# Genetic mapping of genotype-by-ploidy effects in *Arabidopsis thaliana*

Cris L. Wijnen[1,2], Frank F.M. Becker[1], Andries A. Okkersen[1], C. Bastiaan de Snoo[3], Martin P. Boer[2], Fred A. van Eeuwijk[2], Erik Wijnker[1], Joost J.B. Keurentjes[1]

[1] Laboratory of Genetics, Wageningen University & Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands. [2] Biometris, Wageningen University & Research, Droevendaalsesteeg 1, 6708 PB, Wageningen, The Netherlands. [3] Rijk Zwaan R&D Fijnaart, Eerste kruisweg 9, 4793 RS Fijnaart, The Netherlands.

## Abstract

Plants can express different phenotypic responses following polyploidization, but ploidy-dependent phenotypic variation has so far not been assigned to specific genetic factors. To map such effects, segregating populations at different ploidy levels are required. The availability of an efficient haploid-inducer line in *Arabidopsis thaliana* allows for the rapid development of large populations of segregating haploid offspring. Because Arabidopsis haploids can be self-fertilised to give rise to homozygous doubled haploids, the same genotypes can be phenotyped at both the haploid and diploid ploidy level. Here, we compared the phenotypes of recombinant haploid and diploid offspring derived from a cross between two late flowering accessions to map genotype x ploidy (GxP) interactions. Ploidy-specific quantitative trait loci (QTLs) were detected at both ploidy levels. This implies that mapping power will increase when phenotypic measurements of monoploids are included in QTL analyses. A multi-trait analysis further revealed pleiotropic effects for a number of the ploidy specific QTLs as well as opposite effects at different ploidy levels for general QTLs. Taken together, we provide evidence of genetic variation between different Arabidopsis accessions being causal for dissimilarities in phenotypic responses to altered ploidy levels, revealing a GxP effect. Additionally, by investigating a population derived from late flowering accessions we revealed a major vernalisation specific QTL for variation in flowering time, countering the historical bias of research in early flowering accessions.

## Introduction

Although common in some species, phenotypic effects caused by differences in genome ploidy level have so far been very elusive and difficult to study in the plant model *Arabidopsis thaliana*. Nonetheless, the impact of ploidy is illustrated by strong effects on quantitative traits such as salt and drought tolerance, and relative growth rate (Chao et al., 2013, Del Pozo and Ramirez-Parra, 2014,

Fort et al., 2016). Most attempts to reveal ploidy effects in Arabidopsis have used naturally occurring autotetraploid accessions such as Warschau-1 (Wa-1) (Henry et al., 2005, Orzechowska et al., 2016) or artificially induced tetraploids (Yu et al., 2009), which were compared to their diploid and triploid counterparts (Fort et al., 2016, Henry et al., 2005). However, to identify genetic factors that are causal for the observed differences in response to altered ploidy levels, segregating populations are required. While a biparental mapping population was developed using Wa-1 as one of the parental genotypes, it was only later discovered that this genotype was tetraploid and that the inbred lines were derived from triploids (Schiff et al., 2001, Henry et al., 2005). Therefore, the ploidy level segregated in this population and many of the genotypes are not explicitly diploid or tetraploid according to flowcytometry (Henry et al., 2005, Chao et al., 2013). Notwithstanding this unstable population, a mapping resource at different stable ploidy levels has not been developed systematically in Arabidopsis so far.

Monoploids (*i.e.,* individuals consisting of somatic cells containing only the basic number of chromosomes) are usually not taken into account in studies that investigate the effect of ploidy, although exceptions exist in maize (Guo et al., 1996), yeast (Galitski et al., 1999), potato (Stupar et al., 2007) and Chinese cabbage (Gu et al., 2016). These studies focused on transcriptional changes induced by ploidy changes in a single or a few genotypes. For instance, Stupar *et al.* (Stupar et al., 2007) demonstrated that more than 50% of the analysed genes displayed expression differences between monoploids and diploids or tetraploids, suggesting large developmental differences between plants of different ploidy levels. The discovery of a genome elimination mutant in Arabidopsis allowed the quick generation of haploid lines from diploid individuals (Ravi and Chan, 2010), enabling the analysis of ploidy effects in a genetic model species. The generation of haploids in Arabidopsis occurs through elimination of the mutant haploid inducer genome in the offspring of a cross between the mutant and a wild type diploid. Diploid Arabidopsis somatic cells contain 2n = 2x = 10 chromosomes, while haploids contain n = x = 5 chromosomes and thus are equivalent to monoploids. Arabidopsis monoploids are predominantly sterile and cannot be maintained as such. Haploid plants do, however, occasionally set seed, giving rise to homozygous doubled haploids (DHs). Spontaneous diploidization occurs during sexual reproduction due to fusion of euploid gametes resulting from incidental non-disjunction of all homologs at meiosis I or through somatic doubling of haploid cell lines, resulting in chimeric plants with fertile diploid branches. While the ploidy level of the maternally derived seed coat is determined by that of the mother plant, the embryo and endosperm ($2n = 2x$ and $n = 3x$, respectively) contain equal chromosome numbers in seeds derived from mono- or diploids. Doubled haploids thus contain a duplicated genome and consist again of diploid somatic cells containing 2n = 2x = 10 chromosomes, identical to wild type diploids.

When the genome elimination mutant is crossed with an $F_1$ hybrid of two distinct accessions, only the recombinant gametes of the hybrid will contribute to the genomes of the resulting monoploid offspring. By allowing the monoploids to produce DH seeds, the monoploid genome is immortalized in homozygous diploids, resembling recombinant inbred lines (RILs). The generation of such a diploid mapping population using genome elimination thus has the advantage that initially large amounts of segregating monoploids are produced, which except for the ploidy level are genetically identical to the DHs obtained in the next generation (Seymour et al., 2012, Wijnker et al., 2014, Filiault et al., 2017).

These monoploids may provide a useful additional resource for genetic mapping and allow assessment of ploidy effects in comparisons with their subsequent isogenic diploid offspring.

The generation of DH mapping populations has an advantage over the more commonly used RILs, which are typically generated from an $F_1$ individual through eight to ten generations of self-fertilization. This contrasts to DHs for which homozygous diploid populations can be obtained from an $F_1$ in only three generations (Crow, 2007, Seymour et al., 2012, Lister and Dean, 1993, Wijnen and Keurentjes, 2014). The advantage of the fast development of DH populations allows for the investigation of natural variation in late-flowering winter annual accessions, whereas most existing experimental biparental mapping populations are derived from summer annual accessions to shorten the generation time due to their early-flowering phenotype (Lister and Dean, 1993, Alonso-Blanco et al., 1998b, El-Lithy et al., 2006, Simon et al., 2008, O'Neill et al., 2008). Summer annuals germinate in spring and flower within a short period of time, while winter annuals germinate in autumn, survive winter as a rosette and typically flower only after vernalisation, a period of cold conditions. These differences may have a large impact on life history traits but despite an increase in genetic resources, including more accessions to represent the huge global genetic diversity of the species, a bias towards the use of early-flowering accessions remains (El-Lithy et al., 2006, O'Neill et al., 2008, Simon et al., 2008). Illustratively, although the haploid-inducer approach eliminates the need for a lengthy inbreeding process to obtain homozygous lines, the DH populations reported for Arabidopsis so far also originate from early-flowering accessions (Seymour et al., 2012, Wijnker et al., 2012, Fulcher and Riha, 2016, Filiault et al., 2017).

Here, we describe the development and phenotyping of a monoploid, and subsequent diploid mapping population derived from a cross between the two late-flowering accessions, T540 (Kävlinge, Sweden) and Ge-0 (Geneva, Switzerland). These accessions display large phenotypic differences in a number of life history traits. We investigated the diploid generation for the presence of a genotype-by-environment (GxE) effect by mapping variation in flowering time with and without vernalisation. We demonstrate that exploiting genetic variation in late-flowering accessions can increase our knowledge even in a well-studied trait like flowering time. Secondly, we investigated the possibility of detecting genotype-by-ploidy (GxP) interactions by performing a combined analysis across monoploids and diploids, using a multi-trait quantitative trait locus (QTL) model approach. As such we were able to detect ploidy specific QTLs and reveal genotype-by-ploidy interactions. Finally, we analysed all traits for pleiotropic QTLs, and demonstrate that most detected QTLs affect multiple traits at both ploidy levels, while only a minor number of QTLs affect predominantly a single trait at a specific ploidy level. Taken together, this study advocates the use of late flowering mapping populations to analyse so far unexploited genetic variation and provides evidence for genotype-by-ploidy interactions in Arabidopsis.

