Phenotypic variation in plants

Roles for epigenetics

Lauss, K.

Citation for published version (APA):


General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 3

Epigenetic divergence is sufficient to trigger heterosis in *Arabidopsis thaliana*

Kathrin Lauss¹, René Wardenaar², Marieke H.A. van Hulten³, Victor Guryev⁴, Joost J.B. Keurentjes³, Maike Stam¹§, Frank Johannes²,⁵,⁶§

¹ University of Amsterdam, Swammerdam Institute for Life Sciences, Science Park 904 1098XH Amsterdam, The Netherlands.
² University of Groningen, Groningen Bioinformatics Centre, Faculty of Mathematics and Natural Sciences, Nijenborgh 7, 9747 AG Groningen, The Netherlands.
³ University of Wageningen, Laboratory of Genetics, Droevendaalsesteeg 1, 6708PB Wageningen, The Netherlands.
⁴ Genome structure aging, European Research Institute for the Biology of Ageing, University Medical Centre Groningen and University of Groningen, Antonius Deusinglaan 1, Building 3226, 9713 AV Groningen, The Netherlands
⁵ Current address: Population epigenetics and epigenomics, Department of Plant Sciences, Technical University Munich, Liesel-Beckmann-Str. 2, 85354 Freising, Germany
⁶ Current address: Institute for Advanced Study, Technical University Munich, Lichtenbergstr. 2a, 85748 Garching, Germany
§ Corresponding co-last authors: m.e.stam@uva.nl (MS), frank@johanneslab.org (FJ)

Pre-print publication on bioRxiv. (doi: http://dx.doi.org/10.1101/059980)
Chapter 3

Abstract
Despite the importance and wide exploitation of heterosis in commercial crop breeding, the molecular mechanisms behind this phenomenon are not well understood. Interestingly, there is growing evidence that beside genetic also epigenetic factors contribute to heterosis. Here we used near-isogenic but epigenetically divergent parents to create epigenetic F1 hybrids (epiHybrids) in Arabidopsis, allowing us to quantify the contribution of epigenetics to heterosis. We measured traits such as leaf area (LA), growth rate (GR), flowering time (FT), main stem branching (MSB), rosette branching (RB) and final plant height (HT) and observed several strong positive and negative heterotic phenotypes among the epiHybrids. For LA and HT mainly positive heterosis was observed, while FT and MSB mostly displayed negative heterosis. Heterosis for FT, LA and HT could be associated with several heritable, differentially methylated regions (DMRs) in the parental genomes. These DMRs contain 35 (FT and LA) and 14 (HT) genes, which may underlie the heterotic phenotypes observed. In conclusion, our study indicates that epigenetic divergence can be sufficient to cause heterosis.

Author Summary
Crossing two genetically distinct parents generates hybrid offspring. Sometimes hybrids are performing better than their parents in particular traits and this is referred to as heterosis. Hybridization and heterosis are naturally occurring processes and crop breeders intentionally cross genetically different parental lines in order to generate hybrids with maximized traits such as yield or stress tolerance. So far, the mechanisms behind heterosis are not well understood. In this study we focused on the effect of epigenetic variation onto heterosis in hybrids, and for this purpose we created epigenetic hybrids (epiHybrids) by crossing wildtype plants with a selection of genetically very similar but epigenetically
divergent lines. An extensive phenotypic analysis of the epiHybrids and their parental lines showed that epigenetic divergence between parental genomes can be a major determinant of heterosis. Importantly, multiple heterotic phenotypes could be associated with meiotically heritable differentially methylated regions (DMRs) in the parental genomes, allowing us to map epigenetic quantitative trait loci (QTLs) for heterosis. Our results indicate that epigenetic variation can contribute to heterosis and suggests that heritable epigenetic variation could be exploited for the improvement of crop traits.

**Introduction**

Heterosis describes an F1 hybrid phenotype that is superior compared to the phenotype of its parent varieties. The phenomenon has been exploited extensively in agricultural breeding for decades and has improved crop performance tremendously [62,67]. Despite its commercial impact, knowledge of the molecular basis underlying heterosis remains incomplete. Most studies mainly focused on finding genetic explanations, resulting in the classical dominance [67,74,118] and overdominance [118,119] models describing heterosis. In line with genetic explanations it has been observed that interspecies hybrids often show a higher degree of heterosis than intraspecies hybrids, indicating that genetic distance correlates with the extent of heterosis [62,69]. However, genetic explanations do often not sufficiently explain nor predict heterosis. There is growing evidence that also epigenetic divergence plays a role in heterosis [58,72,103]. It has, for example, been shown that altered epigenetic profiles at genes regulating circadian rhythm play an important role in heterotic Arabidopsis hybrids [120]. Moreover, heterotic hybrids of Arabidopsis, maize and tomato are shown to differ in levels of small regulatory RNAs and/or DNA methylation (5mC) relative to their parental lines [45,47,48,77]. Processes such as the transfer of 5mC between alleles
(trans chromosomal methylation, TCM), or a loss of 5mC at one of the alleles (trans chromosomal demethylation, TCDM) have been indicated to contribute to the observed remodeling of the epigenome [72,77,79]. Strikingly, some of these changes in 5mC levels have been shown to be stable over multiple generations [79,121].

In this study, we demonstrate that heterotic phenotypes occur in *A. thaliana* F1 epigenetic hybrids (epiHybrids) that were generated from near-isogenic but epigenetically very divergent parental lines. Moreover, we found that some of those heterotic phenotypes could be associated with differentially methylated regions (DMRs) in their parental genomes, allowing us to map QTLs for heterosis.

**Results and Discussion**

**Construction of epigenetic Hybrids**

Hybrids are usually generated from parental lines that vary at both the genomic and epigenomic level and disentangling those two sources of variation is challenging. To overcome this limitation, we generated epigenetic *A. thaliana* F1 hybrids (epiHybrids) from near-isogenic but epigenetically divergent parental lines by crossing Col-0 wildtype (Col-wt) as maternal parent to 19 near-isogenic *ddm1-2*-derived epigenetic recombinant inbred lines (epiRILs) [52] as the paternal parents (Fig 1a).

DDM1 (*DECREASE IN DNA METHYLATION 1*) is a nucleosome remodeler and a *ddm1-2* deficiency leads to a severe loss of 5mC [122], primarily in long transposable elements and other repeat sequences [18]. EpiRILs carry chromosomes that are a mosaic of Col-wt and hypomethylated *ddm1-2*-derived genomic regions [52,56,106] (Fig 1a). Nineteen epiRIL parental lines were selected that sample a broad range of 5mC divergence from the Col-wt reference methylome (Fig 1b, S1 Table). Besides, lines were chosen that have a wildtype methylation profile at *FWA* (S1 Fig, S1 Table), as loss of DNA methylation at the *FWA*
(FLOWERING LOCUS WAGENINGEN) locus is known to affect flowering time [36]. Furthermore, we selected for a range of phenotypic variation in two traits that have previously been monitored in the epiRILs, flowering time and root length (S1 Table); outliers were excluded [52]. With our experimental design we could demonstrate, as proof-of-principle, the extent to which divergence in 5mC profiles in parental lines can contribute to heterosis.
Chapter 3

A

Phenotyping
Col-0 wildtype (wt)
19 selected epiRILs
19 F1 populations
• # replicates per line: ~28
• # plants total: ~ 1000
• # traits: 7

Analysis
• Tests for heterosis
• Search for heterosis QTLs

B

% methylated probes
371H 492H 195H 14H
parental lines used in cross

C

Additive

D

MPH

E

LPH/HPH

Legend
DNA sequence
Methylated
TE insertion
epiHybrid
wt epiRIL
GR: Growth rate
RB: Rosette branching
LA: Leaf area (14 DAS)
HT: Height
MSB: Main stem branching
FT: Flowering time

G

GR
RB
LA
HT
MSB
FT

GR
RB
LA
HT
MSB
FT

371H 492H 195H 14H

H

I

J

LA (mm²)
HT (cm)
FT (days)

232H 195H 193H
500H 344H 64H 193H 371H
232H 208H 344H
Heterosis occurs in epiHybrids. (A) Experimental setup. Lines are depicted schematically as one chromosome with the numbers indicating the epiRIL ID (e.g. 371 & 492) and the respective epiHybrid (e.g. 371H & 492H). (B) Genome-wide 5mC levels (y-axis) of the Col-wt line in green and the epiRIL parental lines in salmon. Numbers indicate the epiRIL IDs. The 5mC levels were calculated as the proportion of methylated MeDIP probes with respect to the total amount of probes. (C-E) Three classes of phenotypic effects monitored in the epiHybrids. The black dashed line indicates the mid-parent value. The green and salmon dashed lines indicate the mean performance of the parental lines. The white dashed lines indicate the mean performance of the epiHybrids. (F) Col-wt, epiHybrid 232H and epiRIL 232 at 13 days after sowing as an example for high-parent heterosis. (G) Phenotypic effects in six traits monitored across the 19 epiHybrids. The right panel summarizes positive and negative heterotic effects per trait. (H-J) Examples of epiHybrids exhibiting high-parent heterosis in leaf area and height (LA and HT; H and J), and low-parent heterosis in flowering time (FT; J) Error bars, ± 1 SEM. Deviation from high parent or low parent is shown in percent.

Heterotic phenotypes occur in the epiHybrids

The phenotypic performance of the 19 epiHybrids and their parental lines was assessed by monitoring about 1090 plants (~28 replicates per line) for a range of quantitative traits: LA, GR, FT, MSB, RB, HT and SY (S2-S7 Tables). The phenotypic observations for SY were inconsistent in a replication experiment, therefore those datasets were excluded from further analysis. The hybrids and parental lines were grown in parallel in a climate-controlled chamber with automated watering. The plants were randomized throughout the chamber to level out phenotypic effects caused by plant position. LA was measured up to 14 days after sowing (DAS), using an automated camera system (Fig 1f), and growth rate (GR) was determined based on this data (SI text). FT was scored manually as opening of the first flower. After all plants started flowering, the plants
were transferred to the greenhouse and grown to maturity. MSB, RB and HT were scored manually after harvesting of the plants.

The extent of heterosis was evaluated by comparing the hybrid performance with its parental lines. We distinguished five effects (Fig 1c-e): additivity, positive mid-parent heterosis (positive MPH), negative mid-parent heterosis (negative MPH), high-parent heterosis (HPH) and low-parent heterosis (LPH). An additive effect describes a hybrid performance that is equal or close to the average performance of the two parents (the mid-parent value, MPV). MPH refers to deviations in percent from the MPV in positive or negative direction. Hybrids displaying MPH are further tested for HPH and LPH, which describe hybrid performance exceeding the high parent, or falling below the lowest parent, respectively. In crop breeding, the focus is usually on obtaining HPH and LPH as these present novel phenotypes that are outside the parental range. Depending on the trait monitored and commercial application, either HPH or LPH can be considered superior. For instance, early flowering may be preferable over late flowering; in such cases maximizing LPH may be desirable. For other traits, such as yield or biomass, it is more important to maximize HPH. However, in order to obtain a comprehensive view of hybrid performance it is informative to also track MPH in addition to LPH and HPH, because many mature traits may be affected by other traits that do not display fully penetrant heterotic effects.

We observed a remarkably wide range of heterotic phenotypes among the epiHybrids (Fig 1g, S2-19 Tables). The magnitude of these phenotypic effects was substantial (Fig 1h-j, S2 Fig, S8-19 Tables) and similar to that typically seen in hybrids of Arabidopsis natural accessions[123,124]. Many epiHybrids (16/19) exhibited significant MPH in at least one of the six monitored traits (FDR = 0.05, Fig 1g). Across all hybrids and traits, we observed 30 cases of positive MPH and negative MPH. Among those, four cases show LPH and nine cases show HPH (Fig
Interestingly, in 11 out of the 17 cases of MPH the phenotypic means of the epiHybrids were in the direction of the phenotypic means of the epiRIL parent rather than in the direction of the Col-wt parent (S2-7 Tables, F1 trend). Also all four LPH and two of the HPH cases were in the direction of the epiRIL parent (Fig 1i-j, S2 Fig). This observation illustrates that ddm1-2-derived hypomethylated epialleles are often (partially) dominant over wild-type epialleles, which contrasts the situation seen in EMS screens where novel mutations typically act recessively.

We observed cases of HPH for LA, HT and MSB, and cases of LPH for FT and MSB. HPH for LA occurred in epiHybrids 232H, 195H and 193H (3/19 epiHybrids). Those epiHybrids significantly exceeded their best parent (Col-wt) by 17%, 18% and 15%, respectively (Fig 1h, S19 Table). Interestingly, although growth rate (GR) is developmentally related to LA, hybrid effects in GR were only moderately, albeit positively, correlated with LA (rho = 0.57, P = 0.02), which implies that LA heterosis is determined by other traits besides GR.

For HT we detected five cases of significant HPH with up to 6% increases in HT (Fig 1i, S14 Table). One may expect LA HPH to strongly correlate with HT HPH, as the rosette is providing nutrients for the developing shoot[125]. However, HPH for both LA and HT occurred only in one epiHybrid (193H; Fig 1g).

For MSB, we detected one case of HPH (64H; Fig 1g and S2 Fig).

Besides positive heterosis, our phenotypic screen revealed strong negative heterotic effects for FT (earlier flowering) and MSB (less main stem branching). Significant LPH occurred in the epiHybrids 232H, 208H and 344H (FT) and 438H (MSB) (Fig 1j, S2 Fig, S15 and S17 Tables). In the most prominent case for FT (232H), FT was about 10% earlier than that of the earliest flowering parent. 208H and 244H flowered 3% and 4% earlier than their lowest parent (epiRIL 208 and epiRIL 344), respectively. 438H showed 14% less MSB than the lowest parent (S2 Fig).
The reproducibility of our findings was tested by performing replicate experiments, using seeds from newly performed crosses and the same climate controlled growth chamber as before. We focused on epiHybrids that exhibited relatively strong positive or negative heterotic phenotypes in the initial screen (193H, 150H, 232H; Fig 1g), and measured LA, FT and HT. We found that the direction of the heterotic effects in LA, FT and HT was reproducible in all cases tested (Fig 2a and b). Importantly, the LA and HT HPH observed for 193H, and the strong FT LPH for 232H were perfectly reproducible, while LA HPH observed for 232H became positive MPH (Fig 2a). Taken together, these results show that the heterotic effects observed in the epiHybrids are relatively stable for LA, HT and FT, even across fresh parental seed batches and independently performed crosses, which is not always the case for Arabidopsis phenotypes [107].
Chapter 3

A

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Replication</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>

% change

-90 0 120

P < 0.001 NS P = 0.006 P = 0.022

B

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>Replication</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>

% change

-90 0 120

F1 plant mean F1 mean epiRIL mean wt

I = Initial screen R = replication
Figure 2: Confirmation of mid-parent (MP) divergence in the initial screen and replicate experiment for epiHybrids 150H, 193H and 232H. A) Results for cases of HPH and LPH for LA, HT and FT in initial experiment. B) Results for traits showing less eminent phenotypic effects for LA, HT and FT. The mid-parent value (MPV) is shown as a dashed horizontal line and the MP divergence is shown as change from MPV in percent. To illustrate the F1 epiHybrid distribution for each trait, the individual replicate plants are depicted as dots. C) F1 MP divergence for LA, HT and FT for all
epiHybrids. The MPV is shown as a horizontal dashed line and MP divergence is shown as change from MPV in percent. The epiHybrids are ordered from highest (left) to lowest (right) F1 MP divergence. To illustrate the F1 epiHybrid distribution for each trait, the individual replicate plants are depicted as dots. Variance component analysis was used to estimate how much of the total variation in MP divergence can be explained by between-cross variation. The F-statistic from this analysis is shown in the boxes.

**Heterotic phenotypes are associated with QTLs**

To understand the sources of the LA, HT and FT heterotic effects observed among the ~530 epiHybrid plants, we calculated the phenotypic divergence of each epiHybrid plant from its respective mid-parent value. Using variance component analysis we estimated that 17%, 28% and 51% of the total variation in mid-parent divergence for FT, LA and HT, respectively, can be attributed to (epi)genomic differences between the Col-wt and epiRILs used for the crosses (Fig 2c, S20 Table, SI text). Global 5mC divergence between the Col-wt and the epiRILs parental lines could not account for this variation (S3 Fig). We therefore reasoned that heterotic phenotypes are due to (partial) dominance effects caused by specific regions being epi-heterozygous for an epiRIL-inherited hypomethylated epiallele ($U$) and a Col-wt-inherited methylated epiallele ($M$). To test this possibility, we used the methylomes of Col-wt and the epiRIL parents[106] to predict epi-homozygous ($MM$) and epi-heterozygous ($MU$) regions in the genomes of the epiHybrids (Fig 3a, SI text), and assessed whether heritable epigenetic differences at specific loci could explain the variation in MPH among crosses (S4 Fig). The analysis revealed two QTLs on chromosome (chr) 3 contributing to the between-cross variation in MPH in FT (QTL 1: LOD=3.12, 37.62 cM; QTL 2: LOD=3.33, 101.44 cM, Fig 3b; S21 Table). EpiHybrids epi-heterozygous ($MU$) at these
loci showed significant negative MPH compared to their epi-homozygous (MM) counterparts (Fig 3c). While not significant at the genome-wide scale (Fig 3b), the same two QTLs had substantial suggestive effects on LA heterosis in the opposite direction than FT (Fig 3b and c), indicating that both QTLs act pleiotropically.

