay et al., 2003; McKay et al., 2008; Tisné et al., 2010), few recent studies focused on a combining drought with other stresses such as light intensity (Giraud et al., 2008), and prolonged high temperature (Vile et al., 2012).

One of the major environmental factors controlling flowering time (FT) in *Arabidopsis thaliana* (Arabidopsis) is day length, with plants growing in long days (LD) flowering earlier than plants growing in short days (SD) (Andres and Coupland, 2012). The genetics of FT natural variation in Arabidopsis is well dissected via quantitative trait loci (QTL) and genome wide association (GWA) studies (Alonso-Blanco et al., 1998; Loudet et al., 2002; Brachi et al., 2010; Salomé et al., 2011; Méndez-Vigo et al., 2013) and the impact of the underlying genes on other traits such as growth, meristem determinacy, seed germination, leaf flattening and branching has been reported (Melzer et al., 2008; Chiang et al., 2009; Huang et al., 2012; Kozuka et al., 2013). In addition, some QTL studies addressed the Arabidopsis response to drought using SD conditions (McKay et al., 2003; Tisné et al., 2010) to avoid the complication of induction of flowering, which may considerably change plant physiology. However, comparing the Arabidopsis response to drought when grown under long and short photoperiods, which mimics temperate summer and early spring day length conditions, respectively, has so far not been studied. The interaction with different photoperiods is expected to reveal either phenotypic plasticity response, in which significant but similar changes in all genotypic mean values occur due to environmental changes, or genotype by environment interaction (GxE), in which the phenotypic response shows variation depending on the genotype. The GxE response can be reflected by QTL by environment interaction (QxE) (Tétard-Jones et al., 2011; Weigel, 2012; Assmann, 2013; Juenger, 2013;

Tardieu, 2013). Based on the effect of each QTL in all tested environments, QTL can be classified as: (i) A constitutive QTL, where the effect of a QTL is similar in all environments. (ii) An environment specific QTL where the QTL shows an effect in one environment but no effect in another. Such QTL is also known as a conditional neutrality effect QTL. These generally and obstructs the transferability of QTL mapping results from one case to another (Kamoshita et al., 2002; Malosetti et al., 2004; Tuberosa and Salvi, 2006; Holland, 2007; Tuberosa, 2012; Zhao and Xu, 2012; Malosetti et al., 2013).

In the present study we investigated the drought response of the Arabidopsis Sha x Col recombinant inbred line (RIL) population grown in short and long photoperiods including the analysis of the root systems under control and drought environments in SD conditions. Our results showed significant GxE and QxE effects for all traits in response to drought in both photoperiods.

Materials and methods

Mapping population and experimental setup

A population of 164 F6 RILs, comprising a core set of the Sha x Col RIL population genotyped with 86 SNP markers (Simon et al., 2008), was used for the QTL analysis. The RIL population was grown under controlled greenhouse conditions, once under short day (SD) conditions (10 h light) and once under long day (LD) conditions (20 h light). Under LD, the whole RIL population, except five lines, flowered before the final harvest of rosettes. Except for the day length, the environmental conditions in both experiments were comparable, with an average temperature of 21.3 °C and 17.5 °C, and an average relative humidity of 69% and 74%, during day and night respectively.

The experiments were performed in a completely randomized block design with one plant per RIL per block and four or three blocks in the LD and SD experiments respectively. In the SD experiment, plants were grown on fine (beach) sand covered with a thin layer of sieved peat in 4 x 4 x 7 cm squared plastic pots, to facilitate measuring root systems at the end of the experiment. Because plants growing in this set-up dry out faster under LD conditions than under SD conditions, plants in the LD experiment were grown on a 1:1 mix of sand and peat in 4 x 4 x 5 cm

squared plastic pots. This way, the drying rates in terms of days until drying completely were comparable for both experimental set-ups.

All pots in both experiments were watered with nutrient solution at the start of the experiment, until saturation, i.e. 100% soil water holding capacity (SWHC). Nutrient solutions contained 1, 1.1 and 5.9 mM N, P and K, respectively. Per pot, two seeds were sown and three days after germination, seedlings were thinned to one. Nine (under LD) and twelve days (under SD) after germination, water was withheld as drought treatment, while the control treatment was watered every 2 or 3 days, when SWHC was less than 80%, until the end of the experiment. After 27 days (under LD) or 34 days (under SD), rosettes were harvested and fresh weight (RosFW) was measured. Rosettes were oven-dried at 65 °C for three days for dry weight (RosDW) measurements. Water content (WC) was calculated as WC = (RosFW - RosDW) / RosDW, and the water ratio (WR) was calculated as WR = (RosFW - RosDW) / RosFW. For the SD experiment, roots were washed carefully, placed in a plastic tray filled with water and scanned with a flatbed scanner. This turned out to be impossible for the LD experiment, because of the used sand-soil mixture. With the scans, total root length (RL), root volume (RV) and root diameter (DIAM) were measured using WinRhizo (Regent Instruments Inc., Quebec, Canada) and thereafter root DW (RDW) was measured, before the RDW or RL to RosDW ratios RDW/RosDW and RL/RosDW were calculated. An additional SD experiment was carried out to determine FT and RosFW, hereafter referred to as the SD flowering experiment.

Statistical and quantitative trait loci analysis

Statistical analysis was performed on raw data of the three experiments 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 GxE. The heritability was calculated as the ratio between the genetic variation (Vg), i.e. variance between the average values of all RILs, and the total variation (Vt), with Vt = Vg + Ve, where Ve is the environmental variation estimated as the variance between the replicates of all lines.

Data from the analysed traits in the three experiments were used for QTL mapping using the multi environment analysis (MEA) approach (Malosetti et al., 2004; Boer et al., 2007; Malosetti et al., 2013) as implemented in the QTL library of GenStat. The QTL search was done at a step

size of 10 cM, and for composite interval mapping (CIM) a minimum cofactor proximity of 50 cM was allowed. For automatic selection of QTL, a minimum separation of selected QTL of 25 cM was used. A multiple testing threshold value of -log10 = 2.8 was calculated based on the approach implemented in GenStat (Li and 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 QxE 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 QTL from the last composite interval mapping round (two rounds of CIM performed). Phenotypic plasticity for all measured traits was calculated as the difference in the mean phenotypic values per line between treatments (Tétard-Jones et al., 2011).

Validation of QTL

Based on all possible allelic combinations at the four detected QTL for RL, the RILs were classified into 16 genotypic groups. Thereafter, RILs from the two phenotypically most extreme groups, AABB with longest roots and BBAA with shortest roots under drought, were grown again to confirm their root response to drought. In addition, RosFW, RosDW, WC, and RL were determined for those two groups.

Results

Phenotyping the RIL population

FT, rosette traits and root traits were analysed for the 164-F6 RIL population grown in control and drought environments under SD and LD conditions. Figure 1 shows the frequency distributions of the measured traits in both experiments. Transgression beyond both parental lines values was observed for all traits except for RosFW and WR in the SD condition where the transgression was only in one direction. The drought treatment had no significant effect on FT in the LD condition (Figure 1 and Table 1). The effect of drought treatment was not tested in the SD condition, as flowering of most of the RILs took very long under SD. In general RosFW and RosDW in the LD condition was higher than in the SD condition in the control environment, whereas WC and WR were higher in both SD conditions. RL was higher in control than in

drought environments, while RDW, RL/RosDW and RDW/RosDW values were higher in drought environments.

Correlation analysis of all measured traits in both experiments (Table 2) showed that FT was significantly positively correlated with RosFW in both LD and SD conditions. FT in SD condition was also significantly positively correlated with FT in LD conditions. In addition, RosDW, WC and WR were positively correlated when comparing SD control and drought environments, but only RosDW, and not WC or WR, was positively correlated when comparing both LD environments. RL was positively correlated with RosFW, RosDW and WC in both SD environments. Furthermore, RL was positively correlated with WR in the control environment but negatively correlated with WR in the drought environment. In control and in drought environments, RL was positively correlated with RDW and negatively correlated with DIAM, while RDW and RosDW were positively correlated in both SD environments.

Mapping QTL for main effects, GxE and phenotypic plasticity

QTL were classified as main effect QTL or QTL with significant QxE based on the absolute effect of each allele on the trait value in every environment. The effect is expressed in the same units used to measure the trait itself, as mentioned in Table 1. In total 48 QTL, co-locating in 11 clusters, were mapped for the measured traits. Out of these, 28 QTL were mapped for shoot traits of which 23 QTL showed significant QxE (Table 3 and Figure 2). The other 20 QTL were mapped for root traits, of which only five QTL showed significant QxE (Table 4 and Figure 2). Seven QTL for FT (*QFT*) were mapped of which *QFT2, QFT3, QFT4* and *QFT5* showed significant QxE. Variation at the *QFT2* locus in SD conditions had about a four times stronger effect on increase or decrease of flowering time, than in LD conditions. *QFT3* was only mapped under SD condition, with the Col allele reducing FT. *QFT4* was mapped in both conditions with the Col allele contributing to early flowering in SD and Sha allele contributing to early flowering in LD. *QFT5* was only found under LD conditions, at which the Sha allele reduced FT. For *QFT1* no QxE was found and Sha provided the allele increasing FT, as well as the alleles contributing to higher values for the co-locating *FW1, WC1* and *WR1* QTL. Col was the allele increasing FT for *QFT6,* the major FT QTL explaining around 30 % of FT variance in LD condition. In SD condition, *QFT2* co-localised with five root related QTL, *RL1, RV1, RDW1, RL/RosDW* and *RDW/RosDW1,* with the Col allele increasing the trait values. Similarly, *QFT4* co-localised with

QTL for *RL/RosDW2* and *RDW/RosDW2*, both with the Col allele contributing to higher trait values. *QFT7* co-located with four other QTL, *FW6, DW3, WC6* and *WR6,* with the Col allele contributing to higher values.

Besides QTL for FT, some other co-locating QTL were environment specific. For example, *FW4* and *WC4* were only detected in the control environment, not under drought. *FW6, DW3* and *WR4* were only detected under SD, but not LD, conditions. In addition, 21 QTL were mapped for plasticity (Table 5) and apart from *DW4* and *WR7,* all other loci mapped for phenotypic plasticity were also identified as main effect QTL or QTL with QxE (Tables 3 and 4).

Validation of QTL

Based on all possible allelic combinations at the four detected QTL for RL, the RILs were classified into 16 genotypic groups (Figure 3A). For *RL1* and *RL2*, the Col alleles increased RL, while for *RL3* and *RL4*, the Col alleles decreased RL. ANOVA showed there was significant epistasis between *RL2* and *RL4,* with the Col allele at *RL2* resulting in longer roots regardless of the nature of the allele at *RL4*. In contrast, the Sha allele at *RL2* results in shorter roots only if allele at *RL4* is Col. Out of all 16 possible allelic combinations the longest and shortest roots are respectively caused by the AABB and BBAA genotypes (Figure 3B-E). To confirm this in an independent experiment, and to determine the consequences for the related traits, seven RILs with an AABB genotype and six RILs with a BBAA genotype were regrown under drought conditions. Of these plants not only RL was measured, but also RosFW, RosDW and WC (Figure 3F-I). The ANOVA showed a significant difference for RosFW, RosDW and RL between AABB and BBAA, but not for WC.

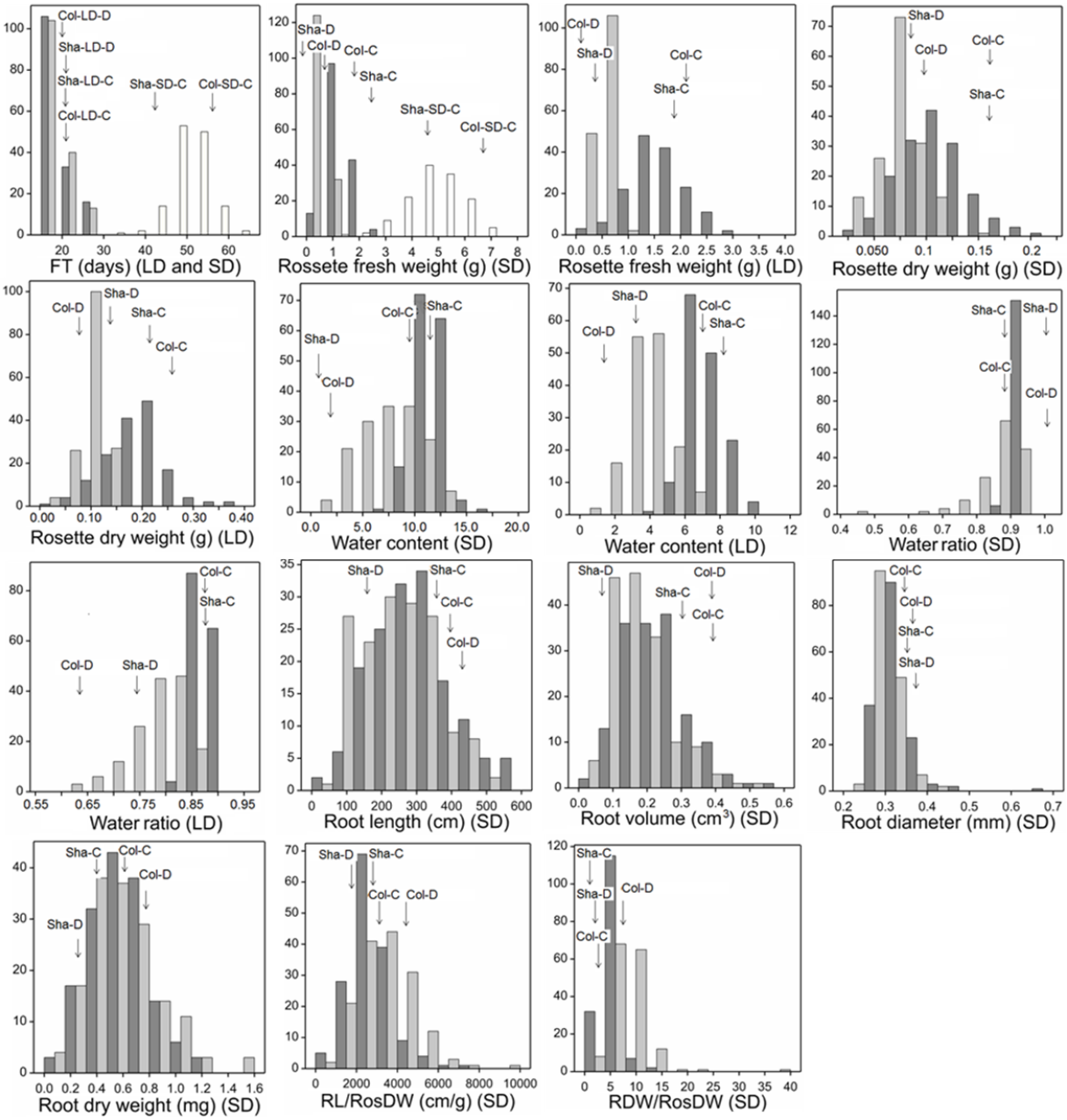


Figure 1: Frequency distributions of the measured traits for the Sha x Col RIL population under control (C, dark grey bars) and drought (D, light grey bars) environments. The white bars refer to flowering time and fresh weight in the SD flowering experiment (only control environment). The *vertical axes* indicate the number of lines per trait value class, and the horizontal axes indicate the different trait value classes. The parental values are indicated with arrows labelled "Col" and "Sha". SD refers to the short day condition, LD refers to the long day condition, FT refers to flowering time, RL/RosDW refers to the ratio between root length (RL) and rosette dry weight (RosDW), RDW/RosDW refers to the ratio between root dry weight (RDW) and RosDW, water content is calculated as (FW - DW) / DW, water ratio is calculated as (FW - DW) / FW.

Table 1: Parental line means for the analyzed traits and performance of the Sha x Col RIL population under control (C) and drought (D) environments in long and short day conditions. FT=flowering time, RosFW=rosette fresh weight, RosDW= rosette dry weight, WC=water content, calculated as (FW-DW)/DW, WR=water ratio, calculated as (FW-DW)/FW, RL=root length, RV= root volume, DIAM=root diameter, RL/RosDW=ratio between RL and RosDW, RDW=root dry weight, RDW/RosDW= ratio between RDW and RosDW. "Min" and "Max" indicate the lowest and the highest value of the RIL population respectively. "Mean" is the average value for all RILs, with standard deviation (sd), and broad sense heritability (h^2). All values are averages of three replications. Under "ANOVA", T refers to treatment, R refers to RILs, and G refers to GxE, ns = non-significant, s = significant.

	Long day experiment										Short day experiment										Short day flowering experiment						
Trait	Parents		RIL population				ANOVA			h^2	Parents		RIL population				ANOVA			h^2	Parents		RIL population				h^2
	Col	Sha	Min	Max	Mean	sd^b	T	R	G		Col	Sha	Min	Max	Mean	sd^b	T	R	G		Col	Sha	Min	Max	Mean	sd^b	
FT-C (days)	18.58	19.33	15.00	26.00	19.62	3.05	ns	s	s	0.90	-	-	-	-	-	-	-	-	-	-	53.33	43.16	35.00	61.00	50.00	4.71	0.90
FT-D (days)	18.00	18.83	15.25	26.00	19.49	2.94				0.85	-	-	-	-	-	-				-	-	-	-	-	-	-	-
RosFW-C (g)	2.03	1.89	0.247	3.066	1.612	0.534	s	s	s	0.58	1.82	2.11	0.37	2.61	1.40	0.44	s	s	s	0.55	6.51	4.49	1.45	6.93	4.63	1.05	0.76
RosFW-D (g)	0.25	0.42	0.133	0.820	0.467	0.135				0.34	0.32	0.23	0.12	1.22	0.58	0.26				0.62	-	-	-	-	-	-	-
RosDW-C (g)	0.25	0.20	0.037	0.366	0.193	0.057	s	s	s	0.56	0.16	0.16	0.033	0.210	0.110	0.032	s	s	s	0.53	-	-	-	-	-	-	-
RosDW-D (g)	0.09	0.11	0.023	0.159	0.100	0.024				0.45	0.098	0.078	0.023	0.155	0.071	0.021				0.42	-	-	-	-	-	-	-
WC-C	7.13	8.45	4.74	10.64	7.33	1.09	-	-	-	-	10.28	12.35	6.09	17.17	11.68	1.39	-	-	-	-	-	-	-	-	-	-	-
WC-D	1.64	2.93	0.64	7.10	3.78	1.18				-	2.28	1.92	0.80	13.89	7.29	2.94				-	-	-	-	-	-	-	-
WR-C	0.88	0.89	0.827	0.914	0.878	0.016	-	-	-	-	0.91	0.93	0.859	0.945	0.920	0.010	-	-	-	-	-	-	-	-	-	-	-
WR-D	0.62	0.75	0.611	0.876	0.781	0.054				-	0.69	0.66	0.444	0.933	0.856	0.077				-	-	-	-	-	-	-	-
RL-C (cm)	-	-	-	-	-	-	-	-	-	-	414.5	372.7	30.0	556.6	301.8	114.0	s	s	s	0.54	-	-	-	-	-	-	-
RL-D (cm)	-	-	-	-	-	-				-	173.6	357.9	59.9	522.8	236.7	106.2				0.70	-	-	-	-	-	-	-
RV-C (cm3)	-	-	-	-	-	-	-	-	-	-	0.43	0.40	0.032	0.540	0.230	0.092	s	s	s	0.59	-	-	-	-	-	-	-
RV-D (cm3)	-	-	-	-	-	-				-	0.17	0.33	0.041	0.500	0.164	0.085				0.63	-	-	-	-	-	-	-
DIAM-C (mm)	-	-	-	-	-	-	-	-	-	-	0.360	0.360	0.267	0.669	0.329	0.046	s	s	s	0.38	-	-	-	-	-	-	-
DIAM-D (mm)	-	-	-	-	-	-	-	-	-	-	0.350	0.035	0.223	0.413	0.295	0.030				0.43	-	-	-	-	-	-	-
RDW-C (mg)	-	-	-	-	-	-	-	-	-	-	0.737	0.540	0.085	1.270	0.580	0.222	s	s	s	0.56	-	-	-	-	-	-	-
RDW-D (mg)	-	-	-	-	-	-				-	0.248	0.440	0.130	1.553	0.598	0.276				0.71	-	-	-	-	-	-	-
RL/RosDW-C (cm/g)	-	-	-	-	-	-	-	-	-	-	2570	2359	320.1	7217.0	2746.0	1042.0	-	-	-	-	-	-	-	-	-	-	-
RL/RosDW-D (cm/g)	-	-	-	-	-	-				-	1778	4604	711.9	9194.0	3393.0	1337.0				-	-	-	-	-	-	-	-
RDW/RosDW-C	-	-	-	-	-	-	-	-	-	-	4.569	3.418	1.47	14.79	5.42	1.89	-	-	-	-	-	-	-	-	-	-	-
RDW/RosDW-D	-	-	-	-	-	-				-	2.541	5.660	1.98	37.20	8.56	3.72				-	-	-	-	-	-	-	-

Table 2: Pearson correlations for the analysed traits of the Sha x Col RIL population grown in control (C) and drought (D) environments under long day (LD) and short day (SD) conditions. FT= flowering time, RosFW= rosette fresh weight, RosDW= rosette dry weight, RosFWFT= rosette fresh weight measured at flowering time, WC= water content and calculated as (FW-DW) / DW, WR= water ratio and calculated as (FW-DW)/FW, RL=root length, RV= root volume, DIAM= root diameter, RL/RosDW= ratio between RL and RosDW, RDW=root DW, RDW/RosDW= ratio between RDW and RosDW. The correlation was performed using the averages of three replications per trait. Dark gray and light gray cells indicate significance level at 0.01 and 0.05 respectively.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
FT-LD-C	1																															
RosFW-LD-C	2	0.158	-																													
RosDW-LD-C	3	.046	0.91	-																												
WC-LD-C	4	0.255	0.489	0.264	-																											
WR-LD-C	5	.054	0.199	0.297	0.656	-																										
FT-LD-D	6	0.674	.057	-.057	0.213	.053	-																									
RosFW-LD-D	7	.084	0.37	0.347	0.243	.137	.071	-																								
RosDW-LD-D	8	-.126	0.605	0.593	0.206	.039	-0.340	0.581	-																							
WC-LD-D	9	0.234	-0.234	-0.24	.027	.104	0.459	0.487	-0.401	-																						
WR-LD-D	10	0.245	-0.198	-0.201	.071	0.16	0.438	0.481	-0.359	0.956	-																					
FT-SD-C	11	0.478	.120	.054	0.238	0.224	0.455	0.178	.062	.122	.121	-																				
RosFWFT-SD-C	12	0.38	0.434	0.367	0.341	0.319	0.297	0.254	0.406	-.096	-.113	0.613	-																			
RosFW-SD-C	13	-.114	0.531	0.516	.157	.021	-0.318	.147	0.555	-0.423	-0.411	0.185	.056	-																		
RosDW-SD-C	14	-.079	0.556	0.556	.139	.029	-0.261	0.184	0.551	-0.382	-0.380	0.179	.103	0.927	-																	
WC-SD-C	15	-.120	.118	.058	.155	.002	-0.257	.030	0.181	-0.167	-.121	.146	-.065	0.422	.073	-																
WR-SD-C	16	-.113	.117	.067	.142	.000	-0.239	.010	0.171	-0.181	-.127	.130	-.078	0.401	.047	0.968	-															
RL-C	17	.062	0.424	0.386	0.177	.037	-.082	.084	0.313	-0.244	-0.272	0.194	.150	0.556	0.549	0.203	0.161	-														
RV-C	18	.025	0.459	0.412	0.159	.002	-.143	.117	0.385	-0.291	-0.315	0.166	.221*	0.619	0.617	0.221	0.165	0.887	-													
DIAM-C	19	-.070	-.007	-.028	-.002	-.035	-.040	.042	.046	.025	.008	-.065	-.086	-.117	-.090	-.100	-.092	-0.242	-.024	-												
RDW-C	20	.057	0.468	0.399	.145	-.055	-.085	.101	0.351	-0.277	-0.294	.156	.211*	0.579	0.576	0.207	.154	0.86	0.954	-.004	-											
RL/ShDW-C	21	.125	.001	-.053	.071	.016	.119	-.097	-.142	.032	.014	.064	.074	-.116	-0.192	0.209	0.164	0.638	0.516	-0.289	0.517	-										
RDW/ShDW-C	22	.117	-.029	-.097	-.006	-.088	.108	-.138	-.151	-.014	-.023	.030	.070	-0.169	-0.242	0.179	.137	0.469	0.49	-.029	0.588	0.856	-									
RosFW-SD-D	23	-.026	0.417	0.324	0.210	-.012	-.083	0.192	0.313	-.144	-.198*	0.300	.102	0.293	0.231	0.233	0.228	0.238	0.253	-.102	0.251	.094	.064	-								
RosDW-SD-D	24	-.103	0.482	0.490	.098	.009	-.144	0.211	0.540	-0.359	-.333**	0.180	.082	0.554	0.562	.108	.107	0.311	0.352	-.062	0.337	-.109	-.129	0.524	-							
WC-SD-D	25	.051	.127	.031	.151	-.021	.007	.075	-.037	.109	.031	0.244	.032	-.072	-.147	.179*	0.185	.067	.042	-.114	.059	0.224	0.215	0.776	-.067	-						
WR-SD-D	26	.117	.147	.077	.144	-.001	-.021	.101	.057	.045	-.031	0.279	.034	.015	-.041	.153	.145	.106	.078	-.153	.064	0.205	.144	0.594	.013	0.751	-					
RL-D	27	-.057	0.4	0.393	.149	.071	-0.157	.154	0.351	-0.236	-.199*	0.217	.030	0.447	0.428	.156	0.157	0.322	0.365	-.023	0.324	.033	-.010	0.371	0.529	.044	.015	-				
RV-D	28	-.058	0.356	0.393	.063	.052	-.153	.113	0.328	-0.248	-.221**	.144	.046	0.458	0.45	.118	.123	0.298	0.348	-.021	0.31	-.019	-.044	0.243	0.463	-.059	-.062	0.886	-			
DIAM-D	29	.106	.069	0.158	-.140	.002	.113	-.039	.019	-.076	-.066	-0.23	.085	.061	.084	-.086	-.090	-.026	-.008	-0.158	.032	-.076	.004	-.052	.068	-.116	-0.158	.089	0.34	-		
RDW-D	30	-.011	0.422	0.455	.076	.053	-.093	.142	0.390	-0.263	-.230**	.111	.066	0.517	0.503	.130	.140	0.311	0.33	-.041	0.315	-.064	-.082	0.237	0.529	-.113	-.126	0.808	0.896	0.343	-	
RL/ShDW-D	31	-.032	.103	.096	.088	.071	-.114	-.024	.014	-.068	-.038	.151	-.047	.102	.080	.081	.094	.157	0.173	.007	.153	.131	.129	.014	-.062	.047	-.070	0.756	0.697	.137	0.572	-
RDW/ShDW-D	32	.083	.020	.059	-.070	.023	.122	-0.249	-0.239	-.039	-.022	-.122	-.075	.146	.135	.006	.026	.131	.102	-.088	.118	.105	.116	.035	.181*	-.080	-0.178	0.6	0.688	0.376	0.756	0.611

Table 3: Overview of QTL detected for shoot traits of the Sha x Col RIL population grown under long day (LD) and short day (SD) conditions using the multi environment analysis. 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 QxE respectively. Positive "effect" values indicate that the Col allele contributes to an increase in the trait value, while negative values indicate that the Sha allele increases the trait value. Units for effect are the same as trait units. R^2 lists the percentage of total phenotypic variance explained by each QTL. R^2 lists the percentage of total phenotypic variance explained by each QTL.

Trait	QTL					LD (control)		LD (drought)		SD (control)		SD (drought)		SD flowering	
(unit)	name	Chr.	cM	-log10(P)	QxE	effect	R^2	effect	R^2	effect	R^2	effect	R^2	effect	R^2
Flowering time (days)	*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	-	-	-	-	-	-	-	-	1.104	5.5
	QFT4	3	0.0	3.0	s	-0.199	1.0	-0.089	0.7	-	-	-	-	0.756	2.6
	QFT5	4	47.9	4.9	s	-0.773	6.4	-0.786	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
Rosette fresh weight (g)	*FW1*	1	51.2	4.0	ns	-0.036	0.5	-0.036	7.2	-0.036	0.7	-0.036	2.0	-0.036	0.1
	FW2	2	30.7	3.8	s	-0.169	10.1	-0.023	3.6	-0.113	6.7	-	-	-0.103	1.0
	FW3	4	0.0	3.1	s	-0.155	8.4	-0.026	3.7	-0.052	1.4	-0.036	2.0	-0.212	4.1
	FW4	4	30.8	2.2	s	-0.092	3.0	-	-	-0.055	1.6	-	-	-0.326	9.7
	FW5	5	30.5	4.4	s	-0.103	3.5	0.020	2.3	-0.118	7.3	-0.080	9.6	-	-
	FW6	5	87.3	8.6	s	0.159	8.9	0.024	4.0	-	-	-	-	0.389	13.8
Rosette dry weight (g)	*DW1*	2	30.7	3.0	s	-0.017	8.6	-0.005	5.0	-0.008	7	-0.003	1.9	-	-
	DW2	5	11.8	3.6	s	-0.009	2.3	-0.008	10.7	-0.007	4.5	-0.002	1.0	-	-
	DW3	5	87.3	5.1	s	0.013	5.5	0.004	2.8	-	-	-	-	-	-
water content	*WC1*	1	65.7	5.0	ns	-0.235	4.7	-0.235	4.0	-0.235	2.9	-0.235	0.6	-	-
	WC2	2	12.2	1.8	s	-0.115	1.1	0.200	2.9	-0.205	2.2	-	-	-	-
	WC3	3	63.4	2.7	s	0.148	1.9	-0.248	4.4	-0.178	1.6	-0.391	1.8	-	-
	WC4	4	35.5	6.3	s	-0.446	16.9	-	-	-0.189	1.8	-	-	-	-
	WC5	5	30.5	8.7	s	-0.255	5.5	0.445	14.2	-0.262	3.5	-0.904	9.4	-	-
	WC6	5	80.9	4.1	s	0.288	7.0	-	-	-0.229	2.7	0.598	4.1	-	-
Water ratio	*WR1*	1	65.7	5.4	s	-0.004	5.7	-0.015	7.4	-	-	-0.013	2.9	-	-
	WR2	2	12.2	2.8	s	-0.002	1.5	0.012	5.2	-0.001	1.8	-0.008	1.1	-	-
	WR3	3	63.4	3.2	s	0.002	1.2	-0.012	5.0	-0.002	2.4	-0.010	1.7	-	-
	WR4	4	35.5	6.5	s	-0.007	17.8	-0.005	1.0	-	-	-	-	-	-
	WR5	5	30.5	8.1	s	-0.004	5.0	0.021	15.0	-0.002	6	-0.018	5.5	-	-
	WR6	5	87.3	5.3	s	0.005	8.5	-	-	-0.002	4.2	0.011	2.2	-	-

Table 4: Overview of QTL detected for root traits of the Sha x Col RIL population grown under short day (SD) conditions using the multi environment analysis. Chromosome numbers (Chr.) and chromosomal positions (in centiMorgans, cM) are indicated. In traits name, RL/RosDW = ratio between root length and rosette dry weight, RDW/RosDW = ratio between root dry weight and rosette dry weight. The trait name column ("name") lists the identified QTL arranged according to chromosome number and position. -log10(P) indicates the QTL significance level. "s" and "ns" refer to significant and non-significant QxE respectively. Positive "effect" values indicate that the Col allele contributes to an increase in the trait value, while negative values indicate that the Sha allele increases the trait value. Units for effect are the same as trait units. R^2 lists the percentage of total phenotypic variance explained by each QTL. R^2 lists the percentage of total phenotypic variance explained by each QTL.

Trait	QTL					SD (control)		SD (drought)	
Unit	name	Chr.	cM	-log10(P)	QxE	effect	R^2	effect	R^2
Root length (cm)	*RL1*	1	31.9	3.8	ns	25.343	4.9	25.343	5.6
	RL2	1	83.1	2.2	ns	18.173	2.5	18.173	2.9
	RL3	4	68.3	2.7	ns	-21.28	3.5	-21.28	4.0
	RL4	5	63.6	3.6	s	12.498	1.2	-27.74	6.8
Root volume (cm^3)	*RV1*	1	31.9	3.7	ns	0.02	4.7	0.02	5.5
	RV2	1	83.1	2.8	ns	0.017	3.4	0.017	3.9
	RV3	4	68.3	3.3	ns	-0.019	4.3	-0.019	5.0
Root diameter (mm)	*RDIAM1*	5	9.0	3.1	s	-0.004	0.8	0.009	8.1
	RDIAM2	5	43.1	2.2	ns	-0.005	1.4	-0.005	3.2
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)	*RLROSDW1*	1	31.9	3.4	ns	223.47	4.6	223.47	2.8
	RLROSDW2	3	3.8	2.2	ns	174.34	2.8	174.34	1.7
	RLROSDW3	4	68.3	4.4	ns	-263.6	6.4	-263.6	3.9
	RLROSDW4	5	63.6	3.6	s	236.95	5.2	-278.7	4.3
RDW/RosDW	*RSR1*	1	31.9	2.4	s	0.003	3.0	0.725	3.8
	RSR2	3	5.9	3.0	ns	0.005	6.6	0.005	1.0
	RSR3	4	68.3	2.6	s	-0.003	3.1	-0.785	4.5

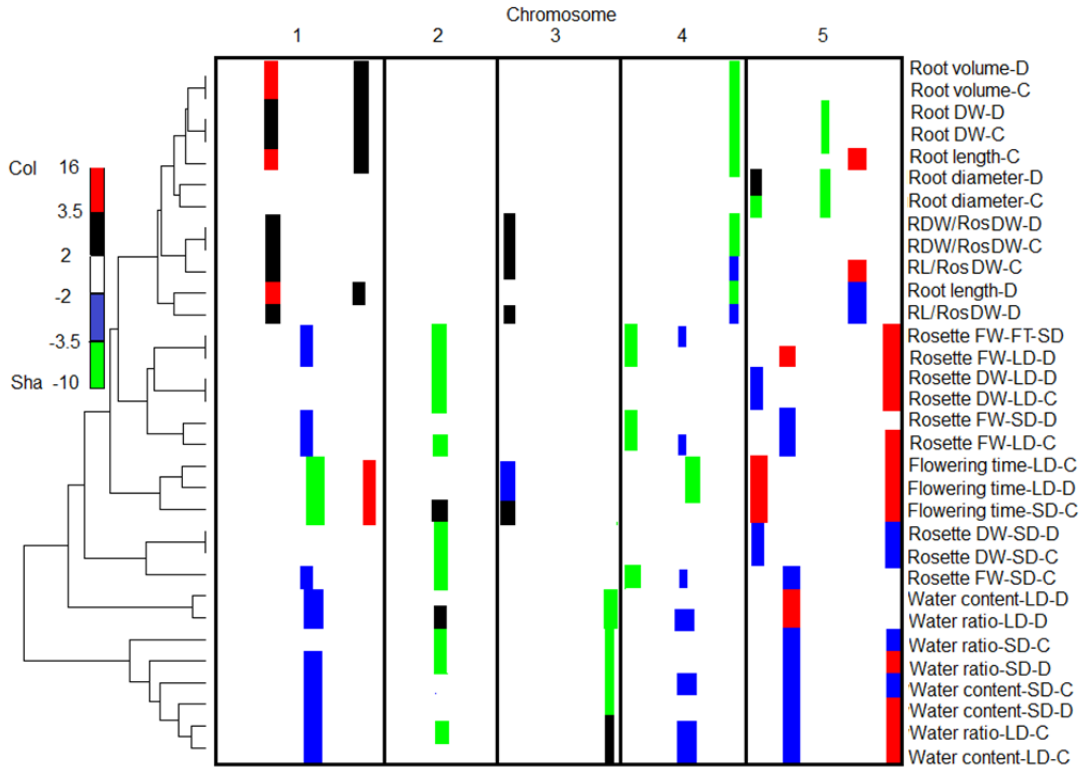


Figure 2: A clustered heat map showing the QTL significance level based on the -log10(P) profiles of the measured traits in the Sha x Col RIL population. The five columns indicate the five chromosomes, scaled in centiMorgans, ascending from the left to right. Rows indicate individual trait profiles. 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 threshold. Red and black indicate a positive effect on the trait by the Col allele, blue and green indicate 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 environments respectively. Rosette FW-F-SD = fresh weight at flowering time under short days, RL/RosDW = ratio between root length and rosette dry weight (RosDW), RDW/RosDW = ratio between root dry weight and RosDW.

Table 5: Overview of QTL detected for phenotypic plasticity in the Sha x Col RIL population grown under short day (SD) conditions using the single environment analysis. Plasticity was calculated as the difference in the mean phenotype between different treatments in long day (LD) and short day (SD) conditions (Tétard-Jones et al., 2011). Chromosome numbers (Chr.) and chromosomal positions (in centiMorgans, cM) are indicated. In traits name, RL/RosDW = ratio between root length and rosette dry weight. The trait name column ("name") lists the identified QTL arranged according to chromosome number and position. -log10(P) indicate the QTL significance level. Positive "effect" values indicate that the Col allele contributes to an increase in the trait value, while negative values indicate that the Sha allele increases the trait value. Units for effect are the same as trait units. R^2 lists the percentage of total phenotypic variance explained by each QTL. R^2 lists the percentage of total phenotypic variance explained by each QTL. QTL highlighted in grey are found only for plasticity and not for main effect or QxE QTL.

Trait	Environment	name	Chr.	cM	effect	-log10(P)	R^2
Flowering time	LD control - SD control	*QFT2*	1	91.5	1.02	2.6	6.1
		QFT4	3	0	0.99	2.6	5.8
		QFT5	4	47.9	0.98	2.5	5.6
Rosette fresh weight	LD control - LD drought	*FW2*	2	12.2	-0.16	4.3	9.8
		FW3	4	0	-0.15	4.2	9.2
	LD control - SD control	*FW3*	4	0	-0.1	2.5	4.6
		FW5	5	87.3	0.21	7.7	19.4
	LD drought - SD drought	*FW4*	5	21.7	0.1	5.7	14.5
	SD control - SD drought	*FW2*	2	30.7	-0.1	2.6	5.7
		FW5	5	80.9	-0.1	2.5	5.4
	SD flowering - SD control	*FW4*	4	30.8	-0.29	2.9	6.9
		FW6	5	87.3	0.44	5.4	15.9
	SD flowering - SD drought	*FW3*	4	7	-0.28	2.7	7.3
		FW6	5	74.7	0.35	3.8	11.1
Rosette dry weight	LD control - LD drought	*DW1*	2	25	-0.014	2.9	8.5
		DW4	4	0	-0.012	2.6	5.5
	LD control - SD control	*DW1*	2	19.3	-0.012	3.1	6.4
		DW4	4	0	-0.012	3	6.2
		DW3	5	87.3	0.015	4.1	9.3
	SD control - LD drought	*DW3*	5	87.3	0.007	2.8	6.8
	LD drought - SD drought	*DW2*	5	0	-0.005	2.7	6.2
Water content	LD control - LD drought	*WC4*	4	22.2	-0.43	3.1	6.5
		WC5	5	30.5	-0.63	4.8	13.9
	LD control - SD control	*WC4*	4	15.9	-0.48	3.5	7.9
		WC6	5	87.3	0.43	2.9	6.4
	LD drought - SD drought	*WC5*	5	21.7	1.51	9.2	23.7
		WC6	5	80.9	-0.71	2.7	5.3
	SD control - SD drought	*WC1*	1	43.2	0.78	3.2	6.6
		WC5	5	21.7	0.92	3.9	9.2
		WC6	5	80.9	-0.96	4.3	10
Water ratio	LD control - LD drought	*WR5*	5	30.5	-0.029	3.9	11.6
	LD drought - SD drought	*WR5*	5	21.7	0.035	3.8	9.4
	SD control - SD drought	*WR7*	1	43.7	0.019	2.8	6..1
		WR5	5	21.7	0.022	3.4	8.2
Root length	SD control - SD drought	*RL4*	5	63.6	39.1	3.8	9.328
Root diameter	SD control - SD drought	*RDIAM1*	5	9	-0.014	3.2	7.7
RL/RosDW	SD control - SD drought	*RLROSDW4*	5	63.6	515.2	4.1	10.2

Discussion

As the severity of drought on worldwide agriculture escalates, drought resistance is expected to play a more important role in improving crop yield than yield potential only (Tuberosa, 2012). Here, we studied the response of the Sha x Col RIL population to drought, when grown under SD and LD conditions.