## Results

### Development and phenotyping of a mapping population at two ploidy levels

To explore the effect of ploidy on genetic mapping in Arabidopsis, a segregating population was generated from a cross between two late flowering accessions, T540 and Ge-0 (Figure 1). Briefly, the late flowering accessions T540 and Ge-0 were crossed to produce an $F_1$ hybrid. This hybrid was subsequently manually crossed to a haploid inducer line (Ravi and Chan, 2010), from which approximately 250 seeds were obtained. These seeds were stratified and pre-germinated, after which seedlings were transferred to Rockwool and grown for three weeks under long day conditions in a climate-controlled growth chamber. After visual inspection, 210 potentially haploid plants were transferred to a cold room for eight weeks vernalisation under short-day conditions. Once vernalised, plants were transferred to a greenhouse under long day conditions and subsequently formed inflorescences, flowered and set seeds. At the end of the growth period non-destructive phenotypes were measured, i.e., main stem length, branching from rosette and branching from the main inflorescence (Supplemental Table 1), allowing the monoploids to produce doubled haploid seeds. These seeds formed the subsequent diploid generation. The diploid seeds harvested from monoploid plants were also analysed for average seed size (Supplemental Table 1).

In a second experiment the 210 potential DH lines were grown in a climate chamber under similar conditions as described for the monoploids. Ten replicates of each line of the diploid population were grown in a completely randomised design. After three weeks, five of these were transferred to a greenhouse to record the time to flowering. The other five replicates were transferred to a cold room and vernalised for eight weeks at 4 °C. These plants were thereafter transferred back to the climate chamber with long day conditions and phenotyped for flowering time in addition to the life-history traits quantified in the monoploids (Supplemental Table 1). Assuming all replicate plants were isogenic, one plant of each genotype was selected for genotyping, which was successful for 195 lines (Supplemental Table 2). After analysis of genotypic and phenotypic data, 171 genotypes, for which phenotypic data at both ploidy levels could be obtained, were selected. The phenotypic data of these lines were used for all further analyses. The genotype data of these lines were used for the construction of a genetic map (Supplemental Table 3 and Supplemental Figure 1) and the QTL mapping of the analysed traits using standard methods.

In addition to the artificial haploid and DH mapping populations, a classical $F_2$ population of 400 lines derived from the same T540 x Ge-0 $F_1$ hybrid was generated and grown simultaneously with the doubled haploids in the second experiment. Half the population was subjected to vernalisation again, while the other half was left to flower (Supplemental Table 4). Moreover, a small set of 71 vernalised $F_2$s was genotyped with the same 123 markers as used to genotype the DHs and their linkage patterns were compared with those in the DH population to confirm no anomalies occurred during the DH development (Supplemental Table 5 and Supplemental Figure 1). With the exception of a slight genotype distortion at the top of chromosome 1 in the DH population (Supplemental Figure 2), no systematic differences were observed between the $F_2$ and the DH population. Importantly, the genetic maps generated from the two populations displayed an almost identical marker order, consistent with the known physical position of markers (Supplemental Figure 1).

**Detection of genetic variation controlling flowering time conditional on vernalisation**

While a vernalisation treatment can have a large overall phenotypic effect on the morphology and inflorescence structures of late flowering accessions (Huang et al., 2013, Lempe et al., 2005), mapping experiments in Arabidopsis have focussed on detecting QTLs in either early flowering populations or on mapping specific QTLs involved in vernalisation requirement with populations derived from parental accessions differing in this aspect (Simon et al., 2008, Alonso-Blanco et al., 1998b, O'Neill et al., 2008, El-Lithy et al., 2006, Ungerer et al., 2002). Here we have the opportunity to compare and map natural variation in flowering time with and without vernalisation in a late flowering segregating DH and $F_2$ population.

Although both parents of the DH and $F_2$ population are late flowering, they do not per se require vernalisation to flower. In our experiment without vernalisation T540 flowered on average after 101.8 days after sowing (DAS), while this was 86.5 days for Ge-0 (Table 1 and Supplemental Table 1). With vernalisation these accessions flowered on average after 19.1 and 14.2 days after transfer (DAT) from the cold, respectively. Similar data were obtained for the $F_2$ population (Table 1 and Supplemental Table 4). The variation in flowering time between the two accessions segregated in the diploid populations with only minor transgression in both conditions. Without vernalisation, the earliest line of the DH population flowered after 63 DAS, while the latest flowered at 123 DAS. With vernalisation the difference between extreme lines reduced to only ten days (12 and 22 DAT, respectively). The correlation in flowering time between the vernalised and non-vernalised plants was positive but far from absolute ($R^2 = 0.39$) (Supplemental Figure 4).

The data for flowering time of the doubled haploids under different vernalisation conditions allowed a multi-environment composite interval mapping (CIM) where the effect of vernalisation was investigated. Additionally, the $F_2$ population was screened for QTLs in vernalised conditions in a separate analysis. A total of seven QTLs spread over the genome were detected for variation segregating in the DH population (Table 2). Of these seven QTLs, three revealed an interaction with the environment, providing evidence for GxE effects of vernalisation. One QTL with a GxE effect was located on chromosome 4 and the other two were detected on chromosome 5. The QTL for flowering time after vernalisation in the middle of chromosome 4 had a normalized effect-size of 0.33 (indicating a positive contribution of the T540 allele), while this QTL was not significantly ($P = 0.653$) detected in the non-vernalised DH population. In contrast, both GxE QTLs on chromosome 5 were significant in both environments but with different effect-sizes (Table 2). The major QTL detected at the bottom of chromosome 5 revealed an additive normalized effect-size of 0.65 when the plants were vernalised while this was only 0.19 in the non-vernalised set. This major QTL was also detected in the vernalised $F_2$ population. The other GxE QTL detected in the middle of chromosome 5 in the DH population had a normalized effect-size of 0.43 in non-vernalised plants, while this was substantially lower (0.21) in vernalised plants. Additional QTLs without GxE effects were detected in the middle and at the bottom of chromosome 1 and chromosome 2 and at the top of chromosome 3 (Table 2). These results indicate that, conditional on the environment, genetic variation can have a variable impact on the time to flower in Arabidopsis.

**Effects of ploidy level on morphological variation**

To investigate if differences in ploidy have an effect on the phenotype, various morphological traits were quantified in the mono- and diploid generation of the recombinant lines generated from the cross T540 x Ge-0 (Figure 2 and Supplemental Table 6). Monoploid recombinant lines were on average much taller than their diploid counterparts (65 versus 46 cm, respectively) (Figure 2A and Supplemental Table 6). Illustrative for this difference in length is that more than 60% of the monoploids grew taller than the tallest diploid, which measured only 61 cm. In addition, branching from the rosette occurred much more frequently in monoploids (95.8%) than in diploids (34.9%), resulting in a larger average number of branches sprouting from the rosette and, similar to main stem length, more pronounced variation (Figure 2B; Supplemental Table 6). Illustratively, a maximum of only three rosette branches was observed in diploids, while monoploids developed on average seven branches from the rosette, with an exceptional maximum of twenty-three branches. In contrast to variation in main stem length and branching from rosette, the variation in branching from the stem spread around almost identical mean values at both ploidy levels, although a larger transgression was observed in the monoploids as compared to the diploids (Figure 2C). Despite the differences in phenotypic variation between the number of branches from the rosette and from the stem, these traits correlated positively in the monoploid population ($R^2$ = 0.32) (Supplemental Figure 3). This resulted in monoploids with up to a total number of thirty-two branches, giving rise to a bushy phenotype.

Similar to branching from the stem, the phenotypic variation in the size of seeds harvested from mono- or diploid plants centred around a comparable mean for both types of population, although the between-line variation was somewhat larger for seeds derived from diploids than for those derived from monoploids (Figure 2D and Supplemental Table 6). Positive Pearson correlations between mono- and diploids were observed for all traits (Supplemental Figure 3), but values remained moderate ($0.3 < R^2 < 0.4$). The replicate measurements of the diploid genotypes also allowed the assessment of trait heritabilities (Supplemental Figure 3). For most traits segregating in the diploid population moderate to high broad sense heritabilities were obtained ($0.30 < H^2 < 0.83$). This suggests that differences between mono- and diploids can be partly explained by simple additive ploidy effects but that the larger part of variation might be the result of more complex genotype-by-ploidy interactions.