We also detected a single QTL locus on chr 4 (LOD=3.33, 56.00 cM) that contributes to the between-cross variation in MPH for HT (Fig 3b, S21 Table). In this case, MU epiHybrids showed significant positive MPH compared to MM epiHybrids (Fig 3c). Interestingly, the HT QTL overlaps with a previously identified QTL$^{\text{epi}}$ for root length in the epiRILs[56]. The same study identified QTLs$^{\text{epi}}$ associated with FT [56] that we did not detect here (Fig 3b), implying that different regions may play a role in FT trait variation than in FT heterosis.
Chapter 3

A

19 epiHybrid populations

Genomic position

Predicted epigenotypes

MM

MU

No info

B

126 markers

MM405 MM547

C

% change from mid-parent value

FT

FT

LA

LA

HT

P = 0.0004

P = 0.0003

P = 0.0019

P = 0.0026

P = 0.0031

MM405

MM547

MM405

MM547

MM698

MM

MU

MM

MU

MM

MU

MM

MU
Figure 3: Interval mapping approach detects significant QTLs for mid-parent divergence. A) Genome-wide patterns of Col-wt and ddm1-2 inherited epi-haplotypees in the (epi)genomes of the parental epiRILs used in this study. B) QTL profiles for FT, HT and LA. Published QTLsep for root length and flowering time are shown. C) Effect direction of the QTLs. Error bars, ± 1 SE of the Estimate (SEE). D) Zoom in of one of the QTL intervals of FT. The top panel shows the annotations along the genome. The bottom panel shows the locations of candidate DMRs and the average methylation level along the genome for epiRIL parents that are either methylated (MM) or unmethylated (UU) at the peak marker.
Heterotic phenotypes are associated with DMRs in the parental genomes

The detection of heterosis QTLs for FT, LA and HT provided a rationale to search for causal variants in the QTL confidence intervals. TE-associated structural variants (TEASVs) are known to occur at low frequency in a $ddm1-2$-derived DNA hypomethylated background [52,56,57,126], hence we re-analyzed whole-genome sequencing data from the epiRIL parents [56] for TEASVs but did not detect any that could account for the QTL effects, suggesting that the QTLs most likely have an epigenetic basis (SI text). Indeed, a thorough analysis of the methylomes of the parental epiRILs, using the available MeDIP tiling array data [106], identified 55 and 18 potentially causal differentially methylated regions (DMRs) in the FT, LA and HT QTL regions, mapping to 35 and 14 unique genes, respectively (Fig 3d, S5–S9 Figs, S22-S26 Tables, SI text). Potentially interesting genes in the candidate regions of the FT/LA QTLs (S25 Table) include for example RPL5A, which was shown to affect development through regulating auxin and influencing leaf shape and patterning [127,128], and AT3G26480, a protein that shows partial homology to GTS1, which has been implemented in biomass accumulation [129]. Another potentially interesting candidate is Chup1, which is crucial for chloroplast movement in leaves in response to light [130]. These candidate genes provide excellent targets for follow-up studies.

Conclusions

In a recently published study, heterosis for rosette area was reported in an epigenetic F1 hybrid generated by crossing a met1-derived epiRIL with Col-wt [58]. $DNA-METHYLTRANSFERASE1$ ($MET1$) is involved in maintenance of DNA methylation at cytosines in CG sequence context and a mutation in this gene causes a severe loss of DNA methylation in the CG and CHH context [131]. Heterosis was observed in a parent-of-origin manner; the
reciprocal cross did not result in heterosis [58]. This suggests that the heterosis detected may be due to an effect of the maternal cytoplasm rather than differences in epigenetic marks in the parental genomes. Here, we used Col-wt as maternal parent in all crosses to specifically monitor phenotypic effects associated with the epiRIL methylomes. We observed a wide range of heterotic effects, and our proof-of-principle QTL mapping approach indicated that these phenotypic effects are very likely attributable to methylation differences between Col-wt and the epiRILs. Moreover, our results, together with those of Dapp et al. [58], indicate that heterosis in F1 hybrids generated from epigenetically divergent lines may be a more general phenomenon. A more recent study described widespread DNA methylation changes in an epiHybrid derived from Col-wt and a met1-mutant [78]. Remarkably, the formation of spontaneous non-parental epialleles was observed in the epiHybrid, mostly at pericentromeric transposon sequences, but also at genic loci [78]. This demonstrates that novel epigenetic variation, which is not readily predictable from the parental methylomes, can be created during hybridization. Future research needs to address if and how these methylome changes relate to phenotypic variation. This study also stresses that for a refined understanding of the effect of epigenetic QTLs as described in this study, methylation changes should be thoroughly analyzed.

**Material and Methods**

**Plant Material**

The epigenetic recombinant inbred lines (epiRILS) in our study were generated by Johannes et al [52]. The epiRILs were constructed as follows: An *Arabidopsis thaliana* Col-0 line deficient for *ddm1-2* (*DECREASE IN DNA METHYLATION 1*) was crossed to an isogenic Col-0 wildtype line (Col-wt) and the resulting F1 was backcrossed as female parent to Col-wt.
Subsequently about 500 progeny plants with a wildtype *DDM1* allele were selected and propagated through six more rounds of selfing, generating a population of 500 different epiRILs. We selected 19 different epiRILs as paternal plants for generating epiHybrids (Line IDs: 14, 232, 92, 208, 438, 195, 350, 500, 150, 118, 432, 202, 344, 64, 193, 508, 260, 579, 371). Our selection criteria were as follows: 1) Wide range of DNA methylation divergence from Col-wt and among the selected lines; 2) Wildtype DNA methylation state at the FWA locus in order to avoid that differences in DNA methylation at this locus give rise to differences in flowering time [36] in the hybrids; 3) Wide range of phenotypic variation in flowering time and root length among the selected lines. The epiRIL lines were purchased from the Arabidopsis Stock center of INRA Versailles (http://publiclines.versailles.inra.fr/).

**Crosses**

To generate F1 hybrids from the selected epiRIL lines and Col-wt, all parental plants were grown in parallel in soil (Jongkind 7 from Jongkind BV, http://www.jongkind.com/) in pots (Danish size 40 cell, Desch Plantpak, http://www.desch-plantpak.com/en/Home.aspx). The plants were grown at 20°C, 60% humidity, in long day conditions (16h light, 8h dark), and were watered 3 times per week. All crosses were performed in parallel in a time frame of two weeks to avoid phenotypic effects in the F1 progeny due to differences in growing conditions. To exclude that differences in maternal cytoplasm affect the phenotypes of the F1 plants, Col-wt plants were used as a maternal parent and the epiRILs as paternal parents. In parallel, all parental lines, Col-wt and epiRILs, were propagated by manual selfing. This to 1) ensure that parental and F1 hybrid seeds were generated under the same growing conditions and 2) exclude potential phenotypic effects derived from hand pollination[117].
Phenotypic Screen

The seeds were stratified at 4°C for 3 days on petri-dishes containing filter paper and water before transferring them onto Rockwool/Grodan blocks (soaked in Hyponex NPK: 6.5 – 6.19 medium) in a climate controlled chamber (20°C, 70% humidity, long day conditions (16h light, 8h dark)). The transfer of the seeds onto the Rockwool blocks is defined as time point 0 days after sowing (DAS). Seeds from each parental and hybrid line were sown in 28 replicates and their positions were randomized throughout the growth chamber to level out phenotypic effects caused by plant position. The plants were watered two or three times per week depending on their size. After the plants started flowering, they were transferred to the greenhouse (20°C, 60% humidity, long day conditions (16h light, 8h dark)). In the greenhouse, the plants were watered 3 times per week and stabilized by binding them to wooden sticks at later developmental stages. The plants were harvested once the siliques of the main inflorescence and its side branches were ripe.

Rosette Leaf Area (LA): LA was monitored by an automated camera system (Open Pheno System, WUR) from 4 days after sowing (DAS). The system consists of 14 fixed cameras that can take pictures of up to 2145 plants daily, every two hours. We monitored LA until 14 DAS since at later time points leaves start overlapping hampering the correct detection of LA. Leaf area in mm2 was calculated by an ImageJ based measurement setup (http://edepot.wur.nl/169770).

Flowering time (FT): FT was defined as the DAS at which the first flower opened. FT was scored manually each day before 12am.

Height (HT): HT was scored manually in cm on dried plants. The measurement was taken at the main inflorescence, from the rosette to the highest flowerhead.
Branching: Branching was scored on the dried plants by counting the branches emerging from the rosette (RB) and from the main stem (MSB).

Total Seed Yield (SY): Seeds were harvested from the dried plants, cleaned by filtering and seed yield was subsequently determined by weighing (resulting in mg seeds per plant).

**Data analysis**
For the data analysis see the Supplementary Information.

**Replication experiment with selected hybrids**
Freshly ordered seeds of epiRILs (Line IDs: 92, 150, 193, 232) from the Arabidopsis Stock center Versailles were used for the replication experiment with the hybrids selected. The crosses with the epiRILS and the phenotypic screen were performed as described above with the exception that more replicates were monitored for each parental and hybrid line: 60 replicates for LA and 30 replicates for the traits FT and HT. Furthermore, branching was not examined in the replication experiment.

**Acknowledgements**
We thank F. Becker, I. Hövel, D. Angorro, R. Kooke, J.A. Bac-Molenaar, M. Tark-Dame, P. Sanderson, M. Koini, T. Bey, B. Weber, L. Tikovsky and Unifarm Wageningen for technical support during sowing or phenotyping. We thank H. Westerhoff for discussion and critically reading the manuscript.

**Funding**
K.L. was supported by the Centre for Improving Plant Yield (CIPY)(part of the Netherlands Genomics Initiative and the Netherlands Organization for Scientific Research).
Author Contributions
K.L., M.S. and F.J. designed the study, interpreted the data and wrote the manuscript with contributions from J.J.B.K. and R.W.; K.L. and M.H.A.v.H. planned and performed the phenotypic screen; F.J. and R.W. performed the data analysis; V.G. analyzed sequencing data of the epiRIL parents.
**Supporting Information**

**S1 Fig. Methylation profile at the FWA locus.** Methylation level of tiling array probes located within the gene promoter (GP; red rectangles) and gene body (GB; green rectangles) of the FWA gene. The methylation profiles are shown for the wild-type parent, the ddm1 mutant parent and the 19 selected epiRILs (line IDs on the right side). The methylation level was calculated with the use of the HMM results (construction methylomes; see SI text section 2.6.3).
Chapter 3

S2 Fig. Detected cases of high-parent heterosis (HPH; A) and low-parent heterosis (LPH; B) for MSB. In case of HPH percent increase is calculated with respect to the parent with higher phenotypic values. In case of LPH percent increase is calculated with respect to the parent with lower phenotypic values. The corresponding numerical results for statistical tests for HPH and LPH can be found in S15 Table.

S3 Fig. Relationship between genome-wide methylation level of paternal epiRILs (x-axis) and level of mid-parent heterosis in F1 epiHybrids derived from these epiRILs (y-axis). Each blue dot represents one F1 epiHybrid population (N = 19).
S4 Fig. Shown are frequency histograms of the percent change from mid-parent value for the 19 epiHybrid crosses. The percent change values are quantitatively distributed among the 19 epiHybrid crosses and can be treated as a phenotype for QTL mapping.
S5 Fig. Conservation score of probes and methylation differences between epiRIL, Col-wt and ddm1-2 founder lines. The conservation score distribution of all 711,320 probes (A) and methylation differences between the epiRIL Col-wt and ddm1-2 founder lines (B). Shown is the conservation score cutoff that was used (A). The number in the gray rectangle indicates the total number of probes.
(genome-wide) with acceptable quality (probes that are less likely to cross-hybridize). (B) shows the methylation difference between the epiRIL Col-wt and ddm1-2 founder lines for probes that are of acceptable quality (U: unmethylated; I: intermediate methylation; M: methylated; Col-wt → ddm1-2). The number in the gray rectangle indicates the total number of probes (genome-wide) that lost methylation as a result of the ddm1-2 mutation.
S6 Fig. Selection of candidate probes based on correlation with peak marker. Shown are the distributions of the correlation values of marker (A and C) and
non-marker probes (B and D) upstream or downstream from the peak marker, and the cutoff that was used for the selection of candidate probes for the Flowering Time (FT), Leaf Area (LA) and Height (HT) QTL intervals. All marker probes were selected. The cutoff for non-marker probes was based on the 5th percentile of the distribution of the marker probes upstream or downstream from the peak marker depending on the location of the non-marker probes [21]. The FT and LA QTL intervals did not contain any marker and non-marker probes downstream from the peak marker. The HT QTL interval did not contain any marker and non-marker probes upstream from the QTL interval. For all QTL intervals the interval started or ended with a peak marker.

**S7 Fig. Annotation categories of the DMRs.** Shown are the number of DMRs that have an overlap with the different combinations of annotations indicated by the colored rectangles. The left barplot shows the results for Flowering Time (FT) and Leaf Area (LA). The right barplot shows the results for Height (HT).
A = FLCHR03REG01DMR0027  
B = FLCHR03REG01DMR0028  
C = FLCHR03REG01DMR0029  
D = FLCHR03REG01DMR0030  
E = FLCHR03REG01DMR0037  
F = FLCHR03REG02DMR0002  
G = FLCHR03REG02DMR0003  
H = FLCHR03REG02DMR0004  
I = FLCHR03REG02DMR0005
Chapter 3

S8 Fig. Methylation profile of the epiRIL parents around the Flowering Time (FT) and Leaf Area (LA) candidate DMRs. Shown is the average methylation level of the epiRIL parents around the DMRs that have an overlap with a gene body or a gene promoter. A separation was made for epiRILs that have the wild type epigenotype (methylated; MM) at the peak marker (dark gray) or the ddm1-2 epigenotype (unmethylated; UU) at the peak marker (light gray). The methylation level was calculated with the use of the HMM results (construction methylomes; see SI text section 2.6.3). At the top of the panels the positions of gene bodies (green), gene promoters (red), transposable elements (blue) and intergenic regions (gray) are shown. The brown rectangles below the horizontal line indicate positions of DMRs that have an overlap with genes (body or promoter). Letters refer to DMR IDs which can found at the bottom of the figure.
S9 Fig. Methylation profile of the epiRIL parents around the Height (HT) candidate DMRs. Shown is the average methylation level of the epiRIL parents around the DMRs that have an overlap with a gene body or a gene promoter. A separation was made for epiRILs that have the wild type epigenotype (methylated; MM) at the peak marker (dark gray) or the ddm1-2 epigenotype (unmethylated; UU) at the peak marker (light gray). The methylation level was calculated with the use of the HMM results (construction methylomes; see SI text section 2.6.3). At the top of the panels the positions of gene bodies (green), gene promoters (red), transposable elements (blue) and intergenic regions (gray) are shown. The brown rectangles below the horizontal line indicate positions of DMRs that have an overlap with genes (body or promoter). Letters refer to DMR IDs which can found at the bottom of the figure.
<table>
<thead>
<tr>
<th>Line ID</th>
<th>Number of probes</th>
<th>Percentage of probes</th>
<th>Loss meth. at FWA</th>
<th>FT (days)</th>
<th>RL (cm)</th>
<th>Outlier</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>U</td>
<td>I</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>501188</td>
<td>123340</td>
<td>86792</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>70.46</td>
<td>17.34</td>
<td>12.20</td>
<td>NO</td>
<td>43.02</td>
<td>59.51</td>
</tr>
<tr>
<td>232</td>
<td>507776</td>
<td>111848</td>
<td>91696</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>71.39</td>
<td>15.72</td>
<td>12.89</td>
<td>NO</td>
<td>39.47</td>
<td>57.38</td>
</tr>
<tr>
<td>92</td>
<td>478672</td>
<td>128169</td>
<td>104479</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>67.29</td>
<td>18.02</td>
<td>14.69</td>
<td>NO</td>
<td>38.92</td>
<td>54.28</td>
</tr>
<tr>
<td>208</td>
<td>511974</td>
<td>93980</td>
<td>105366</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>71.98</td>
<td>13.21</td>
<td>14.81</td>
<td>NO</td>
<td>41.48</td>
<td>54.05</td>
</tr>
<tr>
<td>438</td>
<td>523734</td>
<td>81102</td>
<td>106484</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>73.63</td>
<td>11.40</td>
<td>14.97</td>
<td>NO</td>
<td>39.47</td>
<td>56.18</td>
</tr>
<tr>
<td>195</td>
<td>475069</td>
<td>126735</td>
<td>109516</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>66.79</td>
<td>17.82</td>
<td>15.40</td>
<td>NO</td>
<td>41.50</td>
<td>65.73</td>
</tr>
<tr>
<td>350</td>
<td>513998</td>
<td>84165</td>
<td>113157</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>72.26</td>
<td>11.83</td>
<td>15.91</td>
<td>NO</td>
<td>40.15</td>
<td>55.62</td>
</tr>
<tr>
<td>500</td>
<td>491867</td>
<td>106096</td>
<td>113357</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>69.15</td>
<td>14.92</td>
<td>15.94</td>
<td>NO</td>
<td>39.88</td>
<td>50.58</td>
</tr>
<tr>
<td>150</td>
<td>468401</td>
<td>128733</td>
<td>114186</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>65.85</td>
<td>18.10</td>
<td>16.05</td>
<td>NO</td>
<td>40.01</td>
<td>57.67</td>
</tr>
<tr>
<td>118</td>
<td>497059</td>
<td>99933</td>
<td>114328</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>69.88</td>
<td>14.05</td>
<td>16.07</td>
<td>NO</td>
<td>40.85</td>
<td>60.40</td>
</tr>
<tr>
<td>432</td>
<td>479227</td>
<td>117478</td>
<td>114615</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>67.37</td>
<td>16.52</td>
<td>16.11</td>
<td>NO</td>
<td>40.91</td>
<td>54.92</td>
</tr>
<tr>
<td>202</td>
<td>503503</td>
<td>93006</td>
<td>114811</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>70.78</td>
<td>13.08</td>
<td>16.14</td>
<td>NO</td>
<td>40.18</td>
<td>58.73</td>
</tr>
<tr>
<td>344</td>
<td>504283</td>
<td>92213</td>
<td>114824</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>70.89</td>
<td>12.96</td>
<td>16.14</td>
<td>NO</td>
<td>39.96</td>
<td>65.21</td>
</tr>
<tr>
<td>64</td>
<td>496686</td>
<td>99596</td>
<td>115038</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>69.83</td>
<td>14.00</td>
<td>16.17</td>
<td>NO</td>
<td>39.96</td>
<td>59.48</td>
</tr>
<tr>
<td>492</td>
<td>493583</td>
<td>102625</td>
<td>115112</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>69.39</td>
<td>14.43</td>
<td>16.18</td>
<td>NO</td>
<td>41.06</td>
<td>59.86</td>
</tr>
<tr>
<td>193</td>
<td>493558</td>
<td>97478</td>
<td>120284</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>69.39</td>
<td>13.70</td>
<td>16.91</td>
<td>NO</td>
<td>39.00</td>
<td>NA</td>
</tr>
<tr>
<td>260</td>
<td>499492</td>
<td>90694</td>
<td>121134</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>70.22</td>
<td>12.75</td>
<td>17.03</td>
<td>NO</td>
<td>43.38</td>
<td>57.36</td>
</tr>
<tr>
<td>579</td>
<td>497286</td>
<td>92790</td>
<td>121244</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>69.91</td>
<td>13.04</td>
<td>17.04</td>
<td>NO</td>
<td>41.02</td>
<td>65.69</td>
</tr>
<tr>
<td>371</td>
<td>483236</td>
<td>106279</td>
<td>121805</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>67.94</td>
<td>14.94</td>
<td>17.12</td>
<td>NO</td>
<td>39.84</td>
<td>56.31</td>
</tr>
<tr>
<td>wt</td>
<td>499673</td>
<td>89513</td>
<td>122134</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>70.25</td>
<td>12.58</td>
<td>17.17</td>
<td>NO</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**S1 Table. Selection epiRIL parental lines.** Provided are the number (and percentage) of genomewide unmethylated (U), intermediately methylated (I) and methylated (M) probes as well as the phenotypic values for flowering time (FT) and root length (RL) for each of the 19 selected epiRIL parental lines (Line ID). The table also indicates whether there was a loss of methylation observed at the *FWA* locus (Loss meth. at *FWA*; see also Figure S1) and whether the phenotypic values were considered to be outliers or not (Outlier). Values that deviated more than two standard deviations from the mean were considered to be outliers. NA means not available.
Chapter 3