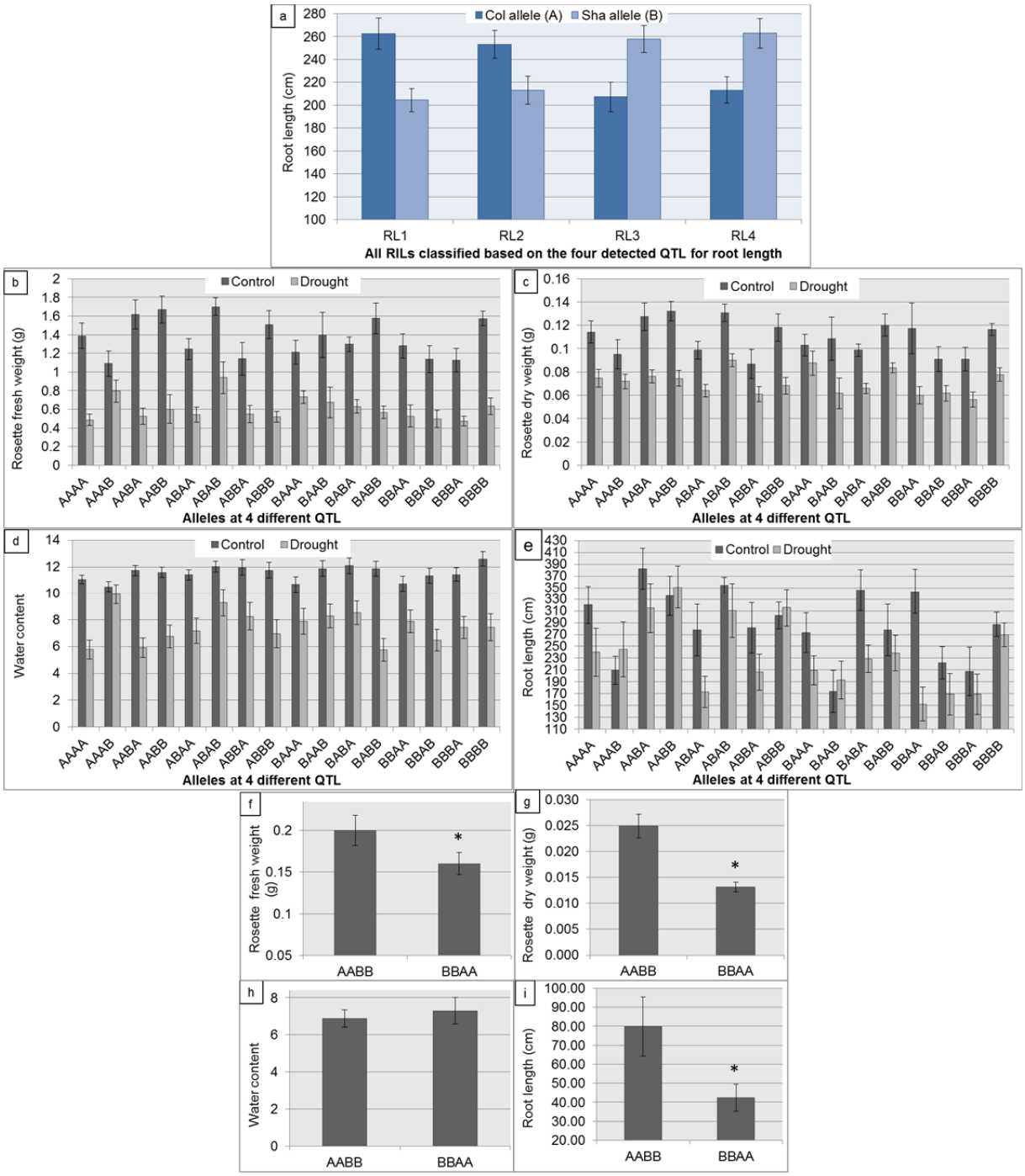


Figure 3: Analysis of the phenotypic effect of variation at the *RL1, RL2, RL3* and *RL4* QTL detected for root length in the Sha x Col RIL population (Table 5), on rosette fresh weight, rosette dry weight, water content and root length. (a) Root length for RILs classified based on the allelic effect of the Col allele (A) and the Sha allele (B) for each QTL. (b-e) All RILs were classified based on the four possible allelic combinations at the *RL* QTL and examined for rosette fresh weight (b) and dry weight (c), water content (d) and root length (e). (f-i) Results of a validation experiment using 13 RILs, seven RILs with AABB, and six RILs with BBAA genotypes, with three replications each, showing phenotypic comparison between both sets for rosette fresh weight, rosette dry weight water content and root length. Significant differences are indicated with * ($p<0.05$).

In general, the LD-grown Arabidopsis exhibited higher RosDW than the SD-grown Arabidopsis in control and in drought environments, but their WC was significantly less. This observation can be explained by an evolutionary selection for leaf cooling, via transpiration, rather than for water conservation, a process which is enhanced in continuous light condition (Crawford et al., 2012). Another mechanism to cope with drought is by increasing the proportion of roots relative to total biomass, as expressed in the RDW/RosDW ratio, upon drought exposure, which was observed here. This has also been found a population of epigenetic recombinant inbred lines (epiRIL) when grown on a mix of sand and field soil (Zhang et al., 2013). The authors suggested that upon drought, there will be positive selection on high root biomass allocation and in well-watered conditions, this will be selected against.

By comparing the two parental lines Col and Sha, we observed that Col had longer roots than Sha in the control environment, while Sha had longer roots in the drought environment. This is similar to what was reported for their response to potassium starvation (Kellermeier et al., 2013). The response of the parental lines is the same as we found for the effect of *RL4,* which is the only RL QTL with significant QxE. For this QTL the Sha allele increases RL in the drought environment. The effect of drought on root length taking all RL QTL together, was further shown by grouping all RIL RL data based on the *RL* QTL genotypes. Comparing the two extreme groups selected based on their RL (Figure 3 F-I) showed a significant decrease in one extreme group, not only for RL but also for RosFW and RosDW, whereas WC was not significantly different. This can be explained by assuming that a lower total root length means plants are likely to take up less water and as a consequence less nutrients leading to low RosFW and RosDW However, the WC of RIL's with shorter roots was almost similar to RIL's with longer root suggesting that stomatal closure, or any other physiological response affecting water household, may reduce water loss.

Some of the root associated QTL we found, co-locate with root trait QTL mapped in previous studies. For example, Galpaz and Reymond (2010), using the same RIL population, mapped six QTL for RL under control and salt-stress conditions, from which two co-located with the main effect QTL *RL1* and *RL2* we found. In addition, two QTL, *RLROSDW2* and *RSR2,* mapped for RL/RosDW and RDW/RosDW, respectively, were mapped together to the top of chromosome 3 and co-located with a RL QTL (Galpaz and Reymond, 2010), and with a QTL for the same traits mapped in the Bay-0 x Sha RIL population when grown in a hydroponics system (Bouteillé et al., 2012). For all loci mapping in this region, the Sha allele contributed to a decrease in the trait values However, no co-location with *RL3* was found in earlier studies.

In a study where different traits are analysed, co-location of QTL for different traits may suggest pleiotropy. However, with the relatively low resolution of QTL studies, close linkage of different genes can never be excluded. Once the molecular basis of one of the co-locating QTL has been identified, it may indicate the gene involved. The latter is of special interest with FT QTL of which many, including several that appear to segregate in the Sha x Col population, have been identified molecularly (reviewed in (Alonso-Blanco et al., 2009). For example, convincing evidences were shown for the pleiotropic effect of *FRI* on flowering time and water use efficiency (WUE) (Lovell et al., 2013). In the current study, *RL2, RV2* and *RDW2* were mapped in the same confidence interval of *QFT2*, which co-located with *FLOWERING LOCUS M* (*FLM*), mapped in the Nd x Col RIL population (Werner et al., 2005). This QTL showed significant QxE in response to day length, affecting the flowering time more in SD than in LD. A comparable effect was reported for the action of *FLM* (Werner et al., 2005), suggesting indeed both FT loci to be similar. In addition, the *QFT4* locus which was mapped earlier to the top of chromosome 3 in the same population (Simon et al., 2008), and in the Landsberg *erecta* (L*er*) x Sha population (El-Lithy et al., 2004), co-located with the *RLROSDW2* and *RSR2* loci. *QFT4* was the only FT QTL that showed QxE in response to drought, with the Col allele contributing to earlier flowering in SD and later flowering in LD.
QFT2 and *QFT4* were not the only loci co-locating with other traits, suggesting possible pleiotropic effects. For example, *QFT1,* with the Sha allele reducing flowering time, co-located with *FW1, WC1* and *WR1*, with the trait value enhancing effect coming from the Sha allele. *QFT1* was mapped very closely to the *FLOWERING LOCUS T* (*FT*) gene (Schwartz et al., 2009). The paralogue of *FT*, the *TWIN SISTER OF FT (TSF)* gene (Brachi et al., 2010), is found in the confidence interval of *QFT5*, a QTL with significant QxE where it was mapped only in LD. It appears to co-locate with another LD FT QTL, previously mapped in the Nd x Col population (Werner et al., 2005). Both genes play important roles in plant growth and development under SD and LD conditions (Hiraoka et al., 2013). *QFT3*, which was mapped only in SD, co-located with *WC2* and *WR2*. The *QFT3* mapping interval comprises the *EARLY FLOWERING 3* (*ELF3*) gene, which was shown to have pleiotropic effects on bolting date, FT, rosette diameter and leaf number in response to shade avoidance, as examined in the Bay x Sha RIL population (Jiménez-Gómez et al., 2010). The Bay and Sha alleles of *ELF3* are differentially regulating developmental time and circadian clock period length in a light-dependent manner. Both *QFT3* and *QFT5* had an effect on FT only in SD, respectively LD conditions. These seem to be typical examples of loci with conditional neutrality, meaning the QTL has a phenotypic effect in one environment, but no effect in another environment,

(Anderson et al., 2011; Fournier-Level et al., 2011; Fournier-Level et al., 2013; Leinonen et al., 2013).

A large percentage of the FT variance was explained by *QFT6*, which co-located with the major FT locus *FLOWERING LOCUS C* (*FLC*) (Lempe et al., 2005). *QFT6* co-located with *DW2*, for RosDW. *FLC* was shown earlier to have pleiotropic effects on WUE and nitrogen content (Loudet et al., 2003; McKay et al., 2003). The *FLC* gene, with a functional allele in Col and not Sha, is interacting with the other main flowering time gene *FRIGIDA (FRI),* which is probably functional in Sha, but not Col (McKay et al., 2003; Lempe et al., 2005; Moore and Lukens, 2011). The *QFT7* locus, was found to co-locate with the *FLC* paralogue, *MADS AFFECTING FLOWERING 2 (MAF2-5*) (Ratcliffe et al., 2003; Caicedo et al., 2009; Salomé et al., 2011). It also co-located with a QTL previously mapped in the L*er* x An-1 RIL population when grown in control and drought environments, which showed to have pleiotropic effects on leaf area and leaf number (Tisné et al., 2010). That agrees very well with the co-location we found of *QFT7* with the *FW6* and *DW3* biomass loci, and the *WC6* and *WR6* loci for the related traits WC and WR.

We mapped six QTL for RosFW of which three (*FW2, FW5* and *FW6*) also co-located with RosDW loci (*DW1*, *DW2* and *DW3*) and four (*FW1*, *FW4*, *FW5* and *FW6*) with water content loci (*WC1*, *WC4*, *WC5*, and *WC6*). These co-locations explain the significant correlation observed between these related traits. Two RosFW QTL, *FW2* and *FW3,* mapped in SD and LD conditions, co-localised with two rosette diameter QTL mapped previously in the same population when grown in 16 hrs day length (LD) (Simon et al., 2008), suggesting that the higher FW was due to larger rosette size.

Next to QTL with main effects and/or QxE, we mapped loci for phenotypic plasticity. Mapping QTL underlying plasticity that overlap with main effect or QxE QTL together with a few plasticity specific QTL, as was found before for plant-insect interaction in barley (Tétard-Jones et al., 2011), is comparable to what we found. Also the root and shoot traits we examined in the plant drought response of Arabidopsis, are likely to be important for environmental plasticity. The co-location of plasticity with main effect and GxE loci, supports the idea of a pleiotropic model suggesting that differential expression of constitutive genes / QTL across environments is responsible for plasticity (Lacaze et al., 2009; Tétard-Jones et al., 2011). Only two loci were mapped that were not also found as main effect or QxE loci. These were the *DW4* locus for RosDW and the *WR7* locus for WR. The only way to explain such loci is that at each environment or condition there is insufficient genetic variation to identify a QTL at the locus, but there is a phenotypic difference between both environments, for which a

QTL at the locus can be mapped. So these loci specifically explain the different phenotypic responses of drought and day length on rosette biomass and rosette water ratio.

In conclusion, this study investigated the differential genotypic response of Arabidopsis to drought stress when grown in two day length conditions. We found that Arabidopsis was more drought tolerant in LD than in SD, as reflected in the higher rosette biomass and the lower water content in LD, and noticed co-location of a number of rosette and root traits with flowering time, which suggests pleiotropy. We also found significant QxE for RL with differential allelic response similar to the response of the parental lines to drought. Longer roots is a desirable trait to acquire more water from deep soil, and therefore, breeding for this trait in drought tolerant breeding programs is demanded as long as there is no trade-offs with aboveground biomass. Further confirmation and fine mapping for the QTL detected in the current study is the next step to identify genes underlying those QTL and will be an important step to understand the molecular genetic networks underlying plant tolerance to drought and subsequently implement this in breeding programs.

Acknowledgements

We would like to thank Charles N. Moreira, Pingping Huang and Nihal Erol-Öztolan for their help in sowing and harvesting the plants used in the described experiments.

References

Alonso-Blanco C, Aarts MGM, Bentsink L, Keurentjes JJB, 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.

Alonso-Blanco C, El-Assal SED, Coupland G, Koornneef M. 1998. Analysis of natural allelic variation at flowering time loci in the Landsberg erecta and Cape Verde islands ecotypes of *Arabidopsis thaliana*. *Genetics* **149**, 749-764.

Anderson JT, Willis JH, Mitchell-Olds T. 2011. Evolutionary genetics of plant adaptation. *Trends in 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 SM. 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.

Boer MP, Wright D, Feng L, Podlich DW, Luo L, Cooper M, van Eeuwijk FA. 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.

Bouchabke O, Chang F, Simon M, Voisin R, Pelletier G, Durand-Tardif M. 2008. Natural variation in *Arabidopsis thaliana* as a tool for highlighting differential drought responses. *PLoS ONE* **3**, e1705.

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, Bergelson J, Cuguen J, Roux F. 2010. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. *PLoS Genetics* **6**, e1000940.

Caicedo AL, Richards C, Ehrenreich IM, Purugganan MD. 2009. Complex rearrangements lead to novel chimeric gene fusion polymorphisms at the Arabidopsis thaliana MAF2-5 flowering time gene cluster. *Molecular Biology and Evolution* **26**, 699-711.
Chiang GCK, Barua D, Kramer EM, Amasino RM, Donohue K. 2009. Major flowering time gene, FLOWERING LOCUS C, regulates seed germination in *Arabidopsis thaliana*. *Proce Nat Acad Sci USA* **106**, 11661-11666.
Crawford AJ, McLachlan DH, Hetherington AM, Franklin KA. 2012. High temperature exposure increases plant cooling capacity. *Current Biology* **22**, R396-R397.
El-Lithy ME, Clerkx EJ, Ruys GJ, Koornneef M, Vreugdenhil D. 2004. Quantitative trait locus analysis of growth-related traits in a new Arabidopsis recombinant inbred population. *Plant Physiol* **135**, 444-458.
Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM. 2011. A map of local adaptation in *Arabidopsis thaliana*. *Science* **334**, 86-89.
Fournier-Level A, Wilczek AM, Cooper MD, *et al.* 2013. Paths to selection on life history loci in different natural environments across the native range of *Arabidopsis thaliana*. *Molecular Ecology* **doi:10.1111/mec.12285**, n/a-n/a.
Galpaz N, Reymond M. 2010. Natural variation in *Arabidopsis thaliana* revealed a genetic network controlling germination under salt stress. *PLoS ONE* **5**, e15198.
Giraud E, Ho LHM, Clifton R, *et al.* 2008. The absence of ALTERNATIVE OXIDASE1a in Arabidopsis results in acute sensitivity to combined light and drought stress. *Plant Physiol* **147**, 595-610.
Hiraoka K, Yamaguchi A, Abe M, Araki T. 2013. The florigen genes FT and TSF modulate lateral shoot outgrowth in Arabidopsis thaliana. *Plant and Cell Physiology* **54**, 352-368.
Holland JB. 2007. Genetic architecture of complex traits in plants. *Curr Opin Plant Biol* **10**, 156-161.
Huang X, Effgen S, Meyer RC, Theres K, Koornneef M. 2012. Epistatic natural allelic variation reveals a function of AGAMOUS-LIKE6 in axillary bud formation in Arabidopsis. *The Plant Cell* **24**, 2364-2379.
Jiménez-Gómez JM, Wallace AD, Maloof JN. 2010. Network analysis identifies *ELF3* as a QTL for the shade avoidance response in Arabidopsis. *PLoS Genetics* **6**, e1001100.
Juenger TE. 2013. Natural variation and genetic constraints on drought tolerance. *Curr Opin Plant Biol* **16**, 274-281.
Kamoshita A, Zhang J, Siopongco J, Sarkarung S, Nguyen HT, Wade LJ. 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 Physiol*.
Kozuka T, Suetsugu N, Wada M, Nagatani A. 2013. Antagonistic regulation of leaf flattening by phytochrome B and phototropin in Arabidopsis thaliana. *Plant and Cell Physiology* **54**, 69-79.
Lacaze X, Hayes PM, Korol A. 2009. Genetics of phenotypic plasticity: QTL analysis in barley, *Hordeum vulgare*. *Heredity (Edinb)* **102**, 163-173.
Leinonen PH, Remington DL, Leppálá J, Savolainen O. 2013. Genetic basis of local adaptation and flowering time variation in *Arabidopsis lyrata*. *Molecular Ecology* **22**, 709-723.
Lempe J, Balasubramanian S, Sureshkumar S, Singh A, Schmid M, Weigel D. 2005. Diversity of flowering responses in wild *Arabidopsis thaliana* strains. *PLoS Genetics* **1**, e6.
Li J, Ji L. 2005. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity (Edinb)* **95**, 221-227.
Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F. 2002. Bay-0 x Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in Arabidopsis. *Theoretical and Applied Genetics* **104**, 1173-1184.
Loudet O, Chaillou S, Krapp A, Daniel-Vedele F. 2003. Quantitative trait loci analysis of water and anion contents in interaction with nitrogen availability in *Arabidopsis thaliana*. *Genetics* **163**, 711-722.
Lovell JT, Juenger TE, Michaels SD, Lasky JR, Platt A, Richards JH, Yu X, Easlon HM, Sen S, McKay JK. 2013. Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. *Phil. Trans. R. Soc. B* **280**, 20131043.
Malosetti M, Ribaut JM, van Eeuwijk FA. 2013. The statistical analysis of multi-environment data: modeling genotype-by-environment interaction and its genetic basis. *Frontiers in Physiology* **4**, 44.
Malosetti M, Voltas J, Romagosa I, Ullrich SE, van Eeuwijk FA. 2004. Mixed models including environmental covariables for studying QTL by environment interaction. *Euphytica* **137**, 139-145.
McKay JK, Richards JH, 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 JK, Richards JH, Nemali KS, Sen S, Mitchell-Olds T, Boles S, Stahl EA, Wayne T, Juenger TE. 2008. Genetics of drought adaptation in *Arabidopsis thaliana* II. QTL analysis of a new mapping population, KAS-1 × TSU-1. *Evolution* **62**, 3014-3026.

Melzer S, Lens F, Gennen J, Vanneste S, Rohde A, Beeckman T. 2008. Flowering-time genes modulate meristem determinacy and growth form in *Arabidopsis thaliana*. *Nature Genetics* **40**, 1489-1492.
Méndez-Vigo B, Gomaa NH, Alonso-Blanco C, Xavier Picó F. 2013. Among- and within-population variation in flowering time of Iberian *Arabidopsis thaliana* estimated in field and glasshouse conditions. *New Phytologist* **197**, 1332-1343.
Mittler R. 2006. Abiotic stress, the field environment and stress combination. *Trends in Plant Science* **11**, 15-19.
Moore S, Lukens L. 2011. An evaluation of Arabidopsis thaliana hybrid traits and their genetic control. *G3: Genes, Genomes, Genetics* **1**, 571-579.
Ratcliffe OJ, Kumimoto RW, Wong BJ, Riechmann JL. 2003. Analysis of the Arabidopsis MADS AFFECTING FLOWERING gene family: MAF2 prevents vernalization by short periods of cold. *The Plant Cell* **15**, 1159-1169.
Salomé PA, Bomblies K, Laitinen RAE, Yant L, Mott R, Weigel D. 2011. Genetic architecture of flowering-time variation in *Arabidopsis thaliana*. *Genetics* **188**, 421-433.
Schwartz C, Balasubramanian S, Warthmann N, *et al.* 2009. Cis-regulatory changes at FLOWERING LOCUS T mediate natural variation in flowering responses of *Arabidopsis thaliana*. *Genetics* **183**, 723-732.
Simon M, Loudet O, Durand S, Bérard A, Brunel D, Sennesal F-X, Durand-Tardif M, Pelletier G, 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.
Smith S, De Smet I. 2012. Root system architecture: insights from Arabidopsis and cereal crops. *Phil. Trans. R. Soc. B* **367**, 1441-1452.
Tardieu F. 2013. Plant response to environmental conditions: assessing potential production, water demand and negative effects of water deficit. *Frontiers in Physiology* **4:17. 10.3389/fphys.2013.00017**.
Tétard-Jones C, Kertesz MA, Preziosi RF. 2011. Quantitative trait loci mapping of phenotypic plasticity and genotype–environment interactions in plant and insect performance. *Phil. Trans. R. Soc. B* **366**, 1368-1379.
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 and Environment* **33**, 1875-1887.
Tuberosa R. 2012. Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology* **3:347. doi:10.3389/fphys.2012.00347**.
Tuberosa R, Salvi S. 2006. Genomics-based approaches to improve drought tolerance of crops. *Trends in Plant Science* **11**, 405-412.
Vile D, Pervent M, Belluau M, Vasseur F, Bresson J, Muller B, Granier C, Simonneau T. 2012. Arabidopsis growth under prolonged high temperature and water deficit: independent or interactive effects? *Plant, Cell and Environment* **35**, 702-718.
Weigel D. 2012. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiol* **158**, 2-22.
Werner JD, Borevitz JO, Warthmann N, Trainer GT, Ecker JR, 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. *Proce Nat Acad Sci USA* **102**, 2460-2465.
Zhang Y-Y, Fischer M, Colot V, Bossdorf O. 2013. Epigenetic variation creates potential for evolution of plant phenotypic plasticity. *New Phytologist*, 314-322.
Zhao F, Xu S. 2012. Genotype by environment interaction of quantitative traits: a case study in Barley. *G3: Genes, Genomes, Genetics* **2**, 779-788.

Chapter 3: Genome wide association mapping for genotype by environment interaction and plasticity associated with drought response in *Arabidopsis thaliana*

Mohamed El-Soda[1,2] , Willem Kruijer[3], Maarten Koornneef[1,4], Mark G.M. Aarts[1]

[1]Laboratory of Genetics, Wageningen University, Wageningen, the Netherlands; [2]Department of Genetics, Faculty of Agriculture, Cairo University, Egypt; [3]Biometris - Applied Statistics, Department of Plant Science, Wageningen University, Wageningen, the Netherlands; [4]Max Planck Institute for Plant Breeding Research, Cologne, Germany.

Abstract

Genome wide association mapping is a recently developed approach in plants for identifying genes underlying natural genetic variation. We applied this approach to test the feasibility of mapping SNPs associated with main effect, genotype by environment interaction (GxE) and plasticity in *Arabidopsis thaliana.* We grew a so-called HapMap population of 350 genotypes on sand in control and drought environments and compared them for various rosette and root traits. We mapped 16 significant SNPs with main effect, 58 SNPs underlying GxE, and 100 SNPs affecting plasticity. Of these, 23 SNPs were commonly associated between GxE and plasticity. To find confirmation of candidate genes we compared amino acid sequences of two genes between extreme differential accessions from the assessed accession panel. One of those candidate genes co-located with a previously mapped QTL in the Sha x Col recombinant inbred line population. Additional candidate genes known to be affected in their differential expression by abiotic stress as well as novel genes associated with rosette and root growth traits in control and drought environments are evaluated.

Introduction

Environmental abiotic stresses, and especially drought, affect plant growth and limit crop productivity worldwide. With the predicted global temperature increase, drought is likely to become more threatening (Smith and De Smet, 2012; Assmann, 2013). Therefore, plants have evolved three adaptive strategies, known as drought escape, drought avoidance and drought tolerance, to colonize terrestrial ecosystems with limited water availability and to optimise their growth under such conditions (Tuberosa, 2012; Assmann, 2013; Juenger, 2013; Tardieu, 2013). Evaluating such responses in many genotypes across several environments is expected

to reveal phenotypic plasticity, which is the ability of a genotype to produce distinct phenotypes in different environments (Nicotra et al., 2010; Assmann, 2013; Juenger, 2013). When plasticity differs between genotypes, this is described as genotype by environment interaction (GxE) (Weigel, 2012; Assmann, 2013). A key development in the genetic dissection of such complex responses is the analysis of the associations between those responses and genetic markers via traditional QTL mapping and genome wide association studies (GWAS) (Flint-Garcia et al., 2003; Gupta et al., 2005; Stich and Melchinger, 2010). QTL mapping is robust and can detect the effects of rare alleles when the parents differ for such alleles. However, generating the mapping populations as well as fine-mapping is time consuming, the mapping resolution is relatively low, and only those genes that segregate in the cross can be identified. Therefore, GWAS, which does not require crosses, but uses the historical recombination events in the pedigree of a genotype, has emerged as an alternative for fine-mapping the causal genes underlying genomic regions associated with phenotypic variation. (Aranzana et al., 2005; Zhao et al., 2007; Nordborg and Weigel, 2008; Atwell et al., 2010; Bergelson and Roux, 2010; Weigel, 2012). However, GWAS is often not powerful enough to detect the effect of rare alleles or GxE (Eichler et al., 2010; Thomas, 2010; Gibson, 2012). Nevertheless, considering GxE in GWAS will help to discover novel genes that act synergistically with environment (Thomas, 2010) and to identify superior and stable genotypes across different environments (Filiault and Maloof, 2012). Recently, a multi-trait mixed model (MTMM) approach was proposed for multi-trait or multi-environment association mapping (Korte et al., 2012). In this approach, a marker can have different effects in different environments, therefore explaining at least part of the GxE interaction in terms of QTL by environment interactions (QxE).

For such studies, *Arabidopsis thaliana* (Arabidopsis) is a suitable model plant, mainly as there are populations of natural accessions originating from a large geographic distribution, including sequence information for over 450 such accessions. In addition, there is high density, genome-wide, single-nucleotide polymorphism information available for Arabidopsis, which showed that linkage disequilibrium (LD) extends between 5 to 10 kb, or only one to two genes (Bergelson and Roux, 2010; Hancock et al., 2011; Weigel, 2012; Assmann, 2013).

Here we report on screening a haplotype map (HapMap) population of ~350 genetically diverse Arabidopsis accessions collected from all over its natural distribution range (Li et al., 2010) and on mapping putative candidate genes associated with main effects on growth

parameters and with a differential morphological response in control and drought environments.

Materials and Methods

Mapping population and experimental setup

Three hundred fifty accessions of the HapMap population genotyped with 214.051 bi-allelic (Col or non-Col) non-singleton SNPs (Li et al., 2010) were used for genome wide association (GWA) mapping. This population was phenotyped under short days (SD), 10 hrs light, and temperature-controlled greenhouse conditions, together with the Sha x Col RIL population described in chapter 4. The environmental conditions were 69% relative humidity, 21.3 °C during the day and 74% relative humidity, 17.5 °C during the night. The experimental setup was a randomized block design, with one replicate per accession per block and three blocks per treatment. As described for the SD experiment in chapter 4, plants were grown on silver sand covered with a thin layer of sieved peat in 7 cm deep, squared plastic pots, to facilitate measuring roots at the end of the experiment. All pots were watered with nutrient solution at the start of the experiment, until saturation, i.e. 100% soil water holding capacity (SWHC). Nutrient solutions contained 1, 1.1 and 5.9 mM N, P and K, respectively. Per pot, two seeds were sown and three days after germination, seedlings were thinned to one. Twelve days after germination, water was withheld as drought treatment, while the control treatment was continuously watered every 2 or 3 days, when SWHC had dropped below 80%, until the end of the experiment. After 34 days rosettes were harvested and fresh weight (RosFW) was measured. Rosettes were oven-dried at 65 °C for three days for dry weight (RosDW) measurements. Roots were washed carefully, placed in a plastic tray filled with water and scanned with a flatbed scanner. With the scans, total root length (RL) and root diameter (DIAM) were measured using WinRhizo (Regent Instruments Inc., Quebec, Canada) and thereafter root DW (RDW) was measured. Finally RL/RosDW was calculated.

Statistical and GWA mapping analysis

Statistical analysis was performed on the raw data using SPSS for Windows Rel. 19 (Chicago: SPSS Inc). Analysis of variance (ANOVA) was used to test for the significance of treatment differences, lines and for the GxE. The broad sense heritability (h^2) was calculated as

$$h^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_e^2/r},$$

where r is the number of replicates, and the genetic variance σ_g^2 and environmental variance σ_e^2 are estimated using the ANOVA mean sums of squares: $\hat{\sigma}_g^2 = (MS(G) - MS(E))/r$ and $\hat{\sigma}_e^2 = MS(E)$. Genotypic means from the analysed traits were used for for both GWA mapping of plasticity and bivariate GWA-mapping using the MTMM approach (see below). For each accessions and for each measured trait, phenotypic plasticity was calculated as the difference in genotypic mean between treatments (Tétard-Jones et al., 2011). For mapping plasticity, univariate GWA mapping (Kang et al., 2010) was performed. In both uni- and bivariate GWA mapping, only SNPs with a minor allele frequency above 0.05 were considered, and a general significance threshold of $-^{10}\log(P)=4$ was used. For all reported genes, the description was from The Arabidopsis Information Resource (TAIR) vs. 10 (www.arabidopsis.org).

Multi Trait Mixed Model approach

Korte et al. (2012) proposed multi-trait mixed model (MTMM) association mapping, which can model both genetic and environmental correlations between traits. Here we use MTMM for bivariate association mapping on traits measured under drought and control conditions. Marker effects are either common to both environments or environment-specific. The latter case is referred to as QTL by environment (Q x E) interaction. Following the notation of Korte et al., (2012), it is assumed that the trait under stress y_1 and under control y_2 follow a multivariate normal distribution defined by

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = s_1\mu_1 + s_2\mu_2 + x\beta + (x \times s_1)\alpha + v, \qquad (1)$$

$$Var(v) = \begin{pmatrix} \sigma_{g,1}^2 K & \rho_g \sigma_{g,1}\sigma_{g,2} K \\ \rho_g \sigma_{g,1}\sigma_{g,2} K & \sigma_{g,1}^2 K \end{pmatrix} + \begin{pmatrix} \sigma_{e,1}^2 I_n & 0 \\ 0 & \sigma_{e,1}^2 I_n \end{pmatrix}, \quad (2)$$

Where s_1 is the column vector which is 1 for the observations taken under stress and 0 elsewhere; similarly s_2 is the column vector which is 1 for the observations taken under control. x is the vector of $2n$ marker scores, n being the number of accessions. The parameters μ_1 and μ_2 represent the environment-specific means, β is the main effect of the genetic marker and α its environment specific effect. The covariance structure of the errors depends on environment-specific genetic variances ($\sigma_{g,1}^2$ and $\sigma_{g,2}^2$) and residual variances ($\sigma_{e,1}^2$

and $\sigma^2_{e,2}$), as well as the genetic correlation ρ_g. When different traits are observed on the same individual, there is an environmental correlation as well (Korte et al., 2012), this is however not the case here. The kinship matrix K is the identity-by-state matrix based on all SNPs.
Korte et al. (2012) showed that it is straightforward to extend this model to situations where some observations are missing (supplementary files available at http://www.nature.com/ng/journal/v44/n9/full/ng.2376.html#supplementary-information), 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.
For each marker we test the full model against the null model ($\alpha = \beta = 0$). On those markers with a p-value smaller than 10^{-4}.

- We first test for marker by environment interaction: the full model against the model where $\alpha = 0$
- When there is no significant marker by environment interaction (the p-value in the preceding test being larger than 0.05), we report an estimate $\hat{\beta}$ for model (1) without the term $(x \times s_1)\alpha$.
- If there is significant marker by environment interaction, then our estimates of the marker effects are environment specific: $\hat{\beta}_1 = \hat{\beta} + \hat{\alpha}$ and $\hat{\beta}_2 = \hat{\beta}$. P-values are in this case obtained from testing the hypotheses $\beta_1 = 0$ and $\beta_2 = 0$ in the closely related model

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = s_1\mu_1 + s_2\mu_2 + (x \times s_1)\beta_1 + (x \times s_2)\beta_2 + v.$$

Univariate GWA-mapping

Using the methodology of Kang et al., (2010), we performed GWA-mapping on differences of trait means (i.e. plasticity). In the first step the model:

$$Y = \mu 1_n + g + e \qquad (3)$$

is assumed, where $Y = (Y_1, \ldots, Y_n)'$ is the vector of genotypic means of the plasticity of the $n = 350$ genotypes, μ is the population mean, $1_n = (1, \ldots, 1)'$ is the vector of ones. The vector $e = (e_1, \ldots, e_n)' \sim N(0, \sigma^2_e I_n)$ consists of independent environmental noise, and the genotypic effects $g = (g_1, \ldots, g_n)'$ have a zero mean Gaussian distribution with covariance matrix $\sigma^2_g K$. The kinship matrix K is the identity-by-state matrix based on all available

SNPs. For this model, we obtain REML-estimates $\hat{\sigma}_g^2$ and $\hat{\sigma}_e^2$. Conditional on these estimates, we test whether the marker effects are zero. This is done by adding a marker effect to (3) :

$$Y = \mu 1_n + x\beta + g + e, \qquad (4)$$

where x = (x_1,...., x_n) $^{/}$ is the vector of marker score. We then tested, for each marker in turn, the hypothesis $\beta = 0$. This is only done for the markers whose minor allele frequency is at least 0.05.

Analysis of extreme accessions

Based on a trait ratio between drought and control environments, the whole population was classified twice: once based on RosDW ratio, and the second time based on RL ratio. For each classification, two contrasting groups, each composed of 10 accessions, one with the smallest (drought sensitive) and another with the largest (drought tolerant) RosDW or RL ratios, were made. Both sets were compared regarding the amino acid sequence of proteins encoded by candidate genes identified by mapping relevant traits, using the Arabidopsis 1001 genomes browser http://signal.salk.edu/atg1001/3.0/gebrowser.php. To predict the function of candidate genes, their protein domains were compared with those of other well–described, related genes using the Simple Molecular Architecture Research Tool (SMART) browser http://smart.embl-heidelberg.de/.

Results

Phenotyping the HapMap population

RosFW, RosDW, RL, RDW and RL/RosDW were analysed for the 350 accessions grown in control and drought environments. When trait averages of all accessions were considered together, apart from root length and root dry weight in the drought environment, the values of RosFW, RosDW and RL/RisDW were significantly higher in the control environment (Figure 1). ANOVA (Table 1) showed a significant difference between control and drought environment for all traits except for RosDW.

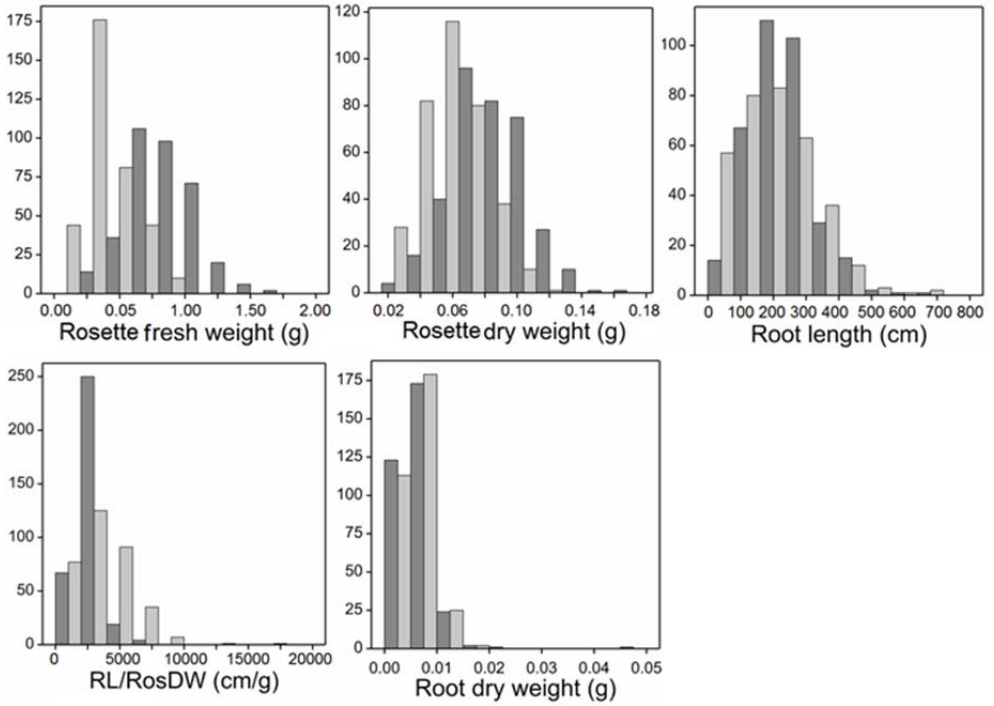


Figure 1: Frequency distributions of the non-normalized reported traits for the 350 accessions grown in control (dark grey bars), and drought (light grey bars) environments. The vertical axes indicate the number of accessions per trait value class, and the horizontal axes indicate the different trait value classes. RL/RosDW refers to the ratio between root length and rosette dry weight.

Table 1: ANOVA table showing sum of squares of the mean for three replications of rosette fresh weight and dry weight and two replications of root length and root dry weight of HapMap plants grown under well-watered conditions (control) and under drought conditions (drought). GxE is the genotype by environment interaction. * = significant at 0.001 level. h^2 = broad sense heritability.

	ANOVA			h^2
Trait	Accessions	Environment	GxE	
Fresh weight control	63.4 *	99.2 *	27.7 *	0.60
Fresh weight drought				0.55
Dry weight control	0.62 *	0.32 *	0.18	0.54
Dry weight drought				0.51
Root length control	9442071*	273204*	5826771*	0.69
Root length drought				0.80
Root dry weight control	0.047*	0.000053*	0.028*	0.70
Root dry weight drought				0.75

Correlation analysis (Table 2) showed a significant, positive correlation between all measured traits in both environments except for RL/RosDW which showed a significant, but negative, correlation with RosFW and RosDW in the drought environment.

Table 2: Pearson correlations for the analysed traits in control (C) and drought (D) environments. RL/RosDW refers to the ratio between root length and rosette dry weight. Fresh and dry weight values are averages of three replications, while root traits are averages of two replications. Dark grey refers to a significance level of 0.01, while light grey refers to a significant level of 0.05.

Trait		1	2	3	4	5	6	7	8	9
Rosette fresh weight -C	1	1								
Rosette dry weight –C	2	.921	1							
Root length –C	3	.660	.657	1						
RL/RosDW – C	4	.133	.139	.762	1					
Root dry weight –C	5	.442	.439	.519	.324	1				
Rosette fresh weight - D	6	.387	.333	.364	.189	.244	1			
Rosette dry weight - D	7	.554	.512	.430	.133	.291	.841	1		
Root length –D	8	.153	.194	.179	.061	.126	.227	.244	1	
RL/RosDW – D	9	-.126	-.078	-.027	.027	-.005	-.188	-.264	.798	1
Root dry weight –D	10	.314	.319	.289	.105	.236	.391	.439	.528	.272

Genome wide association mapping, GxE and candidate genes

The MTMM approach used here enabled us to identify 74 SNPs we considered associated with a trait, of which 16 SNPs indicated a main effect on the trait and 58 SNPs indicated a QxE effect (Figure 2). The 74 SNPs are in or closest to 69 genes, from which Table 3 present genes with significant QxE effect that have functions related to drought or abiotic stresses based on the biological function in TAIR 10 (www.arabidopsis.org). For example, a SNP in *TARGET OF RAPAMYCIN* (*TOR*) (Dobrenel et al., 2011; Ren et al., 2012; Caldana et al., 2013) was associated with RosDW where the non-Col allele increased RosDW in both environments, but the increase was three times higher in the drought environment than in the control (Table 3). Another SNP with significant QxE, was associated with RDW and with its plasticity response. This SNP was mapped in the *SNF1-RELATED PROTEIN KINASE 2.2* (*SnRK2.2*) gene (Fujii et al., 2007; Fujii et al., 2011; Kulik et al., 2011). The effect of the Col

allele of this SNP was the opposite, and four times higher, in the drought environment compared to the control.

An additional association was found between a SNP mapped in the *SHORT ROOT HAIR 1* (*SRH1*) gene (Huang et al., 2013) and RDW. The effect of the non-Col allele in the drought environment was twice the effect of the Col allele in the control environment. Another significant SNP was mapped in the *RIBOSOMAL PROTEIN S3* (*RPS3*) gene (Barakat et al., 2001) and was associated with RL/RosDW with significant QxE effect. The effect of the non-Col allele of this SNP in the drought environment was twice the reversed effect of the allele in the control environment. This SNP was in the interval of the previously mapped RDW QTL, *RDW4*, in chapter 2.

Two relevant SNPs were found to be associated with –log10(p) values below the threshold of 4 (respectively at = 3.6 and 3.7), but because they resided in the confidence intervals of two QTL mapped in the Sha x Col RIL population (chapter 2), *DW2,* for RosDW*,* and *RL4,* for RL, respectively, they were still considered to indicate candidate genes, potentially underlying the QTL. The SNP associated with RosDW located in the *FLOWERING LOCUS C* (*FLC*) gene (Andres and Coupland, 2012), while the SNP associated with RL located with the *YELLOW LEAF SPECIFIC 7* (*YLS7*) gene (Yoshida et al., 2001), also known as *TRICHOME-BIREFRINGENCE-LIKE 17 (TBL17*).