**Effects of genotype-by-ploidy interaction on the detection of QTLs**

To determine whether differences in ploidy level had an effect on mappable genetic variation, each of the four traits measured in both the mono- and diploid population were subjected to trait specific dual-trait CIM, in which measurements at the two ploidy levels were considered to be different traits. Significant QTLs could be detected for each trait in both generations. In total fifteen QTLs were detected for the various traits, of which six displayed a significant interaction with the ploidy level (Table 3). Three genotype-by-ploidy QTLs were detected for main stem length, while one GxP QTL was detected for each of the other traits.

For main stem length five QTLs were detected in total, with a major QTL on the top of chromosome 5 and minor QTLs on chromosomes 3 and 4 (Table 3). The Ge-0 allele at the major QTL at chromosome 5 increased the stem length in the monoploids (normalized effect-size 0.49), whereas genotypic variation at this locus had no significant influence on the length of the diploids (effect-size

0.03). The QTL on chromosome 3 showed a similar pattern with a significant genotype-effect in the monoploids, although with smaller effect-size than the QTL on chromosome 5, but not in the diploids. Finally, on chromosome 4, three QTLs with overlapping support intervals spanning the entire chromosome and similar effect-signs were detected.

For both variation in branching from the rosette and branching from the main stem three QTLs were detected (Table 3). For variation in the number of branches from the rosette, QTLs were detected on the bottom of chromosomes 3 and 5 and the top of chromosome 5. The QTL on the bottom of chromosome 5 revealed a clear GxP interaction, as it was highly significant in the monoploids ($P < 0.001$) while it was not detected in the diploid generation ($P = 0.642$). The Ge-0 genotype at this QTL explained an increase in the number of branches in the monoploids, while a Ge-0 genotype at the two other QTLs decreased the number of branches from the rosette at both ploidy levels. Another GxP QTL was detected for variation in branching from the main stem on the middle of chromosome 5. This QTL was significantly detected in the diploids ($P < 0.001$) but not in the monoploids ($P = 0.204$). Similar to an increase in main stem length, Ge-0 alleles at any of these three QTLs increases the number of branches.

Finally, four QTLs were detected for variation in seed area, of which a GxP interaction was identified for the QTL on the middle of chromosome 3 (Table 3). This QTL was significantly detected in the monoploids ($P = 0.002$) but not in the diploids ($P = 0.224$). However, this QTL exerted only a minor effect. Another QTL on chromosome 3 was significantly detected in both generations, although it was much weaker in the diploids ($P = 0.043$) and a large difference in the effect-size of the QTL was observed (0.42 and 0.17 for monoploids and diploids, respectively). The results of the dual ploidy QTL analysis clearly indicate that differences in ploidy do not affect every genotype and trait equally. Indeed, strong GxP QTLs explain for a large part the phenotypic differences observed between genotypes and ploidy levels.

**Pleiotropic effects of genotype-by-ploidy interactions**

A weak to moderate correlation could be observed between values of the different morphological traits measured in the two isogenic populations of different ploidy (Supplemental Figure 3). These relationships suggest a partial co-regulation of traits. Indeed, we detected QTLs at similar positions for multiple traits (Table 3). We, therefore, subjected the various traits measured in the monoploids and diploids after vernalisation to a single multi-trait CIM analysis to identify possible co-location of QTLs. A total of nine QTLs were detected using this approach (Figure 3 and Supplemental Table 7). None of these QTLs were trait-specific and only the minor QTLs on the bottom of chromosomes 1 and 4 were ploidy-specific ($P < 0.01$), although suggestive QTLs ($P < 0.05$) were detected for other traits or at the other ploidy level as well (Supplemental Table 7).

A minor QTL on the lower arm of chromosome 3 significantly ($P < 0.01$) explained variation in all monoploid traits but only in branching from the rosette in the diploids. The T540 allele at this locus increases the number of branches from the rosette in the monoploids and diploids, even though the diploids did not display a large variation for this trait. Additionally, the same T540 allele causes an increase in branches from the stem in monoploids. However, the same allele decreases main stem

length and seed size of the monoploids. Additional minor to moderate QTLs co-locating on the lower arm of chromosomes 2 and 4 and in the middle of chromosome 3 were detected, explaining variation in multiple traits in both the mono- and diploids. The sign and effect-size of these coinciding QTLs was in line with the observed correlation between these traits (Supplemental Figure 3).

By far the strongest and largest number of QTLs was detected on chromosome 5. Strong QTLs for variation in main stem length and rosette branching in the monoploids coincided at the top of the chromosome, although with opposite effect-sign (Supplemental Table 7). Another strong QTL for variation in the size of seeds derived from monoploids at 61.2 cM coincided with highly significant QTLs for variation in stem branching, flowering time after vernalisation and main stem length of diploids. Finally, close to the end of the chromosome (121.7 cM), a strong QTL for variation in main stem length and branching of the monoploids co-located with a QTL for variation in flowering time after vernalisation and branching from the rosette of diploids. The Ge-0 allele at this locus increased all trait values except flowering time after vernalisation, which was delayed by the T540 allele.

Since genetic variation at the two QTLs at the top and bottom of chromosome 5 has the strongest effect on branching and main stem length (in addition to flowering time in the diploids) we analysed the effect of each of the four possible haplotypes in both the monoploid and diploid populations. Reflecting the absence of a significantly detected QTL for variation in stem length and branching at the top of chromosome 5 in the diploids, genotypic variation at the two QTLs had a much stronger effect on the monoploids (Supplemental Figure 4). This clearly indicates that the effect of genetic variation can be much stronger in monoploids than in diploids (Figure 2).

## Discussion

### Application of a late flowering doubled haploid mapping population

It is well known that different accessions of Arabidopsis respond differently to environmental conditions (Lempe et al., 2005, Koornneef et al., 2004). For instance, day-length sensitivity and vernalisation requirement determine for a large part the discrimination between winter- and summer-annuals (Romera-Branchat et al., 2014, Andres and Coupland, 2012). Moreover, when mapping populations are subjected to short or long day-length conditions with or without vernalisation, differences in the number and strength of detected flowering time QTLs can be observed (Alonso-Blanco et al., 1998a). The use of a haploid-inducer line in this study allowed the generation of a homozygous mapping population from underexploited late flowering accessions. As such, a diploid population could be developed in only three generations. For this population, QTL mapping for variation in flowering time in two different environments (i.e., with and without vernalisation) was performed.

In addition to a number of minor QTLs, a major QTL for variation in flowering time of vernalised plants was detected near the previously described and identified *VERNALISATION INSENSITIVE 3* (*VIN3*; At5g57830) locus at the bottom of chromosome 5 (Dittmar et al., 2014, Grillo et al., 2013, Alonso-Blanco et al., 1998a). Previously, variation in flowering time associated with this

locus was explained by an indel of three nucleotides within an exon of *VIN3* (Grillo et al., 2013). However, this indel is not polymorphic for Ge-0 and T540, although multiple other SNPs differentiate the intronic and promotor region of *VIN3* of these accessions, including 28 nucleotides deleted from the T540 *VIN3* promotor compared to Ge-0 (Supplemental Figure 5).

A second gene, *REDUCED VERNALISATION RESPONSE 2* (*VRN2*; At4g16845), related to response to vernalisation (Gendall et al., 2001), is located within the support interval of a QTL for variation in flowering time after vernalisation, detected on chromosome 4. The *VRN2* protein mediates vernalisation through interaction with the Polycomb Group (PcG) protein complex including *VIN3* (Bastow et al., 2004, Sung and Amasino, 2004). This PcG complex is known to interact with, and cause the stable reduction of the expression levels of, the floral repressor *FLOWERING LOCUS C* (*FLC*; At5g10140) (Bastow et al., 2004, Sung and Amasino, 2004), which collocates with the position of a flowering time QTL on the top of chromosome 5. This QTL was also detected for variation in main stem length, which strongly suggests a pleiotropic effect on the inflorescence architecture and flowering pathways, previously attributed to *FLC* (Huang et al., 2013).

The detection of flowering time QTLs in a segregating mapping population of late flowering accessions, especially after vernalisation, clearly identifies major QTLs other than those usually associated with flowering time variation in early accessions. This suggests that the regulation of flowering time in late accessions is controlled by variation at other loci than those in early flowering accessions (*e.g.*, *FRI* and *FLC*). It is likely that flowering time is not the only trait that discriminates summer annuals from winter annuals, which advocates the analysis of traits in late flowering populations in addition to the abundantly available early flowering populations.