Table. Phenotype summary for Height (HT). Provided are sample sizes, means and variances for the Col-wt parents, epiRIL parents and the epiHybrids. The sample sizes for the Col-wt parents, epiRIL parents and the epiHybrids (F1) are denoted with \( N_{\text{Col-wt}} \), \( N_{\text{epiRIL}} \) and \( N_{\text{F1}} \), respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with \( \mu_{\text{Pl}} \) and \( \sigma^2_{\text{Pl}} \), \( \mu_{\text{Ph}} \) and \( \sigma^2_{\text{Ph}} \), and \( \mu_{\text{F1}} \) and \( \sigma^2_{\text{F1}} \), respectively. The different plant lines are denoted according to their epiHybrid ID; Ph denotes whether the Col-wt or the epiRIL parental line had a higher phenotypic mean; F1 trend indicates whether the phenotypic mean of the epiHybrids are in the direction of the Col-wt or in the direction of the epiRIL parental line; outliers > ± 2 SD from the mean were removed.
**S3 Table. Phenotype summary for Main Stem Branching (MSB).** Provided are sample sizes, means and variances for the Col-wt parents, epiRIL parents and the epiHybrids. The sample sizes for the Col-wt parents, epiRIL parents and the epiHybrids (F1) are denoted with $N_{\text{Col-wt}}$, $N_{\text{epiRIL}}$ and $N_{\text{F1}}$, respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with $\mu_{\text{Pl}}$ and $\sigma^2_{\text{Pl}}$, $\mu_{\text{Ph}}$ and $\sigma^2_{\text{Ph}}$, and $\mu_{\text{F1}}$ and $\sigma^2_{\text{F1}}$, respectively. The different plant lines are denoted according to their epiHybrid ID; Ph denotes whether the Col-wt or the epiRIL parental line had a higher phenotypic mean; F1 trend indicates whether the phenotypic mean of the epiHybrids are in the direction of the Col-wt or in the direction of the epiRIL parental line; outliers $> \pm 2$ SD from the mean were removed.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>$N_{\text{Col-wt}}$</th>
<th>$N_{\text{epiRIL}}$</th>
<th>$N_{\text{F1}}$</th>
<th>$\mu_{\text{Pl}}$</th>
<th>$\mu_{\text{Ph}}$</th>
<th>$\mu_{\text{F1}}$</th>
<th>$\sigma^2_{\text{Pl}}$</th>
<th>$\sigma^2_{\text{Ph}}$</th>
<th>$\sigma^2_{\text{F1}}$</th>
<th>Ph</th>
<th>F1 trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSB</td>
<td>14H</td>
<td>25</td>
<td>26</td>
<td>28</td>
<td>4.62</td>
<td>4.88</td>
<td>4.96</td>
<td>0.65</td>
<td>0.53</td>
<td>0.48</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>MSB</td>
<td>232H</td>
<td>25</td>
<td>22</td>
<td>27</td>
<td>3.68</td>
<td>4.88</td>
<td>4.19</td>
<td>0.42</td>
<td>0.53</td>
<td>0.93</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>92H</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>3.72</td>
<td>4.88</td>
<td>3.56</td>
<td>0.21</td>
<td>0.53</td>
<td>0.42</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>208H</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>3.72</td>
<td>4.88</td>
<td>4.54</td>
<td>1.04</td>
<td>0.53</td>
<td>0.74</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>MSB</td>
<td>438H</td>
<td>25</td>
<td>25</td>
<td>28</td>
<td>4.80</td>
<td>4.88</td>
<td>4.14</td>
<td>0.58</td>
<td>0.53</td>
<td>0.57</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>195H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>4.72</td>
<td>4.88</td>
<td>4.85</td>
<td>0.46</td>
<td>0.53</td>
<td>0.36</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>MSB</td>
<td>350H</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>4.24</td>
<td>4.88</td>
<td>4.83</td>
<td>0.52</td>
<td>0.53</td>
<td>0.42</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>MSB</td>
<td>500H</td>
<td>25</td>
<td>19</td>
<td>28</td>
<td>3.95</td>
<td>4.88</td>
<td>4.18</td>
<td>1.39</td>
<td>0.53</td>
<td>0.37</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>150H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>3.15</td>
<td>4.88</td>
<td>3.56</td>
<td>0.46</td>
<td>0.53</td>
<td>0.41</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>118H</td>
<td>25</td>
<td>28</td>
<td>28</td>
<td>4.14</td>
<td>4.88</td>
<td>4.21</td>
<td>0.35</td>
<td>0.53</td>
<td>0.40</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>432H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>4.27</td>
<td>4.88</td>
<td>4.65</td>
<td>0.36</td>
<td>0.53</td>
<td>0.40</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>MSB</td>
<td>202H</td>
<td>25</td>
<td>26</td>
<td>24</td>
<td>4.08</td>
<td>4.88</td>
<td>4.00</td>
<td>0.31</td>
<td>0.53</td>
<td>0.43</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>344H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>4.15</td>
<td>4.88</td>
<td>4.30</td>
<td>0.44</td>
<td>0.53</td>
<td>0.52</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>64H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>4.67</td>
<td>4.88</td>
<td>5.26</td>
<td>0.46</td>
<td>0.53</td>
<td>0.43</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>MSB</td>
<td>492H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>3.72</td>
<td>4.88</td>
<td>3.96</td>
<td>0.46</td>
<td>0.53</td>
<td>0.58</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>193H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>3.88</td>
<td>4.88</td>
<td>4.07</td>
<td>0.53</td>
<td>0.53</td>
<td>0.30</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>260H</td>
<td>25</td>
<td>25</td>
<td>28</td>
<td>4.28</td>
<td>4.88</td>
<td>4.82</td>
<td>0.63</td>
<td>0.53</td>
<td>0.45</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>MSB</td>
<td>579H</td>
<td>25</td>
<td>23</td>
<td>27</td>
<td>4.88</td>
<td>5.17</td>
<td>5.15</td>
<td>0.53</td>
<td>0.60</td>
<td>0.44</td>
<td>epi</td>
<td>epi</td>
</tr>
<tr>
<td>MSB</td>
<td>371H</td>
<td>25</td>
<td>25</td>
<td>23</td>
<td>4.68</td>
<td>4.88</td>
<td>4.74</td>
<td>0.23</td>
<td>0.53</td>
<td>0.47</td>
<td>wt</td>
<td>epi</td>
</tr>
</tbody>
</table>
**Chapter 3**

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epHybrid ID</th>
<th>N_Col-wt</th>
<th>N_epiRIL</th>
<th>N_F1</th>
<th>( \mu_{pl} )</th>
<th>( \mu_{ph} )</th>
<th>( \sigma^2_{pl} )</th>
<th>( \sigma^2_{ph} )</th>
<th>( \sigma^2_{F1} )</th>
<th>Ph</th>
<th>F1 trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>RB</td>
<td>14H</td>
<td>23</td>
<td>28</td>
<td>27</td>
<td>3.13</td>
<td>4.07</td>
<td>3.89</td>
<td>0.75</td>
<td>2.66</td>
<td>epi</td>
<td>epi</td>
</tr>
<tr>
<td>RB</td>
<td>232H</td>
<td>23</td>
<td>27</td>
<td>27</td>
<td>3.13</td>
<td>4.68</td>
<td>4.30</td>
<td>0.75</td>
<td>1.66</td>
<td>epi</td>
<td>1.29</td>
</tr>
<tr>
<td>RB</td>
<td>92H</td>
<td>23</td>
<td>27</td>
<td>26</td>
<td>3.13</td>
<td>4.44</td>
<td>3.73</td>
<td>0.75</td>
<td>1.26</td>
<td>epi</td>
<td>1.08</td>
</tr>
<tr>
<td>RB</td>
<td>208H</td>
<td>23</td>
<td>25</td>
<td>27</td>
<td>3.13</td>
<td>4.24</td>
<td>3.63</td>
<td>0.75</td>
<td>1.86</td>
<td>epi</td>
<td>1.24</td>
</tr>
<tr>
<td>RB</td>
<td>438H</td>
<td>23</td>
<td>25</td>
<td>27</td>
<td>3.13</td>
<td>4.13</td>
<td>3.54</td>
<td>0.75</td>
<td>2.18</td>
<td>epi</td>
<td>1.39</td>
</tr>
<tr>
<td>RB</td>
<td>350H</td>
<td>23</td>
<td>25</td>
<td>25</td>
<td>3.13</td>
<td>4.04</td>
<td>3.68</td>
<td>0.75</td>
<td>2.44</td>
<td>epi</td>
<td>1.81</td>
</tr>
<tr>
<td>RB</td>
<td>500H</td>
<td>23</td>
<td>19</td>
<td>28</td>
<td>3.13</td>
<td>3.89</td>
<td>3.50</td>
<td>0.75</td>
<td>2.54</td>
<td>epi</td>
<td>1.74</td>
</tr>
<tr>
<td>RB</td>
<td>150H</td>
<td>23</td>
<td>25</td>
<td>26</td>
<td>3.13</td>
<td>4.12</td>
<td>4.35</td>
<td>0.75</td>
<td>0.86</td>
<td>epi</td>
<td>1.12</td>
</tr>
<tr>
<td>RB</td>
<td>118H</td>
<td>23</td>
<td>26</td>
<td>28</td>
<td>3.13</td>
<td>3.31</td>
<td>3.50</td>
<td>0.75</td>
<td>1.34</td>
<td>epi</td>
<td>1.59</td>
</tr>
<tr>
<td>RB</td>
<td>432H</td>
<td>23</td>
<td>27</td>
<td>28</td>
<td>3.13</td>
<td>3.93</td>
<td>3.86</td>
<td>0.75</td>
<td>2.23</td>
<td>epi</td>
<td>2.35</td>
</tr>
<tr>
<td>RB</td>
<td>202H</td>
<td>23</td>
<td>25</td>
<td>25</td>
<td>3.13</td>
<td>4.20</td>
<td>3.56</td>
<td>0.75</td>
<td>1.17</td>
<td>epi</td>
<td>0.92</td>
</tr>
<tr>
<td>RB</td>
<td>344H</td>
<td>23</td>
<td>27</td>
<td>28</td>
<td>3.13</td>
<td>3.70</td>
<td>3.64</td>
<td>0.75</td>
<td>0.91</td>
<td>epi</td>
<td>1.94</td>
</tr>
<tr>
<td>RB</td>
<td>64H</td>
<td>23</td>
<td>27</td>
<td>28</td>
<td>3.13</td>
<td>3.85</td>
<td>3.18</td>
<td>0.75</td>
<td>1.13</td>
<td>epi</td>
<td>2.00</td>
</tr>
<tr>
<td>RB</td>
<td>492H</td>
<td>23</td>
<td>26</td>
<td>28</td>
<td>3.13</td>
<td>4.23</td>
<td>3.93</td>
<td>0.75</td>
<td>0.90</td>
<td>epi</td>
<td>1.33</td>
</tr>
<tr>
<td>RB</td>
<td>193H</td>
<td>23</td>
<td>24</td>
<td>27</td>
<td>3.13</td>
<td>5.88</td>
<td>4.30</td>
<td>0.75</td>
<td>1.77</td>
<td>epi</td>
<td>1.52</td>
</tr>
<tr>
<td>RB</td>
<td>260H</td>
<td>23</td>
<td>26</td>
<td>28</td>
<td>3.13</td>
<td>3.65</td>
<td>3.11</td>
<td>0.75</td>
<td>1.92</td>
<td>epi</td>
<td>1.06</td>
</tr>
<tr>
<td>RB</td>
<td>579H</td>
<td>23</td>
<td>23</td>
<td>27</td>
<td>3.13</td>
<td>4.17</td>
<td>3.44</td>
<td>0.75</td>
<td>1.24</td>
<td>epi</td>
<td>2.41</td>
</tr>
<tr>
<td>RB</td>
<td>371H</td>
<td>23</td>
<td>26</td>
<td>25</td>
<td>3.13</td>
<td>4.04</td>
<td>3.76</td>
<td>0.75</td>
<td>1.64</td>
<td>epi</td>
<td>2.61</td>
</tr>
</tbody>
</table>

**S4 Table. Phenotype summary for Rosette Branching (RB).** Provided are sample sizes, means and variances for the Col-wt parents, epiRIL parents and the epHybrids. The sample sizes for the Col-wt parents, epiRIL parents and the epHybrids (F1) are denoted with \( N_{\text{Col-wt}} \), \( N_{\text{epiRIL}} \) and \( N_{\text{F1}} \), respectively. The means and variances for the low parents (Pl), high parents (Ph) and epHybrids (F1) are denoted with \( \mu_{pl} \), \( \mu_{ph} \) and \( \sigma^2_{pl} \), \( \sigma^2_{ph} \) and \( \sigma^2_{F1} \), respectively. The different plant lines are denoted according to their epHybrid ID; Ph denotes whether the Col-wt or the epiRIL parental line had a higher phenotypic mean; F1 trend indicates whether the phenotypic mean of the epHybrids are in the direction of the Col-wt or in the direction of the epiRIL parental line; outliers > ± 2 SD from the mean were removed.
S5 Table. Phenotype summary for Flowering Time (FT). Provided are sample sizes, means and variances for the Col-wt parents, epiRIL parents and the epiHybrids. The sample sizes for the Col-wt parents, epiRIL parents and the epiHybrids (F1) are denoted with $N_{\text{Col-wt}}$, $N_{\text{epiRIL}}$ and $N_{\text{F1}}$, respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with $\mu_{\text{Pl}}$ and $\sigma^2_{\text{Pl}}$, $\mu_{\text{Ph}}$ and $\sigma^2_{\text{Ph}}$, and $\mu_{\text{F1}}$ and $\sigma^2_{\text{F1}}$, respectively. The different plant lines are denoted according to their epiHybrid ID; Ph denotes whether the Col-wt or the epiRIL parental line had a higher phenotypic mean; F1 trend indicates whether the phenotypic mean of the epiHybrids are in the direction of the Col-wt or in the direction of the epiRIL parental line; outliers $> \pm 2$ SD from the mean were removed.
Chapter 3

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>N_{Col-wt}</th>
<th>N_{epiRIL}</th>
<th>N_{F1}</th>
<th>μ_{Pl}</th>
<th>μ_{Ph}</th>
<th>μ_{F1}</th>
<th>σ^2_{Pl}</th>
<th>σ^2_{Ph}</th>
<th>σ^2_{F1}</th>
<th>Ph</th>
<th>F1 trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR</td>
<td>14H</td>
<td>25</td>
<td>27</td>
<td>26</td>
<td>0.4015</td>
<td>0.4162</td>
<td>0.4120</td>
<td>0.000804</td>
<td>0.002092</td>
<td>0.001200</td>
<td>epi</td>
<td>epi</td>
</tr>
<tr>
<td>GR</td>
<td>232H</td>
<td>25</td>
<td>23</td>
<td>26</td>
<td>0.3417</td>
<td>0.4015</td>
<td>0.4010</td>
<td>0.003646</td>
<td>0.00804</td>
<td>0.002400</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>92H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>0.4015</td>
<td>0.4036</td>
<td>0.4020</td>
<td>0.000804</td>
<td>0.00968</td>
<td>0.001400</td>
<td>epi</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>208H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>0.3908</td>
<td>0.4015</td>
<td>0.3990</td>
<td>0.001424</td>
<td>0.00804</td>
<td>0.001100</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>438H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>0.3903</td>
<td>0.4015</td>
<td>0.3930</td>
<td>0.001582</td>
<td>0.00804</td>
<td>0.002600</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>GR</td>
<td>195H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>0.3981</td>
<td>0.4015</td>
<td>0.4100</td>
<td>0.000548</td>
<td>0.00804</td>
<td>0.000700</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>350H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>0.3995</td>
<td>0.4015</td>
<td>0.4050</td>
<td>0.002061</td>
<td>0.00804</td>
<td>0.000800</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>500H</td>
<td>25</td>
<td>17</td>
<td>25</td>
<td>0.3902</td>
<td>0.4015</td>
<td>0.4060</td>
<td>0.002656</td>
<td>0.00804</td>
<td>0.000800</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>150H</td>
<td>25</td>
<td>24</td>
<td>25</td>
<td>0.4015</td>
<td>0.4096</td>
<td>0.3900</td>
<td>0.000804</td>
<td>0.002040</td>
<td>0.001300</td>
<td>epi</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>118H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>0.3964</td>
<td>0.4015</td>
<td>0.4000</td>
<td>0.001640</td>
<td>0.00804</td>
<td>0.000800</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>432H</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>0.4015</td>
<td>0.4029</td>
<td>0.4090</td>
<td>0.000804</td>
<td>0.001296</td>
<td>0.001500</td>
<td>epi</td>
<td>epi</td>
</tr>
<tr>
<td>GR</td>
<td>202H</td>
<td>25</td>
<td>24</td>
<td>22</td>
<td>0.4015</td>
<td>0.4078</td>
<td>0.4070</td>
<td>0.000804</td>
<td>0.002066</td>
<td>0.001000</td>
<td>epi</td>
<td>epi</td>
</tr>
<tr>
<td>GR</td>
<td>344H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>0.4015</td>
<td>0.4099</td>
<td>0.4100</td>
<td>0.000804</td>
<td>0.001282</td>
<td>0.000900</td>
<td>epi</td>
<td>epi</td>
</tr>
<tr>
<td>GR</td>
<td>64H</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>0.3969</td>
<td>0.4015</td>
<td>0.3950</td>
<td>0.002778</td>
<td>0.00804</td>
<td>0.000100</td>
<td>wt</td>
<td>epi</td>
</tr>
<tr>
<td>GR</td>
<td>492H</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>0.4015</td>
<td>0.4211</td>
<td>0.4180</td>
<td>0.000804</td>
<td>0.002629</td>
<td>0.001500</td>
<td>epi</td>
<td>epi</td>
</tr>
<tr>
<td>GR</td>
<td>193H</td>
<td>25</td>
<td>24</td>
<td>27</td>
<td>0.3885</td>
<td>0.4015</td>
<td>0.4000</td>
<td>0.002730</td>
<td>0.00804</td>
<td>0.001100</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>260H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>0.4015</td>
<td>0.4147</td>
<td>0.3960</td>
<td>0.000804</td>
<td>0.001070</td>
<td>0.001800</td>
<td>wt</td>
<td>wt</td>
</tr>
<tr>
<td>GR</td>
<td>579H</td>
<td>25</td>
<td>22</td>
<td>26</td>
<td>0.4015</td>
<td>0.4185</td>
<td>0.4080</td>
<td>0.000804</td>
<td>0.001496</td>
<td>0.001800</td>
<td>epi</td>
<td>epi</td>
</tr>
<tr>
<td>GR</td>
<td>371H</td>
<td>25</td>
<td>25</td>
<td>24</td>
<td>0.3845</td>
<td>0.4015</td>
<td>0.4020</td>
<td>0.000976</td>
<td>0.00804</td>
<td>0.001300</td>
<td>wt</td>
<td>wt</td>
</tr>
</tbody>
</table>