Mapping plasticity

The ability of a genotype to produce distinct phenotypes in different environments as a way to adapt to the environment, is known as phenotypic plasticity. Out of 100 significant SNPs mapped for plasticity, the difference between trait values in control and drought environments (Figure 3), 23 SNPs were in common with the SNPs mapped with QxE effect (as indicated in Table 3). Table 4 lists SNPs associated with plasticity, selected based on a predicted biological function related to drought stress response. The SNP previously mapped in *SnRK2.2* for its QxE effect on RDW, was also found for plasticity in RDW and in RL. A SNP associated with RosFW was mapped in the *GLYCINE-RICH RNA-BINDING PROTEIN 4* (*GR-RBP4*) gene, related to cold adaptation (Kwak et al., 2011). This SNP was specific for FW plasticity . Another SNP associated with RosFW was mapped in the *GERANYLLINALOOL SYNTHASE 4 (GES4)* gene (Attaran et al., 2008). A SNP in a gene encoding ARABIDOPSIS THALIANA Delta(3) (Goepfert et al., 2008), was associated with RL/RosDW. In addition to a SNP that was associated with both RosDW and mapped in the *ASCORBATE PEROXIDASE 5* (*APX5*) (Panchuk et al., 2002), two more SNPs were

associated with RDW and mapped in *CYTOCHROME P450 710A2* (*CYP710A2*) (Morikawa et al., 2006) and *NUCLEOSIDE DIPHOSPHATE KINASE 3* (*NDPK3*) (Hammargren et al., 2008) genes with functions related to the response to oxidative stress. A SNP in the *RESPONSIVE TO DEHYDRATION 21B* (*RD21B*) gene (Shindo et al., 2012), a gene which is highly expressed in pollen and in roots, was associated with RDW. Seven SNPs mapping to the *FORMS APLOID AND BINUCLEATE CELLS 1C* (*FAB1C*) gene (Whitley et al., 2009) are associated with RL/RosDW.

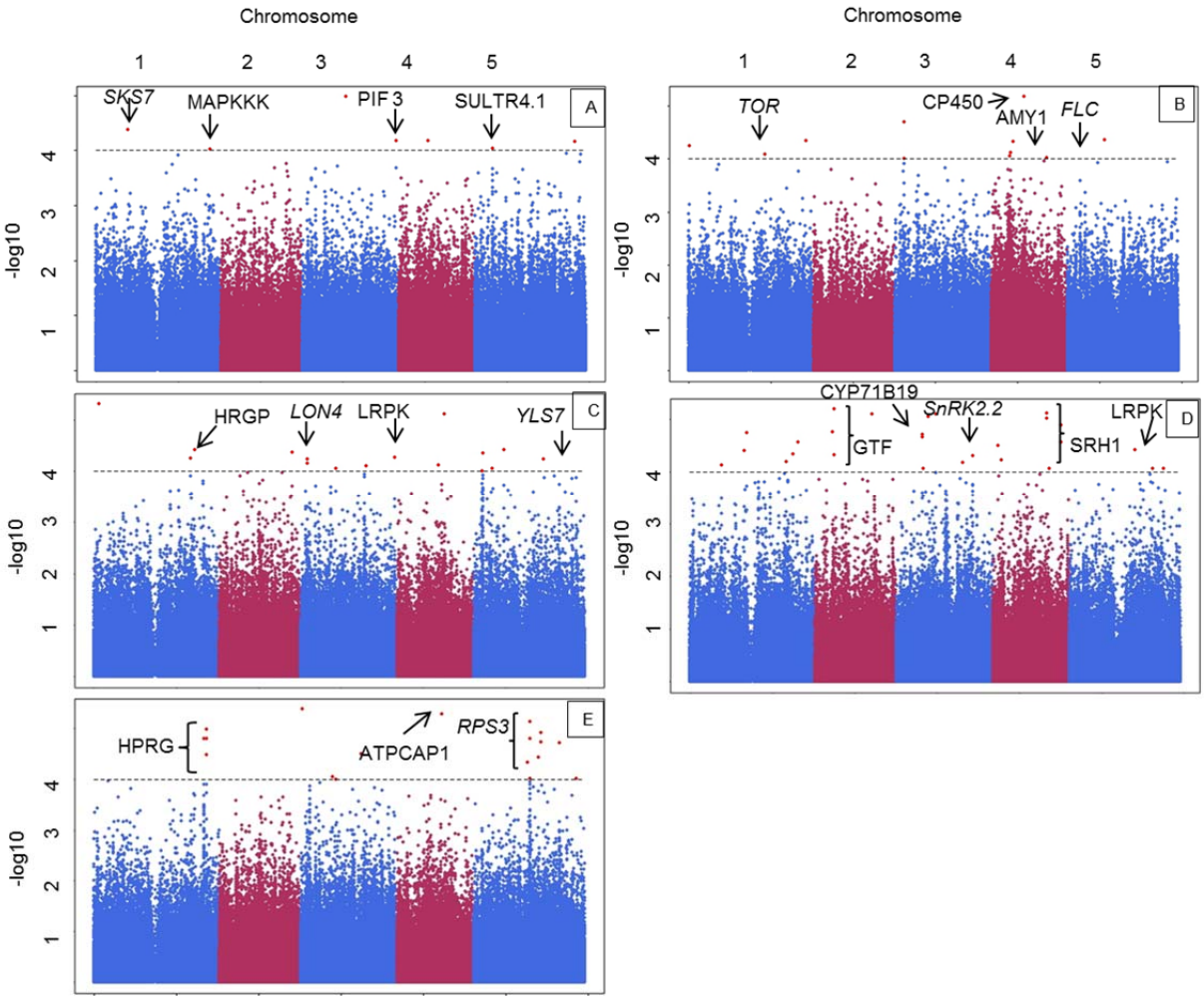


Figure 2: Manhattan plots displaying GWAS results of the MTMM approach showing the -log10(p) values for SNPs associated with rosette fresh weight (A), rosette dry weight (B), root length (C), root dry weight (D), and the ratio between root length and rosette dry weight (E). Arrows indicates significant SNPs in genes with annotated functions, as reported in table 3, next to SNPs in *YELLOW LEAF SENECENCE7* and *FLOWERING LOCUS C*, which were below the –log10(p) threshold of 4.

Table 3: List of candidate genes with a QxE effect on shoot and root traits, selected based on GWA analysis of the Arabidopsis HapMap population. RL/RosDW refers to the ratio between root length and rosette dry weight. MAF is minor allele frequency. –log10(p) indicates the QTL significance level, with the values between parentheses those obtained for univariate mapping of plasticity. β indicates the effect of a SNP in drought or in control environments with positive values when the positive effect on the trait value is coming from the Col allele. Chromosome numbers (Chr.) and SNP positions in kb are given. Both SNP position and description is based on TAIR v.10 (www.arabidopsis.org). LD lists genes found to be in linkage disequilibrium within 10 kb on both sides of the significant SNP.

Trait (unit)	Chr.	Gene	SNP Pos. (Kb)	MAF	-log10(p)	β drought	β control	Description	LD
Rosette fresh weight (g)	1	AT1G21860	7671	0.371	4.4 (2.9)	-0.587	0.214	SKU5 similar 7 (sks7). Involved in response to oxidative stress	
	1	AT1G73660	27691	0.317	4.0 (0.9)	0.592	0.229	Encodes a protein with similarity to MAPKKKs.Involved in photoperiodism, flowering, protein phosphorylation, response to salt stress	AT1G73640-AT1G73650
	3	AT3G62090	22989	0.174	4.2 (2.9)	0.722	-0.272	PHYTOCHROME INTERACTING FACTOR 3-LIKE 2 (PIF 3). Involved in cell wall macromolecule metabolic process, red or far-red light signaling pathway, xylem development .	AT3G62080
	5	AT5G13550	4355	0.446	4.0 (1.4)	-0.656	-0.136	SULFATE TRANSPORTER 4.1 (SULTR4;1). sulfate transport, transmembrane transport	AT5G13510 toAT5G13560
Rosette dry weight (g)	1	AT1G50030	18528	0.154	4.1 (3.7)	-0.76	-0.237	TARGET OF RAPAMYCIN (TOR). Involved in cell adhesion, cell division, cell wall organization	AT1G50010
	4	AT4G13770	7992	0.194	5.2 (5.6)	-0.321	0.615	CYTOCHROME P450 (CYP83A1). Response to UV, response to insect,	
	4	AT4G25000	12852	0.134	4.0 (3.9)	-0.738	-0.023	Alpha-amylase-like (AMY1). Response to abscisic acid stimulus and gibberellin	AT4G24974
Root length (cm)	1	AT1G63540	23567	0.203	4.3 (1.3)	-0.749	-0.161	Hydroxyproline-rich glycoprotein family protein (HRGP) (unknown function)	AT1G63530
	3	AT3G05790	1727	0.414	4.2 (2.8)	0.577	-0.089	LON PROTEASE 4 (LON4). Involved in cellular response to oxidative stress	
	3	AT3G23750	8557	0.140	4.1 (3.2)	0.231	0.802	Leucine-rich repeat protein kinase family (LRPK). Involved in protein phosphorylation	

Table 3 (continued): List of candidate genes with a QxE effect on shoot and root traits, selected based on GWA analysis of the Arabidopsis HapMap population. RL/RosDW refers to the ratio between root length and rosette dry weight. MAF is minor allele frequency. –log10(p) indicates the QTL significance level, with the values between parentheses those obtained for univariate mapping of plasticity. β indicates the effect of a SNP in drought or in control environments with positive values when the positive effect on the trait value is coming from the Col allele. Chromosome numbers (Chr.) and SNP positions in kb are given. Both SNP position and description is based on TAIR v.10 (www.arabidopsis.org). LD lists genes found to be in linkage disequilibrium within 10 kb on both sides of the significant SNP.

Trait (unit)	Chr.	Gene	SNP Pos. (Kb)	MAF	-log10(p)	β drought	β control	Description Tair 10	LD
Root dry weight (g)	2	AT2G32430	13773	0.074	5.1 (5.8)	-1.679	0.903	Galactosyltransferase family protein (GTF) involved in protein glycosylation	AT2G32440 to AT2G32480
	3	AT3G26170	9576	0.074	5.1 (3.7)	-1.886	0.411	Cytochrome P450 (CYP71B19), involved in response to oxidative stress	-
	3	AT3G50500	18742	0.394	4.3 (4.3)	0.866	-0.346	SNF1-Related protein Kinase 2-2 (SnRK2-2). Involved in response to osmotic, salt stresses and water deprivation	-
	4	AT4G34580	16515	0.103	4.9 (3.3)	-1.497	0.811	SHORT ROOT HAIR 1 (SRH1). Involved in lateral root development and root hair elongation	-
	5	AT5G49780	20230	0.063	4.1 (3.7)	-1.872	-0.218	Leucine-rich repeat protein kinase family (LRPK). Involved in protein phosphorylation	-
RL/RosDW (cm/g)	1	AT1G72600	27340	0.32	4.8 (4.0)	0.165	-0.772	Hydroxyproline-rich glycoprotein family protein (HPRG) (unknown function)	AT1G72620
	3	AT3G02250	425	0.06	5.4 (0.3)	-2.942	-0.508	Involved in response to salt stress	-
	4	AT4G20260	10942	0.086	5.3 (2.2)	2.375	-0.329	ARABIDOPSIS THALIANA PLASMA-MEMBRANE ASSOCIATED CATION-BINDING PROTEIN 1 (ATPCAP1). Involved in reponse to cold, salt, osmotic, water stresses and root hair elongation	-
	5	AT5G35530	13711	0.274	4.8 (3.8)	-1.082	0.556	Ribosomal protein S3 family protein. Involved in response to salt stress	AT5G35526, AT5G35535, AT5G35540, AT5G35550
	5	AT5G35580	13760	0.477	5.1 (2.1)	-1.311	0.232	Protein kinase superfamily protein. Involved in protein phosphorylation, expressed in roots	-

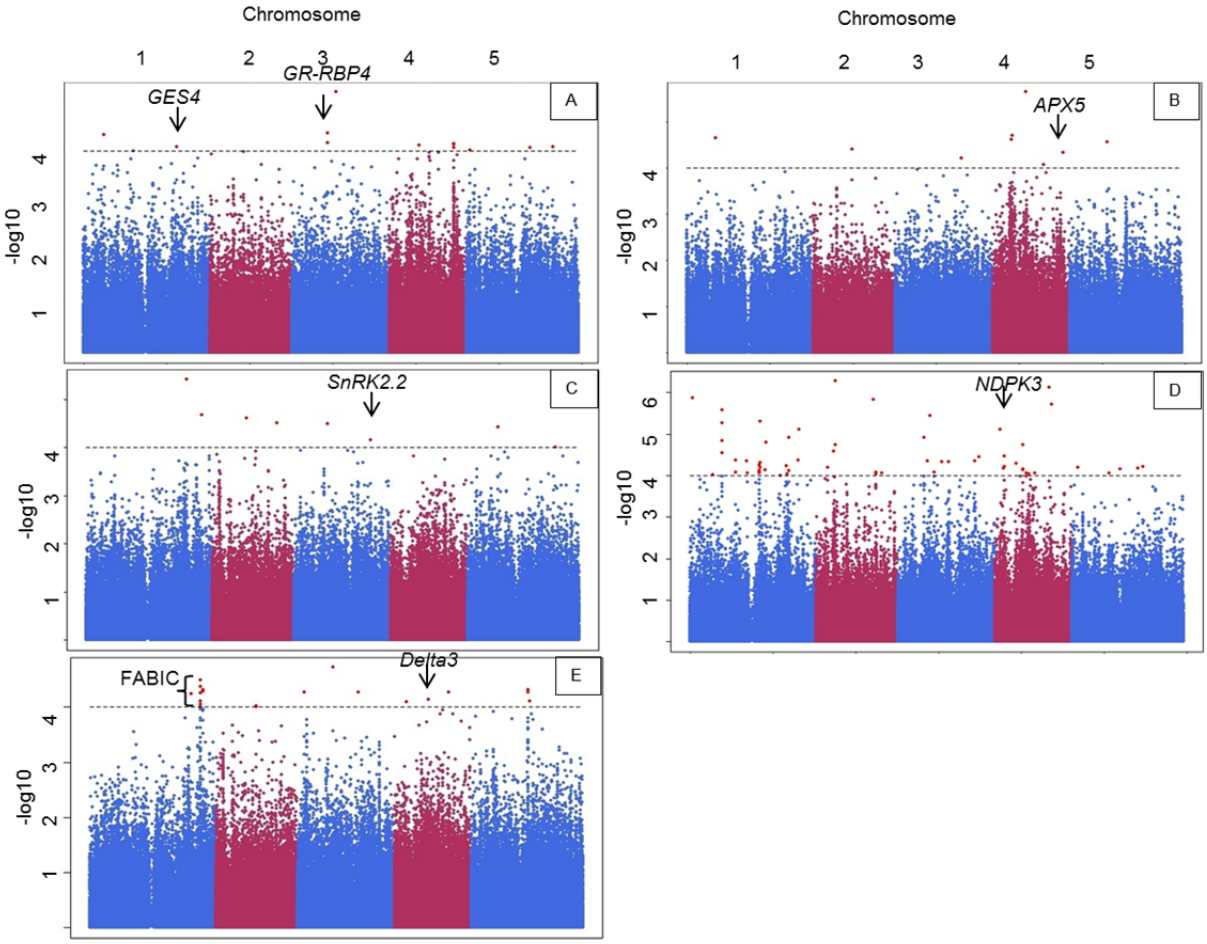


Figure 3: Manhattan plots displaying GWAS results for plasticity using univariate approach showing the -log10(P) values for SNPs associated with rosette fresh weight (A), rosette dry weight (B), root length (C), root dry weight (D), and the ratio between root length and rosette dry weight (E).

Extreme accessions and comparing amino acid sequences of candidate genes

Ten drought tolerant and ten drought sensitive accessions were selected based on their RosDW ratio between drought and control environments (Table 5). The selected extremes did not exhibit any marked geographic distribution. In general, most of the drought sensitive accessions showed a higher total root length in the drought environment compared to the control environment (RL ratio often >1), while no significant difference was observed between RDW of drought tolerant and drought sensitive accessions.

Table 4: List of candidate genes mapped for shoot and root plastic response selected based on GWA analysis of the Arabidopsis HapMap population. RL/RosDW refers to the ratio between root length and rosette dry weight. –log10(p) indicates the QTL significance level. β indicates the effect of a SNP with positive values when the positive effect on the trait value is coming from the Col allele. Chromosome numbers (Chr.) and SNP positions in kb are given. Both SNP position and description is based on TAIR v.10 (www.arabidopsis.org). LD lists genes found to be in linkage disequilibrium within 10 kb on both sides of the significant SNP.

Trait (unit)	Chr.	pos (kb)	Gene	-log10 (p)	β	Function	LD
Rosette fresh weight (g)	1	22523	AT1G61120	4.1	-0.1004	GERANYLLINALOOL SYNTHASE (GES). Involved in abscisic acid mediated signaling pathway and response to water deprivation	-
	3	8608	AT3G23830	4.3	0.1969	GLYCINE-RICH RNA-BINDING PROTEIN 4 4 GR-RBP4. Involved in response to cold, osmotic stress, salt stress, and water deprivation	AT3G23840
	5	819	AT5G03360	4.0	-0.1073	DC1 domain-containing protein. Involved in response to oxidative stress	-
Rosette dry weight (g)	2	9315	AT2G21850	4.4	0.0112	Cysteine/Histidine-rich C1 domain family protein. Intracellular signal transduction, response to oxidative stress	AT2G21830, AT2G21840
	4	17029	AT4G35970	4.3	0.0095	Ascorbate Peroxidase 5 (APX5) involved in the response to oxidative stress	-
Root length (cm)	1	27960	AT1G74380	4.7	-106.7	XYLOGLUCAN XYLOSYL TRANSFERASE 5 (XXT5). Involved in glucose catabolic process, root hair elongation, xyloglucan metabolic process	-
	2	8347	AT2G19230	4.6	76.9	Leucine-rich repeat transmembrane protein kinase protein involved in amino acid phosphorylation,	-
	3	18742	AT3G50500	4.2	60.8	SNF1-RELATED PROTEIN KINASE 2-2 (SnRK2-2). Involved in response to osmotic, salt stresses and water deprivation	-
Root dry weight (g)	1	11092	AT1G31080	4.4	0.0038	F-box family protein with unknown biological function. Expressed in root	AT1G31090
	2 (2)	14536	AT2G34490	4.1	0.0028	Cytochrome P450 (CYP710A2) encodes a protein involved in response to oxidative stress	AT2G34400 to AT2G34500
	4	6732	AT4G11010	4.1	0.0023	NUCLEOSIDE DIPHOSPHATE KINASE 3 (NDPK3). Involved in response to oxidative stress	AT4G11000, AT4G11020
	4	9626	AT4G17140	4.0	0.0026	Involved in root hair cell differentiation	-
	5	17271	AT5G43060	4.2	0.0038	RESPONSIVE TO DEHYDRATION 21B (RD21B). Involved in response to salt stress	-
RL/RosDW (cm/g)	1 (7)	26785	AT1G71010	4.5	-1159.1	FORMS APLOID AND BINUCLEATE CELLS 1C (FAB1C). Involved in cellular protein metabolic process, phosphatidylinositol metabolic process	AT1G71015, AT1G71020
	4	8306	AT4G14430	4.1	-1815.3	ARABIDOPSIS THALIANA DELTA(3). Involved in abscisic acid mediated signaling pathway, response to water deprivation and root hair elongation	-
	5	14934	AT5G37600	4.1	-1155.6	GLUTAMINE SYNTHASE 1.1 (ATGLN1.1). Involved in nitrate assimilation, response to glucose and sucrose stimulus	-

Table 5: The ten most drought sensitive and drought tolerant accessions identified in the HapMap population, ranked according to the lowest, respectively highest rosette dry weight in drought relative to control ratio (RosDW, in bold). The values of the rosette fresh weight (RosFW), root length (RL) and root dry weight (RDW) ratios are also presented. RL-Rank represents the rank position of these accessions when considered for RL ratio in ascending order (see Table 6).

	Accession	Name	Country	RosFW	**RosDW**	RL	RDW	RL-Rank
Drought sensitive	CS76197	Nd-1	Germany	0.16	**0.33**	0.81	0.74	177
	CS76127	Est-1	Estonia	0.22	**0.40**	1.11	1.06	247
	CS76301	Wei-0	Switzerland	0.19	**0.41**	0.63	1.31	132
	CS22689	RRS-10	USA	0.21	**0.45**	0.34	0.54	65
	CS28336	Ha-0	Germany	0.17	**0.46**	0.38	0.60	71
	CS76103	Bu-0	Germany	0.27	**0.47**	1.26	1.03	262
	CS76109	Can-0	Spain	0.26	**0.48**	1.21	1.27	252
	CS76304	Wt-5	Germany	0.19	**0.49**	1.13	1.73	244
	CS76156	Kulturen-1	Sweden	0.48	**0.49**	1.17	0.90	249
	CS76140	Hi-0	Netherlands	0.57	**0.51**	1.47	1.30	287
Drought tolerant	CS76199	NFA-8	UK	0.55	**0.98**	0.91	0.63	205
	CS76167	Lillo-1	Sweden	0.79	**0.98**	0.93	1.42	67
	CS28350	Hn-0	Germany	0.91	**0.98**	0.69	1.12	146
	CS28241	Es-0	Finland	0.80	**1.03**	0.42	0.91	49
	CS28091	Boot-1	UK	0.91	**1.06**	0.66	0.72	139
	CS28193	Com-1	France	0.86	**1.08**	0.33	0.78	62
	CS76214	Pro-0	Spain	0.52	**1.12**	1.07	1.32	230
	CS28013	Alst-1	UK	0.88	**1.13**	0.39	0.70	72
	CS28014	Amel-1	Netherlands	0.81	**1.16**	0.49	1.01	60
	CS28018	Ang-0	Belgium	1.16	**1.66**	0.52	1.06	59

Another twenty accessions were selected based on their RL ratio between drought and control environments (Table 6). 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.

Table 6: Accessions with extreme root length (RL) identified in the HapMap population, ranked according to short and respectively long root length in drought relative to control ratio (RL in bold). The values of rosette fresh weight (RosFW), dry weight (RosDW) and root dry weight (RDW) ratios are also presented. RosDW-Rank represents the rank position of these accessions when considered for RosDW-Rank ratio in ascending order (see Table 5).

	Accession	Name	Country	RosFW	RosDW	**RL**	RDW	RosDW-Rank
Short root	CS28787	Uk-1	UK	0.42	0.80	**0.23**	1.33	250
	CS76198	NFA-10	UK	0.44	0.76	**0.23**	0.74	224
	CS28241	Es-0	Finland	0.80	1.03	**0.24**	0.91	322
	CS28692	Rou-0	France	0.63	0.76	**0.29**	0.95	225
	CS76164	Ler-1	Poland	0.53	0.60	**0.29**	0.59	110
	CS28018	Ang-0	Belgium	1.16	1.06	**0.31**	1.06	350
	CS22689	RRS-10	USA	0.21	0.45	**0.34**	0.54	35
	CS76092	App1-16	Sweden	0.36	0.57	**0.35**	0.59	94
	CS76106	C24	Portugal	0.46	0.58	**0.48**	0.81	96
	CS76113	Col-0	USA	0.47	0.61	**0.61**	0.36	116
Long root	CS76116	Cvi-0	Cape Verde Islands	0.32	0.72	**1.17**	1.28	203
	CS76192	Mt-0	Lybia	0.27	0.69	**1.36**	1.45	184
	CS76142	Hov4-1	Sweden	0.62	0.73	**1.39**	1.13	209
	CS76140	Hi-0	Netherlands	0.57	0.51	**1.47**	1.30	63
	CS76251	Tottarp-2	Sweden	0.32	0.66	**1.51**	1.71	157
	CS76215	Pu2-23	Czech Republic	0.59	0.91	**1.65**	1.76	296
	CS76220	Rmx-A180	USA	0.42	0.80	**1.67**	1.71	252
	CS76219	Rev-2	Sweden	0.55	0.73	**1.69**	1.61	210
	CS76141	Hod	Slovakia	0.86	1.07	**1.72**	1.29	330
	CS76227	Sha	Tajikistan	0.48	0.83	**1.75**	1.37	266

The Arabidopsis 1001 genomes browser http://signal.salk.edu/atg1001/3.0/gebrowser.php (Weigel and Mott, 2009) was used to compare the amino acid of all genes reported in table 3 and 4 and all genes with unknown function between the extreme groups of the accessions listed in tables 5 and 6 and to try to identify genes that show common sequence differences, distinguishing both extreme groups. Only 5 genes showed such amino acid sequence differences (data not shown). The most striking differences were observed when comparing the predicted protein sequences of two genes, the *RIBOSOMAL PROTEIN S3 (RPS3*, AT5G35530) gene mapped for RL/RosDW (Table 3) and the *YELLOW LEAF SENESCENCE 7*(*YLS7,* AT5G51640) gene mapped for RL. The *RPS3* gene was compared between the twenty extreme accessions selected based on RosDW ratio (Table 5) and presented in Figure 4. Five amino acids were consistently different between both groups (Figure 4-a). Based on all possible combinations of those five amino acids, the re-sequenced accessions, that overlapped with the current population, were classified into five haplotypes (Figure 4-b,c):. haplotype "NFA-8", with 31 accessions; haplotype Amel-1, with 14 accessions; haplotype Lilloe-1, including 14 accessions, with the highest RL/RosDW; haplotype Wei-0, with 62

accessions, including Col; and finally a remaining set of haplotypes, comprising 39 accessions, including Sha. The results showed that accessions with amino acids similar to Sha had a higher RL/RosDW than accessions similar to Col. As a control, 10 genes upstream and 10 genes downstream of the *RPS3* gene were also compared regarding their coding regions, but no common differences were observed between both groups (data not shown), indicating the *RPS3* gene to be uniquely highly variable in a relatively large genomic region, supporting its candidacy to be the causal gene identified by the associated SNP.

One more SNP was mapped in *YLS7* gene , with significance level –log10(p) = 3.7. This SNP was in the interval of root length QTL, *RL4,* mapped in chapter 2. For both, the SNP and the QTL, the Col allele increased RL in the control environment, and the non-Col allele, Sha allele, increased RL in the drought environment. Based on the root length ratio, the two contrasting groups, 10 accessions each, one with large root length ratio, included Sha, and one with small root length ratio, included Col (Table 6), were examined for their amino acid sequence of *YLS7* (Figure 5-a). This shows clear common differences for 12 amino acids between both groups. Haplotypes of the *YLS7* (Figure 5-b) using the re-sequenced accessions available in the 1001 genome browser were analysed based on the amino acid of *YLS7.* The re-sequenced accessions were classified into four haplotype groups; 43 accessions with amino acid sequences similar to Col, 6 accessions with long roots and amino acid sequences similar to Sha, 32 accessions with long roots and amino acid sequences slightly different from Sha, and the remaining 85 accessions, with several combinations of amino acid sequences. The results (Figure 5-c) showed that haplotypes similar to Sha have the longest roots in the drought environment, whereas, in the control environment, haplotypes similar to Col have longer roots than haplotypes similar to Sha. Based on the amino acids of twenty genes, ten genes up and ten genes down *YLS7,* no clear differences were found when comparing both groups, except for gene AT5G51630, annotated as a disease resistance protein (Fig. 6). The comparison revealed that accessions similar to Sha exhibited no significant difference between RL in control and drought environments. In addition, no significant difference was observed between RL of accessions similar to Sha and those similar to Col in the drought environment (Figure 6).

Figure 4: (A) Comparison of amino acids (AA) for AT5G35530 (*RPS3*), a gene mapped for RL/RosDW, between accessions with highest and lowest RosDW in table 5. Based on AA sequences, 160 re-sequenced accessions of the Hapmap population available in 1001 genome browser were classified in panel (B) to five haplotypes; 31 accessions with AA sequences similar to NFA-8l, 14 accessions with AA sequence similar to Amel-1, 12 accessions with AA sequence similar to Lilloe-1, 62 accessions with AA sequences similar to Wei-0 (includes Col), and 39 accessions with all other possible AA sequences (includes Sha).

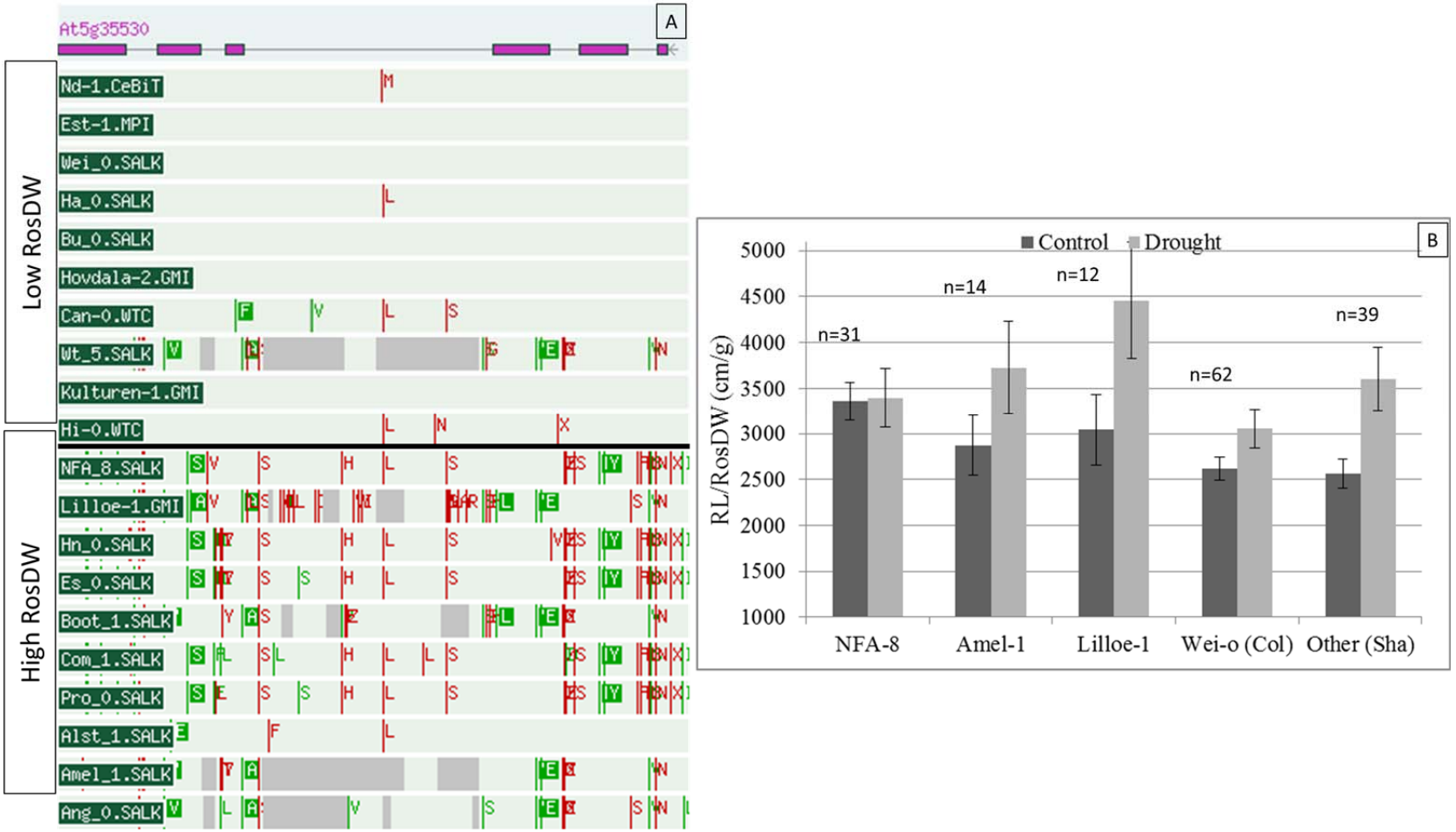


Figure 5: (A) Comparison of amino acids (AA) for AT5G51640 (*YLS7*), a gene mapped for RL, and co-located with *RL4,* between accessions with long and short root in table 6. Based on AA sequences, 160 re-sequenced accessions of the Hapmap population available in 1001 genome browser were classified in panel (B) to four haplotypes; 43 accessions with AA sequences similar to Col, 6 accessions with AA sequence similar to Sha, 32 accessions with AA sequence similar to Cvi, and the 85 accessions with all other possible AA sequences.

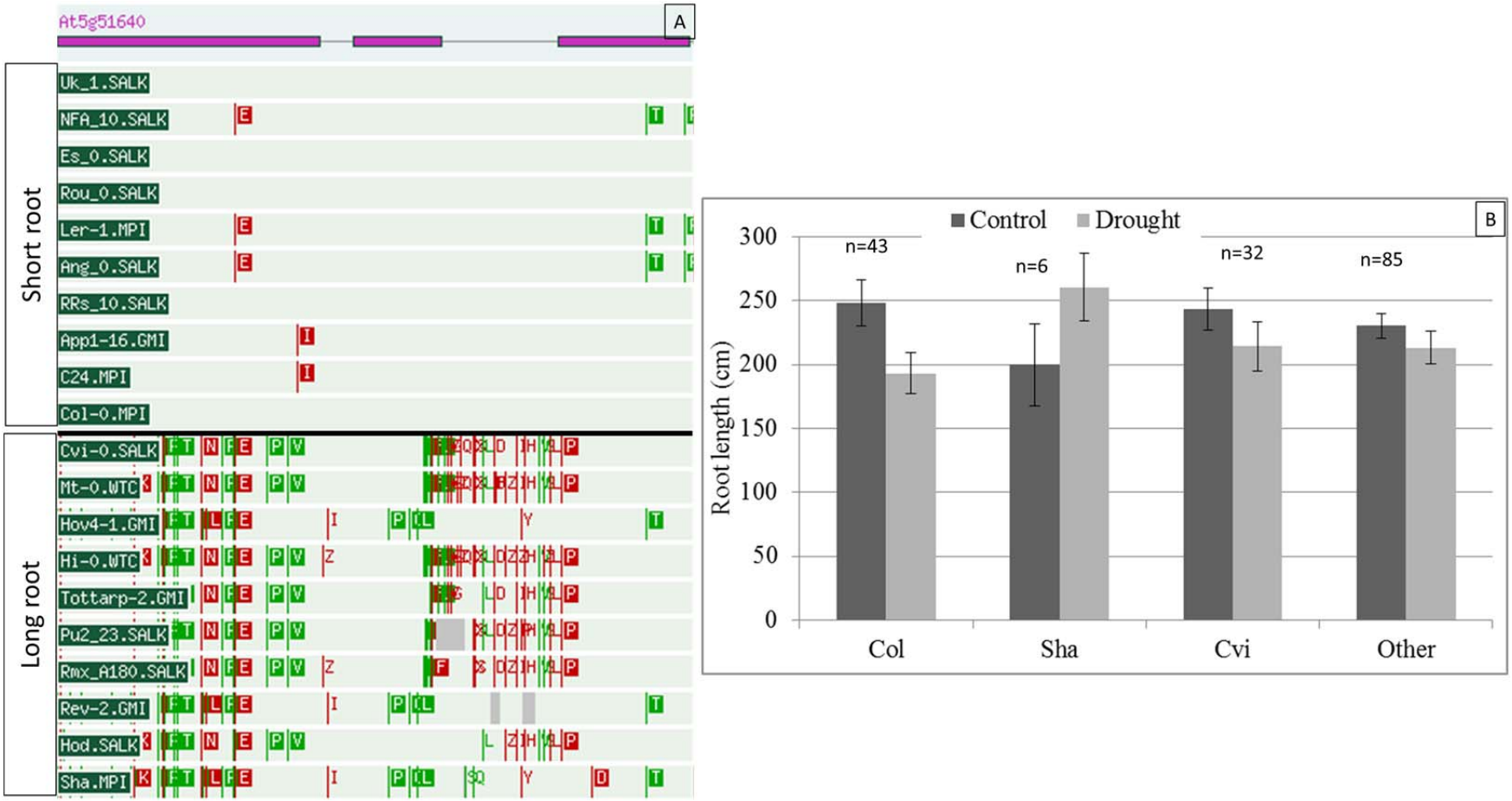
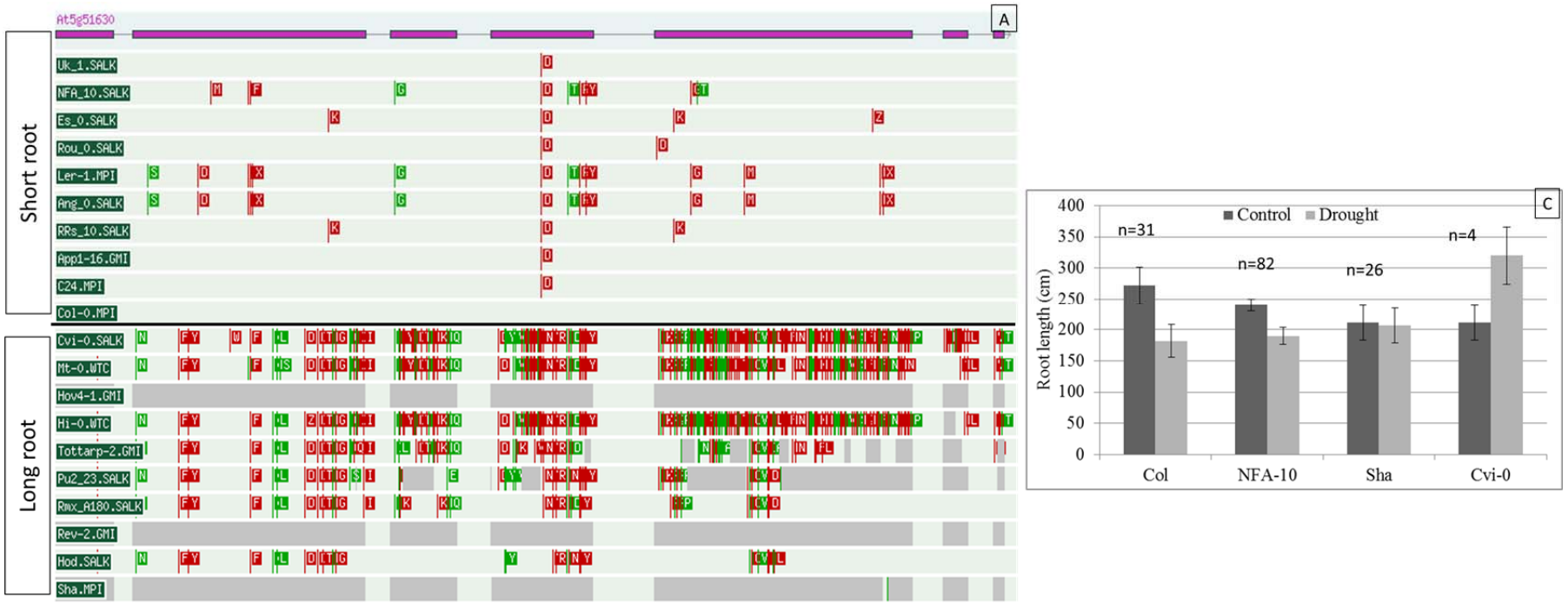


Figure 6: (A) Comparison of amino acids (AA) for AT5G51630, a disease resistance protein mapped for RL, between accessions with long and short root in table 6. Based on AA sequences, 160 re-sequenced accessions of the Hapmap population available in 1001 genome browser were classified in panel (B) to four haplotypes; 31 accessions with AA sequences similar to Col, 82 accessions with short roots but with AA sequence different from Col and similar to NFA-10, 26 accessions with AA sequence similar to Sha, and 4 accessions with long root but with AA sequence different from Sha and similar to Cvi-0.



Discussion

To assess the extent of natural variation in rosette and root morphological responses to drought, 350 Arabidopsis accessions were grown on sand under greenhouse conditions. Such experimental design is a reasonable compromise to avoid the difficulty of phenotyping roots in natural field environments and the unnatural conditions present in hydroponics, aeroponics or agar plates (Tuberosa, 2012). The high heritability of at least 58% for the measured traits indicates homogenous growth of the population in both environments. The highly significant correlation observed between traits and between environments suggests commonalities in their genetic regulation, which is confirmed by finding co-location of significant SNPs for correlated traits.

Earlier studies (Atwell et al., 2010; Filiault and Maloof, 2012; Yano et al., 2013) used different subsets of Arabidopsis collections that overlapped with the 350 accessions used here. These 350 accessions were selected from a much larger collection of 5707 Arabidopsis accessions, based on a maximized genetic diversity and the elimination of population structure (Li et al., 2010; Platt et al., 2010). However, because only 149 SNPs were used to test for population structure some residual population structure will be present in the HapMap set which accounts for false positives. Some statistical approaches were suggested to overcome this problem (Yu et al., 2006; Zhao et al., 2007; Atwell et al., 2010; Brachi et al., 2010), but they introduce the problem of false negatives, meaning that causative genetic markers may not be detected, when they are strongly associated with population structure and corrected for statistically (Brachi et al., 2010).

The commonly used approach for such GWA studies was univariate analysis for each trait and treatment. Although this will identify genes with main effects, it does not account for GxE. Especially for traits that appear to respond to environmental changes, considering GxE will help to discover novel genes that act synergistically with environment (Thomas, 2010), and also to identify superior and stable genotypes across different environments (Filiault and Maloof, 2012). Therefore, the proposed MTMM approach for modelling bi-variant association mapping and for effective mapping of QxE (Korte et al., 2012) was applied here. No less than 58 SNPs were found to be associated with QxE effects, of which 23 SNPs overlapped with SNPs associated with plasticity which validate the results. In addition, 38 SNPs were significant when mapped using bivariate MTMM and slightly below the threshold, between 3.5 and 3.9, when mapped using univariate analysis and vice versa. This is in agreement with earlier reports, e.g. (Filiault and Maloof, 2012; Yano et al., 2013), where the

significance level was depending on the statistical approach used for mapping. In addition, we calculated the correlation between the -log10 values of all 214 k SNPs that resulted from univariate and MTMM approaches and found a positive correlation of between 0.40 and 0.45 for all traits, which supports the expected overlap between them. Apparently, the discrepancy in significance level is due to the applied statistical approach. It is difficult to choose a proper threshold for the association analysis, as it is rather arbitrary. Previously, approaches in which a –log10(p) threshold of 4 was used, gave good enrichment for *a priori* candidates (Atwell et al., 2010; Li et al., 2010), which is why we used the same threshold. However, it has to be noted that so far very few of the candidate genes coming from GWAS in Arabidopsis have been confirmed. One example was the identification of heavy metal atpase3 (*HMA3*) as the primary determinant of natural variation of cadmium in leaves (Chao et al., 2012).