**Effects of haploidisation on phenotypic variation**

Exploiting the availability of a mono- and diploid Arabidopsis mapping population, QTL analyses were applied to map and compare possible ploidy-dependent effects. A dual-trait CIM analysis resulted in the detection of six QTLs with a GxP interaction, while additional QTLs showed large differences in effect-sizes at either ploidy level. An obvious explanation for the GxP QTLs is that monoploid plants are sterile due to unbalanced segregation of the chromosomes during meiosis. Indeed, although not explicitly quantified, monoploids displayed an extended period of flowering compared to fertile diploids, possibly causing the increase in main stem length. Similarly, the development of exceptionally high numbers of rosette branches increases the total number of flowers produced (Ehrenreich et al., 2007). This suggests that the plants attempt to compensate for the lack of viable seed production by an increase in reproductive tissue formation, implying that the QTLs detected specifically for monoploids might be involved in the response to sterility. A similar phenomenon of additional branch formation has been described for the male sterile Landsberg *erecta* mutant (*ms1*-L*er*) (Hensel et al., 1994). Nonetheless, QTLs explaining the observed variation in response to haploidisation were detected, indicating natural variation for the strength of ploidy effects.

The antagonistic effect of the QTL on the top of chromosome 5 for either additional rosette branch formation (inferred by the T540 allele) or taller growth (inferred by the Ge-0 allele) implies that both accessions follow a different morphological approach to achieve a similar increase in the number

of flowers. The fact that a single QTL is identified for variation in rosette branching and main stem length might be due to one of the many pleiotropic genes that function in the control of inflorescence architecture (Rameau et al., 2015). Possible candidate genes may be part of the florigen gene family (Wickland and Hanzawa, 2015) which is known to function as a mobile flowering time switch. For instance, *FLOWERING LOCUS T* (*FT*; At1g65480) and *TWIN SISTER OF FT* (*TSF*; AT4G20370), are known to function in both flower induction and shoot branching pathways (Hiraoka et al., 2012). Another member of the same gene family, *TERMINAL FLOWERING 1* (*TFL1*; At5g03840), is located within the support interval of the QTL at the top of chromosome 5 and has been shown to be involved in flowering architecture (Baumann et al., 2015). Although no variation within the *TFL1* coding sequence could be observed between the two accessions, several SNPs and possibly deletions within the promotor region of the T540 allele might cause a differential expression of this gene (Supplemental Figure 6A). Assuming that flowering architecture is not influenced by *VIN3*, an alternative candidate explaining the effect of the QTL at the bottom of chromosome 5 on both branching and main stem length is *AUXIN RESPONSE FACTOR2* (*ARF2*; At5g62000), which is involved in multiple developmental processes via cell proliferation (Okushima et al., 2005, Schruff et al., 2006). Sequence-based evidence suggests that T540 and Ge-0 possess functionally different alleles (Supplemental Figure 6B). Moreover, a knock-down of *ARF2* leads to an increase in stem length and a sterile phenotype (Okushima et al., 2005).

Other QTLs, such as the one explaining variation in the size of seeds on the top of chromosome 3, coincide with likely candidate genes as well. This QTL has been identified previously as *HAIKU* 2 and was associated to a gene (*IKU2*; At3g19700) in the endosperm growth pathway (Luo et al., 2005). In addition, a monoploid specific QTL on chromosome 2 explaining variation in branching from the stem coincides with the previously identified *AGAMOUS-LIKE 6* gene (*AG6*; a.k.a. *REDUCED SHOOT BRANCHING 1*; AT2G45650), to which pleiotropic phenotypic effects on both the flowering and branching pathways have been previously attributed (Huang et al., 2012).

Although sterility might be causal for some of the GxP interactions of the QTLs, it is possible that other molecular processes are of influence as well. In previous studies on ploidy series including monoploids, performed in maize (Guo et al., 1996), yeast (Galitski et al., 1999), potato (Stupar et al., 2007) and Chinese cabbage (Gu et al., 2016), differentially expressed genes were identified at different ploidy levels, indicating a specific sensitivity to ploidy, instead of sterility. Moreover, in a dosage series (*x* , 2*x*, 4*x*) of maize inbred lines (Riddle et al., 2006), genetic background and ploidy was suggested to interact. Further evidence for GxP interactions independent of sterility come from an RNA-seq comparison of diploid and tetraploid Arabidopsis accessions, in which the accessions Col-0 and L*er*-0 displayed different numbers of upregulated genes at the tetraploid level (Yu et al., 2010). In both studies it was argued that the altered nuclear surface to volume ratio might have caused the differential expression of genes. However, clear mechanisms explaining how these altered ratio's cause gene expression differences are so far elusive. Despite the uncertainty of the possible mechanisms of GxP interactions it is clear that the mapping of quantitative traits in mono- and diploids can reveal additional variation, which might be instrumental in the elucidation of the genetic regulation of complex traits.

## Experimental Procedures

### Population development

Two late flowering accessions, T540 (CS76239) from Sweden and Ge-0 (CS76135) from Switzerland were selected based on phenotypic differences and expected unexplored genotypic differences compared to widely used early flowering accessions. These accessions were crossed to produce a biparental hybrid $F_1$. The $F_1$ (T540 x Ge-0) was used as a pollen donor and crossed to the GFP-tailswap haploid-inducer line to generate monoploid offspring (Ravi and Chan, 2010). From these crosses, 250 viable seeds were sown and 210 putative monoploid lines were selected based on morphology during growth (Wijnker et al., 2014). Spontaneous genome doubling in the monoploids followed by selfing created a set of 171 unique diploid homozygous lines.

In addition to the generation of the doubled haploid lines, the $F_1$ was selfed to generate a batch of $F_2$ seeds.

### Plant growth conditions

All seeds from a cross between the $F_1$ hybrid (T540 x Ge-0) and the GFP-tailswap line were sown on ½ MS agar plates without sucrose. The seeds on these plates were stratified for four days at 4°C in darkness and subsequently placed in a climate chamber at 25°C with a diurnal cycle of 16 hours of light and 8 hours of darkness to induce seed germination. After two days of pre-germination, only potential monoploid seedlings were transplanted to wet Rockwool blocks of 4 x 4 cm in a climate chamber (16h LD, 125 µmolm-2s-1, 70% RH, 20/18˚C day/night cycle). All plants were watered three days per week for 5 min with 1/1000 Hyponex solution (Hyponex, Osaka, Japan) using flooding tables. Here they remained for three weeks to allow growth before vernalisation. Vernalisation was performed for eight weeks (12h LD, 125 µmolm-2s-1, 70% RH, 4˚C constant). After vernalisation, plants were transferred to the greenhouse where they were allowed to flower and mature. Monoploid plants were selected based on morphology as described before (Wijnker et al., 2014). Subsequently, diploid seeds were harvested after recording phenotypic traits of the monoploids.

The second experiment included ten replicates for each of 210 assumed diploids. These were stratified on wet filter paper in similar conditions as the agar plates of the previous experiment. Subsequently, five of the seedlings were grown similar to the monoploids, including three weeks growth in long day conditions and vernalisation for eight weeks, while the five other replicates were transferred to the greenhouse. The five replicates in the greenhouse conditions were allowed to grow in a completely randomized design without vernalisation for a maximum of 100 days after transfer or until flowering or senescence. The five diploids that underwent vernalisation remained in climate chambers with similar conditions as pre-vernalisation (16h LD, 125 µmolm-2s-1, 70% RH, 20/18˚C day/night cycle). The plants were randomized in a completely randomized design where they were allowed to grow for a maximum of ninety days.

Similarly, 400 $F_2$ plants were grown, of which 200 were vernalised as described above. Of the vernalised plants 71 $F_2$s were selected for genotyping.

**Phenotypic measurements**

The monoploids were phenotyped for the number of branches from the rosette and branching from the stem, main stem length (cm) and seed area (approximately 100 seeds were taken three times from the same storage bag for three separate photos, these were analysed for seed area). For the second experiment, the same four phenotypes were measured. However, now also flowering time before and after vernalisation was included as a phenotype. Flowering time without vernalisation was measured as the number of days after planting until the first flower on the main stem opened its petals. Flowering time with vernalisation was measured as the number of days after vernalisation until the first flower on the main stem opened its petals. Plants that did not germinate or that died within the period of the experiment were discarded. For the plants used for genotyping only flowering time was recorded, as taking a flower head, used for extracting DNA, from the plant might influence the other traits. All the monoploid and $F_2$ phenotypes are based on a single observation per genotype, while for the DH population, which were measured with five replicates, the reported values are the means.