S6 Table. Phenotype summary for Growth Rate (GR). Provided are sample sizes, means and variances for the Col-wt parents, epiRIL parents and the epiHybrids. The sample sizes for the Col-wt parents, epiRIL parents and the epiHybrids (F1) are denoted with \( N_{Col-wt}, N_{epiRIL} \) and \( N_{F1} \), respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with \( \mu_{Pl}, \mu_{Ph} \) and \( \mu_{F1} \) and \( \sigma^2_{Pl}, \sigma^2_{Ph} \) and \( \sigma^2_{F1} \), respectively. The different plant lines are denoted according to their epiHybrid ID; Ph denotes whether the Col-wt or the epiRIL parental line had a higher phenotypic mean; F1 trend indicates whether the phenotypic mean of the epiHybrids are in the direction of the Col-wt or in the direction of the epiRIL parental line; outliers > ± 2 SD from the mean were removed.
S7 Table. Phenotype summary for Leaf Area (LA). Provided are sample sizes, means and variances for the Col-wt parents, epiRIL parents and the epiHybrids. The sample sizes for the Col-wt parents, epiRIL parents and the epiHybrids (F1) are denoted with $N_{Col-wt}$, $N_{epiRIL}$ and $N_{F1}$, respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with $\mu_{Pl}$ and $\sigma^2_{Pl}$, $\mu_{Ph}$ and $\sigma^2_{Ph}$, and $\mu_{F1}$ and $\sigma^2_{F1}$, respectively. The different plant lines are denoted according to their epiHybrid ID; Ph denotes whether the Col-wt or the epiRIL parental line had a higher phenotypic mean; F1 trend indicates whether the phenotypic mean of the epiHybrids are in the direction of the Col-wt or in the direction of the epiRIL parental line; outliers $> \pm 2$ SD from the mean were removed.
### S8 Table. Test for mid-parent heterosis in Height (HT).

Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $l_F$ denotes the log-likelihood of the full (unconstrained) model with degrees of freedom given by $df(l_F)$; $l_A$ denotes the log-likelihood of the additive model with degrees of freedom given by $df(l_A)$.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>N_Col-wt</th>
<th>N_epiRIL</th>
<th>N_F1</th>
<th>$l_F$</th>
<th>$df(l_F)$</th>
<th>$l_A$</th>
<th>$df(l_A)$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HT</td>
<td>14H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-175.087</td>
<td>6</td>
<td>-175.313</td>
<td>5</td>
<td>0.500977</td>
</tr>
<tr>
<td>HT</td>
<td>232H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>-193.748</td>
<td>6</td>
<td>-195.101</td>
<td>5</td>
<td>0.099987</td>
</tr>
<tr>
<td>HT</td>
<td>92H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-171.983</td>
<td>6</td>
<td>-172.622</td>
<td>5</td>
<td>0.258605</td>
</tr>
<tr>
<td>HT</td>
<td>208H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>-172.526</td>
<td>6</td>
<td>-172.648</td>
<td>5</td>
<td>0.621807</td>
</tr>
<tr>
<td>HT</td>
<td>438H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-172.738</td>
<td>6</td>
<td>-177.973</td>
<td>5</td>
<td>0.001213</td>
</tr>
<tr>
<td>HT</td>
<td>195H</td>
<td>25</td>
<td>23</td>
<td>26</td>
<td>-170.167</td>
<td>6</td>
<td>-171.357</td>
<td>5</td>
<td>0.122811</td>
</tr>
<tr>
<td>HT</td>
<td>350H</td>
<td>25</td>
<td>25</td>
<td>24</td>
<td>-162.601</td>
<td>6</td>
<td>-169.200</td>
<td>5</td>
<td>0.000280</td>
</tr>
<tr>
<td>HT</td>
<td>500H</td>
<td>25</td>
<td>19</td>
<td>28</td>
<td>-215.088</td>
<td>6</td>
<td>-222.375</td>
<td>5</td>
<td>0.000135</td>
</tr>
<tr>
<td>HT</td>
<td>150H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-183.603</td>
<td>6</td>
<td>-190.431</td>
<td>5</td>
<td>0.000220</td>
</tr>
<tr>
<td>HT</td>
<td>118H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>-167.886</td>
<td>6</td>
<td>-170.941</td>
<td>5</td>
<td>0.013438</td>
</tr>
<tr>
<td>HT</td>
<td>432H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-172.157</td>
<td>6</td>
<td>-176.276</td>
<td>5</td>
<td>0.004103</td>
</tr>
<tr>
<td>HT</td>
<td>202H</td>
<td>25</td>
<td>25</td>
<td>24</td>
<td>-167.340</td>
<td>6</td>
<td>-174.378</td>
<td>5</td>
<td>0.000176</td>
</tr>
<tr>
<td>HT</td>
<td>344H</td>
<td>25</td>
<td>27</td>
<td>25</td>
<td>-161.824</td>
<td>6</td>
<td>-169.780</td>
<td>5</td>
<td>6.64E-05</td>
</tr>
<tr>
<td>HT</td>
<td>64H</td>
<td>25</td>
<td>26</td>
<td>25</td>
<td>-157.053</td>
<td>6</td>
<td>-173.241</td>
<td>5</td>
<td>1.27E-08</td>
</tr>
<tr>
<td>HT</td>
<td>492H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-178.752</td>
<td>6</td>
<td>-180.089</td>
<td>5</td>
<td>0.102026</td>
</tr>
<tr>
<td>HT</td>
<td>193H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>-184.729</td>
<td>6</td>
<td>-227.271</td>
<td>5</td>
<td>2.86E-20</td>
</tr>
<tr>
<td>HT</td>
<td>260H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-183.575</td>
<td>6</td>
<td>-185.550</td>
<td>5</td>
<td>0.046847</td>
</tr>
<tr>
<td>HT</td>
<td>579H</td>
<td>25</td>
<td>23</td>
<td>28</td>
<td>-173.515</td>
<td>6</td>
<td>-174.444</td>
<td>5</td>
<td>0.172689</td>
</tr>
<tr>
<td>HT</td>
<td>371H</td>
<td>25</td>
<td>26</td>
<td>25</td>
<td>-186.759</td>
<td>6</td>
<td>-197.020</td>
<td>5</td>
<td>5.89E-06</td>
</tr>
</tbody>
</table>
### S9 Table. Test for mid-parent heterosis in Main Stem Branching (MSB).
Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $l_F$ denotes the log-likelihood of the full (unconstrained) model with degrees of freedom given by $df\ (l_F)$; $l_A$ denotes the log-likelihood of the additive model with degrees of freedom given by $df\ (l_A)$.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>NCol-wt</th>
<th>N_epiRIL</th>
<th>N_F1</th>
<th>$l_F$</th>
<th>$df\ (l_F)$</th>
<th>$l_A$</th>
<th>$df\ (l_A)$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSB</td>
<td>14H</td>
<td>25</td>
<td>26</td>
<td>28</td>
<td>-86.633</td>
<td>6</td>
<td>-87.452</td>
<td>5</td>
<td>0.200506</td>
</tr>
<tr>
<td>MSB</td>
<td>232H</td>
<td>25</td>
<td>22</td>
<td>27</td>
<td>-84.846</td>
<td>6</td>
<td>-84.949</td>
<td>5</td>
<td>0.649279</td>
</tr>
<tr>
<td>MSB</td>
<td>92H</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>-66.653</td>
<td>6</td>
<td>-77.920</td>
<td>5</td>
<td>2.06E-06</td>
</tr>
<tr>
<td>MSB</td>
<td>208H</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>-94.913</td>
<td>6</td>
<td>-95.558</td>
<td>5</td>
<td>0.256165</td>
</tr>
<tr>
<td>MSB</td>
<td>438H</td>
<td>25</td>
<td>25</td>
<td>28</td>
<td>-86.590</td>
<td>6</td>
<td>-94.303</td>
<td>5</td>
<td>8.59E-05</td>
</tr>
<tr>
<td>MSB</td>
<td>195H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-76.313</td>
<td>6</td>
<td>-76.371</td>
<td>5</td>
<td>0.733909</td>
</tr>
<tr>
<td>MSB</td>
<td>350H</td>
<td>25</td>
<td>25</td>
<td>23</td>
<td>-76.077</td>
<td>6</td>
<td>-77.302</td>
<td>5</td>
<td>0.117456</td>
</tr>
<tr>
<td>MSB</td>
<td>500H</td>
<td>25</td>
<td>19</td>
<td>28</td>
<td>-81.993</td>
<td>6</td>
<td>-82.743</td>
<td>5</td>
<td>0.220798</td>
</tr>
<tr>
<td>MSB</td>
<td>150H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-78.908</td>
<td>6</td>
<td>-83.193</td>
<td>5</td>
<td>0.003419</td>
</tr>
<tr>
<td>MSB</td>
<td>118H</td>
<td>25</td>
<td>28</td>
<td>28</td>
<td>-77.750</td>
<td>6</td>
<td>-79.707</td>
<td>5</td>
<td>0.047878</td>
</tr>
<tr>
<td>MSB</td>
<td>432H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>-74.565</td>
<td>6</td>
<td>-74.696</td>
<td>5</td>
<td>0.608896</td>
</tr>
<tr>
<td>MSB</td>
<td>202H</td>
<td>25</td>
<td>26</td>
<td>24</td>
<td>-71.846</td>
<td>6</td>
<td>-76.181</td>
<td>5</td>
<td>0.003233</td>
</tr>
<tr>
<td>MSB</td>
<td>344H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-82.741</td>
<td>6</td>
<td>-83.566</td>
<td>5</td>
<td>0.198965</td>
</tr>
<tr>
<td>MSB</td>
<td>64H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-80.756</td>
<td>6</td>
<td>-85.391</td>
<td>5</td>
<td>0.002330</td>
</tr>
<tr>
<td>MSB</td>
<td>492H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-82.578</td>
<td>6</td>
<td>-84.399</td>
<td>5</td>
<td>0.056305</td>
</tr>
<tr>
<td>MSB</td>
<td>193H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-75.564</td>
<td>6</td>
<td>-77.719</td>
<td>5</td>
<td>0.037905</td>
</tr>
<tr>
<td>MSB</td>
<td>260H</td>
<td>25</td>
<td>25</td>
<td>28</td>
<td>-84.092</td>
<td>6</td>
<td>-85.150</td>
<td>5</td>
<td>0.145780</td>
</tr>
<tr>
<td>MSB</td>
<td>579H</td>
<td>25</td>
<td>23</td>
<td>27</td>
<td>-80.000</td>
<td>6</td>
<td>-80.261</td>
<td>5</td>
<td>0.469619</td>
</tr>
<tr>
<td>MSB</td>
<td>371H</td>
<td>25</td>
<td>25</td>
<td>23</td>
<td>-66.936</td>
<td>6</td>
<td>-66.966</td>
<td>5</td>
<td>0.807561</td>
</tr>
</tbody>
</table>
### S10 Table. Test for mid-parent heterosis in Rosette Branching (RB).

Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $l_F$ denotes the log-likelihood of the full (unconstrained) model with degrees freedom given by $df (l_F)$; $l_A$ denotes the log-likelihood of the additive model with degrees of freedom given by $df (l_A)$.
### S11 Table. Test for mid-parent heterosis in Flowering Time (FT)

Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; \( I_F \) denotes the log-likelihood of the full (unconstrained) model with degrees of freedom given by \( df(I_F) \); \( I_A \) denotes the log-likelihood of the additive model with degrees of freedom given by \( df(I_A) \).

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>N&lt;sub&gt;Col-wt&lt;/sub&gt;</th>
<th>N&lt;sub&gt;NepRIL&lt;/sub&gt;</th>
<th>N&lt;sub&gt;F1&lt;/sub&gt;</th>
<th>( I_F )</th>
<th>( df(I_F) )</th>
<th>( I_A )</th>
<th>( df(I_A) )</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FT 14H</td>
<td>24 26 26</td>
<td>-143.000</td>
<td>6</td>
<td>-147.990</td>
<td>5</td>
<td>0.001582</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 232H</td>
<td>24 22 27</td>
<td>-132.370</td>
<td>6</td>
<td>-164.510</td>
<td>5</td>
<td>1.07E-15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 92H</td>
<td>24 27 27</td>
<td>-120.600</td>
<td>6</td>
<td>-125.380</td>
<td>5</td>
<td>0.001990</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 208H</td>
<td>24 24 25</td>
<td>-125.480</td>
<td>6</td>
<td>-132.940</td>
<td>5</td>
<td>0.000112</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 438H</td>
<td>24 25 26</td>
<td>-129.960</td>
<td>6</td>
<td>-130.420</td>
<td>5</td>
<td>0.338015</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 195H</td>
<td>24 25 28</td>
<td>-151.190</td>
<td>6</td>
<td>-151.200</td>
<td>5</td>
<td>0.921234</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 350H</td>
<td>24 24 26</td>
<td>-129.930</td>
<td>6</td>
<td>-134.270</td>
<td>5</td>
<td>0.003208</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 500H</td>
<td>24 18 27</td>
<td>-131.930</td>
<td>6</td>
<td>-132.120</td>
<td>5</td>
<td>0.535302</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 150H</td>
<td>24 27 25</td>
<td>-129.100</td>
<td>6</td>
<td>-149.140</td>
<td>5</td>
<td>2.44E-10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 118H</td>
<td>24 26 28</td>
<td>-149.340</td>
<td>6</td>
<td>-149.990</td>
<td>5</td>
<td>0.253293</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 432H</td>
<td>24 26 28</td>
<td>-151.590</td>
<td>6</td>
<td>-152.150</td>
<td>5</td>
<td>0.290411</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 202H</td>
<td>24 25 24</td>
<td>-122.340</td>
<td>6</td>
<td>-123.110</td>
<td>5</td>
<td>0.215080</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 344H</td>
<td>24 26 27</td>
<td>-134.510</td>
<td>6</td>
<td>-141.130</td>
<td>5</td>
<td>0.000276</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 64H</td>
<td>24 27 26</td>
<td>-138.830</td>
<td>6</td>
<td>-138.850</td>
<td>5</td>
<td>0.833507</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 492H</td>
<td>24 25 27</td>
<td>-139.610</td>
<td>6</td>
<td>-139.770</td>
<td>5</td>
<td>0.578003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 193H</td>
<td>24 24 27</td>
<td>-150.030</td>
<td>6</td>
<td>-154.540</td>
<td>5</td>
<td>0.002682</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 260H</td>
<td>24 28 28</td>
<td>-152.570</td>
<td>6</td>
<td>-152.910</td>
<td>5</td>
<td>0.413305</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FT 579H</td>
<td>24 21 27</td>
<td>-134.240</td>
<td>6</td>
<td>-134.430</td>
<td>5</td>
<td>0.544378</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenotype</td>
<td>epiHybrid ID</td>
<td>N_Col-wt</td>
<td>N_epiRIL</td>
<td>N_F1</td>
<td>$l_F$</td>
<td>df ($l_F$)</td>
<td>$l_A$</td>
<td>df ($l_A$)</td>
<td>P-value</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------</td>
<td>---------</td>
<td>---------</td>
<td>-----</td>
<td>------</td>
<td>----------</td>
<td>------</td>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>GR</td>
<td>14H</td>
<td>25</td>
<td>27</td>
<td>26</td>
<td>150.551</td>
<td>6</td>
<td>150.493</td>
<td>5</td>
<td>0.733068</td>
</tr>
<tr>
<td>GR</td>
<td>232H</td>
<td>25</td>
<td>23</td>
<td>26</td>
<td>128.591</td>
<td>6</td>
<td>125.538</td>
<td>5</td>
<td>0.013471</td>
</tr>
<tr>
<td>GR</td>
<td>92H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>157.357</td>
<td>6</td>
<td>157.349</td>
<td>5</td>
<td>0.898575</td>
</tr>
<tr>
<td>GR</td>
<td>208H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>151.884</td>
<td>6</td>
<td>151.826</td>
<td>5</td>
<td>0.733505</td>
</tr>
<tr>
<td>GR</td>
<td>438H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>142.269</td>
<td>6</td>
<td>142.223</td>
<td>5</td>
<td>0.762863</td>
</tr>
<tr>
<td>GR</td>
<td>195H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>168.894</td>
<td>6</td>
<td>167.574</td>
<td>5</td>
<td>0.104150</td>
</tr>
<tr>
<td>GR</td>
<td>350H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>154.404</td>
<td>6</td>
<td>154.245</td>
<td>5</td>
<td>0.572709</td>
</tr>
<tr>
<td>GR</td>
<td>500H</td>
<td>25</td>
<td>17</td>
<td>25</td>
<td>123.444</td>
<td>6</td>
<td>123.021</td>
<td>5</td>
<td>0.357796</td>
</tr>
<tr>
<td>GR</td>
<td>150H</td>
<td>25</td>
<td>24</td>
<td>25</td>
<td>143.266</td>
<td>6</td>
<td>141.757</td>
<td>5</td>
<td>0.082317</td>
</tr>
<tr>
<td>GR</td>
<td>118H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>157.024</td>
<td>6</td>
<td>157.017</td>
<td>5</td>
<td>0.911717</td>
</tr>
<tr>
<td>GR</td>
<td>432H</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>150.435</td>
<td>6</td>
<td>150.126</td>
<td>5</td>
<td>0.431830</td>
</tr>
<tr>
<td>GR</td>
<td>202H</td>
<td>25</td>
<td>24</td>
<td>22</td>
<td>140.537</td>
<td>6</td>
<td>140.505</td>
<td>5</td>
<td>0.800766</td>
</tr>
<tr>
<td>GR</td>
<td>344H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>163.081</td>
<td>6</td>
<td>162.924</td>
<td>5</td>
<td>0.575091</td>
</tr>
<tr>
<td>GR</td>
<td>64H</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>141.311</td>
<td>6</td>
<td>141.232</td>
<td>5</td>
<td>0.690526</td>
</tr>
<tr>
<td>GR</td>
<td>492H</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>141.750</td>
<td>6</td>
<td>141.525</td>
<td>5</td>
<td>0.501849</td>
</tr>
<tr>
<td>GR</td>
<td>193H</td>
<td>25</td>
<td>24</td>
<td>27</td>
<td>145.561</td>
<td>6</td>
<td>145.379</td>
<td>5</td>
<td>0.547169</td>
</tr>
<tr>
<td>GR</td>
<td>260H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>156.114</td>
<td>6</td>
<td>155.174</td>
<td>5</td>
<td>0.170241</td>
</tr>
<tr>
<td>GR</td>
<td>579H</td>
<td>25</td>
<td>22</td>
<td>26</td>
<td>140.752</td>
<td>6</td>
<td>140.719</td>
<td>5</td>
<td>0.796252</td>
</tr>
<tr>
<td>GR</td>
<td>371H</td>
<td>25</td>
<td>25</td>
<td>24</td>
<td>152.041</td>
<td>6</td>
<td>151.537</td>
<td>5</td>
<td>0.315467</td>
</tr>
</tbody>
</table>

**S12 Table. Test for mid-parent heterosis in Growth Rate (GR).** Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $l_F$ denotes the log-likelihood of the full (unconstrained) model with degrees freedom given by $df (l_F)$; $l_A$ denotes the log-likelihood of the additive model with degrees of freedom given by $df (l_A)$. 