GWAS has reduced power in detecting rare alleles even if they have large phenotypic effects. Therefore, combining traditional linkage mapping with GWAS is an alternative for reducing the rate of false positives and for detecting false negatives (Nordborg and Weigel, 2008; Atwell et al., 2010; Bergelson and Roux, 2010; Brachi et al., 2010; Sterken et al., 2012; Weigel, 2012). Therefore, we searched the confidence intervals for QTL presented in chapter 4 for SNPs associated with the same traits, that were at or close to the threshold. Only two of those, *FLC* and *YLS7*, were found, at –log10(p) values of 3.6 and 3.7 respectively. The association involving *FLC* was with RosDW. *FLC* was suggested to be co-locating with RosDW QTL, *DW2,* identified in the Col x Sha RIL population (chapter 2). The pleiotropic effect of *FLC* on increased leaf size, biomass yield and reduced height at flowering time was reported in tobacco plants when transformed with *FLC* (Salehi et al., 2005). Other studies suggested that when *FLC* interacted with *FRIGIDA,* another flowering time gene, it showed pleiotropic effects on water use efficiency and nitrogen content (Loudet et al., 2003; McKay et al., 2003; Tonsor et al., 2005; Lovell et al., 2013) and number of branches, nodes and fruits (Scarcelli et al., 2007). We find this sufficient reason to consider the association of *FLC* with DW to indicate a true candidate explaining the phenotype.

Similarly, we like to consider another association at a –log10(p) value of 3.7, which is the association of a SNP in *YLS7* with RL. This gene was originally found in a screen for senescence induced genes (Yoshida et al., 2001). The protein sequence of *YLS7* was compared with that of the paralogous gene *ESK1*, using the Simple Molecular Architecture Research Tool (SMART) browser, which is involved in the tolerance to freezing (Xin and Browse, 1998; Xin et al., 2007), salt and drought (Lugan et al., 2009) (Figure 7). It shows that the protein domains of the predicted amino acid sequence derived from the following,

homologous genes: *YLS7*; *ESKIMO1* (*ESK1*), also known as *TBL29* (Xin and Browse, 1998; Xin et al., 2007). The comparison revealed that proteins encoded by *YLS7* and *ESK1* shared three main domains, a transmembrane domain, a low complexity region, and a Domain of Unknown Function (DUF).

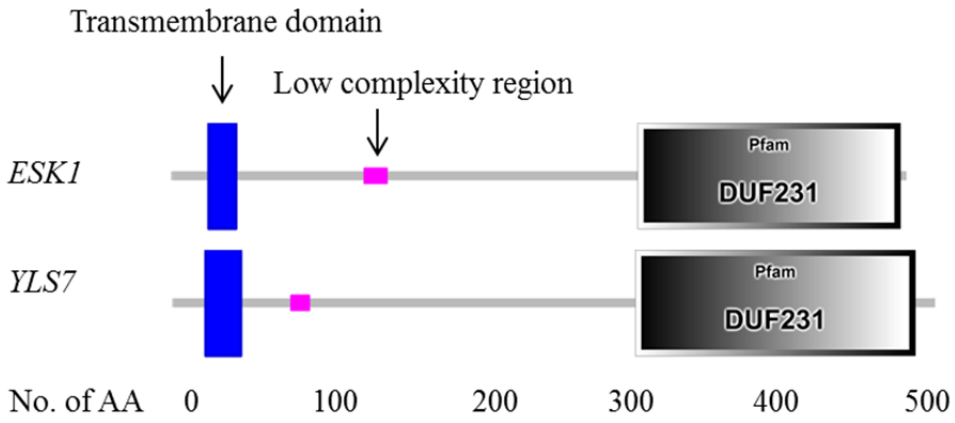


Figure 7: Protein domain based comparison between two TBL genes; *ESKIMO1* gene (*ESK1*) and *Yellow LEAF SPECIFIC GENE 7* (*YLS7*) The comparison is based on three main domains; the transmembrane domain, the low complexity region, and the DUF231 domain. No. of AA refers to number of amino acids per protein.

An earlier study reported that the terminal LCR, is enriched with stress response-related terms (Coletta et al., 2010). The similarity in protein sequence and structure suggests that YLS7 may have a similar funciton as ESK1 and may also be involved in abiotic stress tolerance. Although *YLS7* was not reported to be associated with stress response in Arabidopsis, an earlier study showed that *YLS7* from Ponkan mandarin (*Citrus reticulata*) was down regulated when fruits were stored in the cold (Zhu et al., 2011). *YLS7* showed significant QxE with Col increasing RL in the control environment and the non-Col allele increasing RL in the drought environment. *YLS7,* co-located with the RL QTL, *RL4,* mapped in the Sha x Col RIL population (chapter 5), with the marker with the highest LOD score mapping within 700 Kb of *YLS7*. For both, the *RL4* QTL and the co-locating SNP in *YLS7,* the favourable allele was the Col allele in the control environment and the non-Col allele, which was from Sha in the RIL population, was the favourable allele in the drought environment. *YLS7* also co-located with a QTL for seedling germination under salt stress, also mapped in the Sha x Col population (Galpaz and Reymond, 2010). In addition, the haplotype analysis was in agreement with the results in chapter 5 where Col exhibited a longer total root length in

control environments while Sha exhibited a longer total root length in drought environment. These results together support that *YLS7* is a strong candidate to be the gene underlying the *RL4* QTL, however, final proof will be found by replacing the Col allele with the Sha allele and conferring the Sha phenotype to the transgenic plants.

Many SNPs were significantly associated with the measured traits. We are aware that significant SNPs associated with genes with unknown function are of great importance to understand their functions, however, because no functional analysis confirmation was carried out, only genes that are known to be involved in drought or stress responses were reported. For example, one significant SNP was mapped *SnRK2.2.2* (Fujii et al., 2007) and was associated with RDW and with its plasticity and with RL plasticity. This gene was first reported as a regulator of some of the response to sulphur starvation (Davies et al., 1999). In addition, the Arabidopsis double mutant *snrk2.2/snrk2.3* showed strong ABA insensitive phenotypes in seed germination and root growth inhibition (Fujii et al., 2007). Other studies showed that the Arabidopsis triple mutant snrk2.2/snrk2.3/snrk2.6 was extremely insensitive to ABA and exhibits greatly reduced tolerance to drought (Fujii and Zhu, 2009; Kulik et al., 2011).

Another association was found between RosDW and *TOR*, a growth regulator gene which senses nutrient availability (Dobrenel et al., 2011; Liao et al., 2011; Ren et al., 2012; Caldana et al., 2013). It was shown earlier that overexpression of *TOR* increases shoot biomass and resistant to stresses (Dobrenel et al., 2011). In addition, the inhibition of *TOR* resulted in slower root growth in Arabidopsis, leading to poor nutrient uptake, and slower leaf growth, leading to poor light energy utilization(Ren et al., 2012). In the same way, we think that drought sensitive accessions that showed small RosDW and RL do not express this gene.

The *RPS3* gene (Barakat et al., 2001), which is induced by salt-stress in Arabidopsis roots (McLoughlin et al., 2013), was associated with RL/RosDW. Distinguishing amino acid differences were observed for this gene when comparing accessions with extreme phenotypes for RosDW, grouped according to contrasting phenotype in drought tolerant and drought sensitive accessions. To exclude that this is just one gene of a genetically very variable region of the genome, we also examined 10 genes upstream and downstream of this gene. However, all of these neighbouring genes are much less variable than the candidate gene, strengthening its role as a candidate underlying the observed phenotypic variation. A haplotype analysis showed that in drought environment, Sha and accessions with similar haplotypes, had a higher RL/RosDW than accessions with the Col haplotype, which is very much in line with the QTL

results described in chapter 5. These results, together with the results obtained from *YLS7* increases our confidence to consider both a true association.

Drought stress induces stomatal closure, and leads to excess production of reactive oxygen species (ROSs) and increased oxidative stress (Mittler, 2002; Cruz de Carvalho, 2008; Bhatt et al., 2011). To suppress the effect of ROS under drought conditions, plants produce anti-oxidative enzymes (Cruz de Carvalho, 2008; Mirzaei et al., 2012). In the current study, two SNPs in *SKS7* and *CYP83A1* genes that are involved in the redox process, were associated with RosFW and RDW, respectively. The effect of those SNPs was highly affected by drought treatment which is in agreement with what was reported earlier in drought stressed rice (Mirzaei et al., 2012), where anti-oxidative enzymes were up-regulated in drought grown plants when compared to well-watered plants.

In this chapter we showed promising associations between SNPs in relevant genes and drought tolerance related rosette and root traits. However, additional confirmation using knock outs, reciprocal transformations, and studying differential gene expression and co-expression networks will be required 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, which knowledge can potentially be used to improve drought stress tolerance in crops.

Acknowledgements

We would like to thank Corrie Hanhart, Charles N. Moreira, Roxanne van Rooijen, Christina May and Nihal Erol-Öztolan for their help with sowing and harvesting the experiments.

References

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

Aranzana MJ, Kim S, Zhao K, *et al.* 2005. Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. *PLoS Genetics* **1**, e60.

Assmann SM. 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.

Attaran E, Rostas M, Zeier J. 2008. Pseudomonas syringae elicits emission of the terpenoid (E,E)-4,8,12-trimethyl-1,3,7,11-tridecatetraene in Arabidopsis leaves via jasmonate signaling and expression of the terpene synthase TPS4. *Molecular Plant Microbe Interaction* **21**, 1482-1497.

Atwell S, Huang YS, Vilhjalmsson BJ, *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 IF, Guyot R, Blanc G, Cooke R, Delseny M, Bailey-Serres J. 2001. The organization of cytoplasmic ribosomal protein genes in the Arabidopsis genome. *Plant Physiol* **127**, 398-415.

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

Bhatt D, Negi M, Sharma P, Saxena S, Dobriyal A, Arora S. 2011. Responses to drought induced oxidative stress in five finger millet varieties differing in their geographical distribution. *Physiology and Molecular Biology of Plants* **17**, 347-353.
Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, 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 AR, Willmitzer L, 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.
Chao D-Y, Silva A, Baxter I, Huang YS, Nordborg M, Danku J, Lahner B, Yakubova E, Salt DE. 2012. Genome-wide association studies identify heavy metal ATPase3 as the primary determinant of natural variation in leaf cadmium in *Arabidopsis thaliana*. *PLoS Genetics* **8**, e1002923.
Coletta A, Pinney JW, Solis DYW, Marsh J, Pettifer SR, Attwood TK. 2010. Low-complexity regions within protein sequences have position-dependent roles. *Bmc Systems Biology* **4:43**.
Cruz de Carvalho MH. 2008. Drought stress and reactive oxygen species: production, scavenging and signaling. *Plant Signaling and Behaviour* **3**, 156-165.
Davies JP, Yildiz FH, Grossman AR. 1999. Sac3, an Snf1-like serine threonine kinase that positively and negatively regulates the responses of chlamydomonas to sulfur limitation. *The Plant Cell* **11**, 1179-1190.
Dobrenel T, Marchive C, Sormani R, Moreau M, Mozzo M, Montane MH, Menand B, Robaglia C, Meyer C. 2011. Regulation of plant growth and metabolism by the TOR kinase. *Biochem Soc Trans* **39**, 477-481.
Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH. 2010. Missing heritability and strategies for finding the underlying causes of complex disease. *Nature Reviews Genetics* **11**, 446-450.
Filiault DL, Maloof JN. 2012. A genome-wide association study identifies variants underlying the *Arabidopsis thaliana* shade avoidance response. *PLoS Genetics* **8**, e1002589.
Flint-Garcia SA, Thornsberry JM, Buckler ESt. 2003. Structure of linkage disequilibrium in plants. *Annu Rev Plant Biol* **54**, 357-374.
Fujii H, Verslues PE, Zhu JK. 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 PE, Zhu JK. 2011. Arabidopsis decuple mutant reveals the importance of SnRK2 kinases in osmotic stress responses in vivo. *Proce Nat Acad Sci USA* **108**, 1717-1722.
Fujii H, Zhu JK. 2009. Arabidopsis mutant deficient in 3 abscisic acid-activated protein kinases reveals critical roles in growth, reproduction, and stress. *Proce Nat Acad Sci USA* **106**, 8380-8385.
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 Reviews Genetics* **13**, 135-145.
Goepfert S, Vidoudez C, Tellgren-Roth C, Delessert S, Hiltunen JK, Poirier Y. 2008. Peroxisomal Delta(3),Delta(2)-enoyl CoA isomerases and evolution of cytosolic paralogues in embryophytes. *The Plant Journal* **56**, 728-742.
Gupta P, Rustgi S, Kulwal P. 2005. Linkage disequilibrium and association studies in higher plants: present status and future prospects. *Plant Molecular Biology* **57**, 461-485.
Hammargren J, Rosenquist S, Jansson C, Knorpp C. 2008. A novel connection between nucleotide and carbohydrate metabolism in mitochondria: sugar regulation of the Arabidopsis nucleoside diphosphate kinase 3a gene. *Plant Cell Reports* **27**, 529-534.
Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG, Toomajian C, Roux F, Bergelson J. 2011. Adaptation to climate across the *Arabidopsis thaliana* genome. *Science* **334**, 83-86.
Huang J, Kim CM, Xuan YH, Park SJ, Piao HL, Je BI, Liu J, Kim TH, Kim BK, Han CD. 2013. OsSNDP1, a Sec14-nodulin domain-containing protein, plays a critical role in root hair elongation in rice. *Plant Molecular Biology* **82**, 39-50.
Juenger TE. 2013. Natural variation and genetic constraints on drought tolerance. *Curr Opin Plant Biol* **16**, 274-281.
Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, Sabatti C, Eskin E. 2010. Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics* **42**, 348-354.
Korte A, Vilhjalmsson BJ, 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.
Kulik A, Wawer I, Krzywinska E, Bucholc M, Dobrowolska G. 2011. SnRK2 protein kinases--key regulators of plant response to abiotic stresses. *OMICS* **15**, 859-872.
Kwak KJ, Park SJ, Han JH, Kim MK, Oh SH, Han YS, Kang H. 2011. Structural determinants crucial to the RNA chaperone activity of glycine-rich RNA-binding proteins 4 and 7 in Arabidopsis thaliana during the cold adaptation process. *J Exp Bot* **62**, 4003-4011.

Li Y, Huang Y, Bergelson J, Nordborg M, Borevitz JO. 2010. Association mapping of local climate-sensitive quantitative trait loci in *Arabidopsis thaliana*. *Proce Nat Acad Sci USA* **107**, 21199-21204.
Liao YY, Buckhout TJ, Schmidt W. 2011. Phosphate deficiency-induced cell wall remodeling: linking gene networks with polysaccharide meshworks. *psb* **6**, 700-702.
Loudet O, Chaillou S, Krapp A, Daniel-Vedele F. 2003. Quantitative trait loci analysis of water and anion contents in interaction with nitrogen availability in *Arabidopsis thaliana*. *Genetics* **163**, 711-722.
Lovell JT, Juenger TE, Michaels SD, Lasky JR, Platt A, Richards JH, Yu X, Easlon HM, Sen S, McKay JK. 2013. Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. *Phil. Trans. R. Soc. B* **280**, 20131043.
Lugan R, Niogret MF, Kervazo L, Larher FR, Kopka J, Bouchereau A. 2009. Metabolome and water status phenotyping of Arabidopsis under abiotic stress cues reveals new insight into ESK1 function. *Plant Cell and Environment* **32**, 95-108.
McKay JK, Richards JH, 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.
McLoughlin F, Arisz SA, Dekker HL, Kramer G, de Koster CG, Haring MA, Munnik T, 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.
Mirzaei M, Soltani N, Sarhadi E, Pascovici D, Keighley T, Salekdeh GH, Haynes PA, Atwell BJ. 2012. Shotgun proteomic analysis of long-distance drought signaling in rice roots. *J Proteome Res* **11**, 348-358.
Mittler R. 2002. Oxidative stress, antioxidants and stress tolerance. *Trends in Plant Science* **7**, 405-410.
Morikawa T, Mizutani M, Ohta D. 2006. Cytochrome P450 subfamily CYP710A genes encode sterol C-22 desaturase in plants. *Biochem Soc Trans* **34**, 1202-1205.
Nicotra AB, Atkin OK, Bonser SP, *et al.* 2010. Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**, 684-692.
Nordborg M, Weigel D. 2008. Next-generation genetics in plants. *Nature* **456**, 720-723.
Panchuk, II, Volkov RA, Schoffl F. 2002. Heat stress- and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in Arabidopsis. *Plant Physiol* **129**, 838-853.
Platt A, Horton M, Huang YS, *et al.* 2010. The scale of population structure in *Arabidopsis thaliana*. *PLoS Genetics* **6**, e1000843.
Ren M, Venglat P, Qiu S, *et al.* 2012. Target of rapamycin signaling regulates metabolism, growth, and life span in Arabidopsis. *The Plant Cell* **24**, 4850-4874.
Salehi H, Ransom CB, Oraby HF, Seddighi Z, Sticklen MB. 2005. Delay in flowering and increase in biomass of transgenic tobacco expressing the Arabidopsis floral repressor gene FLOWERING LOCUS C. *J Plant Physiol* **162**, 711-717.
Scarcelli N, Cheverud JM, Schaal BA, Kover PX. 2007. Antagonistic pleiotropic effects reduce the potential adaptive value of the FRIGIDA locus. *Proce Nat Acad Sci USA* **104**, 16986-16991.
Shindo T, Misas-Villamil JC, Horger AC, Song J, van der Hoorn RAL. 2012. A role in immunity for Arabidopsis cysteine protease RD21, the ortholog of the tomato immune protease C14. *PLoS ONE* **7**, e29317.
Smith S, De Smet I. 2012. Root system architecture: insights from Arabidopsis and cereal crops. *Phil. Trans. R. Soc. B* **367**, 1441-1452.
Sterken R, Kiekens R, Boruc J, *et al.* 2012. Combined linkage and association mapping reveals CYCD5;1 as a quantitative trait gene for endoreduplication in Arabidopsis. *Proc Natl Acad Sci U S A* **109**, 4678-4683.
Stich B, Melchinger AE. 2010. An introduction to association mapping in plants. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources* **5**, 1-9.
Tardieu F. 2013. Plant response to environmental conditions: assessing potential production, water demand and negative effects of water deficit. *Frontiers in Physiology* **4:17. 10.3389/fphys.2013.00017**.
Tétard-Jones C, Kertesz MA, Preziosi RF. 2011. Quantitative trait loci mapping of phenotypic plasticity and genotype–environment interactions in plant and insect performance. *Phil. Trans. R. Soc. B* **366**, 1368-1379.
Thomas D. 2010. Gene--environment-wide association studies: emerging approaches. *Nature Review Genetics* **11**, 259-272.
Tonsor SJ, Alonso-Blanco C, Koornneef M. 2005. Gene function beyond the single trait: natural variation, gene effects, and evolutionary ecology in *Arabidopsis thaliana*. *Plant, Cell and Environment* **28**, 2-20.
Tuberosa R. 2012. Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology* **3:347. doi:10.3389/fphys.2012.00347**.
Weigel D. 2012. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiol* **158**, 2-22.
Weigel D, Mott R. 2009. The 1001 genomes project for *Arabidopsis thaliana*. *Genome Biology* **10**, 107.
Whitley P, Hinz S, Doughty J. 2009. Arabidopsis FAB1/PIKfyve proteins are essential for development of viable pollen. *Plant Physiol* **151**, 1812-1822.

Xin Z, Browse J. 1998. Eskimo1 mutants of Arabidopsis are constitutively freezing-tolerant. *Proce Nat Acad Sci USA* **95**, 7799-7804.
Xin Z, Mandaokar A, Chen J, Last RL, Browse J. 2007. Arabidopsis ESK1 encodes a novel regulator of freezing tolerance. *The Plant Journal* **49**, 786-799.
Yano R, Takebayashi Y, Nambara E, Kamiya Y, Seo M. 2013. Combining association mapping and transcriptomics identify HD2B histone deacetylase as a genetic factor associated with seed dormancy in *Arabidopsis thaliana*. *The Plant Journal* **74**, 815-828.
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.
Yu J, Pressoir G, Briggs WH, *et al.* 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* **38**, 203-208.
Zhao K, Aranzana MJ, Kim S, *et al.* 2007. An Arabidopsis example of association mapping in structured samples. *PLoS Genetics* **3**, e4.
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 Bioliology* **53**, 358-374.

Chapter 4: Genetic analysis of morphological traits in a new, versatile, rapid-cycling *Brassica rapa* recombinant inbred line population

Hedayat Bagheri[1,6,†], Mohamed El-Soda[1,8,†], Inge van Oorschot[1], Corrie Hanhart[1], Guusje Bonnema[2], Tanja Jansen-van den Bosch[4], Rolf Mank[4], Joost Keurentjes[1], Lin Meng[7], Jian Wu[7], Maarten Koornneef[1,5] and Mark G.M. Aarts[1]

[1]Laboratory of Genetics and [2]Laboratory of Plant Breeding, Wageningen University, Wageningen, the Netherlands; [4]Keygene N.V., Wageningen, the Netherlands; [5]Max Planck Institute for Plant Breeding Research, Cologne, Germany; [6]Bu-Ali Sina University, Shahid Fahmideh, Hamedan, Iran; [7]Institute of Vegetables and Flowers, Chinese Academy of Agricultural Sciences, Beijing, China; [8]Department of Genetics, Faculty of Agriculture, Cairo University, Giza, Egypt.

†both authors contributed equally to this work

Published in Frontiers in Plant Sciences 3:183. doi: 10.3389/fpls.2012.00183

Abstract

A recombinant inbred line (RIL) population was produced based on a wide cross between the rapid-cycling and self-compatible genotypes L58, a Caixin vegetable type, and R-o-18, a yellow sarson oil type. A linkage map based on 160 F7 lines was constructed using 100 SNP, 130 AFLP®, 27 InDel and 13 publicly available SSR markers. The map covers a total length of 1150 cM with an average resolution of 4.3 cM/marker. To demonstrate the versatility of this new population, 17 traits, related to plant architecture and seed characteristics, were subjected to QTL analysis. A total of 47 QTLs were detected, each explaining between 6 to 54% of the total phenotypic variance for the concerned trait. The genetic analysis shows that this population is a useful new tool for analysing genetic variation for interesting traits in *B. rapa*, and for further exploitation of the recent availability of the *B. rapa* whole genome sequence for gene cloning and gene function analysis.

Keywords: *Brassica rapa,* Recombinant Inbred Line population, QTL analysis, plant breeding

Introduction

Brassica rapa is an important, widely cultivated crop, with various forms or "morphotypes", such as leafy vegetables, turnips and oilseed rape (Zhao et al., 2005). While the use of *B. rapa* as an oilseed crop is relatively modest, it is important as one of the parents of *Brassica napus*, the most important oilseed crop. After the oil has been extracted from the seeds, the remaining seed components (meal) are of economic interest for feeding animals. It has been known for some time that breeding for yellow seed colour is advantageous for meal quality in *B. napus*, because yellow-seeded genotypes have a thinner seed coat associated with a higher protein content and less non-energetic and anti-nutritive fibre components (Liu et al. 2012). Therefore, breeding programs aiming at combining yellow seed colour with yield associated traits such as seed number, seed size, number of siliques per plant, pod shattering, carpel number and vivipary, have been developed in *B. napus* through interspecific crosses with yellow-seeded *Brassica* species (Tang et al., 1997; Badani et al., 2006; Wittkop et al., 2009; Liu et al., 2012). A problem is that the expression of the yellow colour, at least in *B. napus*, is highly dependent on environmental factors (Liu et al. 2012).

Pod shattering, caused by carpel abscission, is an undesirable characteristic in crop breeding as it decreases the yield due to seed loss during harvesting. For *B. napus* the seed yield loss can be as much as 20% of the harvest (Price et al., 1996). The absence of embryonic dormancy during seed development, which prevents seeds to germinate prematurely on the mother plant, can be expressed as vivipary. While this is more commonly observed in cereals, it can also be found in oilseed rape, leading to large economic losses due to significant reduction in seed quality. Resistance to vivipary is therefore a very favourable trait in breeding programs (Zhang et al., 2008). Next to seed related traits, plant height, branch number and leaf number at first flower opening are factors contributing to Brassica plant architecture that differ considerably between genotypes. Plant architecture is of major agronomic importance and has a strong effect on the suitability of a plant species for cultivation, as it affects plant yield and harvest efficiency (Reinhardt and Kuhlemeier, 2002).

With the smallest genome size in the *Brassica* genus, the rapid life cycle of some of its genotypes, and the relatively close relationship to the model plant species *Arabidopsis thaliana*, *B. rapa* is considered to be one of the model dicot crops for genetic studies (Wang et al., 2011). These studies require "immortal" mapping populations, i.e. populations that can be easily propagated through seed without altering their genotypes, as indispensable tools in identifying quantitative trait loci (QTLs) underlying traits of interest (Koornneef et al., 2004).

Doubled haploid (DH) populations are the most commonly used type of immortal mapping populations for *Brassica* species (Pink et al., 2008). However, the poor response of many *B. rapa* genotypes to DH induction (Kole et al., 1997) together with the high degree of segregation distortion often observed in DH populations (Voorrips et al., 1997), limits this use. Instead, when using self-compatible genotypes with short generation times it is feasible to develop Recombinant Inbred Line (RIL) populations through sexual propagation. In this study two *B. rapa* genotypes, corresponding to two distinct morphotypes, the leafy vegetable Cai Xin accession L58, of Chinese ancestry, and the yellow sarson oil seed DH line R-o-18, of Indian ancestry, were crossed to generate a Recombinant Inbred Line (RIL) population. Both parents are early flowering and self-compatible, which facilitates rapid propagation and the ability to maintain the RILs through single seed descent.

Genetic linkage maps are required to properly query DH or RIL populations for the identification of the chromosomal regions or QTLs that harbour the genes controlling important agronomic traits. Single nucleotide polymorphisms (SNPs) represent the most abundant and common type of genetic polymorphisms that can be readily converted into genetic markers for marker assisted selection. Large-scale SNP discovery projects, using high-throughput sequencing techniques, have become a powerful complement to the standard genetic mapping procedures, and the use of resulting markers greatly improves the linkage maps of diploid crops. The Illumina GoldenGate assay is an efficient SNP genotyping tool that has been used already for soybean, tetraploid and hexaploid wheat lines and maize (Hyten et al., 2008; Akhunov et al., 2009; Yan et al., 2010). Currently, SNP genotyping is replacing the use of the AFLP technology, which has previously been very useful for analysing genetic diversity and relationships in many plant species, including *B. rapa*, identifying a large number of polymorphic loci (Zhao et al., 2005).

This paper describes the generation and genetic mapping of a large, versatile, rapid cycling *B. rapa* RIL population dedicated for QTL analysis. As an illustration of the potential importance of this population, we used it to identify 47 QTLs, responsible for most of the observed morphological variation in 17 different traits.

Materials and methods

Plant growth and generation of the RIL population

The two parental genotypes L58 and R-o-18 were crossed reciprocally and from each of the two F1 offspring, one plant was randomly selected to be propagated by subsequent

generations of self-fertilization using a single-seed-descent approach, aimed at minimizing any bias in selecting plants. The seeds of L58 (*B. rapa* ssp. *parachinensis*) were provided by Dr. Xiaowu Wang from the Institute for Vegetables and Flowers of the Chinese Academy of Agricultural Sciences, Beijing, China; and seeds of R-o-18 (*B. rapa* var. *trilocularis*) were obtained from Dr. Lars Østergaard, John Innes Centre, Norwich, UK. One of the two F1 combinations, L58 (♀) × R-o-18 (♂), was propagated until the F7 generation, the other remained at F5 and could be used for future fine-mapping studies. All generations were grown between April 2007 and June 2009 with four replications in a fully randomised design. Individual plants were grown in 19-cm diameter black plastic pots filled with a potting soil consisting of prefertilized peat, obtained from "Lentse potgrond" (www.lentsepotgrond.nl), in a temperature-controlled greenhouse at 21^{o} C with artificial long day light (16 hours). No cold treatment or vernalization was applied for germination or flowering respectively. For every generation, the first flower appeared about four weeks after germination in the early flowering lines. The inflorescences were covered with perforated plastic bags to prevent cross-pollination by insects. In case of poor seed set, hand pollinations were performed. The 160 F7 RILs were multiplied in the same conditions, ensuring homogeneous material for genetic studies.

DNA extraction and genotyping

DNA was extracted from frozen F7 leaves according to a modified CTAB procedure (Beek et al., 1992). The DNA was amplified with the Genomiphi-kit (*Illustra™ GenomiPhi™ V2 DNA Amplification Kit*, GE Healthcare UK) to be suitable for GoldenGate assay analysis (Akhunov et al., 2009). For SNP discovery, two *B. rapa* lines (Kenshin and Chiifu) were compared using CRoPS®-technology (van Orsouw et al., 2007) to reveal more than 1300 putative SNPs. The SNP-harbouring sequences were processed with the Illumina Assay Design Tool (ADT) by Illumina (www.illumina.com). A total of 384 SNPs were selected, all having ADT scores above 0.6. 100-500 ng of genomic DNA (GenomiPhi) per plant was used for Illumina SNP genotyping at Keygene N.V. using the Illumina BeadXpress™ platform and the GoldenGate Assay. Part of the DNA was used for SSR or AFLP detection as described by (Choi et al., 2007) and Vos et al. (1995) respectively. Pre-amplification and selective amplification for AFLP analysis were carried out as described by Zhao *et al*. (2005). For selective amplification seven combinations of EM (*Eco*RI/*Mse*I) primers (E34M15, E34M16, E37M32, E37M49, E37M56, E40M38, and E40M51) and four combinations of PM (P*st*I/*Mse*I) primers (P23M48, P23M50, P21M47and P23M47) were used. The *Pst* I and *Eco*RI primers were

labelled with IRD-700 at their 5' ends (Zhao et al., 2005). The reaction product of selective amplification was mixed with an equal volume of formamide-loading buffer, denatured for 5 minutes at 94° C, cooled on ice and run on a 5.5% denaturing polyacrylamide gel using the LI-COR system 4200 DNA sequencer (Li-Cor, Lincoln, Neb.) (Myburg et al., 2001). The AFLP gel images were analysed by the AFLP-Quantar Pro software. All distinguishable bands ranging from 50 bp to 500 bp were used in the data analysis. AFLP bands were scored as 1 or 0 for presence or absence of the band, respectively. All weak and ambiguous bands were scored as "unknown". In addition, 36 public SSR primer pairs (Choi et al., 2007) were used to screen for polymorphisms using the same LI-COR system to run a 5.5% denaturing polyacrylamide gel. Furthermore, 27 polymorphic InDel markers, based on DNA resequencing information of two parental lines of a DH population, which was used to construct a *B. rapa* reference map for pseudochromosome sequence assembly, were screened as described by (Tuberosa and Salvi, 2006).

Construction of a genetic linkage map and QTL analysis

The genetic map was constructed using JoinMap 4.0 (www.kyazma.nl). Monomorphic markers, markers with a high number of unknown scores and markers with more than 75 % allele skewedness towards either A or B were removed. Recombination frequencies were converted to centiMorgan (cM) distances using Haldane's mapping function. SNP markers positions were confirmed by comparing their primer sequences with the *B. rapa* genome using the *Brassica* database (BRAD) (brassicadb.org) of *Brassica* crops whole genome sequence and genetics data (Cheng et al., 2011). It contains the complete *Brassica* A genome sequence from the reference *B. rapa* genotype Chiifu-401-42 (Wang et al., 2011). InDel markers were compared to the reference map (Tuberosa and Salvi, 2006; Choi et al., 2007), which was previously used for chromosome alignment.

MAPQTL 6.0 (www.kyazma.nl) was used for QTL analysis. First, the interval mapping procedure was performed to detect major QTLs. For each trait a 1,000 X permutation test was performed to calculate the LOD threshold corresponding to a genome-wide false discovery rate of 5% ($P < 0.05$). Markers with LOD scores equal to or exceeding the threshold were used as cofactors in multiple-QTL-model (MQM) mapping. If new QTLs were detected, the linked markers were added to the cofactor list and the MQM analysis was repeated. If the LOD value of a marker dropped below the threshold in the new model, it was removed from the cofactor list and the MQM analysis was rerun. This procedure was repeated until the

cofactor list became stable. The final LOD score for each trait was determined by restricted MQM (rMQM) mapping. In some cases, rMQM mapping showed that some cofactors should be on the same linkage group, but at slightly different positions. In that case, the new marker was selected as a cofactor and the whole procedure was repeated. The linkage map was visualized using Mapchart (Voorrips, 2002).

Trait measurement

The 160 RILs (four replicate plants) and both parents (five replicate plants) were phenotyped for 17 traits. These traits are categorised into two main groups. Seed related traits, including seed colour, seed weight, seed oil, seed germination and seed vivipary; and morphological traits, including flowering time, total height, plant height until the first flower, branch number, silique length, silique beak length, silique number, number of seeds per silique, carpel number, pod shattering, total leaf number and leaf number until the first flower. Seed colour of fully mature F8 seeds was visually scored and ranked into nine different classes ranging from yellow (1) to black (9). Seed germination data were obtained by sowing 30 seeds of each line and scoring the percentage of germination 15 hours after sowing. The seeds were sterilized in 2% sodium hypochlorite for 2 minutes. After rinsing 2 times with sterile distilled water, they were sown in two rows of 15 seeds on square plates containing 50 ml of half MS medium + 1% agar. The plates were placed vertically in a 25ºC growth chamber with a 16/8 hours light/dark photoperiod. Silique length and number of seeds per silique were averaged from three ripe siliques. Seed vivipary was scored as either 0 (no vivipary), 0.5 (medium) or 1 (high) based on visual estimation of the number of seeds with radicles when harvested. Shattering was scored at harvesting time as either 0 (no open siliques), 0.5 (few open siliques) or 1 (many open siliques) (Figure 1). Seed oil was extracted by a crude method of hexane extraction, grinding 10 weighed F7 seeds of each line in 650 µl of hexane, shaking the mix for two minutes followed by one minute of centrifugation at 14,000 rpm in an Eppendorf microfuge. 600 µl of supernatant was transferred to a new tube and left overnight in the fume hood to evaporate the hexane. The oil content was determined in mg oil per mg seed (Goossens et al., 1999). All traits were measured for each of the four replicate plants, and the average values were used for mapping, except for seed colour, seed germination and seed oil content, for which only one replication could be measured. The heritability was calculated as the ratio between the genetic variation (Vg), i.e. variance between the average values of all RILs, and the total variation (Vt), with Vt = Vg + Ve, where Ve is the environmental

variation, i.e. variance between the replications of all lines. All statistical analysis was performed in SPSS 19.

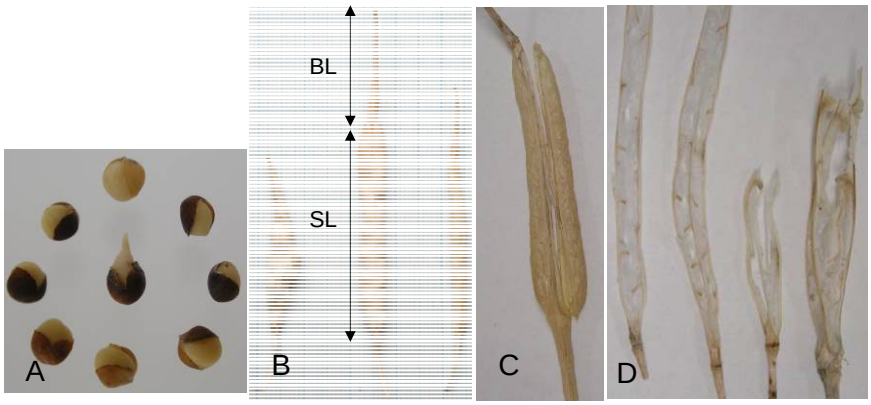


Figure 1: Phenotyping of RIL population of *B. rapa* L58 × R-o-18. A: Seed vivipary; i.e. premature germination of seeds still in the silique, or just after harvesting, B: silique length (SL) and silique beak length (BL), C: pod shattering, corresponding to the fraction of opened siliques at harvesting, D: carpel number, with the left two siliques having two carpels and the two on the right having three.

Results

Genotyping and construction of the linkage map for the RIL population

The availability of the complete genome sequence of *B. rapa* (Wang et al., 2011) and the genome analysis tools provided in the BRAD database (Cheng et al., 2011), were critical for constructing a reliable genetic map of the L58 x R-o-18 RIL population suitable for QTL mapping. Out of the 384 SNPs that could be queried by the Brassica GoldenGate assay we used, 120 SNPs were polymorphic between the parents, of which 100 provided unambiguous genotype calls for mapping. Based on the sequence of the SNP primers, the position of the 100 mapped SNP markers could be linked to their sequence position on the *B. rapa* genome, thus confirming the mapping results and providing anchoring points for chromosome number assignment and proper orientation of the chromosomal linkage maps with the genome sequence. The same was done for the SSR markers previously used to create the *B. rapa* reference linkage map (Choi et al., 2007). In total 94 InDel markers were screened, from which 27 showed polymorphism between the two parental lines. Seven of these polymorphic markers have been mapped on the reference map used for *B. rapa* pseudochromosome assembly (Tuberosa and Salvi, 2006), while the other markers were assigned to the

chromosomes according to the position of their corresponding sequence scaffolds. The final linkage map was constructed for the L58 × R-o-18 F7 RIL population using 100 SNP, 130 AFLP, 27 InDel and 13 SSR markers. It covers a total length of 1150 cM with an average resolution of 4.3 cM per marker (Figure 2).

Phenotyping the RIL population

A total of 17 traits were analysed for the F7 RIL population. Figure 3 shows the frequency distributions of the measured traits over the whole population. Transgression beyond the parental lines values was observed for most of the traits except seed colour, pod shattering, seed germination and vivipary. Broad sense heritabilities ranged from 0.35, for stem thickness, to 0.92, for flowering time (Table 1). Heritabilities could not be determined for seed colour, seed germination and seed oil content, as for these traits only one replication could be measured. Correlation analysis of all measured traits (Table 2) showed that flowering time was highly positively correlated with total leaf number and leaf number until the first flower. Silique number, seed number per silique, pod shattering and silique length were also positively correlated. In general, plants with more siliques had longer siliques with more seeds and higher seed oil content, all contributing to traits favoured for oil seed rape.

QTL analysis

In total 47 QTLs were mapped for the 17 analysed traits (Table 3 and Figure 4). Seed colour was a very prominent phenotype segregating in the population. A major QTL for seed colour (*Sc1*) was mapped to chromosome A9 with a LOD score of 30.8 and explaining 53.7 % of the total seed colour variance. This region on A9 appears to be rich in genetic variation, with several other QTLs co-located with *Sc1*, which are loci for pod shattering (*Sh*), number of seed per silique (*Nsps1*) and seed oil (*So*). The *Sh* QTL also explains a large portion, 18%, of the genetic variance. Another QTL for seed colour (*Sc2*), with a LOD score of 12.1, was mapped to chromosome A3, accounting for 15 % explained variance. Variation in vivipary (Vi) was explained by two loci, one locus on A9 (*Vi1*), with 20% explained variance, and another on A6 (*Vi2*) that explains 13% of the variance. The carpel number QTL (*Cn1*) co-localised with the silique length (*Sil*) QTL on A4, each explaining 15%, respectively 17% of the variance. This region also harbours one of the silique beak length QTLs (*Bl3*). As can be seen from Figure 1, these traits appear to be pleiotropic effects of the same locus, as the increase in carpel number often corresponds with malformed, shorter siliques with shorter beaks.