**Genotyping of the populations**

For 210 doubled haploids and 71 $F_2$s the DNA was extracted from flower heads by applying a CTAB DNA extraction protocol which was adapted for use on 96 well plates. Genotyping was performed using a GoldenGate Assay from Illumina, using 384 SNP markers. Of those, 142 markers were polymorphic for the two parental lines. Of these 142, only 114 markers showed nonredundant recombination patterns for either the diploids or $F_2$s. Nine additional KASPar markers (KBiosciences) were included to a total of 123 markers (Smith and Maughan, 2015). From 210 selected DH lines, 195 were successfully genotyped and only four were discarded because of too much heterozygosity or missing data. Eventually, only 171 DHs were used for the final analyses because of redundant genotypes and lack of data in either mono- or diploid generation.

**Genetic map comparison of the doubled haploid and $F_2$ population**

To confirm no anomalies were present in the doubled haploids, a comparison with an $F_2$ population was performed. Individual lines from both populations were genotyped and genetic maps were generated. A subset of 71 $F_2$s and 171 DHs were successfully genotyped. Genetic maps were constructed for both the $F_2$ and the DHs independently using Kosambi's regression mapping function in JoinMap 6.1 (Kyazma). Segregation distortions were determined by GenStat 19th edition. The DH map was also used for the genetic mapping in monoploids.

**Statistical analyses and QTL mapping**

Pearson correlations between traits were calculated using the cor function in R. The broad-sense heritabilities of the doubled haploids were calculated in R using the repeatability function of the heritability package (Kruijer et al., 2015). QTL analyses were performed using GenStat (19$^{th}$ edition)

(Boer et al., 2015), where mean phenotypic values per DH line were used and single observations in the case of the monoploids and $F_2$s. In order to have a maximum QTL-effect and QTL x E or QTL x Ploidy interaction detection, we first analysed the separate traits using single-trait multiple environment composite interval mapping (where either vernalisation or the ploidy level was considered as the environment). The final analyses encompassed a multi-trait single environment analyses, including all traits measured after vernalisation. First an initial analysis of simple interval mapping was performed with a maximum step size of 5 cM along the genome. Other settings were kept as default (maximum cofactor proximity = 50 cM; minimum distance for QTL selection = 30 cM; threshold for genome-wide significance level = $\alpha = 0.05$). After these first analyses, markers associated with candidate QTLs were automatically set as cofactors for the composite interval mapping. The QTLs that resulted from this scan were tested for interaction effects in the selection of a final QTL model.

## Acknowledgements

We like to express our gratitude to G. Stunnenberg and T. Stoker of Wageningen University & Research for technical assistance during experimental work and E. Kerdaffrec and M. Nordborg, of the Gregor Mendel Institute in Vienna, Austria for performing the golden-gate genotyping assays. This work has been financially supported by the Netherlands Organisation for Scientific Research under grant numbers STW-12425 for which also support was received from Rijk Zwaan B.V..

# References


ALONSO-BLANCO, C., EL-ASSAL, S. E., COUPLAND, G. & KOORNNEEF, M. 1998a. Analysis of natural allelic variation at flowering time loci in the Landsberg erecta and Cape Verde Islands ecotypes of Arabidopsis thaliana. *Genetics,* 149**,** 749-64.

ALONSO-BLANCO, C., PEETERS, A. J. M., KOORNNEEF, M., LISTER, C., DEAN, C., VAN DEN BOSCH, N., POT, J. & KUIPER, M. T. R. 1998b. Development of an AFLP based linkage map of Ler, Col and Cvi Arabidopsis thaliana ecotypes and construction of a Ler/Cvi recombinant inbred line population. *The Plant Journal,* 14**,** 259-271.

ANDRES, F. & COUPLAND, G. 2012. The genetic basis of flowering responses to seasonal cues. *Nat Rev Genet,* 13**,** 627-39.

BASTOW, R., MYLNE, J. S., LISTER, C., LIPPMAN, Z., MARTIENSSEN, R. A. & DEAN, C. 2004. Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature,* 427**,** 164-167.

BAUMANN, K., VENAIL, J., BERBEL, A., DOMENECH, M. J., MONEY, T., CONTI, L., HANZAWA, Y., MADUENO, F. & BRADLEY, D. 2015. Changing the spatial pattern of TFL1 expression reveals its key role in the shoot meristem in controlling Arabidopsis flowering architecture. *Journal of Experimental Botany,* 66**,** 4769-4780.

BOER, M., CAVE, V., JANSEN, H., MALOSETTI, M., MATHEWS, K., MURRAY, D., VAN EEUWIJK, F. & WELHAM, S. 2015. *A Guide to QTL Analysis in Genstat,* Hertfordshire, UK, VSN International

CHAO, D.-Y., DILKES, B., LUO, H., DOUGLAS, A., YAKUBOVA, E., LAHNER, B. & SALT, D. E. 2013. Polyploids Exhibit Higher Potassium Uptake and Salinity Tolerance in Arabidopsis. *Science,* 341**,** 658-659.

CROW, J. F. 2007. Haldane, Bailey, Taylor and recombinant-inbred lines. *Genetics,* 176**,** 729-32.

DEL POZO, J. C. & RAMIREZ-PARRA, E. 2014. Deciphering the molecular bases for drought tolerance in Arabidopsis autotetraploids. *Plant, Cell & Environment,* 37**,** 2722-2737.

DITTMAR, E. L., OAKLEY, C. G., ÅGREN, J. & SCHEMSKE, D. W. 2014. Flowering time QTL in natural populations of Arabidopsis thaliana and implications for their adaptive value. *Molecular Ecology,* 23**,** 4291-4303.

EHRENREICH, I. M., STAFFORD, P. A. & PURUGGANAN, M. D. 2007. The Genetic Architecture of Shoot Branching in *Arabidopsis thaliana*: A Comparative Assessment of Candidate Gene Associations vs Quantitative Trait Locus Mapping. *Genetics,* 176**,** 1223-1236.

EL-LITHY, M. E., BENTSINK, L., HANHART, C. J., RUYS, G. J., ROVITO, D., BROEKHOF, J. L. M., VAN DER POEL, H. J. A., VAN EIJK, M. J. T., VREUGDENHIL, D. & KOORNNEEF, M. 2006. New Arabidopsis Recombinant Inbred Line Populations Genotyped Using SNPWave and Their Use for Mapping Flowering-Time Quantitative Trait Loci. *Genetics,* 172**,** 1867-1876.

FILIAULT, D. L., SEYMOUR, D. K., MARUTHACHALAM, R. & MALOOF, J. N. 2017. The Generation of Doubled Haploid Lines for QTL Mapping. *Methods Mol Biol,* 1610**,** 39-57.

FORT, A., RYDER, P., MCKEOWN, P. C., WIJNEN, C., AARTS, M. G., SULPICE, R. & SPILLANE, C. C. 2016. Disaggregating polyploidy, parental genome dosage and hybridity contributions to heterosis in Arabidopsis thaliana. *New Phytologist,* 209**,** 590-599.

FULCHER, N. & RIHA, K. 2016. Using Centromere Mediated Genome Elimination to Elucidate the Functional Redundancy of Candidate Telomere Binding Proteins in Arabidopsis thaliana. *Frontiers in Genetics,* 6.

GALITSKI, T., SALDANHA, A. J., STYLES, C. A., LANDER, E. S. & FINK, G. R. 1999. Ploidy Regulation of Gene Expression. *Science,* 285**,** 251-254.

GENDALL, A. R., LEVY, Y. Y., WILSON, A. & DEAN, C. 2001. The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis. *Cell,* 107**,** 525-535.

GRILLO, M. A., LI, C., HAMMOND, M., WANG, L. & SCHEMSKE, D. W. 2013. Genetic architecture of flowering time differentiation between locally adapted populations of Arabidopsis thaliana. *New Phytologist,* 197**,** 1321-1331.

GU, A. X., ZHAO, J. J., LI, L. M., WANG, Y. H., ZHAO, Y. J., HUA, F., XU, Y. C. & SHEN, S. X. 2016. Analyses of phenotype and ARGOS and ASY1 expression in a ploidy Chinese cabbage series derived from one haploid. *Breeding Science,* 66**,** 161-168.

GUO, M., DAVIS, D. & BIRCHLER, J. A. 1996. Dosage Effects on Gene Expression in a Maize Ploidy Series. *Genetics,* 142**,** 1349-1355.

HENRY, I. M., DILKES, B. P., YOUNG, K., WATSON, B., WU, H. & COMAI, L. 2005. Aneuploidy and Genetic Variation in the Arabidopsis thaliana Triploid Response. *Genetics,* 170**,** 1979-1988.