86
### S13 Table. Test for mid-parent heterosis in Leaf Area (LA).

Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $l_F$ denotes the log-likelihood of the full (unconstrained) model with degrees freedom given by $df\ (l_F)$; $l_A$ denotes the log-likelihood of the additive model with degrees of freedom given by $df\ (l_A)$.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>N_Col-wt</th>
<th>N_epiRIL</th>
<th>NF1</th>
<th>$l_F$</th>
<th>df ($l_F$)</th>
<th>$l_A$</th>
<th>df ($l_A$)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA</td>
<td>14H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-382.560</td>
<td>6</td>
<td>-384.760</td>
<td>5</td>
<td>0.036036</td>
</tr>
<tr>
<td>LA</td>
<td>232H</td>
<td>25</td>
<td>23</td>
<td>27</td>
<td>-361.760</td>
<td>6</td>
<td>-400.260</td>
<td>5</td>
<td>1.70E-18</td>
</tr>
<tr>
<td>LA</td>
<td>92H</td>
<td>25</td>
<td>28</td>
<td>27</td>
<td>-390.460</td>
<td>6</td>
<td>-392.140</td>
<td>5</td>
<td>0.066879</td>
</tr>
<tr>
<td>LA</td>
<td>208H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>-352.600</td>
<td>6</td>
<td>-355.030</td>
<td>5</td>
<td>0.027650</td>
</tr>
<tr>
<td>LA</td>
<td>438H</td>
<td>25</td>
<td>25</td>
<td>28</td>
<td>-389.870</td>
<td>6</td>
<td>-390.120</td>
<td>5</td>
<td>0.477701</td>
</tr>
<tr>
<td>LA</td>
<td>195H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-389.670</td>
<td>6</td>
<td>-396.390</td>
<td>5</td>
<td>0.000247</td>
</tr>
<tr>
<td>LA</td>
<td>350H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-383.920</td>
<td>6</td>
<td>-386.240</td>
<td>5</td>
<td>0.031481</td>
</tr>
<tr>
<td>LA</td>
<td>500H</td>
<td>25</td>
<td>17</td>
<td>26</td>
<td>-344.270</td>
<td>6</td>
<td>-344.830</td>
<td>5</td>
<td>0.290000</td>
</tr>
<tr>
<td>LA</td>
<td>150H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-397.660</td>
<td>6</td>
<td>-403.140</td>
<td>5</td>
<td>0.000935</td>
</tr>
<tr>
<td>LA</td>
<td>118H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-380.970</td>
<td>6</td>
<td>-381.000</td>
<td>5</td>
<td>0.808696</td>
</tr>
<tr>
<td>LA</td>
<td>432H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-377.930</td>
<td>6</td>
<td>-377.930</td>
<td>5</td>
<td>0.923223</td>
</tr>
<tr>
<td>LA</td>
<td>202H</td>
<td>25</td>
<td>25</td>
<td>24</td>
<td>-357.570</td>
<td>6</td>
<td>-358.360</td>
<td>5</td>
<td>0.208724</td>
</tr>
<tr>
<td>LA</td>
<td>344H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-382.430</td>
<td>6</td>
<td>-383.470</td>
<td>5</td>
<td>0.149790</td>
</tr>
<tr>
<td>LA</td>
<td>64H</td>
<td>25</td>
<td>28</td>
<td>27</td>
<td>-394.490</td>
<td>6</td>
<td>-394.530</td>
<td>5</td>
<td>0.767753</td>
</tr>
<tr>
<td>LA</td>
<td>492H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-389.460</td>
<td>6</td>
<td>-390.820</td>
<td>5</td>
<td>0.098866</td>
</tr>
<tr>
<td>LA</td>
<td>193H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-387.710</td>
<td>6</td>
<td>-398.830</td>
<td>5</td>
<td>2.42E-06</td>
</tr>
<tr>
<td>LA</td>
<td>260H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>-379.510</td>
<td>6</td>
<td>-381.790</td>
<td>5</td>
<td>0.032656</td>
</tr>
<tr>
<td>LA</td>
<td>579H</td>
<td>25</td>
<td>21</td>
<td>28</td>
<td>-372.110</td>
<td>6</td>
<td>-372.150</td>
<td>5</td>
<td>0.761378</td>
</tr>
<tr>
<td>LA</td>
<td>371H</td>
<td>25</td>
<td>27</td>
<td>24</td>
<td>-376.310</td>
<td>6</td>
<td>-376.310</td>
<td>5</td>
<td>0.962719</td>
</tr>
</tbody>
</table>
Chapter 3

S14 Table. Test for high (low)-parent heterosis in Height (HT). Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; \( l_F \) denotes the log-likelihood of the full (unconstrained) model with degrees freedom given by \( df (l_F) \); \( l_{FD} \) denotes the log-likelihood of the full dominance model with degrees of freedom given by \( df (l_{FD}) \); a horizontal line “- -” indicates that this particular line showed no evidence for mid-parent heterosis and was therefore not tested further.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>NCol-wt</th>
<th>NepiRIL</th>
<th>NF1</th>
<th>( l_F )</th>
<th>( df (l_F) )</th>
<th>( l_{FD} )</th>
<th>( df (l_{FD}) )</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HT</td>
<td>14H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-175.090</td>
<td>6</td>
<td>-175.170</td>
<td>5</td>
<td>0.685930</td>
</tr>
<tr>
<td>HT</td>
<td>232H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>-193.750</td>
<td>6</td>
<td>-193.990</td>
<td>5</td>
<td>0.487088</td>
</tr>
<tr>
<td>HT</td>
<td>92H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-171.980</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td>208H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>-172.530</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td>438H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-172.740</td>
<td>6</td>
<td>-174.070</td>
<td>5</td>
<td>0.102791</td>
</tr>
<tr>
<td>HT</td>
<td>195H</td>
<td>25</td>
<td>23</td>
<td>26</td>
<td>-170.170</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td>350H</td>
<td>25</td>
<td>25</td>
<td>24</td>
<td>-162.600</td>
<td>6</td>
<td>-162.720</td>
<td>5</td>
<td>0.620032</td>
</tr>
<tr>
<td>HT</td>
<td>500H</td>
<td>25</td>
<td>19</td>
<td>28</td>
<td>-215.090</td>
<td>6</td>
<td>-220.570</td>
<td>5</td>
<td>0.000933</td>
</tr>
<tr>
<td>HT</td>
<td>150H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-183.600</td>
<td>6</td>
<td>-185.510</td>
<td>5</td>
<td>0.050756</td>
</tr>
<tr>
<td>HT</td>
<td>118H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>-167.890</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td>432H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-172.160</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td>202H</td>
<td>25</td>
<td>25</td>
<td>24</td>
<td>-167.340</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td>344H</td>
<td>25</td>
<td>27</td>
<td>25</td>
<td>-161.820</td>
<td>6</td>
<td>-164.970</td>
<td>5</td>
<td>0.012123</td>
</tr>
<tr>
<td>HT</td>
<td>64H</td>
<td>25</td>
<td>26</td>
<td>25</td>
<td>-157.050</td>
<td>6</td>
<td>-159.300</td>
<td>5</td>
<td>0.034208</td>
</tr>
<tr>
<td>HT</td>
<td>492H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-178.750</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>HT</td>
<td>193H</td>
<td>25</td>
<td>24</td>
<td>26</td>
<td>-184.730</td>
<td>6</td>
<td>-193.540</td>
<td>5</td>
<td>2.70E-05</td>
</tr>
<tr>
<td>HT</td>
<td>260H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-183.570</td>
<td>6</td>
<td>-184.820</td>
<td>5</td>
<td>0.114777</td>
</tr>
<tr>
<td>HT</td>
<td>579H</td>
<td>25</td>
<td>23</td>
<td>28</td>
<td>-173.510</td>
<td>6</td>
<td>-173.920</td>
<td>5</td>
<td>0.370676</td>
</tr>
<tr>
<td>HT</td>
<td>371H</td>
<td>25</td>
<td>26</td>
<td>25</td>
<td>-186.760</td>
<td>6</td>
<td>-192.430</td>
<td>5</td>
<td>0.000761</td>
</tr>
</tbody>
</table>
S15 Table. Test for high(low)-parent heterosis in Main Stem Branching (MSB).
Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $l_f$ denotes the log-likelihood of the full (unconstrained) model with degrees freedom given by $df (l_f)$; $l_{FD}$ denotes the log-likelihood of the full dominance model with degrees of freedom given by $df (l_{FD})$; a horizontal line “- -” indicates that this particular line showed no evidence for mid-parent heterosis and was therefore not tested further.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>NCol-wt</th>
<th>NepiRIL</th>
<th>N$_{F_1}$</th>
<th>$l_f$</th>
<th>$df (l_f)$</th>
<th>$l_{FD}$</th>
<th>$df (l_{FD})$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSB</td>
<td>14H</td>
<td>25</td>
<td>26</td>
<td>28</td>
<td>-86.633</td>
<td>6</td>
<td>-86.726</td>
<td>5</td>
<td>0.666353</td>
</tr>
<tr>
<td>MSB</td>
<td>232H</td>
<td>25</td>
<td>22</td>
<td>27</td>
<td>-84.846</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>92H</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>-66.653</td>
<td>6</td>
<td>-67.158</td>
<td>5</td>
<td>0.314777</td>
</tr>
<tr>
<td>MSB</td>
<td>208H</td>
<td>25</td>
<td>25</td>
<td>26</td>
<td>-94.913</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>438H</td>
<td>25</td>
<td>25</td>
<td>28</td>
<td>-86.590</td>
<td>6</td>
<td>-91.527</td>
<td>5</td>
<td>0.001678</td>
</tr>
<tr>
<td>MSB</td>
<td>195H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-76.313</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>350H</td>
<td>25</td>
<td>25</td>
<td>23</td>
<td>-76.077</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>500H</td>
<td>25</td>
<td>19</td>
<td>28</td>
<td>-81.993</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>150H</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>-78.908</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>118H</td>
<td>25</td>
<td>28</td>
<td>28</td>
<td>-77.750</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>432H</td>
<td>25</td>
<td>26</td>
<td>26</td>
<td>-74.565</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>202H</td>
<td>25</td>
<td>26</td>
<td>24</td>
<td>-71.846</td>
<td>6</td>
<td>-71.944</td>
<td>5</td>
<td>0.657954</td>
</tr>
<tr>
<td>MSB</td>
<td>344H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-82.741</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>64H</td>
<td>25</td>
<td>27</td>
<td>27</td>
<td>-80.756</td>
<td>6</td>
<td>-82.700</td>
<td>5</td>
<td>0.048647</td>
</tr>
<tr>
<td>MSB</td>
<td>492H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-82.578</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>193H</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>-75.564</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>260H</td>
<td>25</td>
<td>25</td>
<td>28</td>
<td>-84.092</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>579H</td>
<td>25</td>
<td>23</td>
<td>27</td>
<td>-80.000</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>MSB</td>
<td>371H</td>
<td>25</td>
<td>25</td>
<td>23</td>
<td>-66.936</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>
Chapter 3

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>N_Col-wt</th>
<th>N_RIL</th>
<th>N_R1</th>
<th>$l_F$</th>
<th>$df (l_F)$</th>
<th>$l_{FD}$</th>
<th>$df (l_{FD})$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>RB</td>
<td>14H</td>
<td>23</td>
<td>28</td>
<td>27</td>
<td>-123.530</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>232H</td>
<td>23</td>
<td>22</td>
<td>27</td>
<td>-106.450</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>92H</td>
<td>23</td>
<td>27</td>
<td>26</td>
<td>-107.240</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>208H</td>
<td>23</td>
<td>26</td>
<td>27</td>
<td>-120.610</td>
<td>6</td>
<td>-120.670</td>
<td>5</td>
<td>0.729701</td>
</tr>
<tr>
<td>RB</td>
<td>438H</td>
<td>23</td>
<td>25</td>
<td>27</td>
<td>-112.350</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>195H</td>
<td>23</td>
<td>26</td>
<td>28</td>
<td>-119.220</td>
<td>6</td>
<td>-119.670</td>
<td>5</td>
<td>0.340869</td>
</tr>
<tr>
<td>RB</td>
<td>350H</td>
<td>23</td>
<td>26</td>
<td>25</td>
<td>-119.270</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>500H</td>
<td>23</td>
<td>19</td>
<td>28</td>
<td>-111.220</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>150H</td>
<td>23</td>
<td>25</td>
<td>26</td>
<td>-99.803</td>
<td>6</td>
<td>-100.130</td>
<td>5</td>
<td>0.415976</td>
</tr>
<tr>
<td>RB</td>
<td>118H</td>
<td>23</td>
<td>26</td>
<td>28</td>
<td>-114.860</td>
<td>6</td>
<td>-115.030</td>
<td>5</td>
<td>0.559295</td>
</tr>
<tr>
<td>RB</td>
<td>432H</td>
<td>23</td>
<td>27</td>
<td>28</td>
<td>-128.700</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>202H</td>
<td>23</td>
<td>25</td>
<td>25</td>
<td>-99.779</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>344H</td>
<td>23</td>
<td>27</td>
<td>28</td>
<td>-113.940</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>64H</td>
<td>23</td>
<td>27</td>
<td>28</td>
<td>-117.340</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>492H</td>
<td>23</td>
<td>26</td>
<td>28</td>
<td>-107.190</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>193H</td>
<td>23</td>
<td>24</td>
<td>27</td>
<td>-112.790</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>260H</td>
<td>23</td>
<td>26</td>
<td>28</td>
<td>-113.820</td>
<td>6</td>
<td>-113.820</td>
<td>5</td>
<td>0.930222</td>
</tr>
<tr>
<td>RB</td>
<td>579H</td>
<td>23</td>
<td>23</td>
<td>27</td>
<td>-113.210</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>RB</td>
<td>371H</td>
<td>23</td>
<td>26</td>
<td>25</td>
<td>-118.660</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

S16 Table. Test for high(low)-parent heterosis in Rosette Branching (RB). Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $l_F$ denotes the log-likelihood of the full (unconstrained) model with degrees freedom given by $df (l_F)$; $l_{FD}$ denotes the log-likelihood of the full dominance model with degrees of freedom given by $df (l_{FD})$; a horizontal line “--” indicates that this particular line showed no evidence for mid-parent heterosis and was therefore not tested further.
**Chapter 3**

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>epiHybrid ID</th>
<th>N_col wt</th>
<th>N_epiRIL</th>
<th>N_F1</th>
<th>l_F</th>
<th>df (l_F)</th>
<th>l_FD</th>
<th>df (l_FD)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>FT</td>
<td>14H</td>
<td>24</td>
<td>26</td>
<td>26</td>
<td>-143.000</td>
<td>6</td>
<td>-144.160</td>
<td>5</td>
<td>0.127591</td>
</tr>
<tr>
<td>FT</td>
<td>232H</td>
<td>24</td>
<td>22</td>
<td>27</td>
<td>-132.370</td>
<td>6</td>
<td>-149.350</td>
<td>5</td>
<td>5.60E-09</td>
</tr>
<tr>
<td>FT</td>
<td>92H</td>
<td>24</td>
<td>27</td>
<td>27</td>
<td>-120.600</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>208H</td>
<td>24</td>
<td>24</td>
<td>25</td>
<td>-125.480</td>
<td>6</td>
<td>-127.450</td>
<td>5</td>
<td>0.047087</td>
</tr>
<tr>
<td>FT</td>
<td>438H</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>-129.960</td>
<td>6</td>
<td>-130.070</td>
<td>5</td>
<td>0.635624</td>
</tr>
<tr>
<td>FT</td>
<td>195H</td>
<td>24</td>
<td>25</td>
<td>28</td>
<td>-151.190</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>350H</td>
<td>24</td>
<td>24</td>
<td>26</td>
<td>-129.930</td>
<td>6</td>
<td>-130.230</td>
<td>5</td>
<td>0.438016</td>
</tr>
<tr>
<td>FT</td>
<td>500H</td>
<td>24</td>
<td>18</td>
<td>27</td>
<td>-131.930</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>150H</td>
<td>24</td>
<td>27</td>
<td>25</td>
<td>-129.100</td>
<td>6</td>
<td>-130.970</td>
<td>5</td>
<td>0.053643</td>
</tr>
<tr>
<td>FT</td>
<td>118H</td>
<td>24</td>
<td>26</td>
<td>28</td>
<td>-149.340</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>432H</td>
<td>24</td>
<td>26</td>
<td>28</td>
<td>-151.590</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>202H</td>
<td>24</td>
<td>25</td>
<td>24</td>
<td>-122.340</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>344H</td>
<td>24</td>
<td>26</td>
<td>27</td>
<td>-134.510</td>
<td>6</td>
<td>-137.740</td>
<td>5</td>
<td>0.011081</td>
</tr>
<tr>
<td>FT</td>
<td>64H</td>
<td>24</td>
<td>27</td>
<td>26</td>
<td>-138.830</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>492H</td>
<td>24</td>
<td>25</td>
<td>27</td>
<td>-139.610</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>193H</td>
<td>24</td>
<td>24</td>
<td>27</td>
<td>-150.030</td>
<td>6</td>
<td>-150.200</td>
<td>5</td>
<td>0.562762</td>
</tr>
<tr>
<td>FT</td>
<td>260H</td>
<td>24</td>
<td>28</td>
<td>28</td>
<td>-152.570</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>579H</td>
<td>24</td>
<td>21</td>
<td>27</td>
<td>-134.240</td>
<td>6</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>FT</td>
<td>371H</td>
<td>24</td>
<td>28</td>
<td>24</td>
<td>-145.990</td>
<td>6</td>
<td>-146.000</td>
<td>5</td>
<td>0.915332</td>
</tr>
</tbody>
</table>

**S17 Table. Test for high(low)-parent heterosis in Flowering Time (FT).** Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $l_F$ denotes the log-likelihood of the full (unconstrained) model with degrees freedom given by $df (l_F)$; $l_{FD}$ denotes the log-likelihood of the full dominance model with degrees of freedom given by $df (l_{FD})$; a horizontal line “- - -” indicates that this particular line showed no evidence for mid-parent heterosis and was therefore not tested further.
### S18 Table. Test for high(low)-parent heterosis in Growth Rate (GR).

Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $I_F$ denotes the log-likelihood of the full (unconstrained) model with degrees freedom given by $df (I_F)$; $I_{FD}$ denotes the log-likelihood of the full dominance model with degrees of freedom given by $df (I_{FD})$; a horizontal line “- -” indicates that this particular line showed no evidence for mid-parent heterosis and was therefore not tested further.
S19 Table. Test for high (low)-parent heterosis in Leaf Area (LA). Summarized are the Likelihood Ratio Test (LRT) results for each of the lines; $l_F$ denotes the log-likelihood of the full (unconstrained) model with degrees freedom given by $df (l_F)$; $l_{FD}$ denotes the log-likelihood of the full dominance model with degrees of freedom given by $df (l_{FD})$; a horizontal line “- -” indicates that this particular line showed no evidence for mid-parent heterosis and was therefore not tested further.
### S20 Table. Variance component analysis for mid-parent heterosis.

This table shows the proportion of variance ($R^2_{adj}$) in mid-parent heterosis among the ~500 F1 plants that can be explained by (epi)genomic differences between the epiRIL parental lines used for the 19 crosses.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>$R^2_{adj}$</th>
<th>$F$-value</th>
<th>df.numerator</th>
<th>df.denominator</th>
<th>$P$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR</td>
<td>0.026272</td>
<td>1.732991</td>
<td>18</td>
<td>471</td>
<td>3.11E-02</td>
</tr>
<tr>
<td>RB</td>
<td>0.017461</td>
<td>1.505481</td>
<td>18</td>
<td>494</td>
<td>8.27E-02</td>
</tr>
<tr>
<td>LA</td>
<td>0.278716</td>
<td>11.841114</td>
<td>18</td>
<td>487</td>
<td>1.15E-28</td>
</tr>
<tr>
<td>HT</td>
<td>0.512668</td>
<td>30.04658</td>
<td>18</td>
<td>479</td>
<td>5.01E-67</td>
</tr>
<tr>
<td>MSB</td>
<td>0.163212</td>
<td>6.439624</td>
<td>18</td>
<td>484</td>
<td>2.07E-14</td>
</tr>
<tr>
<td>FT</td>
<td>0.174109</td>
<td>6.879348</td>
<td>18</td>
<td>484</td>
<td>1.31E-15</td>
</tr>
</tbody>
</table>
Chapter 3

Table. Summary of interval mapping results. Shown are the LOD scores of the peak QTL DMRs (bold) along with lower and upper confidence intervals (see Type; 1.5 LOD drop-off). The genetic (cM) and physical (bps) locations correspond to the DMRs most proximal to the QTL peak, and are indicated as DMR*. Genome-wide significant QTL were only detected for Flowering Time (FT) and Height (HT). However, because the QTL profiles for Leaf Area (LA) appear to trace those of FT (Fig. 3B), we also provide the effects of the FT QTLs on Leaf Area (LA). Genome-wide LOD thresholds corresponding to a 5% false positive rate were obtained from 10,000 permutations of the data. These thresholds were 2.88, 4.34 and 3.24 for FT, LA and HT, respectively. The threshold normalized LOD scores (see LOD/Thr) are plotted in Figure 3B.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>DMR</th>
<th>DMR*</th>
<th>Type</th>
<th>LOD</th>
<th>LOD/Thr</th>
<th>Chr</th>
<th>Region</th>
<th>Position (cM)</th>
<th>DMR* start (bps)</th>
<th>DMR* stop (bps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flowering Time (FT)</td>
<td>MM399</td>
<td>MM399</td>
<td>lower</td>
<td>3</td>
<td>1</td>
<td>34.80</td>
<td>8937118</td>
<td>8938596</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM405</td>
<td>MM405</td>
<td>peak</td>
<td>3.12</td>
<td>1.08</td>
<td>3</td>
<td>37.62</td>
<td>9692187</td>
<td>9698029</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c3.loc40</td>
<td>MM405</td>
<td>upper</td>
<td>3</td>
<td>1</td>
<td>40.00</td>
<td>9692187</td>
<td>9698029</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM546</td>
<td>MM546</td>
<td>lower</td>
<td>3</td>
<td>2</td>
<td>95.64</td>
<td>22232528</td>
<td>22235548</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM547</td>
<td>MM547</td>
<td>peak</td>
<td>3.33</td>
<td>1.16</td>
<td>3</td>
<td>101.44</td>
<td>23204533</td>
<td>23207428</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM547</td>
<td>MM547</td>
<td>upper</td>
<td>3</td>
<td>2</td>
<td>101.44</td>
<td>23204533</td>
<td>23207428</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leaf Area (LA)</td>
<td>MM405</td>
<td>MM405</td>
<td>peak</td>
<td>2.40</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM547</td>
<td>MM547</td>
<td>peak</td>
<td>2.26</td>
<td>0.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Genome-wide LOD thresholds corresponding to a 5% false positive rate were obtained from 10,000 permutations of the data. These thresholds were 2.88, 4.34 and 3.24 for FT, LA and HT, respectively. The threshold normalized LOD scores (see LOD/Thr) are plotted in Figure 3B.
Chapter 3

<table>
<thead>
<tr>
<th></th>
<th>Flowering Time &amp; Leaf Area</th>
<th>Height</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chr 3</td>
<td>Chr 4</td>
</tr>
<tr>
<td></td>
<td>Region 1</td>
<td>Region 2</td>
</tr>
<tr>
<td># Probes in QTL interval</td>
<td>4608</td>
<td>5883</td>
</tr>
<tr>
<td># Probes pass 1 - quality</td>
<td>3719</td>
<td>4891</td>
</tr>
<tr>
<td># Probes pass 2 - loss meth</td>
<td>387</td>
<td>190</td>
</tr>
<tr>
<td># Probes pass 3 - correlation</td>
<td>171</td>
<td>70</td>
</tr>
<tr>
<td># DMRs</td>
<td>39</td>
<td>16</td>
</tr>
<tr>
<td># Unique genes</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td># Gene promoters (GP)</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td># Gene bodies (GB)</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td># Transposable elements (TE)</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td># Intergenic regions (IGR)</td>
<td>9</td>
<td>5</td>
</tr>
</tbody>
</table>

**S22 Table. Number of DMRs detected and number of annotation units overlapping the DMRs.** The number of DMRs detected for each QTL interval after several filtering steps and the number of annotation units (gene promoters, gene bodies, transposable elements and intergenic regions) overlapping the DMRs.
S23 Table. Annotation categories that have an overlap with the Flowering Time (FT) and Leaf Area (LA) DMRs. Indicated are the annotation categories that have an overlap with the Flowering Time (FT) and Leaf Area (LA) DMRs (Unit Class: GP: Gene Promoter; GB: Gene Body; TE: Transposable element; IGR: Intergenic Region). The DMRs of both phenotypes are the same. Therefore only one table is provided. The ID of the DMRs starts with “FL” (Flowering Time and Leaf Area). The inserted table at the bottom shows the number of DMRs with a certain combination of annotations. Genomic locations of the DMRs are in Table S25.

<table>
<thead>
<tr>
<th>Annotation Combination</th>
<th># DMRs</th>
<th>Genomic locations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene Promoter only</td>
<td>1</td>
<td>FLCHR03REG01DMR0001</td>
</tr>
<tr>
<td>Gene Body only</td>
<td>11</td>
<td>FLCHR03REG01DMR0002</td>
</tr>
<tr>
<td>Transposon only</td>
<td>12</td>
<td>FLCHR03REG01DMR0003</td>
</tr>
<tr>
<td>Intergenic only</td>
<td>2</td>
<td>FLCHR03REG01DMR0004</td>
</tr>
<tr>
<td>Gene Promoter &amp; Gene Body</td>
<td>8</td>
<td>FLCHR03REG01DMR0005</td>
</tr>
<tr>
<td>Gene Promoter &amp; Transposon</td>
<td>8</td>
<td>FLCHR03REG01DMR0006</td>
</tr>
<tr>
<td>Gene Promoter &amp; Intergenic</td>
<td>1</td>
<td>FLCHR03REG01DMR0007</td>
</tr>
<tr>
<td>Gene Body &amp; Transposon</td>
<td>1</td>
<td>FLCHR03REG01DMR0008</td>
</tr>
<tr>
<td>Gene Body &amp; Intergenic</td>
<td>1</td>
<td>FLCHR03REG01DMR0009</td>
</tr>
<tr>
<td>Transposon &amp; Intergenic</td>
<td>8</td>
<td>FLCHR03REG01DMR0010</td>
</tr>
<tr>
<td>Gene Promoter &amp; Gene Body &amp; Transposon</td>
<td>2</td>
<td>FLCHR03REG01DMR0011</td>
</tr>
<tr>
<td>Gene Promoter &amp; Gene Body &amp; Intergenic</td>
<td>12</td>
<td>FLCHR03REG01DMR0012</td>
</tr>
<tr>
<td>Gene Promoter &amp; Transposon &amp; Intergenic</td>
<td>2</td>
<td>FLCHR03REG01DMR0013</td>
</tr>
<tr>
<td>Gene Body &amp; Transposon &amp; Intergenic</td>
<td>1</td>
<td>FLCHR03REG01DMR0014</td>
</tr>
<tr>
<td>Gene Body &amp; Intergenic &amp; Transposon &amp; Intergenic</td>
<td>8</td>
<td>FLCHR03REG01DMR0015</td>
</tr>
<tr>
<td>Gene Promoter &amp; Gene Body &amp; Transposon &amp; Intergenic</td>
<td>2</td>
<td>FLCHR03REG01DMR0016</td>
</tr>
<tr>
<td>Gene Promoter &amp; Gene Body &amp; Transposon &amp; Intergenic</td>
<td>1</td>
<td>FLCHR03REG01DMR0017</td>
</tr>
<tr>
<td>Gene Promoter &amp; Gene Body &amp; Transposon &amp; Intergenic</td>
<td>1</td>
<td>FLCHR03REG01DMR0018</td>
</tr>
<tr>
<td>Gene Promoter &amp; Gene Body &amp; Transposon &amp; Intergenic</td>
<td>1</td>
<td>FLCHR03REG01DMR0019</td>
</tr>
<tr>
<td>Gene Promoter &amp; Gene Body &amp; Transposon &amp; Intergenic</td>
<td>1</td>
<td>FLCHR03REG01DMR0020</td>
</tr>
</tbody>
</table>
Chapter 3

<table>
<thead>
<tr>
<th>DMR ID</th>
<th>Unit Class</th>
<th>DMR ID</th>
<th>Unit Class</th>
<th>DMR ID</th>
<th>Unit Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTCHR04REG01DMR0001</td>
<td>X X</td>
<td>HTCHR04REG01DMR0007</td>
<td>X</td>
<td>HTCHR04REG01DMR0013</td>
<td>X</td>
</tr>
<tr>
<td>HTCHR04REG01DMR0002</td>
<td>X X</td>
<td>HTCHR04REG01DMR0008</td>
<td>X</td>
<td>HTCHR04REG01DMR0014</td>
<td>X</td>
</tr>
<tr>
<td>HTCHR04REG01DMR0003</td>
<td>X X</td>
<td>HTCHR04REG01DMR0009</td>
<td>X</td>
<td>HTCHR04REG01DMR0015</td>
<td>X</td>
</tr>
<tr>
<td>HTCHR04REG01DMR0004</td>
<td>X</td>
<td>HTCHR04REG01DMR0010</td>
<td>X</td>
<td>HTCHR04REG01DMR0016</td>
<td>X X</td>
</tr>
<tr>
<td>HTCHR04REG01DMR0005</td>
<td>X X</td>
<td>HTCHR04REG01DMR0011</td>
<td>X</td>
<td>HTCHR04REG01DMR0017</td>
<td>X X X</td>
</tr>
<tr>
<td>HTCHR04REG01DMR0006</td>
<td>X X</td>
<td>HTCHR04REG01DMR0012</td>
<td>X</td>
<td>HTCHR04REG01DMR0018</td>
<td>X X X</td>
</tr>
</tbody>
</table>

**Annotation Combination** | **# DMRs**
--- | ---
Gene Promoter only | 1
Gene Body only | 8
Transposon only | 1
Intergenic only | 0
Gene Promoter & Gene Body | 1
Gene Promoter & Transposon | 0
Gene Promoter & Intergenic | 1
Gene Body & Transposon | 0
Gene Body & Intergenic | 0
Transposon & Intergenic | 4
Gene Promoter & Gene Body & Transposon | 0
Gene Promoter & Gene Body & Intergenic | 0
Gene Promoter & Transposon & Intergenic | 2
Gene Body & Transposon & Intergenic | 0
Gene Promoter & Gene Body & Transposon & Intergenic | 0

**S24 Table.** Annotation categories that have an overlap with the Height (HT) DMRs. Indicated are the annotation categories that have an overlap with the Height (HT) DMRs (Unit Class: GP: Gene Promoter; GB: Gene Body; TE: Transposable element; IGR: Intergenic Region). The ID of the DMRs starts with “HT”. The inserted table at the bottom shows the number of DMRs with a certain combination of annotations. Genomic locations of the DMRs are in Table S26.
<table>
<thead>
<tr>
<th>Chr</th>
<th>Start</th>
<th>Stop</th>
<th>Transposon Type</th>
<th>Family</th>
<th>Superfamily</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>895800</td>
<td>895855</td>
<td>protein_coding_gene</td>
<td>ORC37</td>
<td>Ubiquitin-conjugating enzyme 37</td>
</tr>
<tr>
<td>2</td>
<td>895810</td>
<td>895815</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>3</td>
<td>895820</td>
<td>895825</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>4</td>
<td>895830</td>
<td>895835</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>5</td>
<td>895840</td>
<td>895845</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>6</td>
<td>895850</td>
<td>895855</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>7</td>
<td>895860</td>
<td>895865</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>8</td>
<td>895870</td>
<td>895875</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>9</td>
<td>895880</td>
<td>895885</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>10</td>
<td>895890</td>
<td>895895</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>11</td>
<td>895900</td>
<td>895905</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>12</td>
<td>895910</td>
<td>895915</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>13</td>
<td>895920</td>
<td>895925</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>14</td>
<td>895930</td>
<td>895935</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>15</td>
<td>895940</td>
<td>895945</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>16</td>
<td>895950</td>
<td>895955</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>17</td>
<td>895960</td>
<td>895965</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>18</td>
<td>895970</td>
<td>895975</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>19</td>
<td>895980</td>
<td>895985</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>20</td>
<td>895990</td>
<td>896000</td>
<td>protein_coding_gene</td>
<td>unknown protein</td>
<td>unknown protein</td>
</tr>
<tr>
<td>ID</td>
<td>Chromosome</td>
<td>Gene Symbol</td>
<td>Gene Name</td>
<td>Protein Description</td>
<td>Protein Function</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>-------------</td>
<td>-----------</td>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>LAG1</td>
<td>Longevity assurance gene 1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase A-2</td>
<td>ATPase A-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase B-2</td>
<td>ATPase B-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase C-2</td>
<td>ATPase C-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase D-2</td>
<td>ATPase D-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
</tbody>
</table>

ATPases are involved in various cellular processes, including energy generation and waste disposal.

<table>
<thead>
<tr>
<th>ID</th>
<th>Chromosome</th>
<th>Gene Symbol</th>
<th>Gene Name</th>
<th>Protein Description</th>
<th>Protein Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase E-2</td>
<td>ATPase E-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase F-2</td>
<td>ATPase F-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase G-2</td>
<td>ATPase G-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase H-2</td>
<td>ATPase H-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
</tbody>
</table>

ATPases are involved in various cellular processes, including energy generation and waste disposal.

<table>
<thead>
<tr>
<th>ID</th>
<th>Chromosome</th>
<th>Gene Symbol</th>
<th>Gene Name</th>
<th>Protein Description</th>
<th>Protein Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase I-2</td>
<td>ATPase I-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase J-2</td>
<td>ATPase J-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase K-2</td>
<td>ATPase K-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase L-2</td>
<td>ATPase L-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
</tbody>
</table>

ATPases are involved in various cellular processes, including energy generation and waste disposal.

<table>
<thead>
<tr>
<th>ID</th>
<th>Chromosome</th>
<th>Gene Symbol</th>
<th>Gene Name</th>
<th>Protein Description</th>
<th>Protein Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase M-2</td>
<td>ATPase M-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase N-2</td>
<td>ATPase N-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase O-2</td>
<td>ATPase O-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase P-2</td>
<td>ATPase P-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
</tbody>
</table>

ATPases are involved in various cellular processes, including energy generation and waste disposal.