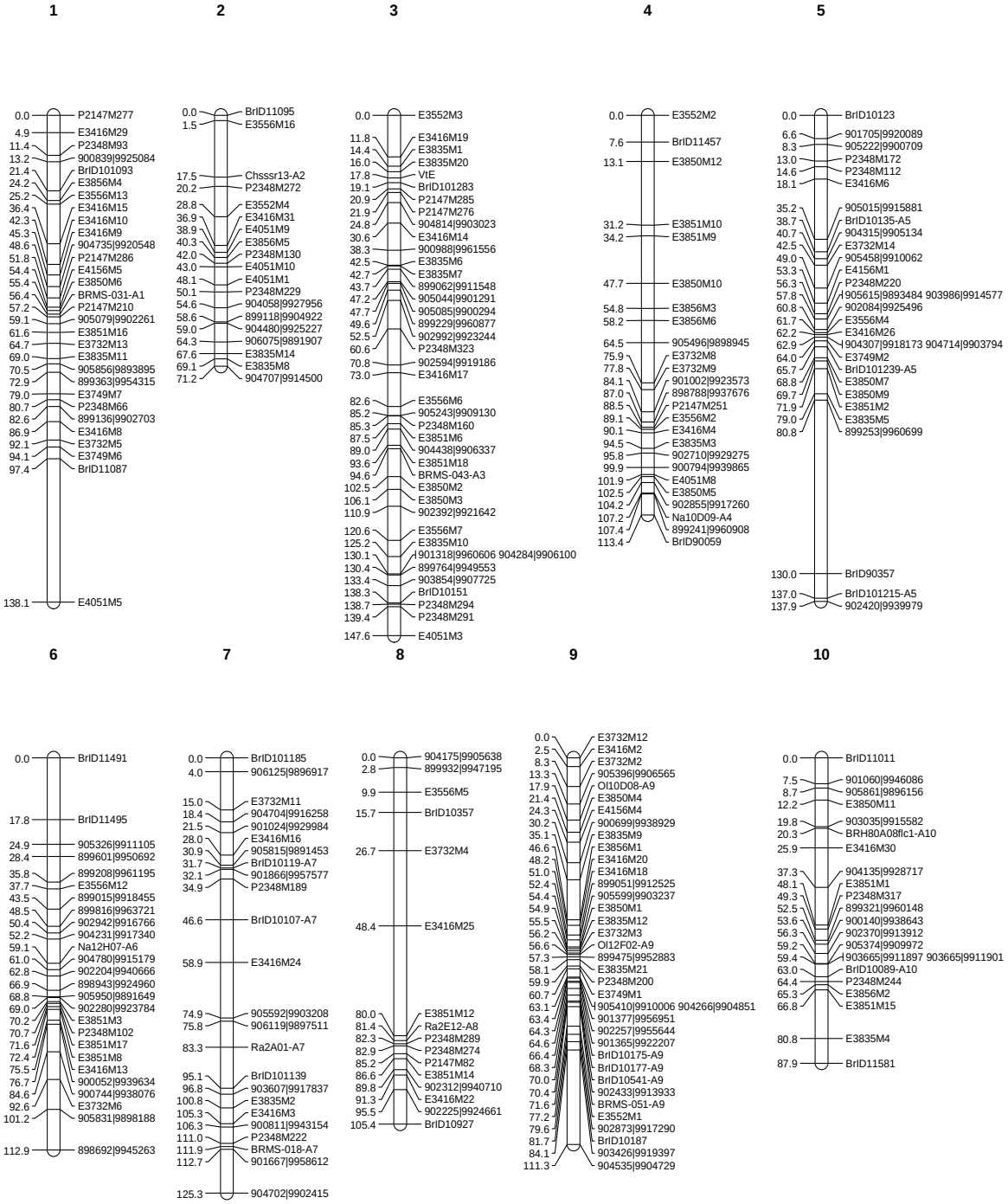


Figure 2: Genetic linkage map of the *B. rapa* L58 × R-o-18 RIL population, showing the positions of 270 markers (100 SNP, 130 AFLP, 27 InDel and 13 SSR markers) distributed over 10 linkage groups corresponding to the 10 chromosomes of the *Brassica* A genome. Markers labelled with [6 digits|7 digits] are SNPs, markers labelled "E…M.." or "P…M.." are respectively EcoRI/MseI or PstI/MseI generated AFLPs, markers labelled "BrID…." are InDels and the remaining markers are SSRs.

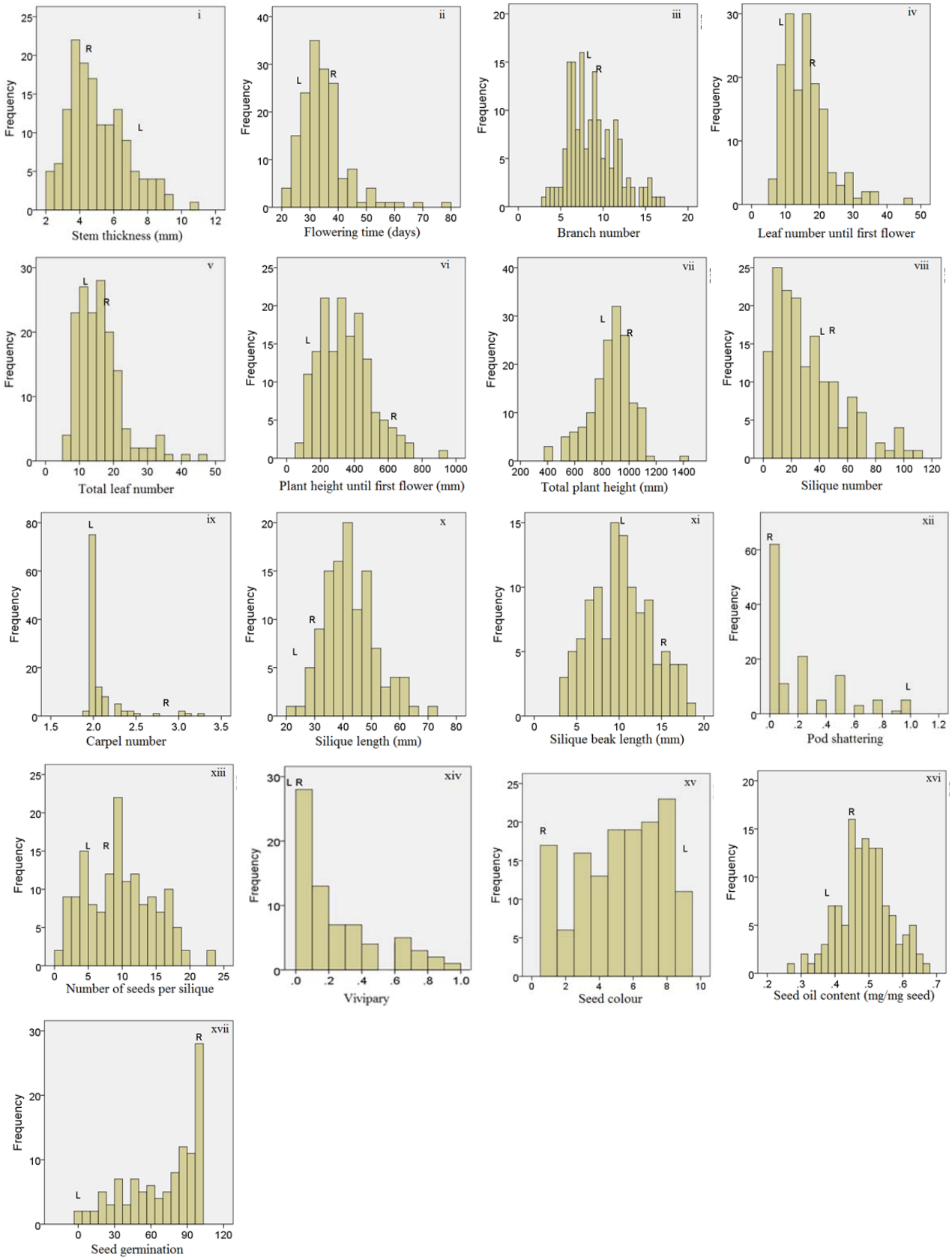


Figure 3: Frequency distributions of non-normalized data of the reported traits for the L58 × R-o-18 RIL population. The vertical axes indicate the number of lines per trait value class and the horizontal axes indicate the different trait value classes. The parental values (indicated with L and R) are the mean of five replicates.

Table 1: Phenotype data for both parental lines (L58 and R-o-18) and the RIL population, for the 17 analysed traits. "%" indicates the relative performance of R-o-18 compared to L58. "Min" and "Max" indicate the values of the RIL with respectively the lowest of the highest value, while "Range" indicates the difference between these values. "Mean" is the average value for all RIL lines, with standard deviation (SD), and h^2 is broad sense heritability. For all traits four replicate samples were measured, except for seed colour, seed germination and seed oil content, for which only one sample could be measured. All 160 lines have been scored.

		Parental lines			RIL population					
Trait (unit)	abbreviation	L58	R-o-18	%	Min	Max	Range	Mean	SD	h^2
Stem thickness (mm)	*St*	7.5	4.5	60.0	2.0	10.8	8.7	5.1	1.7	0.35
Flowering time (days)	*Ft*	28.8	38.4	133.6	22.0	79.8	57.8	34.6	8.2	0.92
Branch number	*Bn*	8.4	8.9	105.9	3.0	17.0	14.0	8.7	2.9	0.69
Leaf number until first flower	*Lnf*	11.8	17.8	151.8	5.0	46.0	41.0	15.9	6.6	0.89
Total leaf number	*Tln*	13.9	17.8	128.1	5.5	46.8	41.3	15.9	6.9	0.89
Plant height until first flower (mm)	*Ph*	171	658	384	88	943	855	344	154	0.81
Total plant height (mm)	*Tph*	878	1078	123	398	1403	1005	864	156	0.76
Silique number	*Sin*	46.5	48.8	104.8	0.0	108.5	108.5	32.3	24.1	0.73
Carpel number	*Cn*	2.0	2.8	141.3	1.9	3.3	1.3	2.1	0.2	0.50
Silique length (mm)	*Sil*	22.9	34.6	151.3	23.3	70.9	47.6	42.5	9.1	0.78
Silique beak length (mm)	*Bl*	10.2	14.8	145.4	3.7	18.5	14.8	10.3	3.6	0.71
Pod shattering	*Sh*	1.0	0.0	0.0	0.0	1.0	1.0	0.2	0.3	0.50
Number of seeds per silique	*Nsps*	6.2	8.1	131.1	1.0	22.9	21.9	9.7	5.0	0.61
Vivipary	*Vi*	0.0	0.0	-	0.0	1.0	1.0	0.2	0.3	0.68
Seed colour	*Sc*	9.0	1.0	11.1	1.0	9.0	8.0	5.3	2.5	-
Seed oil content (mg oil/mg seed)	*So*	0.4	0.4	112.8	0.3	0.7	0.4	0.5	0.1	-
Seed germination	*Sg*	0.0	100.0	-	0.0	100.0	100.0	68.1	29.8	-

Table 2: Pearson correlations for the analysed traits of the L58 × R-o-18 RIL population.

Trait	*St*	*Ft*	*Bn*	*Lnf*	*Tln*	*Ph*	*Tph*	*Sin*	*Cn*	*Sil*	*Bl*	*Sh*	*Nsps*	*Vi*	*Sc*	*So*
St	1															
Ft	.571	1														
Bn	.376	.382	1													
Lnf	.562	.847	.563	1												
Tln	.566	.861	.557	.988	1											
Ph	.183	.275	.354	.315	.289	1										
Tph	.252	-.014	.157	.016	.012	.629	1									
Sin	.079	.084	.025	.067	.074	-.109	-.084	1								
Cn	.127	.064	.032	.179	.157	.174	.127	-.025	1							
Sil	-.191	-.108	-.078	-.105	-.123	.196	.107	.184	-.181	1						
Bl	-.137	-.003	-.053	-.028	-.021	.175	.081	.198	-.006	.680	1					
Sh	.053	.072	.021	.033	.054	-.148	-.028	.419	.045	.047	.186	1				
Nsps	.249	.171	.120	.139	.143	.085	.097	.367	.166	.151	.075	.355	1			
Vi	-.344	-.369	-.256	-.392	-.388	-.188	.018	.147	-.150	-.001	.103	-.015	-.305	1		
Sc	-.094	-.075	-.007	-.121	-.120	.012	.004	.060	.142	-.166	.021	.238	.170	-.095	1	
So	.205	.285	.076	.222	.221	.010	-.112	.345	.020	-.040	-.048	.113	.516	-.228	.179	1
Sg	.017	-.105	.091	-.104	-.080	.052	-.020	.032	.024	-.045	-.154	-.123	.007	.173	-.127	-.004

St: stem thickness, *Ft*: flowering time, *Bn*: branch number, *Lnf*: leaf number until first flower, *Tln*: total leaf number, *Ph*: plant height until first flower, *Tph*; total plant height, *Sin*: silique number, *Cn*: carpel number, *Sil*: silique length, *Bl*: silique beak length, *Sh*: Pod shattering, *Nsps*: number of seed per silique, *Vi*: vivipary, *Sc*: seed colour, *So*: seed oil content, *Sg*: seed germination
Dark gray boxes means significant at $P \leq 0.01$; light gray boxes means significant at $P \leq 0.05$.

Pleiotropy is also the likely cause of the co-localisation of flowering time QTLs *Ft3, Ft4* and *Ft5* with QTLs for total leaf number (*Tln1, Tln2, Tln4*) and leaf number until the first flower (*Lnf1, Lnf3, Lnf4*) on respectively A2, A7 and A8. Tln and Lnf share four of the six QTLs found for these traits, in line with the correlation found between them. The locus on A8 also seems to account for variation for branch number (Bn), harbouring the major Bn QTL (*Bn1*). Plant height until the first flower (Ph) and total plant height (Tph) also share one common QTL, on A10 (*Ph1* and *Tph2*).

Table 3: QTLs detected for the analysed traits in the L58 × R-o-18 RIL population. Per trait, QTLs are numbered according to decreasing LOD score (LOD). LOD thresholds are calculated per trait based on 1000 permutation tests and an experimental error rate of $P < 0.05$. R^2 is the percentage of total phenotypic variance explained by each QTL. For each QTL, the allelic effect is calculated as μA-μB (μ= mean), where A and B are RILs carrying L58 respectively R-o-18 alleles at the QTL.

Trait	QTL	Linkage group	LOD threshold	LOD	Position of LOD peak (cM)	R^2	Effect
Stem thickness	*St1*	A7	3	4.7	58.9	13.1	1.3
	St2	A4		2.8	7.6	8	1.1
Flowering time	*Ft1*	A7	2.8	5.4	34.9	11.2	5.5
	Ft2	A5		4.9	60.8	9	-5.4
	Ft3	A8		4.9	85.2	9	-5.1
	Ft4	A2		3.7	64.3	6.6	4.2
	Ft5	A7		3.1	106.3	6.5	-4.5
Branch number	*Bn1*	A2	3.1	5.5	64.3	13.1	2.1
	Bn2	A3		3.2	38.3	7.4	-1.7
	Bn3	A6		2.8	76.7	6.4	1.5
Leaf number until first flower	*Lnf1*	A8	2.9	8.8	91.3	15	-5.3
	Lnf2	A4		4.9	90.1	8	3.8
	Lnf3	A7		4.8	96.7	8	-3.9
	Lnf4	A2		3.88	64.3	6.2	3.4
	Lnf5	A5		3.2	35.171	5	-3.1
	Lnf6	A9		2.9	79.578	4.5	-2.8
Total leaf number	*Tln1*	A8	2.9	8.3	91.3	14	-5.3
	Tln2	A7		5.7	106.4	10.6	-4.8
	Tln3	A7		4.1	58.9	7.6	4.1
	Tln4	A2		4.6	64.3	7.3	3.8
	Tln5	A4		3.3	77.8	5.1	3.2
	Tln6	A3		3	38.29	4.6	-3.3
Plant height until first flower	*Ph1*	A10	2.8	4	59.2	9.3	95.4
	Ph2	A8		3.6	81.4	8.3	-90.2

	Ph3	A1		2.9	55.5	6.6	-85.2
Total plant height	*Tph1*	A3	3	5	38.3	11.3	-112.3
	Tph2	A10		4.1	64.4	9.5	96.1
	Tph3	A5		3.2	69.7	7.1	87.2
Silique number	*Sin*	A6	3	4	72.4	11	16.0
Carpel number	*Cn1*	A4	2.7	4.4	84.1	15.2	0.2
	Cn2	A2		2.5	17.5	8.5	-0.1
Silique length	*Sil*	A4	3	5	90.1	18.6	-8.1
Silique beak length	*Bl1*	A10	3	5.7	53.6	14.4	2.8
	Bl2	A8		4.4	0	11	-2.4
	Bl3	A4		4.1	90.1	10	-2.3
	Bl4	A1		72.9	3	7.2	-2.0
Pod shattering	*Sh*	A9	3	5.5	60	18	0.2
Number of seeds per silique	*Nsps1*	A9	3	3.5	58.1	10	3.1
	Nsps2	A3		2.7	82.6	7.2	-2.7
Vivipary	*Vi1*	A9	3	3.9	111.3	20.5	0.3
	Vi2	A6		2.6	61.1	13.3	-0.2
Seed colour	*Sc1*	A9	3	30.8	56.6	53.7	3.7
	Sc2	A3		12.1	52.5	15	2.1
Seed oil content	*So*	A9	3	3	58.2	9.1	0.1
Seed germination	*Sg1*	A5	3.1	4.1	137	14.4	-23.2
	Sg2	A3		3.1	147.6	10.5	-22.1

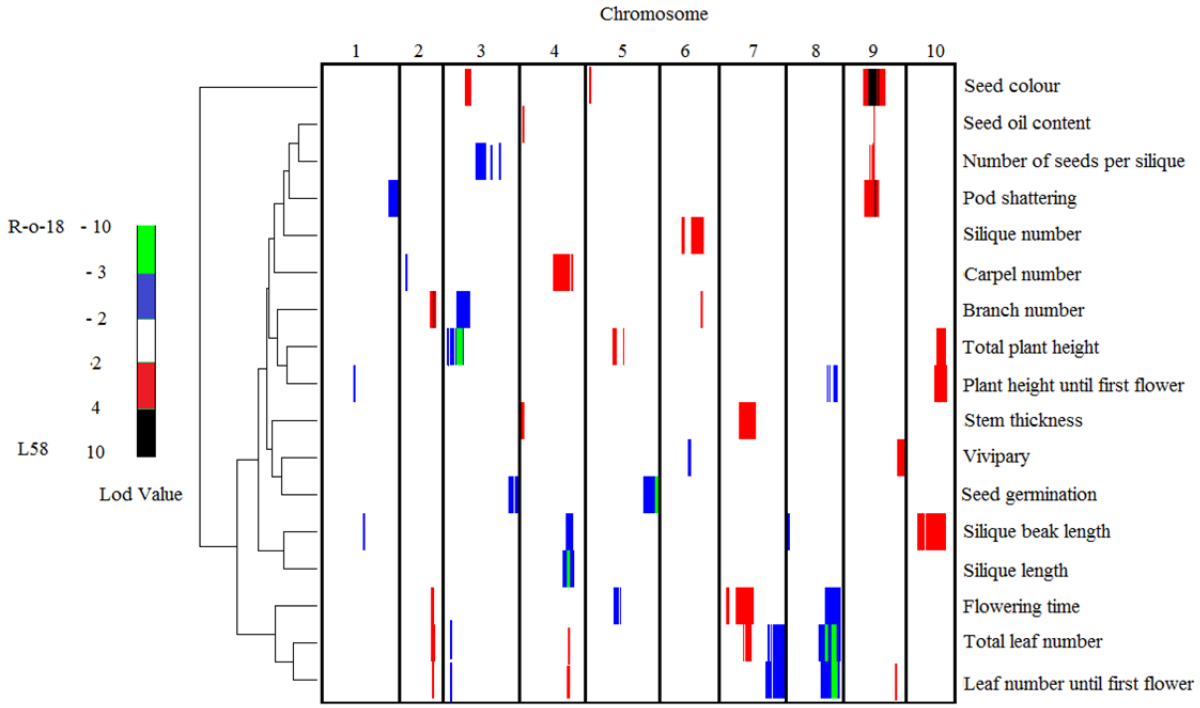


Figure 4: A clustered heat map showing the LOD profiles of the measured traits. Columns indicate the 10 chromosomes in centimorgans, ascending from the left to right; rows indicate

individual trait LOD profiles. A colour scale is used to indicate the QTL significance corresponding to the LOD score. Positive values (red and black) represent a positive effect of the trait by the L58 allele, negative values (blue and green) represent a positive effect on the trait by the R-o-18 allele. The width of a bar indicates the significance interval of the QTL that was calculated by rMQM in MAPQTL 6. Hierarchical clustering, shown on the left, reflects the correlation between traits based on the QTL profiles.

Discussion

The L-58 × R-o-18 population is a new RIL population, designed for general QTL mapping studies. The parents of this population were selected for a number of reasons. Rapid cycling and self-compatibility were two important reasons, as these would permit the rapid construction of the population and easy maintenance through single-seed-descent propagation. These are also the reasons that both parents are more and more used as reference genotypes, expanding their use for other purposes, such as the generation of a TILLING population in R-o-18 (Schmidt et al., 2011), as reference species in micro-array design (Love et al., 2010), as well as being used in setting up a diversity fixed foundation set (DFFS) and as parents in other mapping populations. For the latter purpose, currently the genome sequences and transcript profiles of both parents are being determined (Jian Wu, Xiaowu Wang e.a., unpublished results). There are not many "immortal" *B. rapa* populations available for mapping studies, with immortal meaning that the individual lines are genetically homozygous and can thus be propagated through seeds while maintaining the established genotype in their progeny. There are few other RIL populations (Kole et al., 1997; Iniguez-Luy et al., 2009), although others may still be in development (www.brassica.info). In addition, there are several doubled haploid (DH) populations available (Tuberosa and Salvi, 2006; Zhang et al., 2006; Choi et al., 2007; Lou et al., 2007; Tuberosa, 2012), which are also very useful for genetic mapping studies, although they generally comprise about half the number of recombination events compared to RIL populations and often suffer more from regions with skewedness towards one of the parental alleles.

The transgression beyond the parental lines, which was observed in the F2 generation (Bagheri et al., manuscript under review) was encouraging to produce the F7 RIL family through single-seed-descent. Out of 200 F2 lines, only 160 F7 lines were available for genotyping. This 16% loss from F2 till F7 was mostly due to plant sterility, apparently from reduced pollen production. Although this may have a genetic basis, it was not obviously related to strong skewedness of the population towards one of the two parental alleles at a

particular locus. To reduce the risk of skewedness we started off with a relatively large population. Thus, most of the cross-overs between alleles located in skewed segregation region could be detected and correct linkage distances could be calculated along any skewed marker region. We did find the occasional marker with more than 75 % allele skewedness towards either L58 or R-o-18 alleles, but since all of these markers were flanked by closely linked, non-skewed markers, the skewedness was found to be due to marker scoring problems rather than genetic skewedness, upon which the improperly scored markers were removed.

The residual heterozygosity in the RIL population was not significantly higher than the expected value of 1.56 %. Unintended selection during single-seed-descent propagation, for instance for plant size or fecundity, could lead to increased heterozygosity at some loci (Loudet et al., 2002). Since effort was made to randomly designate which plants would be selected at each propagation cycle, we were able to avoid this type of distortion in this population. The genetic map was constructed for the F7 RILs using a mix of AFLPs, SNPs, InDel and SSR markers. The whole genome sequence information of *B. rapa* (Cheng et al., 2011; Wang et al., 2011) ensured the correct genome location of the SNPs, SSR and InDel markers for which primer sequences were available. This was very efficient in resolving any mapping ambiguities and in assigning chromosome numbers to linkage groups. The current map covers a total length of 1150 cM with an average resolution of 4.3 cM. This map is comparable to two *B. rapa* reference linkage maps based on DH populations, with a total length of 1182 cM (Choi et al., 2007) and 1234.2 cM (Tuberosa and Salvi, 2006), and the map reported for another RIL population, of 1125 cM (Iniguez-Luy et al., 2009). The marker resolution of 4.3 cM per marker found for this population, is also in line with the reported maps. In some cases, composite interval mapping (CIM), which is one of the QTL mapping methods we used, can be affected by an uneven distribution of markers in the genome (Zeng et al., 1999), which is why non-informative markers were omitted if they did not detect additional recombination events, to keep the smallest informative marker set. Simulation studies have shown that the advantages of increasing marker density beyond one marker every 4.3 cM are less significant than those obtained when increasing the size of the population (Darvasi and Soller, 1994; Charmet, 2000). That means that with the current marker density, there is no need to screen for additional markers in order to improve mapping efficiency.

In total 47 QTLs for 17 analysed traits were mapped. Seed coat colour is a very important trait in *Brassica* oilseed crops. A yellow seed colour is known to be highly correlated with meal quality, because of the thinner seed coat, corresponding to less anti-nutritive fibre components, which is also associated with higher protein content (Tang et al., 1997; Badani et

al., 2006; Wittkop et al., 2009; Liu et al., 2012). Seed coat colour is a maternally inherited trait, with the alleles for black seed coat acting dominantly over the alleles for yellow seed coat. In *B. napus*, seed colour is inherited in different ways, probably depending on the source of the genetic variation, and is strongly affected by environmental factors (Liu et al., 2012). Earlier studies (Stringam, 1980) proposed a two locus model for seed colour in *B. rapa*, involving the *Br1* and *Br2* loci, both of which were not mapped at the time. In *B. rapa* studies involving yellow sarson oilseed types, as used in this study, a major seed colour locus is found on chromosome A9 (Lou et al., 2007). In this RIL population two major QTLs were detected for seed colour, *Sc1* and *Sc2*, on A9 and A3 respectively, explaining about 70% of seed coat colour variation. The *Sc1* locus on A9 co-located with previously reported seed colour QTLs reported for both *B. napus* and *B. rapa* (Lou et al., 2007; Wang et al., 2011; Liu et al., 2012). Near-infrared reflectance spectroscopy measurements of acid detergent lignin (ADL) in seeds of both parental lines confirmed the expected difference in ADL corresponding to yellow and black seed (R. Snowdon, personal communication) suggesting that the *B. rapa* locus we mapped to A9 affects the same gene as the A9 locus cloned from *B. napus*. This locus was found to harbour a mutation in the *CCR1* gene, encoding a cinnamoyl co-A reductase involved in lignin biosynthesis (Liu et al., 2012). In the absence of the L58 allele at *Sc1*, seeds containing the L58 allele at *Sc2* are brown, not yellow. Most of the cultivated *B. rapa* is brown-seeded, while for commercial purpose oilseed *B. rapa* with brown seeds is not preferred due to the darker colouring of the oil (Ramchiary et al., 2011). Introgression of the *Sc2* allele of yellow sarson types like R-o-18, could overcome this. In addition to the reported *Sc1* and *Sc2* loci, we found an additional, but very weak, *Sc3* QTL (LOD = 2.23), which mapped to A5 and accounted for 2% of the phenotypic variance. Previously, a QTL controlling yellow seed colour was mapped to A5 (Teutonico and Osborn, 1994), which could concern the same locus.

Of the five QTLs detected for flowering time (Ft), four flowering time QTLs, *Ft1*, *Ft3*, *Ft4* and *Ft5*, co-localised with previously mapped QTLs (Osborn et al., 1997; Lou et al., 2007; Edwards and Weinig, 2011; Lou et al., 2011). Another QTL, *Ft2* co-localised with a previously mapped, non-significant, QTL on A5 for a circadian clock parameter (Lou et al., 2011). Flowering time is highly co-related with plant architecture traits like plant height (Ph), leaf number until first flower (Lnf), total leaf number (Tln), and branch number (Bn). *Ph2, Tln1* and *Lnf1* co-localised with *Ft3*, while *Ft4* co-localised with *Tln4*, *Lnf4* and *Bn1*, the only branch number QTL co-localising with a flowering time locus. Furthermore, *Ft5* co-localised with *Tln2 and Lfn3*; and finally *Ft1* co-localised with *Tln3*. *Lnf3* and *Lnf4* have been

previously mapped by Lou et al. (2007), who also observed the general co-localisation of *Ft* and *Lnf* loci. *Tln6* and *Lnf6* are two separate loci, mapping to A3 and A9 respectively, which did not co-localise with any *Ft* QTLs in this population, but which co-localised with *Ft* QTLs detected by (Edwards and Weinig, 2011).

Resistance to pod shattering is a recessive complex trait, mainly based on data from *B. napus*, which is difficult to assess because it can only be scored at maturity (Morgan et al., 2003). There are no reports related to *Brassica* loci controlling pod shattering, although work has been done on genetic engineering of pod shattering resistance, using ectopic expression of the *FRUITFULL* gene from Arabidopsis (Østergaard et al., 2006). The pod shattering QTL (*Sh*) on A9 is located in the same region as *Sc1*, but if indeed *Sc1* is caused by variation at the *CCR1* gene, as we expect, this is unlikely to be a pleiotropic effect of the same locus. Fortunately the alleles for black seeds and easy shattering are in coupling phase (Table 2), which means that selection for yellow-seeded lines could easily be accompanied by selection for improved shattering resistance. Since there is limited genetic variation for pod shattering resistance within the *B. napus* germplasm (Morgan et al., 2003), introducing pod shattering resistance alleles from *B. rapa* into a *B. napus* breeding program could well be an interesting approach.

Shattering has a significant positive correlation with the number of seeds per silique (Nsps) and the silique number (Sin). A significant *Sh* QTL co-located with *Nsps1*.The number of seeds per silique is also highly positively correlated with other silique related traits such as silique length (Sil) and silique beak length (Bl). Therefore Sil and Bl are likely to have an overall effect on silique related traits. Silique length (Sil) and beak length (Bl) shared one QTL, *Sil* and *Bl3* respectively. This co-localization is supported with high correlation between the two traits. Lou et al. (2007) reported two genomic regions on A1 and A7 and three loci on A5, A7 and A9, controlling silique length and beak length respectively. The *Sil* QTL on A4 reported here, is a new locus that explains 18.6% of the variance.

Vivipary (pre-harvest sprouting) is another important oilseed quality trait. A major QTL explaining 50. 8% of the total variance for vivipary had previously been mapped to chromosome N11 of *B. napus* (Feng et al., 2009). We are not aware of previous work on seed vivipary QTLs in *B. rapa*. The two QTLs we detected on A9 and A6, explain about 30% of the vivipary variance. Vivipary is negatively correlated with seed oil content in our data. Also in *B. napus*, vivipary decreased seed viability and vigour and resulted in lower seed oil content (Ruan et al., 2008).

With the availability of the *Brassica rapa* genome sequence (Wang et al 2011) and the further

development of molecular genetic tools based on the parental genotypes we used for the L58 x R-o-18 RIL population, we anticipate that the population can be a very useful additional tool to improve gene cloning approaches in *B. rapa* and thus contribute to more efficient *B. rapa* breeding.

Acknowledgements

This work was financially supported by a personal grant to Hedayat Bagheri from the Ministry of Science, Research and Technology of Iran, by the IOP Genomics project IGE050010 on Brassica Vegetable Nutrigenomics and by the Graduate School Experimental Plant Sciences. We acknowledge Dr. Xiaowu Wang and Dr. Lars Østergaard for their generous supply of seeds for the L58 and R-o-18 parental lines and we thank Dr. Rod Snowdon and Dr. Benny Wittkop from the Justus Liebig University in Giessen, Germany, for their help in analysing L58 and R-o-18 seeds for lignin content and composition.

References:

Akhunov E, Nicolet C, Dvorak J. 2009. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *TAG Theoretical and Applied Genetics* **119**, 507-517.

Badani AG, Snowdon RJ, Wittkop B, Lipsa FD, Baetzel R, Horn R, De Haro A, Font R, Lühs W, Friedt W. 2006. Colocalization of a partially dominant gene for yellow seed colour with a major QTL influencing acid detergent fibre (ADF) content in different crosses of oilseed rape (Brassica napus). *Genome* **49**, 1499-1509.

Beek JG, Verkerk R, Zabel P, Lindhout P. 1992. Mapping strategy for resistance genes in tomato based on RFLPs between cultivars: Cf9 (resistance to *Cladosporium fulvum*) on chromosome 1. *TAG Theoretical and Applied Genetics* **84**, 106-112.

Charmet G. 2000. Power and accuracy of QTL detection: simulation studies of one-QTL models. *Agronomie* **20**, 309-323.

Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, Li P, Hua W, Wang X. 2011. BRAD, the genetics and genomics database for Brassica plants. *BMC Plant Biology* **11**, 136.

Choi S, Teakle G, Plaha P, *et al.* 2007. The reference genetic linkage map for the multinational *Brassica rapa* genome sequencing project. *TAG Theoretical and Applied Genetics* **115**, 777-792.

Darvasi A, Soller M. 1994. Optimum spacing of genetic markers for determining linkage between marker loci and quantitative trait loci. *TAG Theoretical and Applied Genetics* **89**, 351-357.

Edwards CE, Weinig C. 2011. The quantitative-genetic and QTL architecture of trait integration and modularity in Brassica rapa across simulated seasonal settings. *Heredity (Edinb)* **106**, 661-677.

Feng F, Liu P, Hong D, Yang G. 2009. A major QTL associated with preharvest sprouting in rapeseed (*Brassica napus* L.). *Euphytica* **169**, 57-68.

Goossens A, Dillen W, De Clercq J, Van Montagu M, Angenon G. 1999. The arcelin-5 gene of Phaseolus vulgaris directs high seed-specific expression in transgenic Phaseolus acutifolius and Arabidopsis plants. *Plant Physiol* **120**, 1095-1104.

Hyten D, Song Q, Choi I-Y, Yoon M-S, Specht J, Matukumalli L, Nelson R, Shoemaker R, Young N, Cregan P. 2008. High-throughput genotyping with the GoldenGate assay in the complex genome of soybean. *TAG Theoretical and Applied Genetics* **116**, 945-952.

Iniguez-Luy F, Lukens L, Farnham M, Amasino R, Osborn T. 2009. Development of public immortal mapping populations, molecular markers and linkage maps for rapid cycling *Brassica rapa*& *B. oleracea*. *TAG*

Theoretical and Applied Genetics **120**, 31-43.
Kole C, Kole P, Vogelzang R, Osbom TC. 1997. Genetic Linkage Map of a Brassica rapa Recombinant Inbred Population. *Journal of Heredity* **88**, 553-557.
Koornneef M, Alonso-Blanco C, Vreugdenhil D. 2004. Naturally occuring genetic variation in *Arabidopsis thaliana Annual Review of Plant Biology* **55**, 141-172.
Liu L, Stein A, Wittkop B, Sarvari P, Li J, Yan X, Dreyer F, Frauen M, Friedt W, Snowdon R. 2012. A knockout mutation in the lignin biosynthesis gene *CCR1* explains a major QTL for acid detergent lignin content in *Brassica napus* seeds. *TAG Theoretical and Applied Genetics*, 1-14.
Lou P, Xie Q, Xu X, Edwards CE, Brock MT, Weinig C, McClung CR. 2011. Genetic architecture of the circadian clock and flowering time in Brassica rapa. *Theor Appl Genet* **123**, 397-409.
Lou P, Zhao J, Kim JS, *et al.* 2007. Quantitative trait loci for flowering time and morphological traits in multiple populations of Brassica rapa. *J Exp Bot* **58**, 4005-4016.
Loudet O, Chaillou S, Camilleri C, Bouchez D, Daniel-Vedele F. 2002. Bay-0 x Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in Arabidopsis. *Theoritical and Applied Genetics* **104**, 1173-1184.
Love CG, Graham NS, Ó Lochlainn S, Bowen HC, May ST, White PJ, Broadley MR, Hammond JP, King GJ. 2010. A *Brassica* exon array for whole-transcript gene expression profiling. *PLoS ONE* **5**, e12812.
Morgan C, Bavage A, Bancroft I, Bruce D, Child R, Chinoy C, Summers J, Arthur E. 2003. Using novel variation in Brassica species to reduce agricultural inputs and improve agronomy of oilseed rape—a case study in pod shatter resistance. *Plant Genetic Resources* **1**, 59-65.
Myburg AA, Remington DL, O'Malley DM, Sederoff RR, Whetten RW. 2001. High-throughput AFLP analysis using infrared dye-labeled primers and an automated DNA sequencer. *BioTechniques* **30**, 348-352, 354, 356-347.
Osborn TC, Kole C, Parkin IA, Sharpe AG, Kuiper M, Lydiate DJ, Trick M. 1997. Comparison of flowering time genes in Brassica rapa, B. napus and Arabidopsis thaliana. *Genetics* **146**, 1123-1129.
Østergaard L, Kempin SA, Bies D, Klee HJ, Yanofsky MF. 2006. Pod shatter-resistant Brassica fruit produced by ectopic expression of the FRUITFULL gene. *Plant Biotechnology Journal* **4**, 45-51.
Pink D, Bailey L, McClement S, Hand P, Mathas E, Buchanan-Wollaston V, Astley D, King G, Teakle G. 2008. Double haploids, markers and QTL analysis in vegetable brassicas. *Euphytica* **164**, 509-514.
Price JS, Hobson RN, Neale MA, Bruce DM. 1996. Seed Losses in Commercial Harvesting of Oilseed Rape. *Journal of Agricultural Engineering Research* **65**, 183-191.
Ramchiary N, Lim YP, Schmidt R, Bancroft I. 2011. Genetics of *Brassica rapa* L. *Genetics and Genomics of the Brassicaceae*, Vol. 9: Springer New York, 215-260.
Reinhardt D, Kuhlemeier C. 2002. Plant architecture. *EMBO Rep* **3**, 846-851.
Ruan SL, Hu WM, Duan XM, Ma HS. 2008. Ultrastructural and electrophoretic analyses of viviparous and normal seeds in hybrid rape (Brassica napus L.). *Seed Science and Technology* **36**, 371-378.
Schmidt R, Bancroft I, Ramchiary N, Lim Y. 2011. Genetics of Brassica rapa L. *Genetics and Genomics of the Brassicaceae*, Vol. 9: Springer New York, 215-260.
Stringam GR. 1980. INHERITANCE OF SEED COLOR IN TURNIP RAPE. *Canadian Journal of Plant Science* **60**, 331-335.
Tang ZL, Li JN, Zhang XK, Chen L, Wang R. 1997. Genetic variation of yellow-seeded rapeseed lines (*Brassica napus L.*) from different genetic sources. *Plant Breeding* **116**, 471-474.
Teutonico RA, Osborn TC. 1994. Mapping of RFLP and qualitative trait loci in <i>Brassica rapa</i> and comparison to the linkage maps of <i>B. napus</i>, <i>B. oleracea</i>, and <i>Arabidopsis thaliana</i&gt. *TAG Theoretical and Applied Genetics* **89**, 885-894.
Tuberosa R. 2012. Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology* **3:347. doi:10.3389/fphys.2012.00347**.
Tuberosa R, Salvi S. 2006. Genomics-based approaches to improve drought tolerance of crops. *Trends in Plant Science* **11**, 405-412.
van Orsouw NJ, Hogers RCJ, Janssen A, *et al.* 2007. Complexity Reduction of Polymorphic Sequences (CRoPS™): A Novel Approach for Large-Scale Polymorphism Discovery in Complex Genomes. *PLoS ONE* **2**, e1172.
Voorrips RE. 2002. MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *Journal of Heredity* **93**, 77-78.
Voorrips RE, Jongerius MC, Kanne HJ. 1997. Mapping of two genes for resistance to clubroot (Plasmodiophora brassicae) in a population of doubled haploid lines of Brassica oleracea by means of RFLP and AFLP markers. *Theor Appl Genet* **94**, 75-82.
Wang X, Wang H, Wang J, *et al.* 2011. The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics* **43**, 1035-1039.

Wittkop B, Snowdon R, Friedt W. 2009. Status and perspectives of breeding for enhanced yield and quality of oilseed crops for Europe. *Euphytica* **170**, 131-140.
Yan J, Yang X, Shah T, Sánchez-Villeda H, Li J, Warburton M, Zhou Y, Crouch J, Xu Y. 2010. High-throughput SNP genotyping with the GoldenGate assay in maize. *Molecular Breeding* **25**, 441-451.
Zeng ZB, Kao CH, Basten CJ. 1999. Estimating the genetic architecture of quantitative traits. *Genet Res* **74**, 279-289.
Zhang X-Q, Li C, Tay A, Lance R, Mares D, Cheong J, Cakir M, Ma J, Appels R. 2008. A new PCR-based marker on chromosome 4AL for resistance to pre-harvest sprouting in wheat (*Triticum aestivum* L.). *Molecular Breeding* **22**, 227-236.
Zhang X-w, Wu J, Zhao J-j, *et al.* 2006. Identification of QTLs Related to Bolting in Brassica rapa ssp. pekinensis (syn. Brassica campestris ssp. pekinensis). *Agricultural Sciences in China* **5**, 265-271.
Zhao J, Wang X, Deng B, Lou P, Wu J, Sun R, Xu Z, Vromans J, Koornneef M, Bonnema G. 2005. Genetic relationships within *Brassica rapa* as inferred from AFLP fingerprints. *TAG Theoretical and Applied Genetics* **110**, 1301-1314.

Chapter 5: Genotype by environment interaction for pre-flowering physiological and morphological responses to drought in *Brassica rapa*

Mohamed El-Soda[1,2], Martin P. Boer[3], Hedayat Bagheri[1,4], Corrie J. Hanhart[1], Maarten Koornneef[1,5], Mark G.M. Aarts[1]

[1]Laboratory of Genetics, Wageningen University, Wageningen, the Netherlands; [2]Department of Genetics, Faculty of Agriculture, Cairo University, Egypt; [3]Biometris - Applied Statistics, Department of Plant Science, Wageningen University, Wageningen, the Netherlands, [4]Bu-Ali Sina University, Shahid Fahmideh, Hamedan, Iran, [5]Max Planck Institute for Plant Breeding Research, Cologne, Germany

Submitted to J. Exp. Botany (under review)

Abstract

Plant growth and productivity are greatly affected by drought, which is likely to become more threatening with the predicted global temperature increase. Understanding the genetic architecture of complex quantitative traits and their interaction with water availability may lead to improved crop adaptation to a wide range of environments. Here, we explore the genetic basis of 20 physiological and morphological traits describing plant performance and growth in a *Brassica rapa* RIL population grown on a sandy substrate supplemented with nutrient solution, under control and drought conditions. Altogether we identified 54 QTL of which many co-located in 11 QTL clusters. Seventeen QTL showed significant QTL by environment interaction (QxE), indicating genetic variation for phenotypic plasticity. Of the measured traits, only hypocotyl length did not show significant GxE between both environments in all experiments. Correlation analysis showed that in the control environment, stomatal conductance was positively correlated with total leaves dry weight (DW) and above ground DW, whereas, in the drought environment, stomatal conductance showed a significant negative correlation with total leaves DW and above ground DW. This correlation was explained by conditional neutrality leading to antagonistic fitness effects in the drought environment, controlled by a QTL cluster on chromosome A7. Our results demonstrate that QxE is an important component of the genetic variance and can play a great role in improving drought tolerance in future breeding programs.