HENSEL, L. L., NELSON, M. A., RICHMOND, T. A. & BLEECKER, A. B. 1994. The Fate of Inflorescence Meristems Is Controlled by Developing Fruits in Arabidopsis. *Plant Physiology,* 106**,** 863-876.

HIRAOKA, K., YAMAGUCHI, A., ABE, M. & ARAKI, T. 2012. The Florigen Genes FT and TSF Modulate Lateral Shoot Outgrowth in Arabidopsis thaliana. *Plant and Cell Physiology,* 54**,** 352-368.

HUANG, X., DING, J., EFFGEN, S., TURCK, F. & KOORNNEEF, M. 2013. Multiple loci and genetic interactions involving flowering time genes regulate stem branching among natural variants of Arabidopsis. *New Phytologist,* 199**,** 843-857.

HUANG, X., EFFGEN, S., MEYER, R. C., 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.

KOORNNEEF, M., ALONSO-BLANCO, C. & VREUGDENHIL, D. 2004. Naturally occurring genetic variation in *Arabidopsis thaliana*. *Annual Review of Plant Biology,* 55**,** 141-172.

KRUIJER, W., BOER, M. P., MALOSETTI, M., FLOOD, P. J., ENGEL, B., KOOKE, R., KEURENTJES, J. J. B. & VAN EEUWIJK, F. A. 2015. Marker-based estimation of heritability in immortal populations. *Genetics,* 199**,** 379-98.

LEMPE, J., BALASUBRAMANIAN, S., SURESHKUMAR, S., SINGH, A., SCHMID, M. & WEIGEL, D. 2005. Diversity of flowering responses in wild Arabidopsis thaliana strains. *PLoS Genet,* 1**,** 109-18.

LISTER, C. & DEAN, C. 1993. Recombinant inbred lines for mapping RFLP and phenotypic markers in Arabidopsis thaliana. *The Plant Journal,* 4**,** 745-750.

LUO, M., DENNIS, E. S., BERGER, F., PEACOCK, W. J. & CHAUDHURY, A. 2005. MINISEED3 (MINI3), a WRKY family gene, and HAIKU2 (IKU2), a leucine-rich repeat (LRR) KINASE gene, are regulators of seed size in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 17531-17536.

O'NEILL, C. M., MORGAN, C., KIRBY, J., TSCHOEP, H., DENG, P. X., BRENNAN, M., ROSAS, U., FRASER, F., HALL, C., GILL, S. & BANCROFT, I. 2008. Six new recombinant inbred populations for the study of quantitative traits in Arabidopsis thaliana. *Theoretical and Applied Genetics,* 116**,** 623-634.

OKUSHIMA, Y., MITINA, I., QUACH, H. L. & THEOLOGIS, A. 2005. AUXIN RESPONSE FACTOR 2 (ARF2): a pleiotropic developmental regulator. *The Plant Journal,* 43**,** 29-46.

ORZECHOWSKA, M., GURDEK, S., SIWINSKA, D. & PIEKARSKA-STACHOWIAK, A. 2016. Cytogenetic characterization of the Arabidopsis thaliana natural tetraploid ecotype Warschau stability during in vitro regeneration. *Plant Cell, Tissue and Organ Culture (PCTOC),* 126**,** 553-560.

RAMEAU, C., BERTHELOOT, J., LEDUC, N., ANDRIEU, B., FOUCHER, F. & SAKR, S. 2015. Multiple pathways regulate shoot branching. *Frontiers in Plant Science,* 5.

RAVI, M. & CHAN, S. W. L. 2010. Haploid plants produced by centromere-mediated genome elimination. *Nature,* 464**,** 615-618.

RIDDLE, N. C., KATO, A. & BIRCHLER, J. A. 2006. Genetic variation for the response to ploidy change in Zea mays L. *Theoretical and Applied Genetics,* 114**,** 101-111.

ROMERA-BRANCHAT, M., ANDRES, F. & COUPLAND, G. 2014. Flowering responses to seasonal cues: what's new? *Curr Opin Plant Biol,* 21**,** 120-127.

SCHIFF, C. L., WILSON, I. W. & SOMERVILLE, S. C. 2001. Polygenic powdery mildew disease resistance in Arabidopsis thaliana: quantitative trait analysis of the accession Warschau-1. *Plant Pathology,* 50**,** 690-701.

SCHRUFF, M. C., SPIELMAN, M., TIWARI, S., ADAMS, S., FENBY, N. & SCOTT, R. J. 2006. The AUXIN RESPONSE FACTOR 2 gene of *Arabidopsis* links auxin signalling, cell division, and the size of seeds and other organs. *Development,* 133**,** 251-261.

SEYMOUR, D. K., FILIAULT, D. L., HENRY, I. M., MONSON-MILLER, J., RAVI, M., PANG, A., COMAI, L., CHAN, S. W. L. & MALOOF, J. N. 2012. Rapid creation of Arabidopsis doubled haploid lines for quantitative trait locus mapping. *Proceedings of the National Academy of Sciences,* 109**,** 4227-4232.

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. M. & MAUGHAN, P. J. 2015. SNP genotyping using KASPar assays. *Methods Mol Biol,* 1245**,** 243-56.

STUPAR, R. M., BHASKAR, P. B., YANDELL, B. S., RENSINK, W. A., HART, A. L., OUYANG, S., VEILLEUX, R. E., BUSSE, J. S., ERHARDT, R. J., BUELL, C. R. & JIANG, J. 2007. Phenotypic and Transcriptomic Changes Associated With Potato Autopolyploidization. *Genetics,* 176**,** 2055-2067.

SUNG, S. & AMASINO, R. M. 2004. Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature,* 427**,** 159-164.

UNGERER, M. C., HALLDORSDOTTIR, S. S., MODLISZEWSKI, J. L., MACKAY, T. F. C. & PURUGGANAN, M. D. 2002. Quantitative trait loci for inflorescence development in Arabidopsis thaliana. *Genetics,* 160**,** 1133-1151.

WICKLAND, DANIEL P. & HANZAWA, Y. 2015. The FLOWERING LOCUS T/TERMINAL FLOWER 1 Gene Family: Functional Evolution and Molecular Mechanisms. *Molecular Plant,* 8**,** 983-997.

WIJNEN, C. L. & KEURENTJES, J. J. B. 2014. Genetic resources for quantitative trait analysis: novelty and efficiency in design from an Arabidopsis perspective. *Curr Opin Plant Biol,* 18**,** 103-9.

WIJNKER, E., DEURHOF, L., VAN DE BELT, J., DE SNOO, C. B., BLANKESTIJN, H., BECKER, F., RAVI, M., CHAN, S. W. L., VAN DUN, K., LELIVELT, C. L. C., DE JONG, H., DIRKS, R. & KEURENTJES, J. J. B. 2014. Hybrid recreation by reverse breeding in Arabidopsis thaliana. *Nat. Protocols,* 9**,** 761-772.

WIJNKER, E., VAN DUN, K., DE SNOO, C. B., LELIVELT, C. L. C., KEURENTJES, J. J. B., NAHARUDIN, N. S., RAVI, M., CHAN, S. W. L., DE JONG, H. & DIRKS, R. 2012. Reverse breeding in Arabidopsis thaliana generates homozygous parental lines from a heterozygous plant. *Nat Genet,* 44**,** 467-470.

YU, Z., HAAGE, K., STREIT, V. E., GIERL, A. & TORRES RUIZ, R. A. 2009. A large number of tetraploid Arabidopsis thaliana lines, generated by a rapid strategy, reveal high stability of neo-tetraploids during consecutive generations. *Theoretical and Applied Genetics,* 118**,** 1107-1119.

YU, Z., HABERER, G., MATTHES, M., RATTEI, T., MAYER, K. F. X., GIERL, A. & TORRES-RUIZ, R. A. 2010. Impact of natural genetic variation on the transcriptome of autotetraploid Arabidopsis thaliana. *Proceedings of the National Academy of Sciences,* 107**,** 17809-17814.