<table>
<thead>
<tr>
<th>ID</th>
<th>Chromosome</th>
<th>Gene Symbol</th>
<th>Gene Name</th>
<th>Protein Description</th>
<th>Protein Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase Q-2</td>
<td>ATPase Q-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase R-2</td>
<td>ATPase R-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase S-2</td>
<td>ATPase S-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase T-2</td>
<td>ATPase T-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
</tbody>
</table>

ATPases are involved in various cellular processes, including energy generation and waste disposal.

<table>
<thead>
<tr>
<th>ID</th>
<th>Chromosome</th>
<th>Gene Symbol</th>
<th>Gene Name</th>
<th>Protein Description</th>
<th>Protein Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase U-2</td>
<td>ATPase U-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase V-2</td>
<td>ATPase V-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase W-2</td>
<td>ATPase W-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>protein_coding</td>
<td>ATPase X-2</td>
<td>ATPase X-2 type family protein / hydrophobic-hydrophilic family protein</td>
<td></td>
</tr>
</tbody>
</table>

ATPases are involved in various cellular processes, including energy generation and waste disposal.
### S25 Table. Annotation details of the Flowering Time (FT) and Leaf Area (LA) DMRs.

Details about genes (promoter or body) and transposable elements that overlap with the Flowering Time (FT) and Leaf Area (LA) DMRs. Abbreviations Unit Class: GP: Gene Promoter; GB: Gene Body; TE: Transposable element; IGR: Intergenic Region.
<table>
<thead>
<tr>
<th>DMR</th>
<th>Chr</th>
<th>Start</th>
<th>Stop</th>
<th>Gene</th>
<th>Type</th>
<th>Name</th>
<th>Description</th>
<th>Transposon</th>
<th>Super family</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTRCH048G1DinMHR0001</td>
<td>4</td>
<td>11363314</td>
<td>11367122</td>
<td>CHRNA5</td>
<td>mRNA</td>
<td>MIR87A</td>
<td>microRNA</td>
<td>MIR87A</td>
<td>microRNA</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0002</td>
<td>4</td>
<td>11368134</td>
<td>11369772</td>
<td>TEP4</td>
<td>protein_coding</td>
<td>B120</td>
<td>B120</td>
<td>B120</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0003</td>
<td>4</td>
<td>11319829</td>
<td>11350953</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0004</td>
<td>4</td>
<td>11350643</td>
<td>11401010</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0005</td>
<td>4</td>
<td>11302780</td>
<td>11304790</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0006</td>
<td>4</td>
<td>11373464</td>
<td>11374462</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0007</td>
<td>4</td>
<td>11399329</td>
<td>11396453</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0008</td>
<td>4</td>
<td>11400433</td>
<td>11401010</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0009</td>
<td>4</td>
<td>11400970</td>
<td>11401010</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0010</td>
<td>4</td>
<td>11495150</td>
<td>11495690</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0011</td>
<td>4</td>
<td>11509018</td>
<td>11509018</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0012</td>
<td>4</td>
<td>11509018</td>
<td>11509018</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0013</td>
<td>4</td>
<td>11515467</td>
<td>11515467</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0014</td>
<td>4</td>
<td>11509018</td>
<td>11509018</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0015</td>
<td>4</td>
<td>11533896</td>
<td>11535265</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0016</td>
<td>4</td>
<td>11533895</td>
<td>11535260</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0017</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0018</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0019</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0020</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0021</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0022</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0023</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0024</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0025</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0026</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0027</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0028</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0029</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0030</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0031</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
<tr>
<td>HTRCH048G1DinMHR0032</td>
<td>4</td>
<td>11685714</td>
<td>11686585</td>
<td>GAB1</td>
<td>protein_coding</td>
<td>CRK2</td>
<td>known protein</td>
<td>CRK2</td>
<td>+</td>
</tr>
</tbody>
</table>
S26 Table. Annotation details of the Height (HT) DMRs. Details about genes (promoter or body) and transposable elements that overlap with the Height (HT) DMRs. Abbreviations Unit Class: GP: Gene Promoter; GB: Gene Body; TE: Transposable element; IGR: Intergenic Region.
Supplementary Information

Epigenetic divergence is sufficient to trigger heterosis in A. thaliana

Kathrin Lauss¹, René Wardenaar², Marieke H.A. van Hulten³, Victor Guryev⁴,
Joost J.B. Keurentjes³, Maike Stam¹§, Frank Johannes²⁵⁶§

¹ University of Amsterdam, Swammerdam Institute for Life Sciences, Science Park 904 1098XH Amsterdam, The Netherlands.
² University of Groningen, Population Epigenetics and Epigenomics, Groningen Bioinformatics Centre, Faculty of Mathematics and Natural Sciences, Nijenborgh 7, 9747 AG Groningen, The Netherlands.
³ University of Wageningen, Laboratory of Genetics, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands.
⁴ Genome structure aging, European Research Institute for the Biology of Ageing, University Medical Centre Groningen and University of Groningen, Antonius Deusinglaan 1, Building 3226, 9713 AV Groningen, The Netherlands
⁵ Current address: Population epigenetics and epigenomics, Department of Plant Sciences, Technical University Munich, Liesel-Beckmann-Str. 2, 85354 Freising, Germany
⁶ Current address: Institute for Advanced Study, Technical University Munich, Lichtenbergstr. 2a, 85748 Garching, Germany

§ Corresponding co-last authors: m.e.stam@uva.nl, frank@johanneslab.org
# Contents

**Plant material and hybrid crosses**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Plant material</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Crosses</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Phenotypic screen</td>
<td>4</td>
</tr>
<tr>
<td>1.3.1 Leaf area (LA)</td>
<td>4</td>
</tr>
<tr>
<td>1.3.2 Flowering time (FT)</td>
<td>4</td>
</tr>
<tr>
<td>1.3.3 Height (HT)</td>
<td>4</td>
</tr>
<tr>
<td>1.3.4 Branching (RB and MSB)</td>
<td>4</td>
</tr>
<tr>
<td>1.3.5 Total Seed Yield (SY)</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Replication experiment in selected hybrids</td>
<td>5</td>
</tr>
<tr>
<td>1.4.1 Plant Material</td>
<td>5</td>
</tr>
<tr>
<td>1.4.2 Crosses</td>
<td>5</td>
</tr>
<tr>
<td>1.4.3 Phenotypic Screen</td>
<td>5</td>
</tr>
</tbody>
</table>

**Data analysis**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Growth curve modeling for leaf area</td>
<td>6</td>
</tr>
<tr>
<td>2.2 Analysis of heterosis</td>
<td>6</td>
</tr>
<tr>
<td>2.2.1 General likelihood approach</td>
<td>7</td>
</tr>
<tr>
<td>2.2.2 Hypothesis testing</td>
<td>7</td>
</tr>
<tr>
<td>2.3 Variance component analysis of population-wide mid-parent heterosis in the F1 hybrids</td>
<td>9</td>
</tr>
<tr>
<td>2.4 Mapping QTL for mid-parent heterosis</td>
<td>10</td>
</tr>
<tr>
<td>2.4.1 Defining the phenotype</td>
<td>10</td>
</tr>
<tr>
<td>2.4.2 Predicting F1 epigenotypes from the methylomes of the parental lines</td>
<td>10</td>
</tr>
<tr>
<td>2.4.3 Interval mapping</td>
<td>11</td>
</tr>
<tr>
<td>2.4.4 Explained variance in mid-parent heterosis</td>
<td>11</td>
</tr>
<tr>
<td>2.5 Interpretation of QTL effects</td>
<td>12</td>
</tr>
<tr>
<td>2.6 Detection of candidate DMRs in the QTL intervals</td>
<td>13</td>
</tr>
<tr>
<td>2.6.1 Selection criterion 1</td>
<td>14</td>
</tr>
<tr>
<td>2.6.2 Selection criterion 2</td>
<td>15</td>
</tr>
<tr>
<td>2.6.3 Selection criterion 3</td>
<td>15</td>
</tr>
<tr>
<td>2.6.4 Final definition of candidate DMRs</td>
<td>15</td>
</tr>
<tr>
<td>2.7 Detection of Structural Variants in QTL intervals</td>
<td>15</td>
</tr>
<tr>
<td>2.7.1 Sequence alignment and calling of structural variants</td>
<td>16</td>
</tr>
</tbody>
</table>
106

Plant material and hybrid crosses

1.1 Plant material

The epigenetic recombinant inbred lines (epiRILS) in our study were generated by Johannes et al[1]. The epiRILs were constructed as follows: An *Arabidopsis thaliana* Col-0 line deficient for *ddm1-2* (*DECREASE IN DNA METHYLATION 1*) was crossed to an isogenic Col-0 wildtype line (Col-wt) and the resulting F1 was backcrossed as female parent to Col-wt. Subsequently about 500 progeny plants with a wildtype *DDM1* allele were selected and propagated through six more rounds of selfing, generating a population of 500 different epiRILs. We selected 19 different epiRILs as paternal plants for generating epi-Hybrids (Line IDs: 14, 232, 92, 208, 438, 195, 350, 500, 150, 118, 432, 202, 344, 64, 193, 508, 260, 579, 371). Our selection criteria were as follows: 1) Wide range of DNA methylation divergence from Col-wt and among the selected lines; 2) Wildtype DNA methylation state at the *FWA* locus in order to avoid that differences in DNA methylation at this locus give rise to differences in flowering time [2] in the hybrids; 3) Wide range of phenotypic variation in flowering time and root length among the selected lines. The epiRIL lines were purchased from the *Arabidopsis* Stock center of INRA Versailles (http://publiclines.versailles.inra.fr/).

1.2 Crosses

To generate F1 hybrids from the selected epiRIL lines and Col-wt, all parental plants were grown in parallel in soil (Jongkind 7 from Jongkind BV, http://www.jongkind.com/) in pots (Danish size 40 cell, Desch Plantpak, http://www.desch-plantpak.com/en/Home.aspx). The plants were grown at 20°C, 60% humidity, in long day conditions (16h light, 8h dark), and were watered 3 times per week. All crosses were performed in parallel in a time frame of two weeks to avoid phenotypic effects in the F1 progeny due to differences in growing conditions. To exclude that differences in maternal cytoplasm affected the phenotypes of the F1 plants, Col-wt plants were used as a maternal par-
Chapter 3

ent and the epiRILs as paternal parents. In parallel, all parental lines, Col-wt and epiRILS, were propagated by manual selfing. This was to 1) ensure that parental and F1 hybrid seeds were generated under the same growing conditions and 2) exclude potential phenotypic effects derived from hand pollination [3].

1.3 Phenotypic screen

The seeds were stratified at 4°C for 3 days on petri-dishes containing filter paper and water before transferring them onto Rockwool/Grodan blocks (soaked in Hyponex NPK: 6.5 – 6.19 medium) in a climate controlled chamber (20°C, 70% humidity, long day conditions (16h light, 8h dark)). The transfer of the seeds onto the Rockwool blocks is defined as time point 0 days after sowing (DAS). Seeds from all parental and hybrid lines were sown in 28 replicates and their positions were randomized throughout the growth chamber to level out phenotypic effects caused by plant position. The plants were watered two or three times per week depending on their size. After the plants started flowering, they were transferred to the greenhouse (20°C, 60% humidity, long day conditions (16h light, 8h dark)). In the greenhouse, the plants were watered 3 times per week and stabilized by binding them to wooden sticks at later developmental stages. The plants were harvested once the siliques of the main inflorescence and its side branches were ripe.

1.3.1 Leaf area (LA)

LA was monitored by an automated camera system (Open Pheno System, WUR) from 4 days after sowing (DAS). The system consists of 14 fixed cameras that can take pictures of up to 2145 plants daily, every two hours. We monitored LA until 14 DAS since at later time points leaves start overlapping, hampering the correct detection of LA. Leaf area in mm² was calculated by an ImageJ based measurement setup (http://edepot.wur.nl/169770).

1.3.2 Flowering time (FT)

FT was defined as the DAS at which the first flower opened. FT was scored manually each day before 12am.

1.3.3 Height (HT)

HT was scored manually in cm on dried plants. The measurement was taken at the main inflorescence, from the rosette to the highest flowerhead.

1.3.4 Branching (RB and MSB)

Branching was scored on the dried plants by counting the branches emerging from the rosette (RB) and from the main stem (MSB).
1.3.5 Total Seed Yield (SY)

Seeds were harvested from the dried plants, cleaned by filtering and seed yield was subsequently determined by weighing (resulting in mg seeds per plant).

1.4 Replication experiment in selected hybrids

1.4.1 Plant Material

Freshly ordered seeds of epiRILs (Line IDs: 92, 150, 193, 232) from the Arabidopsis Stock center of INRA Versailles.

1.4.2 Crosses

Performed as described above.

1.4.3 Phenotypic Screen

Performed exactly as described above with the exception that more replicates for each parental and hybrid line were monitored: 60 replicates for LA and 30 replicates for the traits FT and HT.
Data analysis

2.1 Growth curve modeling for leaf area

We considered LA measurements until 14 DAS. While measurements were taken every two hours, we only used the measurements at mid-day as the leaves were most flattened at that time. For each individual plant we modelled LA as a function of time (in DAS) using a generalized logistic growth model, which we parameterized as follows

\[ g(t; k, b, m) = \frac{k}{1 + e^{b(m-t)}}, \]

where \( k, b \) and \( m \) are the unknown model parameters and \( t = 0, 1, 2, \ldots, 14 \). To obtain parameter estimates, we defined the following function

\[ s(t; k, b, m) = \sum_{t=0}^{14} (g(t; k, b, m) - o_t)^2, \]

where \( o_t \) are the observed leaf area measurements in \( \text{mm}^2 \). Minimizing \( s(t; k, b, m) \) with respect to the unknown parameters \( k, b \) and \( m \) is a standard problem in non-linear least squares regression. The use of the growth curve model had two purposes: 1) It provided a growth rate parameter \( b \) that we used as a phenotype for further analysis (see phenotype GR); and 2) The fitted values \( \hat{s}(0), \hat{s}(1), \hat{s}(2), \ldots, \hat{s}(14) \) could be used in place of the observations \( o(0), o(1), o(2), \ldots, o(14) \), providing cleaner measurements, particular toward later time points where measurements were less accurate due to overlapping leaves. For all subsequent analyses we focused on LA measured at 14 DAS (i.e. \( \hat{s}(14) \), see phenotype LA).

2.2 Analysis of heterosis

Below we describe how we tested for positive and negative Mid-parent heterosis as well as for Low- and High-parent heterosis.
2.2.1 General likelihood approach

Let $Y_i$ be the trait value for the $i$th individual ($i = 1, 2, 3, \ldots, N$). Individual $i$ can belong to either one of the two parental populations ($P_1, P_2$) or the hybrid offspring population $F_1$. We arbitrarily assign $P_1 = 1, P_2 = 2$ and $F_1 = 3$. To keep track of population membership, let $\tilde{Z}_i$ be a 3-dimensional component-label vector, where the $j$ element is defined to be one or zero, according to whether the component of origin of $Y_i$ is equal to $j$ or not ($j = 1, 2, 3$). The $\tilde{Z}_i$ is distributed as a multinomial distribution consisting of one draw of 3 categories with probabilities $\lambda_1, \lambda_2, \lambda_3$. Hence,

$$\Pr(\tilde{Z}_i = \tilde{z}_i) = \lambda_1^{z_{i1}} \lambda_2^{z_{i2}} \lambda_3^{z_{i3}}.$$ 

In our case

$$\lambda_1 = \frac{N_{P_1}}{N}, \lambda_2 = \frac{N_{P_2}}{N}, \lambda_3 = \frac{N_{F_1}}{N},$$

where $\sum_j \lambda_j = 1$. Suppose the conditional density of $Y_i$ given $Z_i = j$ is $f_j(y_i; \Omega_j)$, then it can be shown that the log likelihood for individuals $i = 1, 2, 3, \ldots, N$ is

$$\log L(\tilde{Y}|\tilde{y}) = \sum_{i=1}^{N} \sum_{j=1}^{3} z_{ij}(\log \lambda_j + \log f_j(y_i; \Omega_j)).$$

This log likelihood function can be partitioned more intuitively according to the contributions of each of the three populations ($P_1, P_2, F_1$):

$$\log L(\tilde{Y}|\tilde{y}) = N_{P_1} \log \left( \frac{N_{P_1}}{N} \right) + \sum_{i=1}^{N_{P_1}} \log f_{P_1}(y_i; \tilde{\Omega}_{P_1})$$

$$+ N_{P_2} \log \left( \frac{N_{P_2}}{N} \right) + \sum_{i=1}^{N_{P_2}} \log f_{P_2}(y_i; \tilde{\Omega}_{P_2})$$

$$+ N_{F_1} \log \left( \frac{N_{F_1}}{N} \right) + \sum_{i=1}^{N_{F_1}} \log f_{F_1}(y_i; \tilde{\Omega}_{F_1})$$

$$\propto \sum_{i=1}^{N_{P_1}} \log f_{P_1}(y_i; \tilde{\Omega}_{P_1}) + \sum_{i=1}^{N_{P_2}} \log f_{P_2}(y_i; \tilde{\Omega}_{P_2}) + \sum_{i=1}^{N_{F_1}} \log f_{F_1}(y_i; \tilde{\Omega}_{F_1}).$$

2.2.2 Hypothesis testing

Testing for Mid-parent heterosis in the F1 crosses

We tested each trait for midparent heterosis (MPH) by comparing the full (unconstrained) model against an additive (constrained) model. The model parameters of the log likelihood functions are shown in the below table.

110
We used the likelihood ratio statistic \( D \) to test whether the full model provided a better fit to the data than the additive model. \( D \) is defined by

\[
D = -2l_F + 2l_A,
\]

and distributed as a \( \chi^2 \) random variable with degrees of freedom equal to the differences in the number of parameters of the full compared to the additive model. In total we performed 19·7 = 133 tests. We controlled the false discovery rate (FDR) at 0.05 using the method of Benjamini and Hochberg [5].