Key words: antagonistic fitness effect - *Brassica rapa* - drought - genotype by environment interaction – plasticity – root-shoot ratio - stomatal conductance

Introduction

Plant growth is greatly affected by environmental abiotic stresses from which drought is the most common factor impeding crop productivity. Drought is likely to become more threatening with the predicted global temperature increase (Smith and De Smet, 2012). Three categories of plant adaptive strategies to drought have been recognised: drought escape by early flowering, drought tolerance via increasing water use efficiency (WUE), and drought avoidance via reduced transpiration and increasing water uptake (Tuberosa, 2012; Franks, 2011; Edwards *et al.*, 2012; Juenger, 2013; Tardieu, 2013; Assmann, 2013).

Evaluating those responses in many genotypes in several environments may show phenotypic plasticity, which is defined as the ability of an individual organism to alter its physiology/morphology in response to changes in environmental conditions(Schlichting, 1986). When this plasticity differs between genotypes, so when there is genetic variation for it, it is classified as genotype by environment interaction (GxE) (Assmann, 2013; Weigel, 2012). Better understanding of GxE will provide a solid foundation for genetic improvement of stable crop productivity and will help to identify superior and stable alleles/genotypes across different environments (Zhang *et al.*, 2010). The genetic basis of the observed GxE can be identified by genetically dissecting plant physiological and morphological responses to environments via quantitative trait loci (QTL). This specifies the genetic component of GxE and is expressed as QTL by environment interaction (QxE). (Tardieu, 2013; Weigel, 2012; Mackay *et al.*, 2009; Boer *et al.*, 2007; Malosetti *et al.*, 2004). Different QTL effects can occur if the allele underlying the QTL is strongly expressed in one environment, but weakly in another or if the allele has opposite effects on the same trait in different environments (Mackay, 2001; Sukhwinder *et al.*, 2012). A QTL for which one allele has opposite (pleiotropic) effects on the phenotype in two different environments can lead to fitness trade-offs, elevating fitness in one environment but depressing it in the other environment. Trade-offs can be maintained in nature e.g. by antagonistic pleiotropy, when alleles at a locus underlying a fitness component show clear home-site advantages (Juenger, 2013; Anderson *et al.*, 2011; Leinonen *et al.*, 2013; Anderson *et al.*, 2013). Therefore, considering such antagonistic fitness effects is crucial while selecting for desirable QTL during marker assisted breeding (MAB) programs.

To facilitate improving MAB programs, a model crop plant is required. With the smallest genome size in the *Brassica* genus, the availability of complete genome sequence of *Brassica rapa* (Wang *et al.*, 2011), its close relation with the plant model species *Arabidopsis thaliana* and the genome analysis tools provided in the *Brassica* database (BRAD) (Cheng *et al.*, 2011), *B. rapa* is a useful dicot model crop for genetic and comparative studies.

In the present study we focused on drought avoidance, which enables plants to maintain a high fitness level in drought conditions. Therefore we investigated QxE on growth related traits in a *B. rapa* recombinant inbred line (RIL) population grown on a sandy substrate under control and drought environments. We identified several QTL for main effects and QxE. We found an antagonistic fitness effect for a stomatal conductance/shoot biomass QTL with the same allele reducing stomatal conductance under drought and increasing it under normal watering conditions, while contributing to higher shoot biomass in both environments.

Materials and Methods

Plant material and experimental setup

The recombinant inbred line (RIL) population (F7) used here was previously developed in our lab from a cross between a Yellow Sarson (R-o-18) (♂) and a Caixin type (L58) (♀), and genotyped with 270 markers (Bagheri *et al.*, 2012). The RIL population was screened three times under control (continuous watering for three weeks) and drought (normal watering for one week then plants were left to dry out) environments. In all screens, plants were grown in 13 cm deep square black plastic pots. Each pot was filled with 1.5 kg dried river sand and all pots were watered until saturation with 1100 ml nutrient solution (1, 1.1, 5.9 mmol / litre, N, P and K respectively). The same nutrient solution was used for watering plants every 2 days. Two seeds were sown per pot and four days after germination, seedlings were thinned to one per pot. Seven days after germination, watering was withheld as drought treatment, while the control treatment was continuously watered. Initially, a pilot experiment was performed using 30 randomly selected RILs and both parental lines, with three replications per genotype per environment, to test if the drought treatment would reveal significant differences between RILs and between the two environments regarding total leaves fresh and dry weight. Subsequently, a full RIL screening experiment was performed in which 140 RILs and both parents were phenotyped for the 20 studied traits (see "Plant phenotyping") under both environments with 3 replications per RIL and 6 replicates per parental line per treatment. Finally a QTL reproducibility experiment was performed, to confirm the different phenotypes

for contrasting alleles at four identified QTL by screening 27 RILs selected for their discriminating genotypes, with three replicates per RIL per environment.
All experiments were carried out under controlled greenhouse conditions, 16 hrs. light with average temperatures of 22.3°C and 20.3°C and average relative humidity of 77.8 and 81.3 % during day and night respectively. The experimental set-up involved a complete randomized block design with one plant per RIL and two replicates for each parent per block.

Plant phenotyping

In the full RIL screening and QTL reproducibility experiments, 20 traits were analysed under control and drought environments. We chose these traits as being the ones describing as best as possible the different aspects of plant performance in the experiments. Directly before harvesting, when less than 5% of plants had visible flower primordia, the number of leaves (LN) was counted. Chlorophyll content (CHL) was measured (only in the full RIL screening experiment) using a SPAD-502 chlorophyll meter (Minolta Co., Ltd. Japan). For this measurement the average of three leaves per plant per replication per treatment was taken. Leaf stomatal conductance (STC) was measured using a leaf porometer (Decagon Devices Inc., USA), for one fully expanded leaf per plant per replication (either the 3rd or 4th leaf). Thereafter, total leaves fresh weight (LFW) and dry weight (LDW) was measured and the dry weight of the 3rd and 4th (i.e. fully expanded) leaves (3,4DW). Dry weights were determined after drying plant materials at 65°C for four to five days until weight constancy. Leaf area (LA) of the 3rd and 4th leaves was measured using a Licor LI-3100 (Licor Inc., Lincoln, NE, USA) and subsequently their combined specific leaf area (SLA) was calculated as the LA divided by the 3,4DW, as well as the dry weight ratio (DWR) between 3,4DW and LDW. Hypocotyl length (HL) was measured using a ruler and hypocotyl DW (HDW) was determined. The shoot DW (SDW) was calculated as the sum of LDW and HDW. Subsequently, root systems were washed carefully to remove adhering sand, placed in a plastic tray filled with water, spread and scanned with a flatbed scanner. From this, the total root system length (RL), root volume (RV) and root diameter (RD) were measured using WinRhizo (Regent Instruments Inc., Quebec, Canada). This was used to calculate the RL to SDW ratio (RL/SDW), which illustrates the aboveground matter that is supported by a given root length. Thereafter roots were dried to measure root DW (RDW) and to calculate the root to shoot DW ratio (R/S). Similarly, to indicate the relative investment in shoots or roots, the shoot to total plant (shoot + root) DW ratio was calculated (S/SR), for which total plant DW

was calculated as the sum of SDW and RDW. Finally, the leaf water content (LWC) was calculated as (LFW - LDW)/LDW.

Statistical and quantitative trait loci analysis

Statistical analysis was performed on raw data of each experiment using GenStat for Windows, 15th Edition (GenStat; VSN International Ltd., Hemel Hempstead, UK). Analysis of variance (ANOVA) was used to test the significance difference between treatments, lines and for the interaction (GxE). Heritability was estimated as implemented in GenStat. In the linear mixed model, genotypes were fitted as random and blocks as fixed. The generalized heritability measure used, as described by Cullis *et al.* (2006), and in a more general context by Welham *et al.* (2010), is given by:

$$h^2 = 1 - \frac{\text{mean}(\text{pev}(g_i))}{\sigma_g^2}$$

where the set of predicted genotype means (Best Linear Unbiased Predictors, BLUPs) are $g_1 \ldots g_N$ with prediction error variance pev(g_i) and estimated genetics variance component σ_g^2.

Pearson correlations were calculated using GenStat.

Data from the 20 traits analysed in the full RIL screening experiment were used for QTL mapping using a multi-environment analysis (MEA) approach, which accounts for GxE, as implemented in the QTL library in GenStat. A step size of 10 cM, a minimum cofactor proximity of 50 cM, a minimum separation of selected QTL of 30 cM, and a threshold of -log10P = 2.8 were used for QTL analysis. Following the mixed model approach described by (Boer *et al.*, 2007; Malosetti *et al.*, 2004), first the whole genome was scanned using simple interval mapping and based on that cofactors were selected for two rounds of composite interval mapping. Thereafter, a final QTL model was selected using backward selection on the selected cofactors, where it estimated the allelic effect of each of QTL in each environment, the effect of QxE, and the explained phenotypic variance of each QTL per environment. In addition to determining phenotypic plasticity as QxE, a second method to determine plasticity QTL was used as described by (Tétard-Jones *et al.*, 2011), by QTL mapping the difference in the mean phenotypic values per line between treatments.

Confirming reproducibility of four QTL clusters

To confirm the reproducibility of the major QTL detected in the full RIL screening experiment, we selected four QTL clusters: on chr. 3 between 38-42 cM, on chr. 7 between

30-40 cM, on chr. 8 between 85-95 cM, and on chr. 9 between 70-84 cM. The whole population was genetically classified into 16 groups based on all possible allelic combinations at the four selected QTL. Thereafter, for every tested QTL, phenotypic data of RILs with contrasting genotypes for one QTL, but similar genotypes for the other QTL, were compared. E.g., to test for the QTL on chr.3, ANNN RILs were compared with BNNN RILs in paired groups, so AAAA with BAAA, ABAA with BBAA, ABBA with BBBA, etc. The 27 RILs with the highest and lowest average values at each tested QTL were selected and grown as described. For all measured traits, a correlation analysis between traits measured in the control environments and between traits measured in the drought environments of the full RIL screening experiment and the QTL reproducibility experiments was used to test for a significantly similar response to the treatment as a confirmation of the level of reproducibility.

Results

Phenotyping the RIL population

The results obtained from the pilot experiment (data not shown) indicated there was ample phenotypic variation for drought response, which justified phenotyping the whole RIL population. A total of 20 traits related to growth and performance of plants were analysed under control and drought environments. Figure 1 shows the frequency distributions of the measured traits over the whole population. Transgression beyond both parental lines values was observed for most of the traits except for root volume, root DW, hypocotyl DW, shoot DW, S/SR and leaf water content, where transgression was only in one direction. The drought treatment decreased fresh weight, leaf number, leaf area, total leaves DW, root length and stomatal conductance, and increased the root-shoot ratio, R/S (Fig. 1, Table 1). For stomatal conductance the reduction in the L58 parent was minor and not significant as was also the case for some of the RILs. Correlation analysis of all measured traits in this experiment was performed to unveil the genetic and physiological relationships of the various traits (Table 2). The correlations may exist because of similar physiological mechanisms or due to pleiotropy; however, correlations can also be caused by genetic linkage of loci affecting different traits, which are not physiologically related or pleiotropic. For instance, the analysis showed that in the control environment, chlorophyll content was positively correlated with root diameter, which is hard to envision being due to pleiotropy.

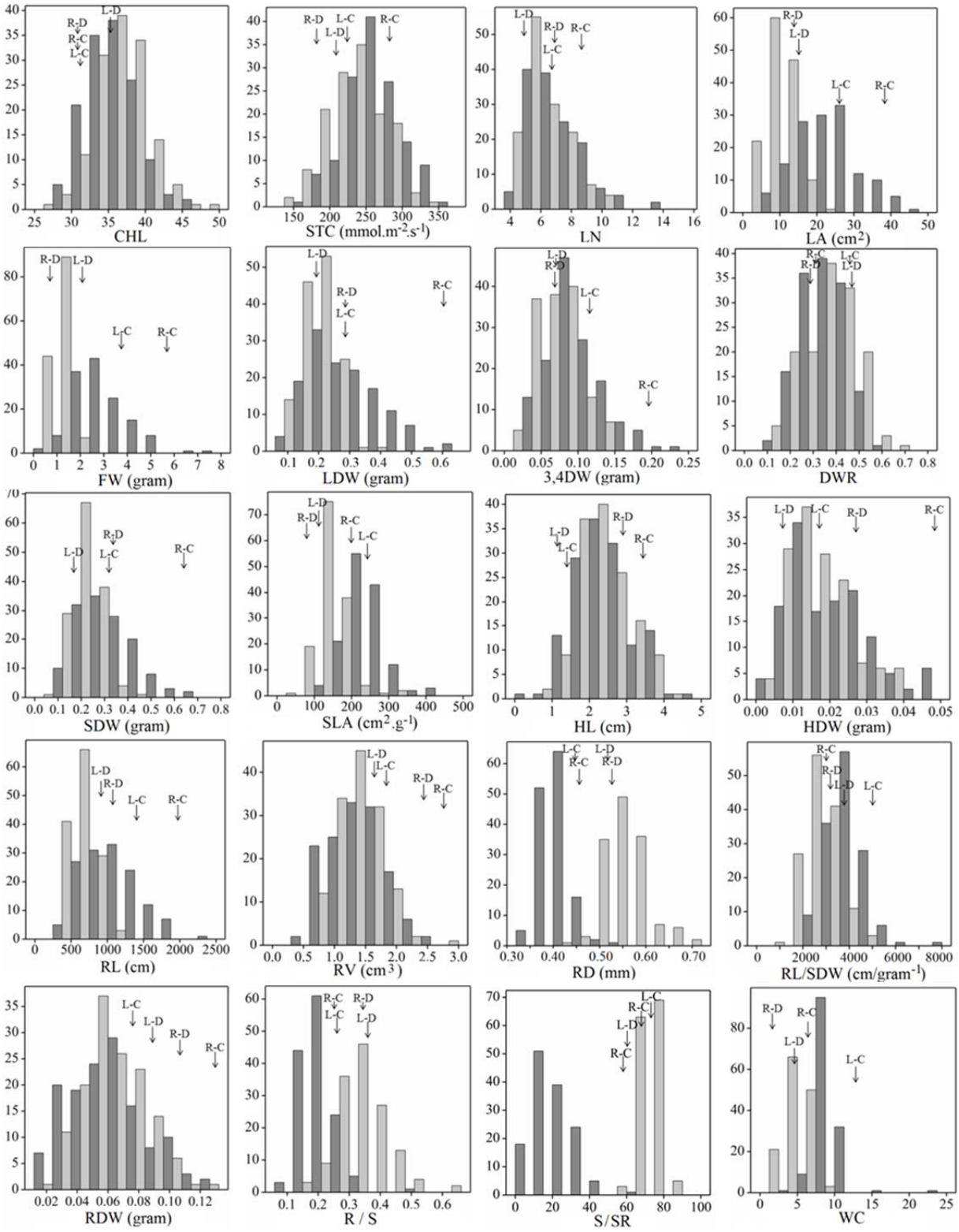


Figure 1: Frequency distributions of the non-normalized trait values for the L58 × R-o-18 RIL population under control (C) (dark grey bars), and drought (D) (light grey bars) conditions. The *vertical axes* indicate the number of lines per trait value class, and the horizontal axes indicate the different trait value classes (with units between brackets). The parental values are indicated with L (L58) and R (R-o-18). The same trait abbreviations are used as described in table 1.

Table 1: Parental line means for the analyzed traits and performance of the phenotyped RIL population under control (C) and drought (D) conditions. The following traits were measured: Chlorophyll content (CHL); Stomatal conductance (STC); Leaf Number (LN); Leaf area (LA); Leaf Fresh Weight (LFW); Leaf Dry Weight (LDW); 3rd and 4th leaf Dry Weight (3,4DW); 3,4DW vs LDW ratio (DWR), Shoot dry weight (SDW); Specific Leaf Area (SLA); Hypocotyl Length (HL); Hypocotyl Dry Weight (HDW); Root system Length (RL); Root Volume (RV); Root Diameter (RD); Root Length to Shoot Dry Weight ratio (RL/SDW); Root Dry Weight (RDW); Root to Shoot dry weight ratio (R/S); Shoot to plant dry weight ratio (S/SR); Leaf Water Content (WC). "Min" and "Max" indicate the values of the RIL with respectively the lowest and the highest value. "Mean" is the average value for all RILs, with standard deviation (SD) and broad sense heritability (h^2). In the ANOVA test, s = significant and ns = non-significant. For all traits three replicates were measured.

Trait	Unit	Parental lines		RIL population				ANOVA 2nd experiment			h^2	ANOVA 3rdnd experiment		
		L58	R-o-18	Min	Max	Mean	SD[b]	Treatment	RILs	GxE		Treatment	RILs	GxE
CHL–C		31.84	31.45	28.20	47.45	35.77	3.47	S	s	ns	0.59			
CHL–D		35.14	30.96	27.40	47.70	36.87	3.58				0.50			
STC–C	mmol/(m²·s)	237.87	266.12	173.95	364.25	265.50	37.59	S	s	s	0.31	s	ns	ns
STC–D		220.50	165.67	125.90	340.70	231.80	39.17				0.81			
LN–C		6.67	7.50	4.50	14.00	7.24	1.75	S	s	s	0.56	s	s	ns
LN–D		5.60	6.60	3.67	10.33	6.20	1.42				0.78			
LA–C	cm^2	26.38	35.79	6.33	47.91	23.73	8.47	S	s	ns	0.26	s	s	s
LA–D		11.87	10.04	1.28	22.24	9.23	3.83				0.46			
LFW–C	gram	3.68	5.11	0.772	7.458	2.994	1.136	S	ns	ns	0.44	s	s	s
LFW–D		1.21	0.77	0.268	2.046	0.979	0.324				0.36			
LDW–C	gram	0.276	0.619	0.087	0.637	0.293	0.114	S	s	ns	0.31	s	s	s
LDW–D		0.196	0.275	0.060	0.370	0.193	0.057				0.37			
3,4DW–C	gram	0.112	0.194	0.032	0.231	0.099	0.039	S	s	ns	0.48	s	s	s
3,4DW–D		0.082	0.081	0.017	0.150	0.069	0.029				0.57			
DWR–C		0.406	0.314	0.106	0.561	0.356	0.096	-	-	-	-	-	-	-
DWR–D		0.419	0.295	0.098	0.670	0.366	0.117				-			
SDW–C	gram	0.288	0.671	0.092	0.665	0.315	0.122	-	-	-	-	-	-	-
SDW–D		0.205	0.300	0.074	0.408	0.210	0.062				-			
SLA–C	$cm^2 \cdot g^{-1}$	235.03	184.48	139.00	426.35	246.20	52.43	-	-	-	-	-	-	-
SLA–D		144.45	123.95	35.42	325.75	137.60	42.89				-			

Table 1 (continued): Parental line means for the analyzed traits and performance of the phenotyped RIL population under control (C) and drought (D) conditions. The following traits were measured: Chlorophyll content (CHL); Stomatal conductance (STC); Leaf Number (LN); Leaf area (LA); Leaf Fresh Weight (LFW); Leaf Dry Weight (LDW); 3rd and 4th leaf Dry Weight (3,4DW); 3,4DW vs LDW ratio (DWR), Shoot dry weight (SDW); Specific Leaf Area (SLA); Hypocotyl Length (HL); Hypocotyl Dry Weight (HDW); Root system Length (RL); Root Volume (RV); Root Diameter (RD); Root Length to Shoot Dry Weight ratio (RL/SDW); Root Dry Weight (RDW); Root to Shoot dry weight ratio (R/S); Shoot to plant dry weight ratio (S/SR); Leaf Water Content (WC). "Min" and "Max" indicate the values of the RIL with respectively the lowest and the highest value. "Mean" is the average value for all RILs, with standard deviation (SD) and broad sense heritability (h^2). In the ANOVA test, s = significant and ns = non-significant. For all traits three replicates were measured.

Trait	Unit	Parental lines		RIL population				ANOVA 2nd experiment			h^2	ANOVA 3rdnd experiment		
HL–C	cm	1.60	3.40	0.500	4.550	2.439	0.738	s	s	ns	0.90	s	s	ns
HL–D		1.10	2.78	0.733	4.300	2.373	0.692				0.91			
HDW–C	gram	0.012	0.052	0.003	0.048	0.021	0.011	s	s	ns	0.33	s	s	s
HDW–D		0.009	0.025	0.004	0.038	0.017	0.008				0.60			
RL–C	cm	1398.50	1970.45	304.40	2300.00	1078.00	388.60	s	s	s	0.48	s	s	s
RL–D		741.80	940.92	254.79	1249.64	616.60	183.40				0.48			
RV–C	cm^3	1.71	2.98	0.442	2.696	1.381	0.455	ns	s	ns	0.38	ns	s	s
RV–D		1.52	2.21	0.697	2.990	1.379	0.361				0.46			
RD–C	mm	0.395	0.424	0.340	0.560	0.411	0.030	s	s	s	0.36	ns	s	s
RD–D		0.500	0.550	0.429	0.686	0.548	0.046				0.26			
RL/SDW–C	$cm.g^{-1}$	4849.20	2936.90	2088.00	7298.00	3545.00	797.40	-	-	-	-	-	-	-
RL/SDW–D		3616.80	3138.90	1502.00	5525.00	3049.00	750.30				-			
RDW–C	gram	0.070	0.157	0.013	0.122	0.060	0.025	s	s	ns	0.29	s	s	s
RDW–D		0.080	0.100	0.020	0.124	0.063	0.020				0.47			
R/S–C		0.242	0.233	0.097	0.410	0.192	0.047	-	-		-	-	-	-
R/S–D		0.369	0.339	0.151	0.557	0.308	0.071				-			
S/SR–C		77.13	74.75	2.04	60.30	21.53	10.84	-	-	-	-	-	-	-
S/SR–D		69.94	68.45	55.32	93.97	70.62	4.66				-			
WC–C		12.31	7.26	0.684	7.362	2.701	1.043	-	-	-	-	-	-	-
WC–D		5.16	1.80	0.110	1.765	0.786	0.300				-			

Table 2: Pearson correlations for the analysed traits of the L58 × R-o-18 RIL population under; control (A) and drought (B) conditions. The same trait abbreviations are used as in Table 1. Highlighted boxes refer to significant correlations at $p<0.01$ (dark grey) and $p<0.05$ (light grey) significance levels.

A

CHL	1	-																	
STC	2	-0.017	-																
LN	3	-0.075	0.115	-															
LA	4	-0.027	0.114	-0.235	-														
LFW	5	-0.010	0.188	0.448	0.571	-													
LDW	6	0.079	0.207	0.441	0.538	0.939	-												
3,4DW	7	0.096	0.129	-0.125	0.851	0.668	0.737	-											
DWR	8	0.009	-0.172	-0.776	0.366	-0.410	-0.412	0.265	-										
SDW	9	0.072	0.207	0.443	0.539	0.943	0.998	0.735	-0.413	-									
SLA	10	-0.260	-0.025	-0.161	0.168	-0.224	-0.404	-0.333	0.130	-0.399	-								
HL	11	-0.128	0.116	-0.092	0.116	0.150	0.102	0.135	0.015	0.147	-0.039	-							
HDW	12	-0.023	0.139	0.329	0.387	0.696	0.671	0.493	-0.296	0.719	-0.228	0.562	-						
RL	13	0.037	0.087	0.471	0.457	0.800	0.831	0.629	-0.339	0.830	-0.352	-0.028	0.569	-					
RV	14	0.100	0.106	0.460	0.464	0.770	0.797	0.619	-0.311	0.797	-0.338	-0.021	0.558	0.929	-				
RD	15	0.161	-0.037	-0.169	-0.181	-0.349	-0.355	-0.250	0.160	-0.357	0.134	-0.033	-0.272	-0.475	-0.190	-			
RL / SDW	16	-0.042	-0.215	0.018	-0.152	-0.259	-0.316	-0.205	0.193	-0.321	0.141	-0.312	-0.276	0.218	0.165	-0.175	-		
RDW	17	0.087	0.105	0.459	0.445	0.778	0.810	0.627	-0.321	0.818	-0.380	0.021	0.649	0.891	0.929	-0.245	0.057	-	
R / S	18	0.068	-0.086	0.140	-0.061	-0.053	-0.102	-0.035	0.082	-0.084	-0.008	-0.088	0.123	0.264	0.358	0.096	0.660	0.441	-
S / SR	19	0.084	0.161	0.410	0.483	0.830	0.905	0.684	-0.370	0.904	-0.393	0.041	0.623	0.820	0.790	-0.324	-0.190	0.819	0.057
WC	20	-0.224	-0.112	-0.078	-0.044	-0.039	-0.336	-0.289	0.114	-0.320	0.559	0.106	-0.066	-0.227	-0.228	0.086	0.299	-0.238	0.247
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Table 2 (continued): Pearson correlations for the analysed traits of the L58 × R-o-18 RIL population under; control (A) and drought (B) conditions. The same trait abbreviations are used as in Table 1. Highlighted boxes refer to significant correlations at $p<0.01$ (dark grey) and $p<0.05$ (light grey) significance levels.

B

CHL	1	-																	
STC	2	0.029	-																
LN	3	0.000	-0.075	-															
LA	4	-0.035	-0.107	-0.320	-														
LFW	5	0.013	-0.080	0.131	0.571	-													
LDW	6	-0.047	-0.189	0.404	0.473	0.494	-												
3,4DW	7	-0.098	-0.213	-0.226	0.773	0.455	0.637	-											
DWR	8	-0.051	-0.062	-0.685	0.557	0.121	-0.141	0.653	-										
SDW	9	-0.066	-0.187	0.405	0.463	0.473	0.995	0.618	-0.160	-									
SLA	10	0.072	0.141	-0.122	0.305	0.123	-0.119	-0.283	-0.218	-0.103	-								
HL	11	-0.141	0.156	-0.090	-0.027	0.041	0.022	-0.013	-0.041	0.083	-0.007	-							
HDW	12	-0.174	-0.107	0.278	0.248	0.175	0.635	0.282	-0.236	0.711	0.040	0.481	-						
RL	13	0.005	-0.169	0.198	0.404	0.130	0.646	0.429	-0.053	0.648	0.032	-0.156	0.453	-					
RV	14	-0.010	-0.280	0.288	0.392	0.273	0.664	0.460	-0.034	0.662	-0.037	-0.099	0.429	0.828	-				
RD	15	-0.042	-0.093	0.063	-0.261	0.061	-0.204	-0.180	-0.061	-0.214	-0.146	0.116	-0.219	-0.556	-0.084	-			
RL / SDW	16	0.067	0.045	-0.289	-0.103	-0.429	-0.452	-0.226	0.175	-0.452	0.147	-0.262	-0.303	0.340	0.116	-0.434	-		
RDW	17	-0.047	-0.181	0.325	0.370	0.224	0.708	0.423	-0.116	0.725	-0.017	-0.029	0.607	0.813	0.845	-0.240	0.029	-	
R / S	18	-0.025	0.001	-0.110	-0.133	-0.326	-0.317	-0.211	0.063	-0.286	0.060	-0.037	0.019	0.217	0.235	-0.057	0.622	0.404	-
S / SR	19	0.080	0.010	0.102	0.122	0.328	0.256	0.212	-0.002	0.196	-0.088	-0.190	-0.280	-0.227	-0.227	0.094	-0.530	-0.420	-0.940	-
WC	20	0.043	0.126	-0.314	0.085	0.449	-0.505	-0.158	0.302	-0.517	0.240	0.047	-0.432	-0.521	-0.389	0.279	0.055	-0.493	0.010	0.043
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19

The correlation observed between root length and S/SR was positive in the control environment (longer roots contributing to relatively more shoots) but negative in the drought environment, indicating a proportionally higher investment in roots. Under both environments, leaf water content was negatively correlated with shoot DW, root length, root volume and root dry weight, while it was positively correlated with leaves FW and negatively correlated with stomatal conductance under drought conditions. Stomatal conductance was negatively correlated with leaves DW in the drought environment, but positively correlated under control conditions. In general, plants with a longer root system had higher plant DW.
As expected, all traits measured in control and drought environments showed a positive correlation, except for leaf water content.

Mapping QTL with main effects and QxE

In total 54 QTL were mapped for the traits analysed under control and drought environments (Table 3, Fig. 2). Six QTL, *STC1, LA4, SLA1, RD3, RL/SDW3,* and *S/SR2*, had opposite allelic effects when comparing both environments. The phenotypic effects of three QTL, *LA1, SLA1* and *S/SR2*, were respectively 9, 101 and 15 times higher in one environment than the other (Table 3). *SLA1* co-located with *3,4DW1* with the alleles increasing the trait values in the control environment were from L58 and R-o-18 respectively. Four QTL were mapped for chlorophyll content of which *CHL1, CHL2,* and *CHL3* showed the highest effect from the L58 allele, while for *CHL4* the R-o-18 allele had the highest effect in both environments. Hypocotyl length was mapped to four loci with the R-o-18 alleles contributing most to increased hypocotyl length. In total 11 QTL clusters were observed of which seven comprised at least three co-locating QTL (Table 3, Fig. 2).

Stomatal conductance QTL (*STC1*) and fitness trade-offs in the drought environment

The correlation analysis showed that in the control environment, stomatal conductance was positively correlated with leaves and shoot DW. On the other hand, in the drought environment, stomatal conductance showed a significant negative correlation with leaves DW and shoot DW. These correlations were associated with altering the trait value enhancing allele for *STC1* from R-o-18 in the control environment to L58 in the drought environment. The trait enhancing alleles for the QTL co-locating with *STC1* (*LFW2*, *LDW2*, *3,4DW2*) were R-o-18 in both environments (Table 3, Fig. 2). This means that the R-o-18 alleles for these

loci were enhancing fitness both under control and drought conditions, although having contrasting phenotypic effects on stomatal conductance when comparing both conditions.

Table 3: QTL detected in the L58 × R-o-18 RIL population for the traits described in Table 1, using the multi environment analysis (MEA) approach. The same trait abbreviations are used as in Table 1. Per trait, QTL are numbered according to chromosome number. R^2 is the percentage of total phenotypic variance explained by each QTL. Effects with positive values represent a positive contribution of the R-o-18 allele to the trait value and those with negative values represent a positive contribution of the L58 allele to the trait value. Highlighted QTL (grey) show significant QxE effects and "Ratio" refers to the ratios between the effects of each QTL in both environments.

Trait	QTL				Control environment		Drought environment		Ratio
	QTL name	Linkage group	Position of highest peak	-log10P	Effect	R^2	Effect	R^2	
CHL	*CHL1*	A1	24.23	4.2	-0.855	5.7	-0.855	5.4	
	CHL2	A6	59.11	4.5	-0.885	5.9	-0.885	5.6	
	CHL3	A9	77.21	8.9	-1.304	14.2	-1.304	13.4	
	CHL4	A10	56.32	5.9	1.011	10.4	1.011	9.7	
STC	*STC1*	A7	96.75	3.0	6.395	1.9	-9.754	7.9	-1.5
LN	*LN1*	A7	40.75	11.1	0.799	15.6	11.103	23.5	
	LN2	A10	62.99	4.6	0.426	5.4	4.650	8.1	
LA	*LA1*	A1	70.54	6.0	-3.343	15.6	-0.358	0.9	9.3
	LA2	A7	32.05	4.8	-1.251	2.2	-1.251	10.7	
	LA3	A8	85.20	5.2	1.321	2.4	1.321	11.9	
	LA4	A9	24.28	3.4	1.865	4.8	-0.394	1.1	-4.7
LFW	*LFW1*	A3	42.66	3.2	0.313	7.6	0.062	3.7	5.0
	LFW2	A7	105.32	3.6	0.092	0.7	0.092	8.1	
	LFW3	A8	85.20	6.1	0.120	1.1	0.120	13.7	
LDW	*LDW1*	A3	42.66	3.5	0.036	9.9	0.010	3.0	3.7
	LDW2	A7	100.81	3.9	0.016	2.1	0.016	8.4	
	LDW3	A8	91.33	7.3	0.023	4.0	0.023	16.5	
3,4DW	*3,4DW1*	A5	69.71	2.8	0.011	8.7	0.003	1.3	3.4
	3,4DW2	A7	34.89	7.4	-0.011	8.5	-0.011	14.9	
	3,4DW3	A7	100.81	3.0	0.007	3.1	0.007	5.5	
DWR	*DWR1*	A3	120.61	3.9	0.024	6.1	0.024	4.2	
	DWR2	A4	75.90	2.9	0.020	4.4	0.020	3.0	
	DWR3	A7	40.75	18.3	-0.054	31.1	-0.076	42.8	1.4
SDW	*SDW1*	A3	42.66	3.2	0.038	9.7	0.011	2.9	3.6
	SDW2	A7	83.27	3.0	0.016	1.8	0.016	7.0	
	SDW3	A8	91.33	7.0	0.025	4.2	0.025	16.3	
SLA	*SLA1*	A5	60.82	4.6	-20.3	15.0	0.201	0.0	-101.4

HL	*HL1*	A3	94.58	6.6	0.245	11.0	0.245	12.5	
	HL2	A4	54.79	3.7	0.176	5.7	0.176	6.5	
	HL3	A6	101.21	6.6	0.265	12.9	0.265	14.7	
	HL4	A7	18.39	4.7	0.200	7.3	0.200	8.3	
HDW	*HDW1*	A3	38.29	4.9	0.003	4.9	0.003	8.8	
	HDW2	A6	62.85	4.0	0.002	3.8	0.002	6.8	
	HDW3	A7	3.99	6.8	0.004	14.0	0.002	5.9	2.0
	HDW4	A8	33.95	4.1	0.003	6.3	0.003	11.3	
RL	*RL1*	A5	69.71	4.1	126.3	10.6	30.215	2.7	4.2
	RL2	A8	21.23	3.3	120.1	9.6	32.583	3.2	3.7
	RL3	A8	86.57	5.0	62.1	2.6	62.124	11.5	
RV	*RV1*	A5	69.71	3.2	0.141	9.6	0.056	2.4	2.5
	RV2	A8	86.57	5.7	0.124	7.4	0.124	11.8	
RD	*RD1*	A3	5.92	5.6	-0.010	10.8	-0.010	4.6	
	RD2	A5	35.17	4.8	-0.008	7.8	-0.008	3.3	
	RD3	A6	48.53	2.6	-0.005	2.8	0.010	5.0	-2.0
	RD4	A8	95.50	4.3	-0.007	6.0	-0.007	2.6	
RL/SDW	*RL/SDW1*	A3	21.88	2.3	-133.0	2.8	-133.0	3.1	
	RL/SDW2	A7	18.39	6.2	-236.8	8.8	-236.8	10.0	
	RL/SDW3	A7	125.27	2.7	160.8	4.1	-142.8	3.6	-1.1
	RL/SDW4	A10	62.99	4.0	182.5	5.2	182.5	5.9	
RDW	*RDW1*	A5	69.71	3.1	0.005	4.4	0.005	6.7	
	RDW2	A8	86.57	4.9	0.006	6.7	0.006	10.3	
R/S	*R/S1*	A7	18.39	3.6	-0.014	8.2	-0.014	3.7	
	R/S2	A9	84.14	3.0	0.004	0.7	-0.021	8.5	5.1
S/SR	*S/SR1*	A4	90.12	3.2	-0.010	6.3	-0.010	4.7	
	S/SR2	A9	69.95	3.4	0.001	0.0	-0.013	8.2	-14.9

Mapping QTL underlying plasticity

Seventeen of the mapped QTL showed a significant QxE effect (table 3) indicating the loci contributing to phenotypic plasticity between both environments. In addition to the GenStat method to determine these plasticity loci, an alternative method to describe QTL that are affected by the environments has been suggested by Tétard-Jones et *al.* (2011). This uses the differences between the trait value averages of the lines in the two environments to determine QTL. Using this procedure 15 plasticity QTL were mapped (Table 4), with nine of them co-locating with previously mapped QTL, six of which were found to show QxE (Table 3). Thus, this analysis detected six new plasticity QTL, which did not exceed the statistical significance levels with the GenStat method.

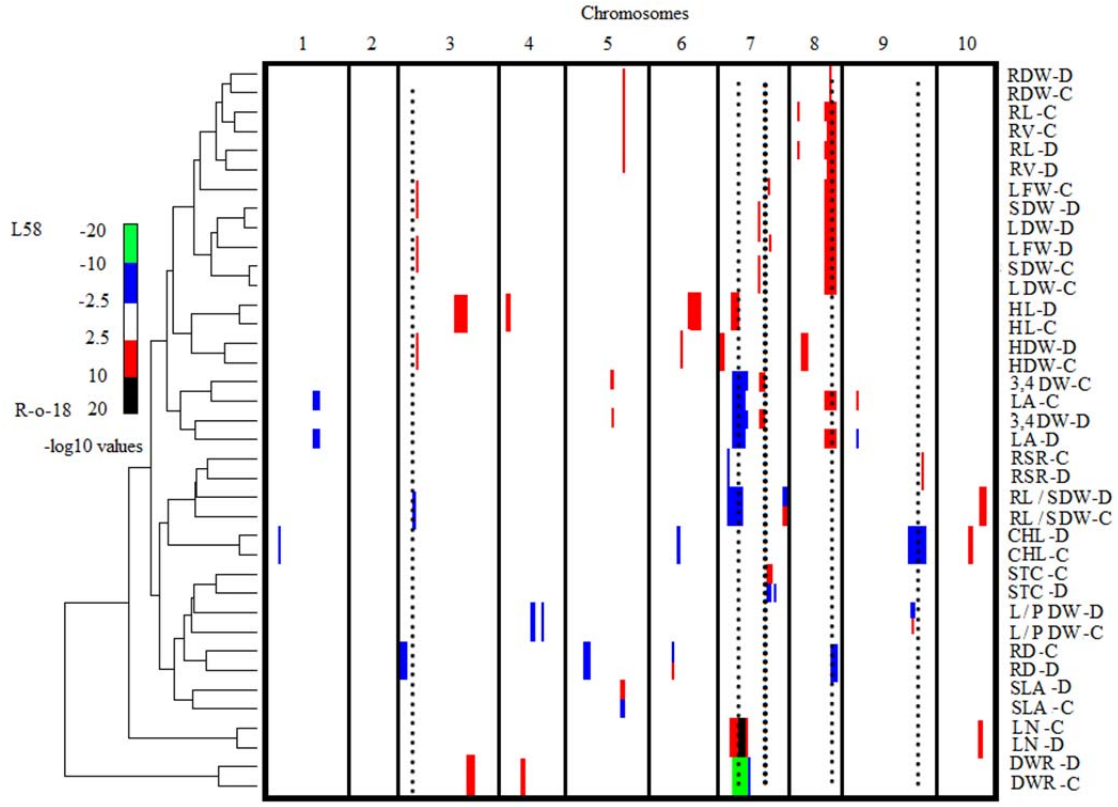


Figure 2: A clustered heat map showing the -log10P profiles of the measured traits. Columns indicate the 10 chromosomes in centiMorgans, ascending from the left to right, rows indicate individual trait -log10P profiles. A colour scale is used to indicate the QTL significance corresponding to the -log10P score. Red and black represent a positive effect on the trait value from the R-o-18 allele, whereas blue and green represent a positive effect on the trait value from the L58 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. The same trait abbreviations are used as described in table 1. C and D refer to control and drought environments respectively. Black dotted lines indicate the QTL with an antagonistic fitness effect at the bottom of A7, and the other four lines refer to the QTL confirmed in the reproducibility experiment.

Table 4: QTL mapped for phenotypic plasticity in the L58 x R-o-18 RIL population. Plasticity was calculated as described by (Tétard-Jones *et al.*, 2011) as the difference in the mean phenotype between different treatments per trait. The same trait abbreviations are used as in Table 1. QTLs are numbered according to chromosome (Chr) number. Effects with positive values represent a positive contribution of the R-o-18 allele to the trait value and those with negative values represent a positive contribution of the L58 allele to the trait value. R^2 is the percentage of total plastic variance explained by each QTL. Highlighted QTL were mapped before using the MEA approach (Table 3).

Trait	QTL name	Locus	Chr.	Position	Effect	-log10P	R^2
STC	*STC1*	903607|9917837	7	96.8	17.22	3.8	11.0
LA	*LA1*	E3835M11	1	69.0	-2.69	4.6	13.3
LDW	*LDW4*	E3850M9	5	69.7	0.03	3.2	9.2
3,4DW	*3,4DW3*	Ra2A01-A7	7	83.3	-0.01	2.9	8.3
SDW	*SDW4*	E3850M9	5	69.7	0.03	3.0	8.8
SLA	*SLA2*	E3749M6	1	94.1	17.26	3.0	7.4
	SLA1	BrID101239-A5	5	65.7	-22.36	4.0	12.4
HL	*HL5*	902225|9924661	8	95.5	-0.08	3.0	8.0
RL	*RL4*	E3732M5	1	92.1	-101.68	3.3	7.9
	RL1	E3850M9	5	69.7	96.41	2.9	7.1
	RL3	E3416M22	8	91.3	112.99	4.0	9.8
RV	*RV3*	E3732M5	1	92.1	-0.12	3.0	8.1
RD	*RD3*	899015|9918455	6	43.5	-0.02	3.6	9.7
RL/SDW	*RL/SDW3*	C7P119	7	119.0	304.08	3.4	11.8
R/S	*R/S2*	BrID10177-A9	9	68.3	-0.02	2.7	7.0

Reproducibility

From the 11 QTL clusters that were mapped, four were selected to be tested for reproducibility in a subsequent experiment. The first cluster mapped to A3, including the *RD1* and *RL/SDW1* QTL for both positively correlated traits, with trait value enhancing effects from the L58 alleles. Moreover, *LFW1*, *LDW1* and *SDW1*, all contributing to shoot biomass, were mapped to the same cluster, with positive alleles coming from R-o-18. The second cluster was mapped to A7, composed of *LN1, LA2, 3,4DW2, DWR3, R/S1,* and *RL/SDW2*, all with a positive contribution of the R-o-18 allele except for *LN1*. This is in line with the negative correlation of leaf number with the other traits. The third cluster, on A8, included eight co-locating QTL, *LFW3, LDW3, SDW3, LA3, RL3, RVl2, RD4* and *RDW2,* of which the *RD4* L58 allele increased the trait value, while for the other QTL, the R-o-18 allele increased the trait value, in line with the negative correlation of RD with the other traits. The fourth cluster included three QTL, *CHL3, R/S2* and *S/SR2*, mapping to A9. The S/SR ratio showed a negative correlation between control and drought environments and therefore the trait value enhancing effect of *S/SR2* in the drought environment came from the L58 allele, whereas in the control environment it came from the R-o-18 allele. In total 27 lines were selected from the RIL population, to properly represent the 16 possible genotypes for all allelic combinations for the four selected QTL clusters. These lines were regrown under similar

conditions, and rephenotyped (supplementary table 1). A correlation analysis (Table 5) between traits measured in the two control environments and between traits measured in the two drought environments of the full RIL screening and QTL reproducibility experiments showed that all traits were positively correlated in at least one environment, but often both, except for fresh weight, leaf water content and root diameter. This indicates that the phenotyping was robust and the detected QTL clusters are reproducible, making them attractive candidates for further gene cloning experiments.