633

**Table 1:** Flowering time of parental accessions and their derived populations with and without vernalisation.

FTv, flowering time after vernalisation (days after transfer); FTnv, flowering time without vernalisation (days after sowing); s.d., standard deviation; Cv, coefficient of variation.

| Trait | Genotype | Mean (n) | s.d. | Min | Max | Cv (%) |
|---|---|---|---|---|---|---|
| FTv | Ge-0 | 14.2 (29) | 1.17 | 13 | 17 | 8.2 |
| | T540 | 19.1 (22) | 1.50 | 15 | 22 | 7.8 |
| | F1 | 17.1 (18) | 0.80 | 16 | 18 | 4.7 |
| | F2 | 15.4 (172) | 1.34 | 13 | 20 | 8.7 |
| | DH | 15.9 (171) | 1.57 | 12 | 22 | 9.8 |
| FTnv | Ge-0 | 86.5 (8) | 16.27 | 64 | 107 | 18.8 |
| | T540 | 101.8 (4) | 14.08 | 89 | 121 | 13.8 |
| | F2 | 87.6 (180) | 16.41 | 60 | 130 | 18.7 |
| | DH | 88.4 (163) | 12.09 | 63 | 123 | 13.7 |

634

635

**Table 2:** QTL detection for flowering time with and without vernalisation in a DH and $F_2$ population. QTL positions are presented in cM with support intervals between brackets. -$LOG_{10}$(P) indicates the significance of the QTL for the combined treatments, while the P-value provides the specific P-value for each treatment. FT_nv, flowering time without vernalisation (days after sowing); FT_v, flowering time after vernalisation (days after transfer). Effect size is given as the normalized additive effect of the QTL, where positive values indicate a positive effect of the T540 allele and negative values indicate a positive effect of the Ge-0 allele; s.e. is the standard error of the mean effect; %EV is the explained variance according to a mixed model. For the $F_2$ population dominance effects could be calculated, which are indicated as Type. Significant QTL effects on flowering with or without vernalisation are indicated in bold.

| Population | Chromosome | Position (cM) | -$LOG_{10}$(P) | Trait | P-value | Effect size | s.e. | %EV | Type |
|---|---|---|---|---|---|---|---|---|---|
| DH | I | 70.8 (5.2-150.6) | 4.5 | **FT_nv** | **<0.001** | **0.31** | **0.08** | **9.8** | **-** |
| | | | | **FT_v** | **0.005** | **0.16** | **0.06** | **2.4** | **-** |
| | I | 149.5 (5.2-150.6) | 4.4 | FT_nv | 0.391 | 0.06 | 0.07 | 0.3 | - |
| | | | | **FT_v** | **<0.001** | **0.21** | **0.05** | **4.4** | **-** |
| | II | 80.4 (0.9-93.5) | 2.3 | FT_nv | 0.120 | -0.11 | 0.07 | 1.1 | - |
| | | | | **FT_v** | **0.002** | **-0.15** | **0.05** | **2.3** | **-** |
| | III | 5.9 (2.9-117.2) | 4.9 | **FT_nv** | **<0.001** | **0.26** | **0.07** | **6.8** | **-** |
| | | | | **FT_v** | **<0.001** | **0.16** | **0.05** | **2.7** | **-** |
| | IV | 61 (5.3-92.8) | 10.9 | FT_nv | 0.653 | -0.03 | 0.07 | 0.1 | - |
| | | | | **FT_v** | **<0.001** | **0.33** | **0.05** | **10.9** | **-** |
| | V | 61.2 (2.4-133.1) | 9.7 | **FT_nv** | **<0.001** | **-0.43** | **0.07** | **18.7** | **-** |
| | | | | **FT_v** | **<0.001** | **-0.21** | **0.05** | **4.3** | **-** |
| | V | 121.7 (75.4-133.1) | 38.7 | **FT_nv** | **0.004** | **0.19** | **0.07** | **3.7** | **-** |
| | | | | **FT_v** | **<0.001** | **0.65** | **0.05** | **42.6** | **-** |
| F2 | IV | 5.3 (5.3-92.8) | 3.2 | **FT_v** | **-** | **0.26** | **0.19** | **2.0** | **Additive** |
| | | | | | **-** | **0.98** | **0.27** | - | **Dominance** |
| | V | 126.4 (109.5-133.1) | 5.7 | **FT_v** | **-** | **1.21** | **0.23** | **44.0** | **Additive** |
| | | | | | - | - | - | - | Dominance |

**Table 3:** QTL detection for phenotypic variation in a monoploid and diploid segregating mapping population.

Positions of detected QTLs are shown in cM with support intervals between brackets. -$LOG_{10}$(P) indicates the significance of the QTL for the combined ploidy levels, while the P-value specifies the significance for each level. Effect size is given as the normalized additive effect of the QTL, where positive values indicate a positive effect of the T540 allele; s.e. is the standard error of the mean effect; %EV is the explained variance according to a mixed model. QTLs with a significant P-value (<0.05) are indicated in bold, while non-significant QTLs for a specific ploidy level are noted in grey. MSL, main stem length (cm); BFR, branching from rosette (nr.); BFS, branching from stem (nr.); SA, seed area ($mm^2$).

| Trait | Chromosome | Position (cM) | -$LOG_{10}$(P) | Population | P-value | Effect size | s.e. | %EV |
|---|---|---|---|---|---|---|---|---|
| MSL | III | 67.4 (2.9-117.2) | 4.1 | Diploid (DH) | 0.234 | 0.10 | 0.08 | 0.9 |
| | | | | **Monoploid** | **<0.001** | **-0.24** | **0.07** | **5.9** |
| | IV | 5.3 (5.3-92.8) | 3.7 | **Diploid (DH)** | **0.033** | **-0.15** | **0.07** | **2.3** |
| | | | | **Monoploid** | **<0.001** | **-0.24** | **0.06** | **5.8** |
| | IV | 57.9 (5.3-92.8) | 3.3 | **Diploid (DH)** | **0.030** | **-0.18** | **0.08** | **3.2** |
| | | | | **Monoploid** | **<0.001** | **-0.26** | **0.07** | **6.9** |
| | IV | 88.8 (5.3-92.8) | 4.0 | **Diploid (DH)** | **<0.001** | **-0.36** | **0.08** | **12.6** |
| | | | | Monoploid | 0.099 | -0.12 | 0.07 | 1.3 |
| | V | 4.7 (2.4-133.1) | 17.5 | Diploid (DH) | 0.698 | 0.03 | 0.07 | 0.1 |
| | | | | **Monoploid** | **<0.001** | **-0.49** | **0.06** | **23.5** |
| BFR | III | 102.5 (2.9-117.2) | 3.5 | **Diploid (DH)** | **0.001** | **0.26** | **0.08** | **7** |
| | | | | **Monoploid** | **0.009** | **0.18** | **0.07** | **3.2** |
| | V | 4.7 (2.4-133.1) | 9.3 | **Diploid (DH)** | **0.006** | **0.20** | **0.08** | **4.2** |
| | | | | **Monoploid** | **<0.001** | **0.41** | **0.07** | **17** |
| | V | 130.5 (2.4-133.1) | 6.6 | Diploid (DH) | 0.642 | -0.03 | 0.07 | 0.1 |
| | | | | **Monoploid** | **<0.001** | **-0.35** | **0.07** | **12.5** |
| BFS | II | 84.3 (0.9-93.5) | 4.1 | **Diploid (DH)** | **0.010** | **-0.17** | **0.07** | **2.9** |
| | | | | **Monoploid** | **<0.001** | **-0.30** | **0.08** | **8.7** |
| | V | 71 (2.4-133.1) | 3.7 | **Diploid (DH)** | **<0.001** | **-0.23** | **0.06** | **5.3** |
| | | | | Monoploid | 0.204 | 0.09 | 0.07 | 0.9 |
| | V | 126.4 (2.4-133.1) | 14.4 | **Diploid (DH)** | **<0.001** | **-0.48** | **0.06** | **23.1** |
| | | | | **Monoploid** | **<0.001** | **-0.30** | **0.07** | **8.8** |
| SA | II | 84.3 (0.9-93.5) | 6.0 | **Diploid (DH)** | **0.004** | **0.23** | **0.08** | **5.3** |
| | | | | **Monoploid** | **<0.001** | **0.31** | **0.06** | **9.8** |
| | III | 21.6 (2.9-117.2) | 8.1 | **Diploid (DH)** | **0.043** | **-0.17** | **0.08** | **2.9** |
| | | | | **Monoploid** | **<0.001** | **-0.42** | **0.07** | **17.4** |
| | III | 62.7 (2.9-117.2) | 2.2 | Diploid (DH) | 0.224 | -0.11 | 0.09 | 1.3 |
| | | | | **Monoploid** | **0.002** | **-0.24** | **0.08** | **5.5** |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | V | 91.8 (2.4-133.1) | 8.0 | **Diploid (DH)** | **0.006** | **-0.21** | **0.08** | **4.5** |
| | | | | **Monoploid** | **<0.001** | **-0.37** | **0.06** | **13.4** |