**Testing for High-parent heterosis in the F1 crosses**

Positive MPH is a necessary condition for high-parent heterosis (HPH). Conditional on having detected MPH we further tested for HPH. If the ordering of the trait means was \( \mu_{F1} > \mu_{Ph} \) (where \( Ph \) is the high parent), we compared the full model against a model that assumes full positive dominance (FPD). We considered the following models

<table>
<thead>
<tr>
<th>Model</th>
<th>log lik</th>
<th>df</th>
<th>( \Omega_{P1k} )</th>
<th>( \Omega_{P2k} )</th>
<th>( \Omega_{F1k} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>( l_F )</td>
<td>6</td>
<td>( \mu_{P1k}, \sigma^2_{P1k} )</td>
<td>( \mu_{P2k}, \sigma^2_{P2k} )</td>
<td>( \mu_{F1k}, \sigma^2_{F1k} )</td>
</tr>
<tr>
<td>Additive</td>
<td>( l_A )</td>
<td>5</td>
<td>( \mu_{P1k}, \sigma^2_{P1k} )</td>
<td>( \mu_{P2k}, \sigma^2_{P2k} )</td>
<td>( \mu_{F1k} = \frac{\mu_{P1k} + \mu_{P2k}}{2}, \sigma^2_{F1k} )</td>
</tr>
</tbody>
</table>

In this case, the likelihood ratio test is

\[
D = -2l_F + 2l_{FPD}.
\]

**Testing for Low-parent heterosis in the F1 crosses**

Analogous to HPH, negative MPH is a necessary condition for low-parent heterosis (LPH), which in our terminology denotes that the F1 means are significantly lower than the phenotypic mean of the low-performing parent. Hence, if the ordering of the trait means was \( \mu_{F1} < \mu_{Pl} \) (where \( Pl \) is the low parent), we compared the full model against a model that assumes full negative dominance (FND). We considered the following models

<table>
<thead>
<tr>
<th>Model</th>
<th>log lik</th>
<th>df</th>
<th>( \Omega_{P1k} )</th>
<th>( \Omega_{P2k} )</th>
<th>( \Omega_{F1k} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>( l_F )</td>
<td>6</td>
<td>( \mu_{P1k}, \sigma^2_{P1k} )</td>
<td>( \mu_{P2k}, \sigma^2_{P2k} )</td>
<td>( \mu_{F1k}, \sigma^2_{F1k} )</td>
</tr>
<tr>
<td>FND</td>
<td>( l_{FND} )</td>
<td>5</td>
<td>( \mu_{P1k}, \sigma^2_{P1k} )</td>
<td>( \mu_{P2k}, \sigma^2_{P2k} )</td>
<td>( \mu_{F1k} = \mu_{Plk}, \sigma^2_{F1k} )</td>
</tr>
</tbody>
</table>

In this case, the likelihood ratio test is

\[
D = -2l_F + 2l_{FND}.
\]
2.3 Variance component analysis of population-wide mid-parent heterosis in the F1 hybrids

An important question is to which extent inter-individual variation in mid-parent heterosis can be attributed to between-line (i.e. between-cross) differences. In the context of our experimental design, such an estimate quantifies the amount of variation that can be attributed to (epi)genetic differences between the paternal epiRILs that were used for the crosses. To test this, we calculated the mid-parent value for trait $y$ in the $k$th cross as

$$mid_k = \frac{\overline{y}_{P1,k} + \overline{y}_{P2,k}}{2}$$

where $\overline{y}_{P1,k}$ and $\overline{y}_{P2,k}$ are the sample phenotypic means for parents $P1$ and $P2$ in cross $k$, respectively. We defined a measure of mid-parent heterosis for plant $i$ in cross $k$ ($z_{ik}$) by

$$z_{ik} = y_{ik} - mid_k,$$

We treated the $z_{iks}$ ($k = 1, 2, 3, \cdots, 19; i = 1, 2, \cdots, 30$) as a quantitative phenotype. Assume the value for the $i$th plant is given by

$$z_i = \mu_0 + \beta_1 C_{i1} + \beta_2 C_{i2} + \cdots + \beta_p C_{ip} + \epsilon_i,$$

where $\mu_0$ is the overall phenotypic mean, $C_{ik}$ is a dummy variable with coding $C_{ik} = 1$ if individual $i$ belongs to epiHybrid cross $j$, and $C_{ik} = 0$ otherwise. The regression parameter $\beta_j = \mu_j - \mu_0$, and thus quantifies the offset of the phenotypic mean of population $j$ with respect to the overall mean. The total phenotypic variance can be partitioned as

$$\sigma^2(y) = \sigma^2(C) + \sigma^2(\epsilon),$$

with $\sigma^2(C)$ and $\sigma^2(\epsilon)$ denoting the between-cross and the pooled within-cross variance components, respectively. In this linear regression framework $R^2$ quantifies the amount of variance explained by the between-crosses component and is given by:

$$R^2 = 1 - \frac{\sigma^2(\epsilon)}{\sigma^2(y)}.$$

Formally this is equivalent to the broad-sense heritability $H^2$:

$$R^2 = H^2 = \frac{\sigma^2(C)}{\sigma^2(y)},$$

but this terminology may be misleading in the context of studying F1 hybrids, as an assessment of the “inheritance” of the heterotic effects is lacking. Replacing the above variance components by their finite sample estimators, we obtain the
adjusted $R^2$ values:

$$R^2_{adj} = 1 - \frac{\frac{1}{n-(p+1)} s^2(\epsilon)}{\frac{1}{n-1} s^2(y)} = 1 - \frac{n-1}{n-(p+1)} \frac{\sum^n_i (y_i - \hat{y}_i)^2}{\sum^n_i (y_i - \bar{y})^2} = 1 - \frac{n-1}{n-(p+1)} \frac{\sum^n_i (y_i - \mu_0 + \sum_{k=1}^p \hat{\beta}_j C_{ik})^2}{\sum^n_i (y_i - \bar{y})^2}$$

(5)

2.4 Mapping QTL for mid-parent heterosis

2.4.1 Defining the phenotype

As shown in TableS, we detected highly significant $R^2$ for most traits. Next, we sought to search for quantitative trait loci (QTL) underlying mid-parent heterosis. For this QTL-based approach we defined the phenotype as

$$\text{mid}_k = \frac{\bar{y}_{P1,k} + \bar{y}_{P2,k}}{2}$$

where $\bar{y}_{P1,k}$ and $\bar{y}_{P2,k}$ are the sample phenotypic means for parents $P1$ and $P2$ in cross $k$, respectively. We defined a measure of mid-parent heterosis for line $k$ ($z_k$) by

$$z_k = \bar{y}_k - \text{mid}_k,$$

where $\bar{y}_k$ is the phenotypic mean of the $k$th epiHybrid population. As shown in Fig. 2A, the $z_k$s ($k = 1, 2, 3, ..., 19$) are distributed quantitatively among the 19 epiHybrids lines.

2.4.2 Predicting F1 epigenotypes from the methylomes of the parental lines

We recently reported a recombination map of the epiRILs that was obtained using 126 differentially methylated regions (DMRs) as physical markers [4]. These markers cover $\sim 81.9\%$ of the Arabidopsis genome (74.7, 77.0, 98.4, 91.1, and 73.0 %, of chromosomes 1, 2, 3, 4 and 5, respectively), with an average inter-marker spacing of $\sim 0.804$ Mb (3.45 cM). The map was based on the DNA methylomes of 123 epiRILs, 19 of which are siblings of the epiRILs used as parents for the epiHybrids. Previous analyses showed that the 126 DMRs are stable for at least 10 sexual generations, and that the epiRILs are epi-homozygous, either for two methylated Col-wt epialleles (which we denote by $MM$) or epi-homozygote for two $ddm1-2$-derived hypo-methylated epialleles (which we denote by $UU$) [4].
We used the epigenotypes of the 19 parental epiRILs to predict the epigenotypes of the epiHybrids at these marker locations. That is, epiHybrids could either be $MM$ or $MU$, depending on whether their epiRIL parents were $MM$ or $UU$ at a given locus, respectively. Based on this information, the different epiHybrids can be viewed as a single mapping population with recombination events having been contributed by the chromosomes of the parental epiRILs; the Col-wt chromosome copy being invariable among the epiHybrids.

### 2.4.3 Interval mapping

To search for heterosis QTLs at the genome-wide scale, we performed classical interval mapping [6] as implemented in the `scanone` function in R/qtl [7]. The mapping was performed with a step size of 2 cM and estimates were obtained by Haley-Knott regression. Genome-wide significance was determined empirically for each trait using 1000 permutations of the data. The LOD significance thresholds were chosen to correspond to a genome-wide false positive rate of 5%.

### 2.4.4 Explained variance in mid-parent heterosis

For each detected QTLs we considered the nearest linked DMR (i.e. peak marker) in a regression model. For clarity, we detail this procedure below.

#### Additive (epi)genetic model

We consider an additive (epi)genetic model consisting of $q$ QTLs. For phenotype FT and LA, this is a two-locus model ($q = 2$), and for HT and SY this was a single locus model ($q = 1$). In general, we assume that for epiHybrid line $k$ the mid-parent heterosis value $z_k$ be give by:

$$ z_k = \beta_0 + \beta_1 g_{k1} + \beta_2 g_{k2} + \cdots + \beta_q g_{kq} + \epsilon_k, $$

where $\beta_j (j = 1, \ldots, q)$ are the QTL effects, $\beta_0$ is the intercept, $g_{kj} (j = 1, \ldots, q; k = 1, \ldots, h)$ are the $q$ epigenotypes measured for epiHybrid line $k$, and $\epsilon_k$ is a normally distributed error with a mean of zero. The (epi)genotypes are coded as $g = 0$ and $g = 1$ for $MU$ and and $MM$ cases, respectively. The phenotypic variance, $\sigma^2(z)$, can be partitioned as:

$$ \sigma^2(z) = \sum_{j} \beta_j^2 \sigma^2(g_j) + 2 \sum_{m<j} \beta_j \beta_m \sigma(g_m, g_j) + \sigma^2(\epsilon) $$

$$ \sigma^2(z) = \sigma^2(G) + \sigma^2(\epsilon), $$

where $\sigma^2(G)$ is the total contribution of the (epi)genetic variance component.
Variance explained by QTLs

We calculate $R^2(G)$ to quantify the proportion of phenotypic variance explained by the (epi)genetic component:

$$R^2(G) = \frac{\sigma^2(G)}{\sigma^2(y)} = \frac{\sigma^2(y) - \sigma^2(\epsilon)}{\sigma^2(y)} = 1 - \frac{\sigma^2(\epsilon)}{\sigma^2(y)}.$$ 

Replacing these two variance terms with their unbiased sample estimators, we have:

$$R^2(G)_{adj} = 1 - \frac{n - 1}{n - (k + 1)} \frac{\sum_i^n(y_i - [\hat{\beta}_0 + \sum_j^k \hat{\beta}_j g_{ij}])^2}{\sum_i^n(y_i - \bar{y})^2},$$

where $\hat{\beta}_0, \hat{\beta}_j (j = 1, \ldots, q)$ are the OLS multiple regression estimates.

### 2.5 Interpretation of QTL effects

In the construction of the epiHybrid populations we employed an asymmetrical cross design, insofar that all epiRIL parental lines were crossed to a recurrent Col-wt parent. Moreover, for QTL mapping we defined the phenotype as the divergence from the mid-parent value and subsequently treated the different F1 crosses as a single mapping population. This raises the question to whether the detected QTL effects are due to dominance action of the underlying loci, or due to effects such as additivity or epistasis.

To explore this issue analytically, suppose there are $Q$ independent loci determining mid-parent heterosis value $z$. Let $N = Q - 1$ be the number of loci excluding locus $l$, which we will consider as the focal QTL whose phenotypic effects we wish to evaluate. We assume that a proportion $1 - p$ of the $N$ background loci are $UU$ in a randomly chosen epiRIL parent, and $p$ are $MM$. The expected midparent value, $mid$, conditional on the fact that a randomly chosen epiRILs parent is $MM$ at locus $l$ is

$$E(mid | l = MM) = \frac{E(y_{wt}) + E(y_{epi} | l = MM)}{2}.$$  \hspace{1cm} (8)

and conditional on locus $l$ being $UU$ it is

$$E(mid | l = UU) = \frac{E(y_{wt}) + E(y_{epi} | l = UU)}{2}.$$  \hspace{1cm} (9)

The expected mid-parent heterosis value $z$ for randomly chosen epiHybrid conditional on the fact that locus $l$ was $MM$ in the parental epiRIL is

$$E(z | l = MM) = E(y_{F1} | l = MM) + E(mid | l = MM),$$
and similarly the expected mid-parent heterosis value $z$ for randomly chosen epiHybrid conditional on the fact that locus $l$ was $UU$ is

$$E(z | l = UU) = E(y_{F1} | l = UU) + E(mid | l = UU).$$

The QTL effect in the epiHybrids is given by the contrast

$$QTL_{F1,l} = E(z | l = MM) − E(z | l = UU),$$

where the conditionality refers to the epigenotypes of the epiRIL parents lines, rather than the epigenotypes of the F1 hybrids. Considering the definitions given in Figure 2.1 (below), and assuming equal effect sizes across all of the $N$ background loci, it can be shown that the QTL contrast is

$$QTL_{F1,l} = 2\beta_{l:D}N(\beta_{l:A \times A}(p − 1) − 2(p(\beta_{l:A \times D} + \beta_{l:D \times D}))).$$

Because the parameter $p$ is difficult to determine experimentally and the effect sizes arising from background epistasis are difficult to distinguish from the number of epistatic interactions, we integrate out $p$ and replace $N\beta$ with $\beta^\bullet$, which yields

$$QTL_{F1,l}^\bullet = \int_0^1 QTL_{F1,l}dp = 2\beta_{l:D} − \frac{1}{2}\beta_{l:A \times A} + 2\beta_{l:A \times D} − 2\beta_{l:D \times A}.$$  

This equation means that the QTL contrast contains a dominance effect (via $\beta_{l:D}$), but also additional effects arising from epistatic interactions between locus $l$ and the entire (epi)genomic background (via $\beta_{l:A \times A}$, $\beta_{l:A \times D}$ and $\beta_{l:D \times A}$). Here $A \times A$, $A \times D$ and $A \times D$ refer to additive $\times$ additive, additive $\times$ dominance and dominance $\times$ additive interactions. While the relative contributions of the dominance and epistatic terms can only be determined experimentally, for example, by help of introgression lines, the effect does require that causal variants are present in the QTL intervals. The causal variants can be in the form of Differentially Methylated Regions (DMRs) that are in approximate LD with the peak QTL marker, or else rare structural variants, such as those having arisen from TE mobilization events in the $ddm1-2$ founder parent.

### 2.6 Detection of candidate DMRs in the QTL intervals

To search for candidate DMRs within the QTL confidence intervals we leveraged probe-level methylation data from the MeDIP tiling arrays that were available for the 123 epiRILs and their two founder parents [4]. We previously determined the methylation calls for each probe on these arrays using a Hidden Markov Model (HMM) [8]. As previously described [9], we considered probes as candidates when they met the following criteria:
Figure 2.1: (A) We consider \( N \) background loci (in this example \( N = 7 \)). In Col-wt all loci are assumed to have epigenotype \( MM \), while in the epiRILs a proportion \( 1 - p \) are assumed to be \( UU \) and the remainder \( (p) \) are \( MM \). Hence, in the hybrids \( 1 - p \) of the loci are epi-heterozygous \( MU \) and the remainder \( p \) are epi-homozygous \( MM \). (B) Definition of phenotypic effects for epigenotypes \( MM, MU \) and \( UU \) at a given locus. (C) Definition of phenotypic effects resulting from pairwise epistasis between any two loci.

2.6.1 Selection criterion 1

Probes need to be of high quality: Probes needed to have a conservation score lower than 85. The conservation score of a probe indicates the uniqueness of the probe sequence. The conservation scores were obtained by performing a blast search. Scores are defined as a percentage of identity with the second best hit (score range: 45-100). The best hit is with the genomic location for which the probe was designed. Probes with a high conservation score provide poor
measurements due to cross-hybridization.

2.6.2 Selection criterion 2

Selected probes (from criterion 1) need to be differentially methylated between the Col-wt and ddm1-2 founder parents: Since ddm1-2 results mainly in loss of DNA methylation, we considered all probes for which the methylation level was higher in the Col-wt parent compared to the ddm1-2 founder parent. Hence, we considered the following transitions: M → U, M → I, I → U, where M, I and U refer to fully methylated, intermediately methylated and unmethylated, respectively.

2.6.3 Selection criterion 3

Selected probes (from criterion 2) need to show correlation with the epigenotype of the peak marker: Based on the HMM results we calculated the posterior probability for probe i to be unmethylated or methylated given by \(\text{post}(P_i = U)\) and \(\text{post}(P_i = M)\), respectively. Using this we define the methylation level of probe I as \(ML = \text{post}(P_i = U) \cdot (-1) + \text{post}(P_i = M) \cdot 1\). The correlation between the methylation levels of the probes and the epigenotype of the peak QTL marker was determined using Spearman correlation. An appropriate cutoff for the correlation values was defined using probes that are part of markers (marker probes) inside the QTL interval. Marker probes are in tight LD with the peak marker and should therefore be highly corrected with it. Non-marker probes upstream and downstream from the peak marker were treated separately. The cutoffs for the selection of non-marker probes were based on the 5th percentile of the correlation values of the marker probes upstream of the peak marker and downstream of the peak marker. A non-marker probe was selected if its correlation value was higher than the cutoff. All marker probes were selected.

2.6.4 Final definition of candidate DMRs

Individual probes that met the above criteria were considered as candidate probes. Neighboring candidate probes were subsequently merged into DMRs. Merging was also performed when two candidate probes were separated by one non-candidate probe.

2.7 Detection of Structural Variants in QTL intervals

Previous work has shown that specific TE families are mobilized at relatively low rates in the ddm1-2 background [10, 1, 9]. In the epiRILs these mobilization events occur mostly in a line-specific manner during inbreeding. However,
there are also shared TE insertions originating either from the original \textit{ddm1-2} founder line or from the F1 that was used in the initial epiRIL crossing design. Shared TE insertions are potentially problematic in interpreting the detected QTL in terms of epigenetic effects. We re-analyzed whole-genome mate-pair re-sequencing data of 50 epiRILs [9], which contained many of the epiRILs used in the construction of the epiHybrids.

\subsection*{2.7.1 Sequence alignment and calling of structural variants}

Sequence reads from mate-pair libraries (6kb inserts) were mapped against the TAIR10 reference genome using Bowtie2 version 2.1 [11] using following non-default parameters: \texttt{-rf -X 10000}. Structural variants were called using clustering of discordantly mapping read pairs as implemented in 1-2-3-SV v. 0.9 [12] (http://tools.genomes.nl/123sv.htm) with minimal mapping quality 30 and at least five tag pairs per structural variant. We also explored alternative programs such as Pindel, Delly and TE-tracker. Pindel and Delly runs using the same data were terminated after two weeks of running. It is likely that the large insert size significantly increases computation intensity for these tools.
Bibliography