Table 5: Correlation analysis between the control conditions and the drought conditions of the full RIL screening experiment and the reproducibility experiment. The same trait abbreviations are used as in Table 1. Dark grey means significant at the 0.01 level, and light grey means significant at the 0.05 level. The same trait abbreviations are used as described in Table 1.

Trait	Control	Drought
STC	.260	.494
LN	.852	.758
LA	.504	.220
LFW	.122	-.027
LDW	.270	.258
3,4DW	.474	.239
DWR	.296	.128
SDW	.266	.259
SLA	.496	.159
HL	.850	.847
HDW	.211	.324
RL	.209	.085
RV	-.002	.292
RD	.071	-.107
RL/SDW	.095	.378
RDW	.073	.251
R/S	.310	.352
S/SR	.397	.476
WC	.105	.088

Discussion

The current study was carried out in a greenhouse using pots filled with sand. This type of pot experiments is a reasonable compromise to avoid the difficulty of phenotyping roots in natural

field environments and the unnatural conditions present in hydroponics, aeroponics or agar plates (Tuberosa, 2012). However, aspects of root growth in our pot system will still be substantially different from field conditions.

Upon screening the RIL population, we found significant GxE between control and drought environments for stomatal conductance, leaf number, root length and root diameter. This GxE was reflected in QxE detected using the MEA approach for these traits, except for leaf number. MEA is more powerful than the traditional single environment analysis in detecting more significant QTL with higher explained variance. Additional advantage is that it allows quantification of QxE, because it accounts for GxE and tests all detected QTL in all environments and thus shows their effects in each environment (Crossa and Federer 2012). QxE occurs if the QTL effects are strongly expressed in one environment but weakly in another, or if the QTL has opposite effects on the same trait in two different environments (Zhang *et al.*, 2010; Mackay, 2001; MacMillan *et al.*, 2006). Examples of the first case are *LFW1, LA1, LDW1, 3,4DW1, RL1, RL2, RV1, R/S2, HDW3* and *SDW1*, while examples of the latter case are found for *LA4, SLA1, RD3* and *S/SR2*. The latter kind of QxE obstructs the transferability of QTL mapping results from one environment to another (Mackay, 2001), as selection will be in opposite directions in the two environments. Knowing about the QTL with opposite effects on several traits in different environments, also known as antagonistic pleiotropy, is of great importance in breeding programs because breeding for one trait might negatively affect other traits (Juenger, 2013; Anderson *et al.*, 2011; Leinonen *et al.*, 2013; Anderson *et al.*, 2013). The QTL cluster mapped at the bottom of A7 included a stomatal conductance QTL (*STC1*), shows signs of antagonistic pleiotropy, with the R-o-18 allele increasing stomatal conductance under control conditions, but decreasing it under drought conditions, while having positive effects on biomass under both environments through the co-located *LDW3*, *SDW2* and *3,4DW3* QTL. However, the similar effect on biomass and the contrasting effect on stomatal conductance could also mean these traits are not allelic, but the result of close linkage of two loci. Further analysis should reveal this. Stomatal conductance showed clear plasticity, decreasing significantly in the drought environment. Such response is generally correlated with reduced photosynthesis, but also with reduced water loss as an adaptive response to drought (Chaves et al., 2003; Condon et al., 2004; Aroca et al., 2012; Tardieu, 2013). Due to the co-location or antagonistic pleiotropy of the shoot biomass QTL with *STC1*, when comparing both environments, stomatal conductance was negatively correlated with shoot biomass (Table 2B). This reflects an interesting fitness advantage for plants carrying the R-o-18 allele at this QTL cluster, meaning that under drought conditions,

they show a relatively reduced stomatal conductance (contributing to increased drought tolerance) accompanied with relatively increased shoot biomass, compared to plants carrying the L58 allele. Recently, the plasticity and the evolution of flowering time and WUE has been investigated in *B. rapa* under drought environments (Franks, 2011), and the relationship between circadian rhythm, vegetative and reproductive traits, and leaf gas-exchange with the variation of WUE in different watering regimes has been investigated (Edwards *et al.*, 2012). The negative correlation we found for stomatal conductance and shoot biomass under drought, was also observed by Edwards *et al.* (2012) although this was not significant in their study. It also agrees with the positive correlation between WUE and biomass in the drought environment found by these authors and the co-location of WUE and stomatal conductance QTL mapped in *B. rapa* grown under warm and long day conditions (Edwards *et al.*, 2011). In our study, stomatal conductance in the control environment was not correlated with shoot biomass traits, indicating conditional neutrality of *STC1* in the well-watered environment. Although the preferred targets for crop improvement in MAB are generally constitutively expressed QTL (Bernardo, 2008), this QTL cluster is attractive to select for, even if it is not constitutive in view of the QxE observed for *STC1*, as the allele from R-o-18 contributes to increased drought tolerance without having fitness costs due to reducing biomass.

The leaf area response and the underlying QTL in both environments were confirmed by the positive correlation observed between the full RIL screening and QTL reproducibility experiments. Stomatal closure and limited expansion of young leaves under drought have an indirect negative effect on root growth (Chaves *et al.*, 2003; Roycewicz and Malamy, 2012). This was observed by the reduction in root length, concomitant with an increase in root diameter in the drought treatment, corresponding to similar observations reported before for Brassica and other crops (Edwards *et al.*, 2012; Zhu *et al.*, 2011; Poorter *et al.*, 2012). It thus looks like it that under drought stress in pots, *B. rapa* does not invest in longer roots to take up more water, but in thicker roots to act as a water storage buffer.

Under drought, the root-shoot biomass ratio increased compared to the well-watered conditions in our experiments. Biomass allocation under limiting environments can be explained by a functional biomass equilibrium when plants allocate more biomass to roots when the factor limiting growth is below ground, e.g. water or nutrient shortage, to enhance the uptake of that limiting factor (Poorter *et al.*, 2012). The correlation of the root-shoot ratio with drought tolerance has previously also been documented for Arabidopsis and tobacco (Werner *et al.*, 2010), as well as *B. rapa* (Edwards *et al.*, 2012; Kage *et al.*, 2004).

Of the traits we examined, GxE was found for most of them, either in the full RIL screening or the reproducibility experiment (Table 1). With so many traits for which we found GxE, it is not surprising we also found QTL with QxE indicating plasticity for many traits. We used two ways to detect QTL related to phenotypic plasticity, first using the MEA approach (table 3) and subsequently using the difference between average values per line when comparing both treatments per lines (table 4). As previously found by Tétard-Jones *et al.* (2011), there is considerable overlap between both methods, but the latter method also detects some novel QTL, not found previously. This is probably due to the additional statistical power that can be gained by directly using the phenotypic difference values for mapping, meaning QTL will be detected that did not exceed the threshold in the MEA approach.

Although almost all traits showed a positive correlation between the results from the full RIL screening experiment and the reproducibility experiment, confirming the initial results we obtained, this was not the case for leaf fresh weight and root diameter, suggesting a high level of GxE for those traits, or for water content, where a high environmental effect probably prevented mapping a QTL for this trait. The only trait for which no plasticity QTL was found was chlorophyll content (CHL), which is in line with the inability to detect GxE for this trait (Table 1). There is genetic variation for CHL though, with four detected QTL (Table 3). There is also a difference in CHL between drought and control conditions, which agrees with previous observations for four Brassica species (Ashraf and Mehmood, 1990), but genotypes appear to respond similarly to drought exposure in decreasing CHL, explaining the lack of GxE. The four *CHL* QTL mapped to regions previously identified to contain QTL for chlorophyll a and b content in *B. rapa* (Ge *et al.*, 2012), with *CHL1* co-locating with one of the three QTL previously identified for chlorophyll fluorescence (Edwards *et al.*, 2011).

Increasing crop productivity under drought conditions is the ultimate goal for MAB programs. In that respect, the significant antagonistic effect of relatively reduced stomatal conductance along with relatively higher shoot biomass under drought conditions due to the STC1/shoot biomass locus at the bottom of chromosome A7, is very interesting, as it suggests that selection on reduced water loss during drought, through reduced stomatal transpiration, is expected to have disproportionally little effect on shoot biomass reduction, which is a favourable combination. In addition, we reported many QTL underlying several morphological and physiological traits, which appeared to be robust and thus provided the first step towards identifying genes governing those traits. The availability of the whole *B. rapa* genome sequence (Wang *et al.*, 2011) together with possible comparative alignment with the related model species *A. thaliana* (Schranz *et al.*, 2006), will facilitate fine mapping

and cloning of candidate genes underlying the desired QTL . This approach will not only be useful in breeding *B. rapa*, but also in breeding other closely related species like *B. juncea* and *B. napus* (Cheng *et al.*, 2011; Li *et al.*, 2013).

Acknowledgements

The Dutch graduate school Experimental Plant Sciences is acknowledged for financially supporting this research. We thank Nihal Erol-Öztolan for her help with sowing and harvesting, and Unifarm for taking care of our plants in the greenhouse.

References

Anderson JT, Lee C-R, Rushworth CA, Colautti RI, Mitchell-Olds T. 2013. Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology* **22**, 699-708.

Anderson JT, Willis JH, Mitchell-Olds T. 2011. Evolutionary genetics of plant adaptation. *Trends in Genetics* **27**, 258-266.

Aroca R, Bacelar EVA, Moutinho-Pereira J, Gonçalves BC, Brito CQ, Gomes-Laranjo J, Ferreira HF, Correia C. 2012. Water use strategies of plants under drought conditions. *Plant Responses to Drought Stress*: Springer Berlin Heidelberg, 145-170.

Ashraf M, Mehmood S. 1990. Response of four Brassica species to drought stress. *Environmental and Experimental Botany* **30**, 93-100.

Assmann SM. 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.

Bagheri H, El-Soda M, van Oorschot I, *et al.* 2012. Genetic analysis of morphological traits in a new, versatile, rapid-cycling Brassica rapa recombinant inbred line population. *Frontiers in Plant Science* **3:183. 10.3389/fpls.2012.00183**.

Bernardo R. 2008. Molecular markers and selection for complex traits in plants: learning from the last 20 years *Crop Science* **48**, 1649-1664.

Boer MP, Wright D, Feng L, Podlich DW, Luo L, Cooper M, van Eeuwijk FA. 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.

Chaves MM, Maroco JP, Pereira JS. 2003. Understanding plant responses to drought-from genes to the whole plant. *Functional Plant Biology* **30**, 239-264.

Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, Li P, Hua W, Wang X. 2011. BRAD, the genetics and genomics database for Brassica plants. *BMC Plant Biology* **11**, 136.

Condon AG, Richards RA, Rebetzke GJ, Farquhar GD. 2004. Breeding for high water-use efficiency. *Journal of Experimental Botany* **55**, 2447-2460.

Cullis BR, Smith AB, Coombes NE. 2006. On the design of early generation variety trials with correlated data. *Journal of Agricultural Biological and Environmental Statistics* **11**, 381-393.

Crossa J, Federer WT. 2012. Screening experimental designs for quantitative trait loci, association mapping, genotype-by environment interaction, and other investigations. *Frontiers in Physiology* **3:156. 10.3389/fphys.2012.00156**.

Edwards CE, Ewers BE, McClung CR, Lou P, Weinig C. 2012. Quantitative variation in water-use efficiency across water regimes and its relationship with circadian, vegetative, reproductive, and leaf gas-exchange traits. *Molecular Plant* **5**, 653-668.

Edwards CE, Ewers BE, Williams DG, Xie Q, Lou P, Xu X, McClung CR, Weinig C. 2011. The genetic architecture of ecophysiological and circadian traits in Brassica rapa. *Genetics* **189**, 375-390.

Franks SJ. 2011. Plasticity and evolution in drought avoidance and escape in the annual plant Brassica rapa. *New Phytologist* **190**, 249-257.

Ge Y, Wang T, Wang N, Wang Z, Liang C, Ramchiary N, Choi S-R, Lim YP, Piao ZY. 2012. Genetic mapping and localization of quantitative trait loci for chlorophyll content in Chinese cabbage (Brassica rapa ssp. pekinensis). *Scientia Horticulturae* **147**, 42-48.

Juenger TE. 2013. Natural variation and genetic constraints on drought tolerance. *Curr Opin Plant Biol* **16,** 274-281
Kage H, Kochler M, Stützel H. 2004. Root growth and dry matter partitioning of cauliflower under drought stress conditions: measurement and simulation. *European Journal of Agronomy* **20**, 379-394.
Leinonen PH, Remington DL, Leppälä J, Savolainen O. 2013. Genetic basis of local adaptation and flowering time variation in Arabidopsis lyrata. *Molecular Ecology* **22**, 709-723.
Li X, Ramchiary N, Dhandapani V, Choi SR, Hur Y, Nou I-S, Yoon MK, Lim YP. 2013. Quantitative trait loci mapping in Brassica rapa revealed the structural and functional conservation of genetic loci governing morphological and yield component traits in the A, B, and C subgenomes of Brassica species. *DNA Research* **20**, 1-16.
Mackay TFC. 2001. The genetic architecture of quantitative traits *Annual Review of Genetics* **35**, 303-339.
Mackay TFC, Stone EA, Ayroles JF. 2009. The genetics of quantitative traits: challenges and prospects. *Nature Reviews Genetics* **10**, 565-577.
MacMillan K, Emrich K, Piepho HP, Mullins C, Price A. 2006. Assessing the importance of genotype × environment interaction for root traits in rice using a mapping population II: conventional QTL analysis. *Theoretical and Applied Genetics* **113**, 953-964.
Malosetti M, Voltas J, Romagosa I, Ullrich SE, van Eeuwijk FA. 2004. Mixed models including environmental covariables for studying QTL by environment interaction. *Euphytica* **137**, 139-145.
Poorter H, Niklas KJ, Reich PB, Oleksyn J, Poot P, Mommer L. 2012. Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control. *New Phytologist* **193**, 30-50.
Roycewicz P, Malamy JE. 2012. Dissecting the effects of nitrate, sucrose and osmotic potential on Arabidopsis root and shoot system growth in laboratory assays. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 1489-1500.
Schlichting CD. 1986. The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics* **17**, 667-693.
Schranz ME, Lysak MA, Mitchell-Olds T. 2006. The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends in Plant Science* **11**, 535-542.
Smith S, De Smet I. 2012. Root system architecture: insights from Arabidopsis and cereal crops. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 1441-1452.
Sukhwinder S, Hernandez MV, Crossa J, Singh PK, Bains NS, Singh K, Sharma I. 2012. Multi-trait and multi-environment QTL analyses for resistance to wheat diseases. *PLoS ONE* **7**, 1-12.
Tardieu F. 2013. Plant response to environmental conditions: assessing potential production, water demand and negative effects of water deficit. *Frontiers in Physiology* **4:17. 10.3389/fphys.2013.00017**.
Tétard-Jones C, Kertesz MA, Preziosi RF. 2011. Quantitative trait loci mapping of phenotypic plasticity and genotype–environment interactions in plant and insect performance. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 1368-1379.
Tuberosa R. 2012. Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology* **3:347. 10.3389/fphys.2012.00347**.
Wang X, Wang H, Wang J, *et al.* 2011. The genome of the mesopolyploid crop species Brassica rapa. *Nature Genetics* **43**, 1035-1039.
Weigel D. 2012. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiol* **158**, 2-22.
Welham SJ, Gogel BJ, Smith AB, Thompson R, Cullis BR. 2010. A comparison of analysis methods for late-stage variety evaluation trials. *Australian & New Zealand Journal of Statistics* **52**, 125-149.
Werner T, Nehnevajova E, Köllmer I, Novák O, Strnad M, Krämer U, Schmülling T. 2010. Root-specific reduction of cytokinin causes enhanced root growth, drought tolerance, and leaf mineral enrichment in Arabidopsis and Tobacco. *The Plant Cell* **22**, 3905-3920.
Zhang Y, Li Y-x, Wang Y, *et al.* 2010. Stability of QTL across environments and QTL-by-environment interactions for plant and ear height in Maize. *Agricultural Sciences in China* **9**, 1400-1412.
Zhu JM, Ingram PA, Benfey PN, Elich T. 2011. From lab to field, new approaches to phenotyping root system architecture. *Current Opinion Plant Biology* **14**, 310-317.

Supplementary table 1: Phenotypic data of the selected lines for the analysed traits in two experiments, the full RIL screening experiment (2nd exp.) and the QTL reproducibility experiment (3rd exp.) under control (C) and drought (D) conditions. RIL number are provided (RIL), as well as alleles for the indicated markers (a or b).

	Marker	899062|9911548	901866|9957577	P2147M82	BrID10541	Stomatal conductance				leaf number			
	Chromosome	3	7	8	9								
RIL	Position (cM)	43.732	32.051	85.204	69.953	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D
10		a	b	a	b	259.2	230.5	279.3	247.6	7.5	6.0	6.3	5.5
25		a	b	a	b	276.1	286.2	206.7	288.4	6.5	6.0	4.3	4.7
47		a	b	a	b	315.5	238.7	239.2	292.9	8.0	5.0	5.0	5.0
109		a	b	a	b	280.4	272.0	294.8	292.1	9.5	6.3	9.3	5.5
172		a	b	a	b	219.2	231.9	194.1	303.2	5.5	5.0	5.7	3.5
197		a	b	a	b	337.3	278.0	197.8	284.4	10.5	10.0	10.3	8.0
199		a	b	a	b	218.0	225.7	240.2	265.3	8.0	7.0	6.3	5.2
32		b	a	a	b	228.3	237.5	222.5	259.3	6.0	7.3	5.7	5.7
51		b	a	a	b	287.4	239.3	181.2	297.8	7.0	7.0	6.3	6.5
68		b	a	a	b	271.4	230.7	269.7	332.9	6.0	6.0	6.0	3.5
123		b	a	a	b	279.6	237.1	172.2	214.3	7.0	6.3	6.0	4.0
130		b	a	a	b	227.4	204.4	218.1	288.9	7.0	6.5	6.0	5.0
185		b	a	a	b	294.7	203.1	242.9	258.5	5.5	4.3	5.0	4.0
33		b	a	b	a	313.0	199.1	178.4	250.2	7.5	5.5	5.0	5.3
66		b	a	b	a	241.3	191.8	268.9	323.5	8.0	7.0	6.3	5.3
99		b	a	b	a	262.1	231.9	243.8	275.6	6.0	6.0	5.0	4.0
115		b	a	b	a	296.7	250.0	230.4	293.0	8.0	6.0	5.0	5.5
121		b	a	b	a	236.0	226.6	240.8	259.8	8.0	6.3	6.7	6.5
152		b	a	b	a	273.6	256.2	289.1	288.9	6.5	5.0	6.0	5.7
46		b	a	b	b	243.4	239.7	179.7	247.8	6.0	5.5	5.3	4.3
57		b	a	b	b	342.8	225.8	238.5	270.6	7.0	5.7	5.7	4.3
79		b	a	b	b	317.3	233.5	276.3	303.3	8.0	7.0	6.3	5.0
149		b	a	b	b	250.2	234.6	286.3	312.9	8.5	6.3	6.3	5.0
78		b	b	a	b	322.4	247.2	277.0	317.6	7.5	5.3	6.3	5.0
165		b	b	a	b	280.4	181.3	181.5	265.3	10.0	7.5	9.0	8.7
168		b	b	a	b	280.0	254.8	242.1	252.0	12.0	10.0	10.0	7.0
179		b	b	a	b	243.7	203.4	192.3	162.2	14.0	11.7	9.3	9.5

Supplementary table 1 (continued): Phenotypic data of the selected lines for the analysed traits in two experiments, the full RIL screening experiment (2nd exp.) and the QTL reproducibility experiment (3rd exp.) under control (C) and drought (D) conditions. RIL number are provided (RIL), as well as alleles for the indicated markers (a or b).

3rd and 4th leaf dry weight				Leaf dry weight ratio				Shoot dry weight				Specific leaf area			
2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D
0.039	0.034	0.033	0.058	0.243	0.210	0.185	0.341	0.173	0.101	0.198	0.186	301.102	438.529	164.848	373.756
0.037	0.099	0.022	0.027	0.384	1.028	0.358	0.391	0.108	0.368	0.074	0.080	330.631	421.050	122.093	419.724
0.067	0.053	0.027	0.091	0.409	0.324	0.321	0.368	0.180	0.119	0.091	0.277	302.090	414.313	137.625	279.726
0.098	0.094	0.041	0.067	0.279	0.268	0.197	0.374	0.366	0.202	0.219	0.184	177.925	371.576	95.309	491.803
0.032	0.203	0.052	0.014	0.372	2.333	0.344	0.309	0.093	0.440	0.164	0.048	195.876	330.049	113.333	333.212
0.038	0.199	0.037	0.048	0.200	1.059	0.164	0.219	0.200	0.778	0.244	0.231	411.327	338.164	58.378	394.953
0.040	0.060	0.027	0.026	0.234	0.347	0.169	0.132	0.185	0.280	0.170	0.213	363.140	580.122	134.815	458.984
0.065	0.147	0.072	0.080	0.467	1.052	0.470	0.522	0.144	0.505	0.159	0.229	347.092	416.753	130.694	318.093
0.124	0.152	0.062	0.088	0.345	0.423	0.351	0.497	0.382	0.509	0.189	0.285	261.321	385.985	143.280	345.299
0.107	0.108	0.044	0.009	0.397	0.399	0.427	0.085	0.278	0.327	0.112	0.034	256.791	330.050	186.947	561.272
0.084	0.074	0.119	0.024	0.296	0.261	0.451	0.091	0.301	0.211	0.282	0.069	198.775	438.368	90.866	458.333
0.139	0.265	0.063	0.064	0.335	0.640	0.335	0.338	0.452	0.705	0.211	0.173	232.572	345.181	141.587	393.763
0.137	0.105	0.091	0.040	0.481	0.369	0.446	0.195	0.319	0.243	0.238	0.126	242.659	356.744	121.044	305.029
0.126	0.097	0.048	0.056	0.351	0.268	0.387	0.448	0.380	0.219	0.132	0.145	260.396	484.997	152.083	396.341
0.046	0.046	0.052	0.041	0.408	0.412	0.347	0.272	0.138	0.191	0.163	0.116	426.350	606.358	113.333	409.485
0.113	0.183	0.102	0.038	0.440	0.714	0.487	0.180	0.268	0.470	0.219	0.104	240.976	371.923	129.246	405.956
0.198	0.136	0.139	0.086	0.311	0.213	0.491	0.303	0.665	0.340	0.300	0.274	182.088	371.102	114.330	248.340
0.126	0.090	0.082	0.081	0.321	0.228	0.385	0.381	0.420	0.263	0.230	0.252	193.016	427.743	105.183	252.619
0.154	0.118	0.091	0.066	0.327	0.250	0.453	0.329	0.499	0.279	0.211	0.174	148.877	435.329	136.996	474.369
0.185	0.306	0.138	0.057	0.477	0.788	0.592	0.245	0.417	0.616	0.255	0.153	200.342	341.615	161.135	399.591
0.153	0.129	0.083	0.140	0.357	0.300	0.385	0.646	0.446	0.331	0.239	0.280	268.609	423.906	135.880	232.309
0.154	0.272	0.088	0.054	0.306	0.542	0.312	0.192	0.540	0.675	0.315	0.153	172.950	294.436	105.019	339.938
0.123	0.236	0.055	0.059	0.239	0.457	0.231	0.250	0.545	0.611	0.256	0.171	296.911	390.529	181.455	429.309
0.122	0.132	0.084	0.092	0.371	0.402	0.445	0.488	0.362	0.303	0.209	0.235	221.226	341.929	116.905	273.507
0.091	0.050	0.064	0.085	0.281	0.154	0.262	0.347	0.352	0.192	0.281	0.403	282.491	597.790	139.375	367.138
0.051	0.060	0.044	0.016	0.145	0.173	0.223	0.081	0.380	0.264	0.214	0.067	153.882	376.578	107.471	495.137
0.076	0.081	0.039	0.047	0.166	0.176	0.232	0.279	0.481	0.657	0.180	0.249	187.249	374.017	81.197	384.205

Supplementary table 1 (continued): Phenotypic data of the selected lines for the analysed traits in two experiments, the full RIL screening experiment (2nd exp.) and the QTL reproducibility experiment (3rd exp.) under control (C) and drought (D) conditions. RIL number are provided (RIL), as well as alleles for the indicated markers (a or b).

Hypocotyl length				Hypocotyl dry weight				Root system length				Root volume			
2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D
2.533	2.900	2.550	3.000	0.011	0.008	0.020	0.015	568.887	243.233	456.460	540.509	0.807	0.463	1.009	0.419
3.767	5.000	3.667	4.833	0.012	0.041	0.014	0.012	459.583	558.682	289.226	483.350	0.682	0.615	0.819	0.386
2.367	2.900	1.967	2.850	0.016	0.008	0.008	0.029	595.375	336.435	328.831	478.296	0.873	0.579	0.828	0.436
1.833	1.700	1.567	1.450	0.015	0.005	0.013	0.004	1398.002	422.535	747.970	397.321	1.673	0.636	1.536	0.748
2.300	2.250	2.500	2.250	0.006	0.016	0.013	0.003	304.384	565.706	453.561	162.763	0.442	0.380	1.193	0.599
1.600	1.000	2.233	2.100	0.011	0.016	0.018	0.014	691.691	502.812	711.463	411.009	0.860	0.350	1.301	0.407
2.800	4.250	2.933	4.167	0.012	0.025	0.010	0.019	569.735	202.567	254.788	280.887	0.747	0.759	0.765	0.661
1.400	1.633	1.467	1.767	0.004	0.019	0.006	0.011	516.007	585.910	497.943	513.795	0.765	0.490	1.032	0.327
2.833	3.067	2.333	3.000	0.023	0.029	0.012	0.022	1215.168	520.984	581.770	530.062	1.428	0.551	1.446	0.636
1.300	1.400	1.533	1.700	0.008	0.011	0.009	0.002	835.873	583.688	481.615	189.412	0.932	0.370	1.242	0.606
2.233	2.400	1.933	2.567	0.016	0.012	0.017	0.005	1508.600	379.616	950.346	198.853	1.622	0.556	1.601	0.955
2.600	2.350	2.467	1.900	0.038	0.036	0.023	0.009	1257.257	583.279	611.456	440.897	1.545	0.587	1.347	0.619
4.550	4.333	4.300	5.233	0.034	0.020	0.034	0.016	923.917	418.168	875.330	292.031	1.163	0.577	1.737	0.803
1.633	1.500	1.600	1.867	0.019	0.008	0.008	0.009	1408.333	492.561	627.924	559.696	1.869	0.465	1.352	0.473
2.800	2.867	2.467	3.000	0.026	0.011	0.014	0.009	1008.039	200.823	574.237	450.189	1.324	1.050	1.345	0.564
2.300	2.867	2.400	2.667	0.012	0.021	0.010	0.006	912.519	486.024	666.267	268.869	1.095	0.514	1.279	0.660
2.700	2.900	2.300	3.300	0.028	0.017	0.016	0.031	1781.868	418.002	582.949	415.918	1.720	0.592	1.249	0.915
2.033	2.000	1.933	2.500	0.027	0.013	0.017	0.026	1037.729	344.981	830.862	362.465	1.714	0.467	1.620	0.791
2.000	1.833	2.267	2.167	0.027	0.010	0.011	0.011	1832.121	416.780	559.083	364.178	1.800	0.596	1.284	0.634
2.533	3.500	2.400	3.567	0.029	0.042	0.022	0.012	1582.049	630.937	623.272	478.166	1.928	0.446	1.585	0.431
2.200	2.633	2.367	2.867	0.017	0.020	0.022	0.033	1310.825	393.077	642.428	372.346	1.860	0.654	1.337	0.750
3.833	3.850	3.833	3.733	0.038	0.024	0.034	0.012	1268.560	643.322	678.485	385.005	1.672	0.467	1.675	0.833
2.233	1.967	1.700	2.200	0.030	0.019	0.018	0.011	1878.768	641.397	920.608	466.362	2.315	0.430	2.202	0.513
2.800	3.100	2.533	2.850	0.032	0.021	0.021	0.021	1317.439	554.756	777.230	547.189	1.645	0.609	1.416	0.667
2.333	2.450	2.300	2.733	0.028	0.009	0.036	0.043	1294.230	298.440	982.614	640.895	1.652	0.815	1.766	0.523
3.000	2.750	3.000	3.067	0.031	0.014	0.019	0.007	980.392	240.509	450.762	310.155	1.188	1.111	1.364	0.576
2.433	3.900	2.767	4.500	0.022	0.042	0.012	0.026	1422.536	703.735	473.983	443.432	1.790	0.409	1.691	0.433

Supplementary table 1 (continued): Phenotypic data of the selected lines for the analysed traits in two experiments, the full RIL screening experiment (2nd exp.) and the QTL reproducibility experiment (3rd exp.) under control (C) and drought (D) conditions. RIL number are provided (RIL), as well as alleles for the indicated markers (a or b).

Root diameter				RL/SDM				Root sry weight				R/S			
2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D	2nd exp.-C	3rd exp.-C	2nd exp.-D	3rd exp.-D
0.426	0.490	0.554	0.314	3287.096	2410.632	2300.708	2904.401	0.027	0.021	0.044	0.055	0.168	0.224	0.244	0.321
0.455	0.373	0.627	0.319	4247.534	1519.190	3887.452	6072.232	0.023	0.041	0.032	0.025	0.239	0.124	0.528	0.375
0.435	0.468	0.573	0.340	3301.217	2836.723	3633.489	1729.199	0.033	0.020	0.030	0.068	0.202	0.177	0.361	0.273
0.396	0.437	0.514	0.498	3823.161	2091.758	3415.388	2159.942	0.078	0.024	0.092	0.044	0.221	0.120	0.448	0.243
0.435	0.293	0.588	0.677	3278.818	1286.718	2759.446	3415.801	0.013	0.040	0.054	0.035	0.147	0.095	0.354	0.781
0.403	0.298	0.485	0.356	3463.651	646.620	2921.821	1777.335	0.033	0.085	0.061	0.067	0.178	0.112	0.271	0.311
0.417	0.663	0.654	0.564	3087.438	724.404	1502.287	1319.955	0.025	0.023	0.030	0.031	0.147	0.089	0.189	0.160
0.490	0.326	0.534	0.288	3580.065	1159.337	3125.818	2243.974	0.024	0.058	0.043	0.040	0.172	0.119	0.281	0.183
0.389	0.369	0.559	0.393	3183.290	1022.875	3080.323	1862.155	0.060	0.049	0.059	0.043	0.167	0.102	0.332	0.165
0.384	0.288	0.572	0.640	3010.348	1785.344	4314.254	5546.480	0.035	0.033	0.053	0.013	0.129	0.106	0.517	0.402
0.370	0.439	0.480	0.780	5013.626	1799.695	3372.814	2872.216	0.068	0.023	0.079	0.023	0.240	0.114	0.299	0.357
0.390	0.365	0.530	0.421	2782.980	827.405	2902.481	2546.087	0.086	0.087	0.083	0.061	0.208	0.131	0.442	0.374
0.405	0.412	0.502	0.599	2899.017	1721.330	3671.686	2320.775	0.053	0.030	0.090	0.041	0.186	0.136	0.439	0.376
0.412	0.347	0.528	0.329	3708.092	2248.623	4745.015	3855.541	0.074	0.022	0.056	0.032	0.207	0.105	0.455	0.232
0.436	0.938	0.549	0.395	7297.577	1053.266	3518.612	3868.155	0.057	0.029	0.050	0.033	0.506	0.163	0.335	0.307
0.389	0.379	0.520	0.564	3405.343	1034.975	3047.419	2596.100	0.040	0.037	0.072	0.024	0.154	0.083	0.345	0.245
0.348	0.426	0.530	0.532	2679.904	1228.453	1945.108	1517.672	0.082	0.026	0.052	0.042	0.129	0.080	0.182	0.172
0.426	0.436	0.498	0.561	2473.140	1309.558	3618.737	1440.640	0.106	0.023	0.077	0.058	0.269	0.092	0.359	0.259
0.356	0.430	0.541	0.476	3669.134	1493.657	2647.593	2092.575	0.084	0.027	0.051	0.040	0.177	0.099	0.254	0.244
0.399	0.301	0.553	0.340	3790.548	1024.083	2446.441	3127.994	0.111	0.061	0.084	0.036	0.285	0.107	0.362	0.256
0.424	0.463	0.513	0.514	2942.147	1188.142	2688.732	1330.281	0.062	0.038	0.072	0.059	0.144	0.123	0.331	0.239
0.419	0.305	0.561	0.525	2347.446	952.505	2153.692	2508.721	0.075	0.048	0.079	0.036	0.149	0.074	0.280	0.252
0.399	0.294	0.554	0.388	3444.963	1050.036	3596.592	2732.058	0.119	0.051	0.096	0.050	0.230	0.086	0.405	0.314
0.395	0.373	0.483	0.395	3640.340	1832.289	3713.474	2325.990	0.074	0.024	0.069	0.047	0.223	0.087	0.363	0.221
0.400	0.606	0.488	0.327	3677.836	1553.431	3496.434	1590.178	0.066	0.019	0.091	0.116	0.202	0.105	0.370	0.321
0.396	0.767	0.631	0.491	2581.792	910.674	2107.022	4606.261	0.055	0.015	0.066	0.023	0.158	0.060	0.338	0.379
0.400	0.273	0.665	0.353	2955.817	1071.080	2631.287	1779.778	0.063	0.061	0.043	0.049	0.137	0.100	0.253	0.218

Chapter 6: General discussion

Plants and especially crops often grow in environments that are suboptimal, which prevents them from expressing their full genetic potential for growth and reproduction (Atkinson and Urwin, 2012). To further understand how plants adapt to stress environments, this thesis focused on phenotyping shoot and root traits of *Arabidopsis thaliana* (Arabidopsis) and *Brassica rapa* grown in control and drought environments (chapters 2,3 and 5). We screened morphological traits of Arabidopsis RIL, HapMap and *B. rapa* RIL populations grown on sand in greenhouse conditions. This type of experiments is a reasonable compromise to avoid the difficulty of phenotyping roots in natural soils and the unnatural conditions present in hydroponics, aeroponics or agar plates (Tuberosa, 2012). The main focus of these experiments was to study drought avoidance and drought tolerance. Both mechanisms together with drought escape, e.g. early flowering to finish the life cycle, are the main three strategies by which plants can colonize terrestrial ecosystems and optimise their growth in environments with limited water availability (Tuberosa, 2012; Assmann, 2013; Juenger, 2013; Tardieu, 2013). In Arabidopsis no evidence for drought avoidance via early flowering was observed, however, there was significant genotype by environment interaction (GxE) between control and drought environments (Chapter 2).

Natural variation for plant morphological and physiological traits is genetically controlled by quantitative trait loci (QTL). QTL are very often environmentally sensitive, resulting in QTL by environment interaction (QxE) (Nicotra et al., 2010; Weigel, 2012; Assmann, 2013; Juenger, 2013; Tardieu, 2013). To better understand how plants have adapted naturally to drought stress and to illustrate the effect of QxE, I analysed in this thesis distinct phenotypic responses, phenotypic plasticity, and the differential genotypic responses in different environments (GxE). My main objectives were: (i) to genetically dissect and identify QTL, and preferentially even the causal candidate genes, underlying the plant morphological or physiological response to drought; (ii) to understand how plants differentially interact with their control and drought environments; (iii) to transfer data gained from model plant to a real crop, a well characterised *B. rapa* RIL population was established and phenotyped under drought conditions.

Mapping plasticity and genotype by environment interaction

Plasticity is the ability of a genotype to produce distinct phenotypes in different environments, often as a way to adapt to the environment, and when phenotypic plasticity differs between genotypes this is considered as GxE. Mapping plasticity in chapters 2, 3 and 5 via mapping the difference between the control and the drought environments resulted in one of the following three cases. (i) A QTL for the trait, but not for plasticity. This is the case when a QTL has only a main effect in both environments (Figure 1-A), and the effect is the same. (ii) A plasticity QTL that co-located with a main effect QTL but not showing any QxE (Figure 1-B). In this case, the QTL effect explaining the difference between both environments for each genotype is comparable. (iii) A plasticity QTL that co-locates with QTL with QxE. In this case genotypes show differences in their response to the environment, which adds the GxE component (Figure 1-C).

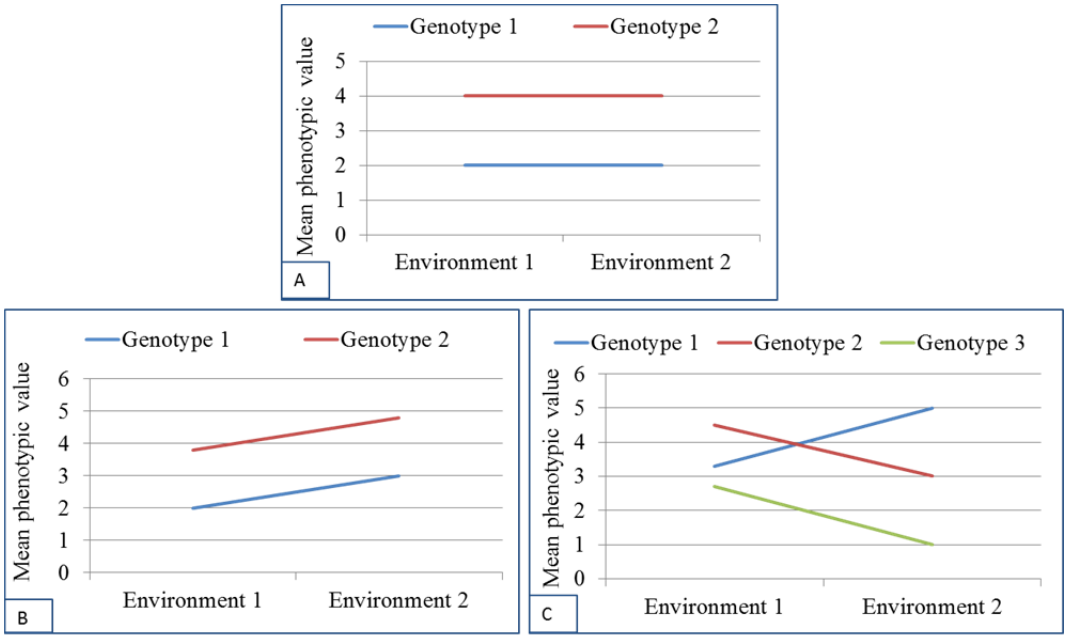


Figure 1: Reaction norms of two and three genotypes illustrating three cases involving response to the environment. (A) No plasticity or GxE, phenotypic values are similar over environments. (B) Phenotypic plasticity, the trait values change across environments but the rank order of genotypes across environments remains unchanged (non-crossed parallel reaction norms). (C) GxE causes significant rank changes among genotypes evaluated in different environments (crossed reaction norms), however some genotypes still show only plasticity.

In general, the results reported in chapters 2 and 3 showed that 90 % and 82 % of the QTL or SNPs associated with plasticity, respectively, were from that kind of response described in Figure 1-C, where, plasticity QTL co-locate with QTL or SNP with QxE. This result is in agreement with what was reported in barley (Lacaze et al., 2009) who reported co-locating of plasticity QTL with only QTL showing QxE. However, another study in barley (Tétard-Jones et al., 2011) showed some QTL mapped for plasticity only and did not co-locate with other main effect QTL or QTL with QxE which is in agreement with the results reported in chapter 5 where 40 % of the QTL underlying plasticity did not co-locate with QTL with main effect or with QxE. These results indicates that the kind of QTL underlying plasticity depend on the studied plant species and/or the tested population.

Towards gene identification

In Arabidopsis, many studies have identified QTL, but only a few genes underlying morphological or physiological responses to drought using natural variation have been identified. For example, the *ERECTA* gene was identified as a regulator for transpiration efficiency (Masle et al., 2005) by affecting stomatal density. *PHYTOCHROME B* (*PHYB*) was proposed to increase drought tolerance by enhancing ABA sensitivity when soil water becomes limiting (Gonzalez et al., 2012).