659

660

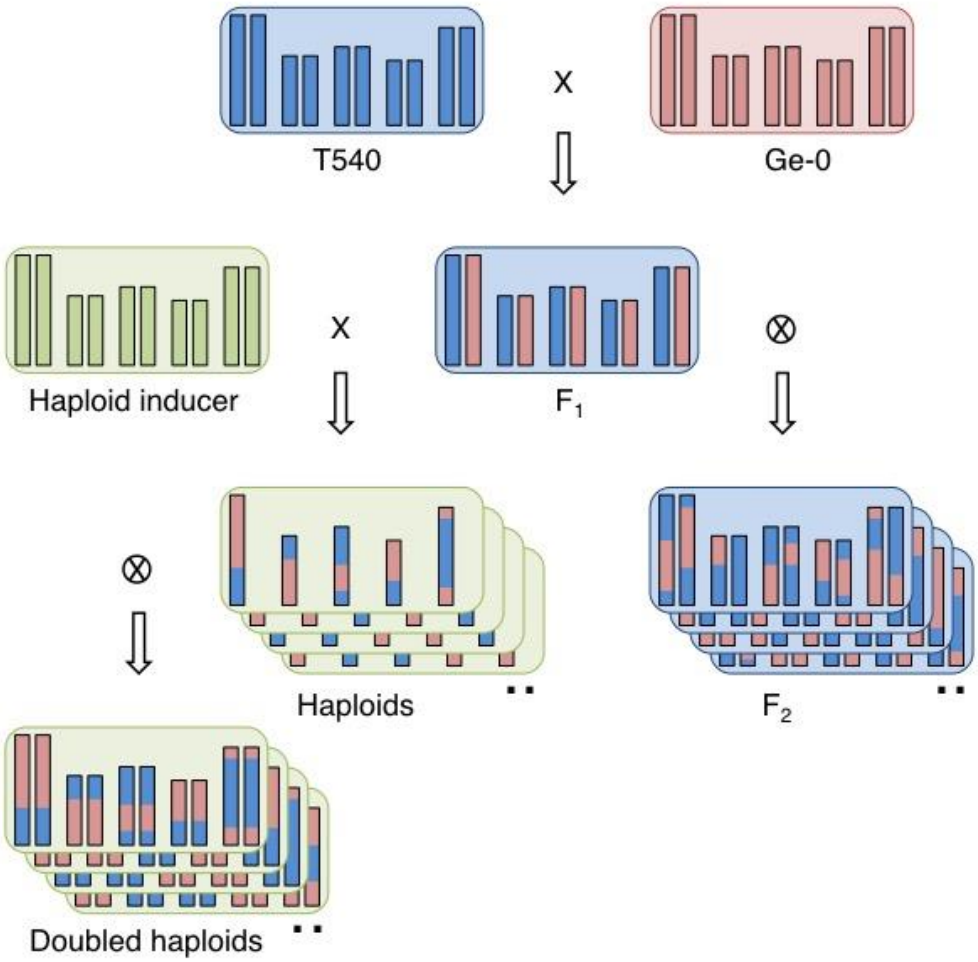



**Figure 1:** Crossing scheme for the development of haploid and doubled haploid recombinant populations.

Each parental genotype (T540 blue; Ge-0 red) is depicted by five double vertical bars, which represent the five chromosomes, while the box indicates the respective genotype of the cytoplasm. The haploid inducer line was obtained in a Col-0 genotypic background (green). Note that the haploids (monoploids) and doubled haploids (diploids) retain the cytoplasm of the haploid inducer line, while the $F_2$ population retains the cytoplasm of the original $F_1$.

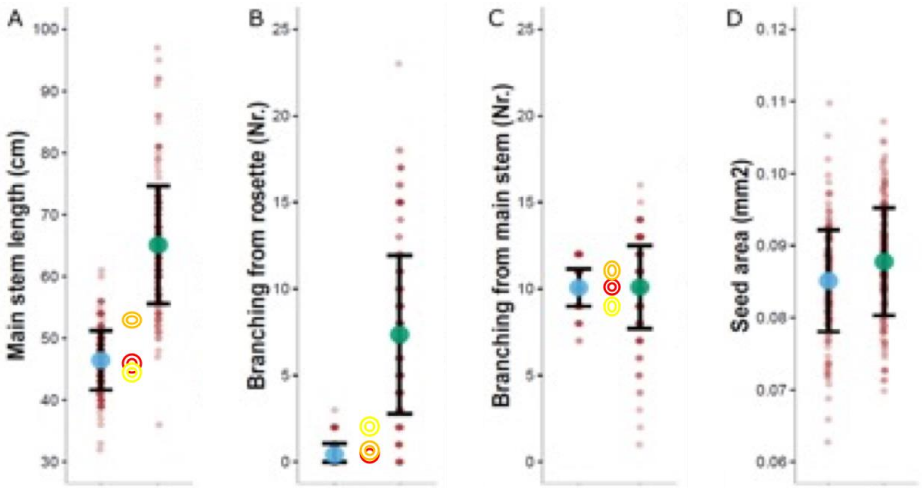
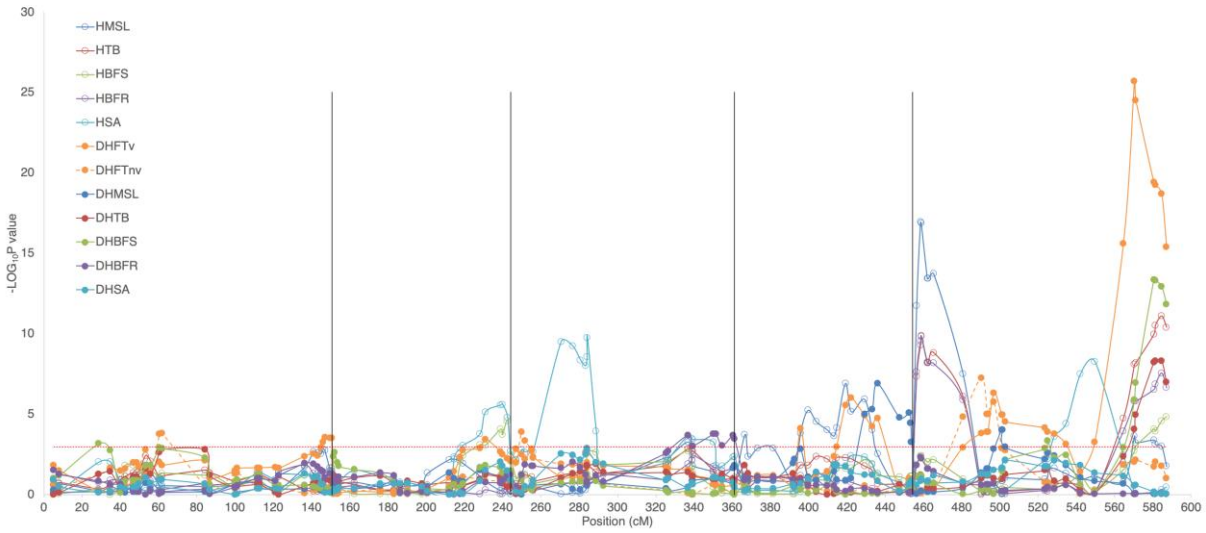



**Figure 2:** Distribution of morphological trait values in monoploid and diploid Arabidopsis plants. The mean value of diploids (DHs) and monoploids is indicated with blue and green dots, respectively. The shaded dots depict the value of individual monoploids and the line average of five replicates for each diploid genotype, respectively. Open circles represent mean trait values of the parental lines Ge-0 (red) and T540 (yellow) and their $F_1$ hybrid (orange). Error bars indicate the standard deviation of the mean.

682




683

684 **Figure 3:** QTL plots of morphological traits mapped in a haploid and diploid segregating population.
685 Traits were mapped in a single multi-trait CIM analysis. Chromosomes are separated by vertical black
686 lines while map positions are indicated on a continuous scale. The horizontal dotted red line indicates
687 the significance threshold ($-LOG_{10}P$ 2.964). Open and closed symbols represent traits measured in the
688 haploid and diploid population, respectively. Solid lines represent traits measured in vernalised plants
689 while the dashed line represents flowering time measured without vernalisation. MSL, main stem
690 length; TB, total branching; BFS, branching from stem; BFR, branching from rosette; SA, seed area;
691 FT, flowering time.

692

693