In chapters 2 and 3, two approaches were followed to propose a candidate gene to be underlying the natural variation of the measured traits. In the first approach (chapter 2), we looked for possible pleiotropic effects of molecularly identified flowering time (FT) QTL, reviewed in (Alonso-Blanco et al., 2009), of which many were segregating in the Sha x Col RIL population. By studying the natural genetic variation of Arabidopsis response to drought under long days (LD) and short days (SD), large number of QTL with QxE in response to light conditions and/or water treatment were mapped. Moreover, seven FT QTL co-located with QTL underlying rosette and root traits. For example, *FLOWERING LOCUS T* (*FT*), *TWIN SISTER OF FT* (*TSF*), *FLOWEING LOCU C* (*FLC*) and the *FLC* paralogue, *MADS AFFECTING FLOWERING 2-5 (MAF2-5*) were suggested to have pleiotropic effect on rosette fresh weight and water content response to drought under long and short days conditions. A recent study on flowering time mutants in Arabidopsis (Riboni et al., 2013) revealed that under long days, drought conditions promote the flower-promoting gene *GIGANTEA* (*GI*) to regulate *FT* and *TSF* via abscisic acid (ABA) signalling. Another study (Lovell et al., 2013) reported the role of *FRIGIDA* in drought escape not only by early flowering but also by fast growth and low water use efficiency (WUE). Fast growth and

reduced life cycle by early flowering are typical drought escape mechanisms which leads plants to put less energy on biomass leading to low WUE or low transpiration efficiency.
The second approach that was followed in chapter 3 is genome wide association (GWA) mapping which proposed several genes to be associated with drought response. This study reported several significant SNPs associated with genes known to be involved in drought stress as well as genes of un-known function. Many of the mapped genes were indirectly involved in drought tolerance such as genes involved in oxidative stress or producing antioxidants (Cruz de Carvalho, 2008; Mirzaei et al., 2012) which shows the power of the GWA mapping approach in identifying drought related genes. This is in turn emphasizes the importance of applying GWA mapping in crop plants. Several recent GWA studies have been conducted in the Brassica genus, using *B. napus* (Ecke et al., 2010; Honsdorf et al., 2010; Xiao et al., 2012) and *B. rapa* (Zhao et al., 2007; Pino Del Carpio et al., 2011). However, the problem in these studies was either using a small population or a limited numbers of markers. A recent study with an attempt to overcome both problems introduced *B. napus* population of 313 inbred lines genotyped with 7367 markers, which provides a valuable resource for further genetic studies (Delourme et al., 2013). In the future, and once the sequence of the *B. oleracea* and *B. napus* genomes are publically available, the availability of the whole *B. rapa* genome sequence information (Wang et al., 2011) will rapidly overcome the problems associated with marker density, as it will allow the creation of high-density SNP-based maps for a sufficiently large collection of inbred genotypes of different origin, and thus facilitate proper GWA mapping.

Combining QTL and GWA mapping

GWA mapping in its current state suffers from the high rate of false negative and false positives (Brachi et al., 2010). In addition, the results reported in chapter 3 on GWA study using the HapMap population, showed that the significant level of a SNP can vary based on the statistical approach used for mapping which may increase the number of false negative. This discrepancy between different statistical approaches is in agreement with what reported recently in Arabidopsis (Yano et al., 2013) where a gene associated with seed dormancy was significant using one statistical approach and not significant using another approach. On the other hand, traditional QTL mapping, as presented in chapters 2, 4 and 5, based on populations derived from bi- or multi-parental crosses, suffers from the relatively low resolution due to restricted recombination, and the strong limitation on the number of segregating alleles. Therefore, combining traditional QTL mapping with GWA mapping was

proposed as the preferred alternative to overcome most of these problems (Nordborg and Weigel, 2008; Atwell et al., 2010; Bergelson and Roux, 2010; Brachi et al., 2010; Sterken et al., 2012; Weigel, 2012). Brachi et al., (2010) already introduced a proof of concept of this assumption by screening flowering time in 13 RIL families and in a collection of 184 accessions. By mapping already known flowering time genes, their results showed clearly that QTL mapping increased their power to distinguish true from false association in GWA mapping and allowed to identify false negatives. In chapter 3 of this thesis, the SNP that mapped in *YELLOW LEAF SENESCENCE 7* (*YLS7*) and was associated with root length (RL), although its probability just did not reach the arbitrary significance threshold we set at –log10(P)=4, was considered as a false negative association. Another reason for considering this association was the significant association between this SNP and the ratio between RL and rosette dry weight (RosDW), when using a different version of the univariate approach. It was therefore proposed as the gene underlying RL natural variation, and the sequence analysis of alleles further confirmed its candidacy, although final proof has not yet been provided. In addition, one non-significant SNP mapped in *FLC* was associated with RosDW (chapter 3), and the same gene was found to be co-locating with a QTL for RosDW in the RIL population (chapter 2). Few more significant SNPs were mapped in the interval of significant QTL, but because the function of the marked gene was not known, or the QTL interval was large, or no common difference in amino acid sequence was found between extreme groups, those co-locations were not reported. It is not ideal to only consider the co-location between QTL and GWA mapping of RIL and HapMap populations grown in the same conditions, but better to also consider other studies that reported on the same traits and/or on similar treatments. That was the reason why in chapter 3 some co-locations were reported, found with QTL identified in the same Sha x Col RIL population, but grown on agar under salt stress (Galpaz and Reymond, 2010). Moreover, it will always be useful and justified to include other populations sharing Col as one of the parents, or populations with parents showing a strongly differential response for the studied trait, in the further analysis and interpretation of GWA mapping results.

This approach will not be limiting to Arabidopsis, as a recent study in *B. napus* (Zou et al., 2010) showed there was overlap between GWA results and several QTL mapping studies. However, for the QTL studies described in chapter 4 and 5 using *B. rapa*, no GWA mapping populations were yet available for comparison and thereby looking for the overlap and identify possible association was not yet possible.

Drought stress, metabolites and anti-oxidants

Many compounds, e.g. plant hormones; specific sugars; anti-oxidants; osmolytes, such as glycine betaine; Reactive Oxygen Species (ROS) and their detoxifying enzymes play a role in plant response to stress. For example, the level of ROS increases in Arabidopsis plants subjected to various stresses leading to significant damage in cell structure. This phenomenon is known as oxidative stress, a shift in the balance between reduced and oxidized biomolecules within cells leading to oxidative stress. In chapter 3, three significant SNPs with significant Q x E were mapped in genes involved in the oxidative stress. Plants respond to oxidative stress by producing proline, antioxidants such as ascorbic acid, anthocyanins and tocophorol as well as by up-regulating enzymes such as superoxide dismutase and catalases, which will scavenge ROS or convert them into less reactive compounds (Mittler, 2002; Cruz de Carvalho, 2008; Bhatt et al., 2011; Setter, 2012; Bhargava and Sawant, 2013). Although proline is well known to be related to abiotic stresses and drought (Verslues and Sharma, 2010; Lv et al., 2011; Sharma et al., 2011; Liang et al., 2013), no significant association was found for any of measured traits with proline biosynthesis, storage or activity. Although it was shown that Sha had fourfold less proline accumulation at low water potential than L*er* or Col (Kesari et al., 2012), in chapter 2, where the Sha x Col RIL population was screened, no QTL known to be related to proline was mapped. Not mapping proline in both studies can be as a result of false negative or imply that although proline accumulates under drought stress, it is not essential for increasing the performance of the morphological traits studied here.

Proline accumulation in plants experiencing abiotic stresses is regulated by ABA (Sharma et al., 2011), which is the most important hormone involved in regulating drought avoidance or tolerance (Setter, 2012; Tuberosa, 2012). Although in the current thesis ABA was not measured, it is likely that the variation in stomatal conductance as was observed in Brassica (chapter 5) relates to variation in ABA levels. Under drought conditions, the direct effect of ABA and the indirect negative effect of stomatal closure and limited expansion of young leaves on root growth was observed in chapter 5. This effect was via reducing the root length and increasing the root dry weigh, which is explained by a trade-off between longer roots to take up water and greater carbon investment in roots (Tardieu, 2013). Therefore, it will be useful to map ABA accumulation in response to drought for the whole population, as it is likely to add additional valuable information on the morphological and physiological mechanisms underlying drought tolerance.

As reported in chapter 5, the only QTL mapped for stomatal conductance co-located with a QTL for WUE, indicating that closing stomata leads to less water loss and consequently better WUE, which is defined as carbon gain at the cost of water loss. In general, under drought stress, a good predictor of stomatal conductance and WUE in C3 plants is $\delta\ ^{13}C$, $^{13}C/^{12}C$, where the variation of $\Delta^{13}C$ is wider than in C4 plants (Tuberosa, 2012). In general, the response of C3 photosynthesis to drought is very well studied, however, it is significantly different from the response of C4 crops, such as maize and sorghum, that are pivotal to the current and the future global food security (Ghannoum, 2009).

In order to investigate the impact of drought on the metabolite composition and their concentrations, it is important to measure the total amount of these compounds and the whole plant biomass in control and drought conditions (Selmar and Kleinwächter, 2013). Measuring the whole plant biomass is not possible unless plants are grown in vitro, in hydroponics or on sandy soil. However, it is well known that the accumulation of metabolites strongly depends on the growth environments (Setter, 2012; Selmar and Kleinwächter, 2013), which raises doubts on the applicable values of *in vitro* metabolic studies on developing drought tolerant crops. Although in this thesis I did not consider metabolite response to drought, the growing system used here introduced a natural system for phenotyping the whole plant response to drought.

In general, it remains a challenge to identify which metabolites provide the most meaningful phenotypic information. Furthermore, to understand how these metabolites interact with the environment, there is a need to repeat each experiment in different environments, which is time consuming, expensive, or technically demanding (Setter, 2012; Joosen et al., 2013). Therefore, using a next generation genetical genomics approach which includes environmental perturbation within a single experimental design (Li et al., 2008; Joosen et al., 2013) will allow testing different environments at low cost. Such an approach aims for the creation of subpopulations of RILs, one for each environment to be tested, with an optimal distribution of parental alleles over all available markers.

Conclusions and recommendations

This thesis illustrated the importance of GxE and how it can be affected by QxE. This was achieved by mapping shoot and root traits using QTL and GWA mapping in Arabidopsis and Brassica populations grown on sand. There is still a need to confirm these results by using heterogenous inbred families, combining QTL and GWAS results, gene expression studies

and genetic transformation of strong alleles in null-allele backgrounds. In general, studying GxE at the genomics, proteomic and transcriptomics levels will facilitate better use of the available information for both crop improvement and the understanding of adaptation to different ecological habitats. In the future this should be done for single as well as for combined stresses.

Reverences:

Alonso-Blanco C, Aarts MGM, Bentsink L, Keurentjes JJB, 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.

Assmann SM. 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.

Atkinson NJ, Urwin PE. 2012. The interaction of plant biotic and abiotic stresses: from genes to the field. *J Exp Bot* **63**, 3523-3543.

Atwell S, Huang YS, Vilhjalmsson BJ, *et al.* 2010. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. *Nature* **465**, 627-631.

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

Bhargava S, Sawant K. 2013. Drought stress adaptation: metabolic adjustment and regulation of gene expression. *Plant Breeding* **132**, 21-32.

Bhatt D, Negi M, Sharma P, Saxena S, Dobriyal A, Arora S. 2011. Responses to drought induced oxidative stress in five finger millet varieties differing in their geographical distribution. *Physiology and Molecular Biology of Plants* **17**, 347-353.

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

Cruz de Carvalho MH. 2008. Drought stress and reactive oxygen species: production, scavenging and signaling. *Plant Signaling and Behaviour* **3**, 156-165.

Delourme R, Falentin C, Fomeju BF, *et al.* 2013. High-density SNP-based genetic map development and linkage disequilibrium assessment in *Brassica napus* L. *BMC Genomics* **14**, 120.

Ecke W, Clemens R, Honsdorf N, Becker HC. 2010. Extent and structure of linkage disequilibrium in canola quality winter rapeseed (Brassica napus L.). *Theoretical and Applied Genetics* **120**, 921-931.

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

Ghannoum O. 2009. C4 photosynthesis and water stress. *Annals of Botany* **103**, 635-644.

Gonzalez CV, Ibarra SE, Piccoli PN, Botto JF, Boccalandro HE. 2012. Phytochrome B increases drought tolerance by enhancing ABA sensitivity in *Arabidopsis thaliana*. *Plant Cell and Environment* **35**, 1958-1968.

Honsdorf N, Becker HC, Ecke W. 2010. Association mapping for phenological, morphological, and quality traits in canola quality winter rapeseed (Brassica napus L.). *Genome* **53**, 899-907.

Joosen RV, Arends D, Li Y, Willems LA, Keurentjes JJ, Ligterink W, Jansen RC, Hilhorst HW. 2013. Identifying genotype-by-environment interactions in the metabolism of germinating Arabidopsis seeds using generalized genetical genomics. *Plant Physiol* **162**, 553-556.

Juenger TE. 2013. Natural variation and genetic constraints on drought tolerance. *Curr Opin Plant Biol* **16**, 274-281.

Kesari R, Lasky JR, Villamor JG, Des Marais DL, Chen Y-JC, Liu T-W, Lin W, Juenger TE, Verslues PE. 2012. Intron-mediated alternative splicing of Arabidopsis P5CS1 and its association with natural variation in proline and climate adaptation. *Proc Natl Acad Sci U S A* **109**, 9197–9202.

Lacaze X, Hayes PM, Korol A. 2009. Genetics of phenotypic plasticity: QTL analysis in barley, *Hordeum vulgare*. *Heredity (Edinb)* **102**, 163-173.

Li Y, Breitling R, Jansen R. 2008. Generalizing genetical genomics: getting added value from environmental perturbation. *Trends in Genetics* **24**, 518 - 524.

Liang X, Zhang L, Natarajan SK, Becker DF. 2013. Proline mechanisms of stress survival. *Antioxidants and Redox Signaling* **doi:10.1089/ars.2012.5074**.

Lovell JT, Juenger TE, Michaels SD, Lasky JR, Platt A, Richards JH, Yu X, Easlon HM, Sen S, McKay JK. 2013. Pleiotropy of FRIGIDA enhances the potential for multivariate adaptation. *Phil. Trans. R. Soc. B* **280**, 20131043.
Lv WT, Lin B, Zhang M, Hua XJ. 2011. Proline accumulation is inhibitory to Arabidopsis seedlings during heat stress. *Plant Physiol* **156**, 1921-1933.
Masle J, Gilmore SR, Farquhar GD. 2005. The ERECTA gene regulates plant transpiration efficiency in Arabidopsis. *Nature* **436**, 866-870.
Mirzaei M, Soltani N, Sarhadi E, Pascovici D, Keighley T, Salekdeh GH, Haynes PA, Atwell BJ. 2012. Shotgun proteomic analysis of long-distance drought signaling in rice roots. *J Proteome Res* **11**, 348-358.
Mittler R. 2002. Oxidative stress, antioxidants and stress tolerance. *Trends in Plant Science* **7**, 405-410.
Nicotra AB, Atkin OK, Bonser SP, *et al.* 2010. Plant phenotypic plasticity in a changing climate. *Trends in Plant Science* **15**, 684-692.
Nordborg M, Weigel D. 2008. Next-generation genetics in plants. *Nature* **456**, 720-723.
Riboni M, Galbiati M, Tonelli C, Conti L. 2013. GIGANTEA enables drought escape response via abscisic acid-dependent activation of the florigens and SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1. *Plant Physiol* **162**, 1706-1719.
Selmar D, Kleinwächter M. 2013. Stress enhances the synthesis of secondary plant products: the impact of stress-related over-reduction on the accumulation of natural products. *Plant and Cell Physiology* **54** 817-826.
Setter TL. 2012. Analysis of constituents for phenotyping drought tolerance in crop improvement. *Frontiers in Physiology* **3**, 180.
Sharma S, Villamor JG, Verslues PE. 2011. Essential role of tissue-specific proline synthesis and catabolism in growth and redox balance at low water potential. *Plant Physiol* **157**, 292-304.
Sterken R, Kiekens R, Boruc J, *et al.* 2012. Combined linkage and association mapping reveals CYCD5;1 as a quantitative trait gene for endoreduplication in Arabidopsis. *Proc Natl Acad Sci U S A* **109**, 4678-4683.
Tardieu F. 2013. Plant response to environmental conditions: assessing potential production, water demand and negative effects of water deficit. *Frontiers in Physiology* **4:17. 10.3389/fphys.2013.00017**.
Tétard-Jones C, Kertesz MA, Preziosi RF. 2011. Quantitative trait loci mapping of phenotypic plasticity and genotype–environment interactions in plant and insect performance. *Phil. Trans. R. Soc. B* **366**, 1368-1379.
Tuberosa R. 2012. Phenotyping for drought tolerance of crops in the genomics era. *Frontiers in Physiology* **3:347. doi:10.3389/fphys.2012.00347**.
Verslues PE, Sharma S. 2010. Proline metabolism and its implications for plant-environment interaction. *The Arabidopsis Book*, e0140.
Wang X, Wang H, Wang J, *et al.* 2011. The genome of the mesopolyploid crop species *Brassica rapa*. *Nature Genetics* **43**, 1035-1039.
Weigel D. 2012. Natural variation in Arabidopsis: from molecular genetics to ecological genomics. *Plant Physiol* **158**, 2-22.
Xiao Y, Cai D, Yang W, Ye W, Younas M, Wu J, Liu K. 2012. Genetic structure and linkage disequilibrium pattern of a rapeseed (Brassica napus L.) association mapping panel revealed by microsatellites. *Theoretical and Applied Genetics* **125**, 437-447.
Yano R, Takebayashi Y, Nambara E, Kamiya Y, Seo M. 2013. Combining association mapping and transcriptomics identify HD2B histone deacetylase as a genetic factor associated with seed dormancy in *Arabidopsis thaliana*. *The Plant Journal* **74**, 815-828.
Zou J, Jiang C, Cao Z, Li R, Long Y, Chen S, Meng J. 2010. Association mapping of seed oil content in Brassica napus and comparison with quantitative trait loci identified from linkage mapping. *Genome* **53**, 908-916.

Summary

Drought is the major abiotic stress affecting plant growth and limiting crop productivity worldwide. To cope with drought, plants have evolved three adaptive strategies namely drought escape, drought avoidance and drought tolerance. Breeding for such complex strategies is not straightforward and better understanding of plant morphological and physiological response to drought is needed. Therefore knowledge on how Quantitative Trait Loci (QTL), or genes underlying them interact with their environments will significantly increase our understanding and the success of breeding for drought tolerance.

Three main goals of this thesis were achieved by phenotyping shoot and root traits of *Arabidopsis thaliana* and *Brassica rapa* grown on sand supplemented with mineral solutions in greenhouses. These goals were: (i) to genetically dissect and identify QTL, and preferentially even the causal candidate genes, underlying the plant morphological or physiological response to drought. (ii) To understand how plants differentially interact with their control and drought environments. (iii) In order to transfer data gained from model plant to a real crop, a *B. rapa* RIL population was established and phenotyped under control and drought conditions.

An already existing Arabidopsis RIL population, selected on the basis of the differential root response to drought of the two parental lines, Sha and Col, was evaluated in control and drought environments both under long and short day photoperiods (**chapter 2**). The examined population exhibited significant genotype by environment interaction (GxE) for all traits in both photoperiods, which was reflected by significant QTL by environment interactions (QxE),using the MEA procedure of Genstat. Besides the 27 QTL mapped with significant QxE, 21 QTL were mapped with main effects. No significant effect of the drought treatment on flowering time (FT) under LD conditions was observed. However, a significant GxE for FT was observed in the SD condition. Two flowering time QTL showed conditional neutrality when comparing day length conditions. Besides, four QTL were mapped for root length, where three of them showed an additive effect and only one showed a significant QxE. Moreover, the plasticity QTL, detected by mapping the difference per line between both environments, either overlapped with QTL with QxE and main effect or were identified only for plasticity response.

In order to identify candidate genes that are underlying shoot and root response to drought, a large population of Arabidopsis accessions collected from all over its natural distribution area, genotyped with > 200,000 SNPs, was used in a Genome Wide Association Study (GWAS)

(**chapter 3**). The results showed that 16 significant SNPs with main effect, 58 SNPs underlying GxE, and 100 SNPs affecting plasticity were associated with the measured traits. Again the QxE and plasticity QTL showed a large overlap, expressed in high correlation R^2, between the significance levels of SNP's mapped with QxE and those mapped for plasticity. One co-location between a significant SNP associated with root length (RL) and a QTL mapped for RL in the Sha x Col RIL was reported. Additional candidate genes known to be affected in their differential expression by abiotic stress as well as novel genes associated with rosette and root growth traits in control and drought environments are evaluated.

In order to apply the same methodology in crop breeding, a contribution was made to the genetic mapping of a new *B. rapa* RIL population, consisting of 160 lines and genotyped with 270 different markers (**chapter 4**). Because this population exhibited segregation for several traits, it was encouraging to evaluate its morphological and physiological responses to drought which was achieved in **chapter 5**. The Brassica RIL population showed significant GxE and QxE. Altogether 54 QTL of which seventeen QTL showed significant QxE were mapped. In addition, 15 QTL for plasticity, from which 9 co-located with QTL with main effect or QxE were mapped. Correlation analysis showed that in the control environment, stomatal conductance was significantly positively correlated with total leaves dry weight (DW) and above ground DW. In contrast and in the drought environment, stomatal conductance showed significant negative correlation with total leaves DW and above ground DW. This correlation was described as conditional neutrality leading to an antagonistic fitness effect in the drought and control environment, controlled by a QTL cluster on chromosome A7.

The results presented in the present thesis demonstrate that QxE is an important component of the genetic variance and can play a great role in improving drought tolerance in future breeding programs. There is still a need to confirm these results by using heterogeneous inbred families, combining QTL and GWAS results, and gene expression studies. In general, this thesis illustrates the effect of considering GxE, while mapping QxE in terms of mapping more significant QTL.

Samenvatting

Droogte is de belangrijkste abiotische stressfactor, die plantengroei negatief beïnvloedt, en die daardoor in vele delen van de wereld de productiviteit van gewassen beperkt. Planten hebben drie strategieën ontwikkeld om zich aan droogte aan te passen namelijk: het ontsnappen aan droogte, het vermijden van droogte en droogtetolerantie. Het veredelen op dergelijke complexe eigenschappen is niet zonder complicaties en daarom is een beter begrip van de morfologische en fysiologische reacties van planten op droogte noodzakelijk.

Het is nodig te weten hoe QTL (loci die kwantitatieve eigenschappen bepalen), of de genen die voor deze eigenschappen coderen, interacteren met de omgevingsfactoren. Deze kennis zal medebepalend zijn voor het succes van veredeling op droogtetolerantie.

Drie belangrijke doelstellingen werden bereikt door scheut- en wortel- eigenschappen bij *Arabidopsis thaliana* (zandraket) en *Brassica rapa* (raapzaad/Chinese kool) te fenotyperen bij planten die in de kas opgeweekt werden op zand. Deze doelstellingen waren: ten eerste het genetisch ontrafelen en identificeren van QTL en liefst ook van de kandidaat genen die de morfologische en fysiologische respons van de planten op droogte bepalen. Ten tweede: te begrijpen hoe planten op verschillende manieren reageren op de droogte- respectievelijk de controle-behandeling. Ten derde: kennis van de modelplant Arabidopsis vertalen naar een gewas. Hiertoe werd een *B. rapa* RIL (recombinante inteeltlijn) populatie ontwikkeld en gefenotypeerd onder droogte-condities.

Een reeds bestaande Arabidopsis RIL populatie werd geselecteerd op basis van de verschillende reactie van het wortelsysteem van de ouders op droogte. Deze populatie met Sha en Col als ouders werd geanalyseerd in controle- en droogte-omstandigheid bij opkweek zowel onder lange als onder korte dag (**hoofdstuk 2**). De onderzochte populatie vertoonde een significante interactie tussen het genotype en de omgevings-condities (GxE) voor alle eigenschappen in beide daglengten, wat weerspiegeld werd in QTL met een significante QTL x omgevings (Q xE) interactie. Er werden naast 27 QTL met een dergelijke interactie component 21 QTL gelokaliseerd met alleen een significant hoofdeffect. Hoewel er geen statistisch significant effect van de droogte op bloeitijd werd gevonden onder lange dagcondities, werd onder korte dagomstandigheden wel een significante interactie met de behandeling waargenomen. Voor twee bloeitijd QTL was de detectie afhankelijk van de daglengte-condities, wat beschreven wordt als conditionele neutraliteit. Daarnaast werden er 4 QTL gevonden voor wortellengte, waarvan er drie een additief effect hadden en slechts één QTL een interactie met de toetsomgeving vertoonde. QTL die geïdentificeerd werden op

basis van het verschil tussen de behandelingen werden als plasticiteits QTL beschreven. Zij overlapten met de bovengenoemde QTL, op twee na die alleen bij deze analyse geïdentificeerd werden.

Om de kandidaat genen die de scheut- en wortel-systeem QTL bepalen te vinden, werd een grote populatie van Arabidopsis accessies, die over de hele wereld verzameld waren en die gegenotypeerd waren met meer dan 200.000 SNPs (polymorfe basepaar posities), gebruikt **(hoofdstuk 3).** In het experiment werden 16 significante SNPs met een hoofdeffect, 58 SNPs met een genotype x omgevings effect en 100 SNPs met een plasticiteits effect gevonden. Zoals verwacht was er een grote overlap tussen de twee laatstgenoemde categorieën QTL, hetgeen te zien was aan een hoge correlatie tussen de significanties uitgedrukt als LOD scores van beide analyses. Er werd een co-lokalisatie gevonden van een QTL die geïdentificeerd was voor wortellengte in de Sha x Col RIL populatie en een significante SNP geassocieerd met wortellengte in de accessie-collectie. Daarnaast werden andere kandidaat-genen bediscussieerd waarvan beschreven was dat hun expressie afhangt van het niveau van abiotische stress. Ook werden genen voor rozet- en wortelgroeieigenschappen als kandidaat geëvalueerd, die daar tot nu toe niet mee geassocieerd werden.

Met het doel dezelfde genetische methoden toe te passen op een cultuurgewas werd bijgedragen aan het genetisch in kaart brengen van een nieuwe *B. rapa* RIL populatie, die uit 160 lijnen bestaat en die gegenotypeerd was met 270 moleculaire merkers **(hoofdstuk 4).**

Omdat deze RIL populatie uitsplitsing voor diverse eigenschappen vertoonde, was het aantrekkelijk om de morfologische en fysiologische respons op droogte in dit materiaal te analyseren, hetgeen beschreven is in **hoofdstuk 5**. In deze Brassica populatie werd interactie van de genetische verschillen met de omgevingscondities aangetoond en werden er conditiespecifieke QTL gevonden. Er werden 54 QTL geïdentificeerd waarvan er 17 een interactie met de testcondities vertoonden. Daarnaast werden 15 QTL gevonden voor plasticiteit, waarvan er 9 co-localiseerden met de QTL die gevonden werden met de Q x E analyse. De correlatie-analyse liet zien dat in de controlecondities de geleidbaarheid van de huidmondjes positief gecorreleerd was met het drooggewicht van de bladeren en met het drooggewicht van alle bovengrondse delen, terwijl in droogtecondities deze correlatie met deze parameters negatief was. Deze correlatie wordt beschreven als door omstandigheden bepaalde neutraliteit die leidt tot een tegengesteld ‘fitness’ effect in de droogtecondities, bepaald door een QTL cluster op chromosoom A7.

De resultaten die in dit proefschrift zijn gepresenteerd laten zien dat Q x E (interactie van QTL met omgevingsfactoren) een belangrijke component is van de genetische variatie en een

grote rol kan spelen bij de verbetering van de droogtetolerantie in toekomstige veredelingsprogramma's. Er blijft de noodzaak om deze resultaten te bevestigen met heterogene inteelt families en het combineren van QTL met genoombrede associatie- en genexpressie studies. In algemene zin laat dit proefschrift zien wat het effect is van het meeanalyseren van de genotype-omgevings interactie waardoor meer significante QTL gevonden kunnen worden.

Acknowledgments

Living five years in Wageningen that were full of academic and social success would not be possible without the support of Almighty ALLAH collaborative efforts by many people whom I would like to acknowledge my sincere thanks to them.

I would like first to thank Dr. Mark Aarts who gave me the opportunity to do this research in his group. Mark, my experience with you was unique as you always prefer to challenge your students to get the best out of them which was not always a pleasant experience for me. In the beginning I thought you do not trust me as you kept asking why and how for every single step, but eventually we both came to know each other better and learnt to work together. I have to admit that I enjoyed the freedom you gave me to design and carryout the research work which made me a very confident and independent researcher. I am deeply grateful for your guidance and your critical comments on manuscripts.

I am most indebted to my promoter Professor Maarten Koornneef whom I considered as a father and loved his attitude. Maarten, I do not know how to thank you and how I would be able to achieve this work without your support and your advices. My deep heartfelt thanks to you and to your invaluable ideas and suggestions you provided in conducting this research. Maarten, you are not only a wonderful and fatherly advisor, but also a genuine critic whose comments and suggestions are greatly reflected in this work. I appreciate the endless hours you spent reviewing and editing this thesis and manuscripts in spite of your busy schedule.

I would like to thank our friendly boss, Professor. Bas Zwaan. I am amazed how you can manage your time to be available for everyone who needs to talk to you and how you are able to solve problems with few words. I am very thankful for the extension I got for writing my thesis. I appreciate a lot the advices I got from Dr. Gerard Van der Linden, Prof. Hans De Jong, Dr. Joost Keurentjes, Dr. Martijn Schenk, Dr. Wilco Ligterink, Dr. Marcos Malosetti, Dr. Martin Boer and Dr.Willem Kruijer for their generous support and ideas every time I needed it. I also would like to thank our group secretary Mrs. Wytske Nijenhuis for her administration help over the five years.

I would not be able to finish this work without the assistance and advices I got from Corrie Hanhart. Corrie, want to thank you for our friendship and for the few but unforgettable social activities we had together. I would like to heartedly offer my special thanks to Diana Zuluaga, Zeshan Hassan, Hedayat Bagheri, Nihal Erol-Öztolan, Pingping Huang, Aina Prinzenberg, Christina May, Charles Moreira, Roxanne van Rooijen, Ya-fen Lin, Mina Tomaz, Yanli Wang, Frank Becker, Padraic Flood, Myriam Olortegui Guzman, Natalia Carreno Quintero, Ana

Carolina and Kaige Zhao and Johanna Molenaar for their help and their ideas. I enjoyed being with all of you. My deep special thanks to Nihal Erol-Öztolan, Pingping Huang, Christina May, Aina Prinzenberg, and Erik Wijnker for giving me the chance to be in multicultural enjoyable and cheerful time in the lab and certainly outside the lab. I would like to thank Kaige Zhao and Pingping Huang for designing the cover of this thesis. Thanks to everyone in the lab of Genetics and plant sciences group. It was a pleasure working with such a great group of people. I also appreciate the help I got from Taede Stoker, Gerrit Stunnenberg, and all people from Unifarm for taking care of preparing nutrient solutions and watering plants.

I would like to thank my friends for their smiling faces and for supporting me both intellectually and emotionally especially Jasper Schouten, Tarek Soliman, Ahmed Abd El-Halim, Fouad Lakhal, Rashid Kazmi, Michaele Bandecchi, Nasr El-Din Mohamed, Wahed Awaad, Ahmed El-Wazeery.

Finally all this would not have been possible without my family and their encouragement and love was always with me.

About the author

Mohamed El-Soda was born on 21st August 1979 in Cairo, Egypt. After completing his Bachelors (1998-2002) with a specialization in Biotechnology from Faculty of Agriculture, Cairo University, Egypt, he was appointed as a Lecturer Assistant in Department of Genetics in the same faculty. In 2005, he had one year training course in Germany 'Plant Oriented Biotechnology and Biosafety' where he got theoretical and practical basis for different biotechnological techniques. He Obtained his master degree with a thesis entitled "QTL mapping on adaptive root and shoot traits under drought stress in malting Barley candidate introgression lines" from Leibniz Universität Hannover, Germany. His master thesis was under supervision of Prof. Ralf Uptmoor and Prof. Hartmut Stützel. On completion of his master degree (September 2008), he started his PhD at the Laboratory of Plant Genetics, Wageningen University and Research Center (WUR), The Netherlands. This thesis presents the outcome of his PhD research work.

List of publication

Mohamed El Soda, Satya Swathi Nadakuduti, Klaus Pillen and Ralf Uptmoor **(2010).** Stability parameter and genotype mean estimates for drought stress effects on root and shoot growth of wild barley pre-introgression lines. **Molecular Breeding** 26:583–593. (equally contributed first author)

Hedayat Bagheri, **Mohamed El-Soda**, Inge van Oorschot, Corrie Hanhart, Guusje Bonnema, Tanja Jansen-van den Bosch, Rolf Mank, Joost J. B. Keurentjes, Lin Meng, Jian Wu, Maarten Koornneef and Mark G. M. Aarts **(2012)**. Genetic analysis of morphological traits in a new, versatile, rapid-cycling *Brassica rapa* recombinant inbred line population. **Frontiers in Plant Sciences**. 3:183. doi: 10.3389/fpls.2012.00183 (equally contributed first author)

Hedayat Bagheri, **Mohamed El-Soda**, Hye Kyong Kim, Steffi Fritsche, Christian Jung, and Mark G.M. Aarts **(2013)**. Genetic analysis of health- related secondary metabolites in the new *Brassica rapa* recombinant inbred line population. **International Journal of Molecular Sciences** 14, 15561-15577.

Mohamed El-Soda, Martin P. Boer, Hedayat Bagheri, Corrie J. Hanhart, Maarten Koornneef and Mark G.M. Aarts. Genotype by environment interactions affecting pre-flowering physiological and morphological traits of *Brassica rapa* grown in control and drought environments. **(Submitted and in review in J of Exp Botany)**.

Mohamed El-Soda, Marcos Malosetti, Bas Zwaan, Maarten Koornneef and Mark G.M. Aarts. Genotype by environment interaction, QTL mapping and plant ecology: Lessons from Arabidopsis thaliana. **(Invited review to be submitted soon to Trends in Plant Science).**

Mohamed El-Soda, Willem Kruijer, Marcos Malosetti, Maarten Koornneef, Mark G.M. Aarts. Combining QTL and genome wide association mapping for genotype by environment interaction and plasticity associated with drought response in Arabidopsis thaliana. **(In preparation).**

Education Statement of the Graduate School

Experimental Plant Sciences

The Graduate School EXPERIMENTAL PLANT SCIENCES

Issued to: **Mohamed El-Soda**
Date: **28 October 2013**
Group: **Genetics - Group Koornneef, Wageningen University & Research Centre**

1) Start-up phase	*date*
► **First presentation of your project**	
BSM-QTL analysis of drought stress response in Arabidopsis thaliana	Mar 02, 2009
► **Writing or rewriting a project proposal**	
BSM-QTL analysis of drought stress response in Arabidopsis thaliana	Aug 2009
► **Writing a review or book chapter**	
Genotype by environment interaction and QTL mapping in plants. Lessons from Arabidopsis thaliana; *invited review, Trends in Plant Science*	Jul 2013
► **MSc courses**	
► **Laboratory use of isotopes**	
Subtotal Start-up Phase	*13.5 credits**

2) Scientific Exposure	*date*
► **EPS PhD student days**	
PhD Student Day, Naturalis Museum, Leiden (NL)	Feb 26, 2009
2nd International Retreat of PhD Students in Plant Sciences, Cologne (Germany)	Apr 15-17, 2010
PhD student day, Wageningen University, Wageningen (NL)	May 20 2011
3rd International Retreat of PhD Students in Plant Sciences, Paris (France)	Jul 05-08, 2011
EPS career day 2011, Wageningen University, Wageningen (NL)	Nov 2011
► **EPS theme symposia**	
EPS Theme 2 symposium 'Interactions between Plants and Biotic Agents', Utrecht University	Jan 22, 2009
EPS Theme I Symposium 'Developmental Biology of Plants', Leiden University	Jan 30, 2009
EPS Theme 3 Symposium 'Metabolism and Adaptation', University of Amsterdam	Feb 18, 2009
EPS Theme 4 Symposium: 'Genome Plasticity', Radboud University Nijmegen	Dec 11, 2009
EPS Theme I Symposium 'Developmental Biology of Plants', Wageningen University	Jan 28, 2010
EPS Theme 3 Symposium 'Metabolism and Adaptation', Wageningen University	Feb 19, 2010
EPS Theme 3 Symposium 'Metabolism and Adaptation', Wageningen University	Feb 09, 2011
EPS Theme 3 Symposium 'Metabolism and Adaptation', Utrecht University	Apr 26, 2012
► **NWO Lunteren days and other National Platforms**	
ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 06-07, 2009
ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 09-10, 2010
ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 04-05, 2011
ALW meeting 'Experimental Plant Sciences', Lunteren (NL)	Apr 02-03, 2012
► **Seminars (series), workshops and symposia**	
Dr. Sjef Smeekens "sugar signaling and stress"	Nov 27, 2008
QTLMAS workshop	Apr 20-21, 2009

Dr. Wallace A. Cowling 'Linkage disequilibrium and association mapping – helping to overcome the paradox of modern plant breeding'	Jun 26, 2009
Dr. Hiro Nonogaki "Seeds, microRNA and Darwin?"	Sep 17, 2009
Dr. Bertrand Hirel, "Improving nitrogen use efficiency in maize and other crops. Importance of plant phenomics for basic and agronomic studies"	Sep 21, 2009
Prof. Ulrich Schurr, "Phenotyping of dynamic structures and function by non-invasive technologies"	Sep 21, 2009
EPS symposium "Ecology and Experimental Plant Sciences 2"	Sep 22, 2009
Dr. Wim Sopp, "The molecular regulation of seed dormancy"	Oct 20, 2009
Dr Xiaowu Wang, "Brassica rapa genome assembly using short read sequencing technology"	Oct 20, 2009
Glenda Willems, "Genome-wide association mapping in Arabidopsis thaliana"	Jul 26, 2010
Dr. Adam Price, 'Studying the genetics of root growth in rice'	Sep 17, 2010
Jose Jimenez-Gomez, 'Next Generation Quantitative Genetics'	Nov 29, 2010
Ales Pecinka, 'Genome and epigenome stability under abiotic stress'	Nov 29, 2010
Eric Visser, 'Exploring roots selective root placement in nutrient-rich hotspots'	Dec 08, 2010
PhD discussion group: 'Plant Soil Interactions' Prof. Fred van Eeuwijk	Mar 15, 2011
Carol Wagstaff: Crop improvement, food processing and their relationship to human health	Oct 27, 2011
EPS Mini-symposium 'Plant Breeding in the genomics era'	Nov 25, 2011
Association mapping	Feb, 23 2012
► **Seminar plus**	
Seminar-plus: with Patrick Hussey and Fabio Fornara	Jan 28, 2010
Carol Wagstaff	Oct 27, 2011
► **International symposia and congresses**	
International Society of Root Research (ISRR) Meeting	Jun 26-29, 2012
Symposium "on Genetics of Plant Mineral Nutrition", Hanover Germany)	Sep 30-Oct 02, 2010
► **Presentations**	
Poster: 2nd International Retreat of PhD Students in Plant Sciences, Cologne (Germany)	Apr 15-17, 2010
Poster: ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 04-05, 2011
Oral presentation 3rd International Retreat of PhD Students in Plant Sciences, Paris (France)	Jul 05-08, 2011
International Society of Root Research (ISRR) Meeting	Jun 26-29 2012
Poster: ALW meeting 'Experimental Plant Sciences', Lunteren	Apr 02-03, 2012
► **IAB interview**	Feb 18, 2011
► **Excursions**	
Subtotal Scientific Exposure	*18.7 credits**

3) In-Depth Studies	*date*
► **EPS courses or other PhD courses**	
ETNA School 'Plant Phynotyping'	Nov 01-10, 2009
Biostatistics (basics)	May 2010
Course 'Mixed model based QTL mapping in Genstat'	May 14-16, 2012
PhD Summer School 'Natural Variation'	Aug 21-24 2012
► **Journal club**	
member of literature discussion group of Plant Genetics	2008-2012
► **Individual research training**	
Subtotal In-Depth Studies	*9.0 credits**

4) Personal development	*date*
► **Skill training courses**	
Course: Project- and Time Management	Jan 12,26 & Feb 20, 2010
Competence Assessment	Aug 2010
Course: Techniques for writing and presenting a scientific paper	Sep 05-07 2012
► **Organisation of PhD students day, course or conference**	
► **Membership of Board, Committee or PhD council**	
Subtotal Personal Development	*3.0 credits**

TOTAL NUMBER OF CREDIT POINTS*	**44.2**

Herewith the Graduate School declares that the PhD candidate has complied with the educational requirements set by the Educational Committee of EPS which comprises of a minimum total of 30 ECTS credits

** A credit represents a normative study load of 28 hours of study.*

Cover page design by **Pingping Huang** and **Kaige Zhao**
Printing: GVO drukkers & vormgevers B.V., Ede, The Netherlands

Propositions

1. Considering GxE in genome wide association mapping does not only allow better understanding of plant genetics, but also better understanding of plant physiology. (This thesis)

2. Studying plants above ground and ignoring plants below ground is never enough to describe plant performance especially under drought stress. (This thesis)

3. There is no single approach to understand the functional genes, interdisciplinary research should be the focus of the future.

4. Focusing on phenotyping public molecular data provides good opportunities for developing countries to participate in science.

5. Working in an open working places switches the morning working hours to social hours and the evening hours to working hours.

6. Having prepositions may distract from the content of a thesis.

7. Journals impact factors as well as university ranking are not good indicators to decide where to publish or where to study.

Propositions belonging to the PhD thesis:

Genetic analysis of drought stress response in *Arabidopsis thaliana* and *Brassica rapa*

Mohamed El-Soda

Wageningen, 28th October, 2013