Phenotypic variation in plants

Roles for epigenetics

Lauss, K.

Publication date
2017
Document Version
Other version
License
Other

Chapter 3

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

Kathrin Lauss[1], René Wardenaar[2], Marieke H.A. van Hulten[3], Victor Guryev[4], Joost J.B. Keurentjes[3], Maike Stam[1§], Frank Johannes[2,5,6§]

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 bioRixV. (doi: http://dx.doi.org/10.1101/059980)

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.

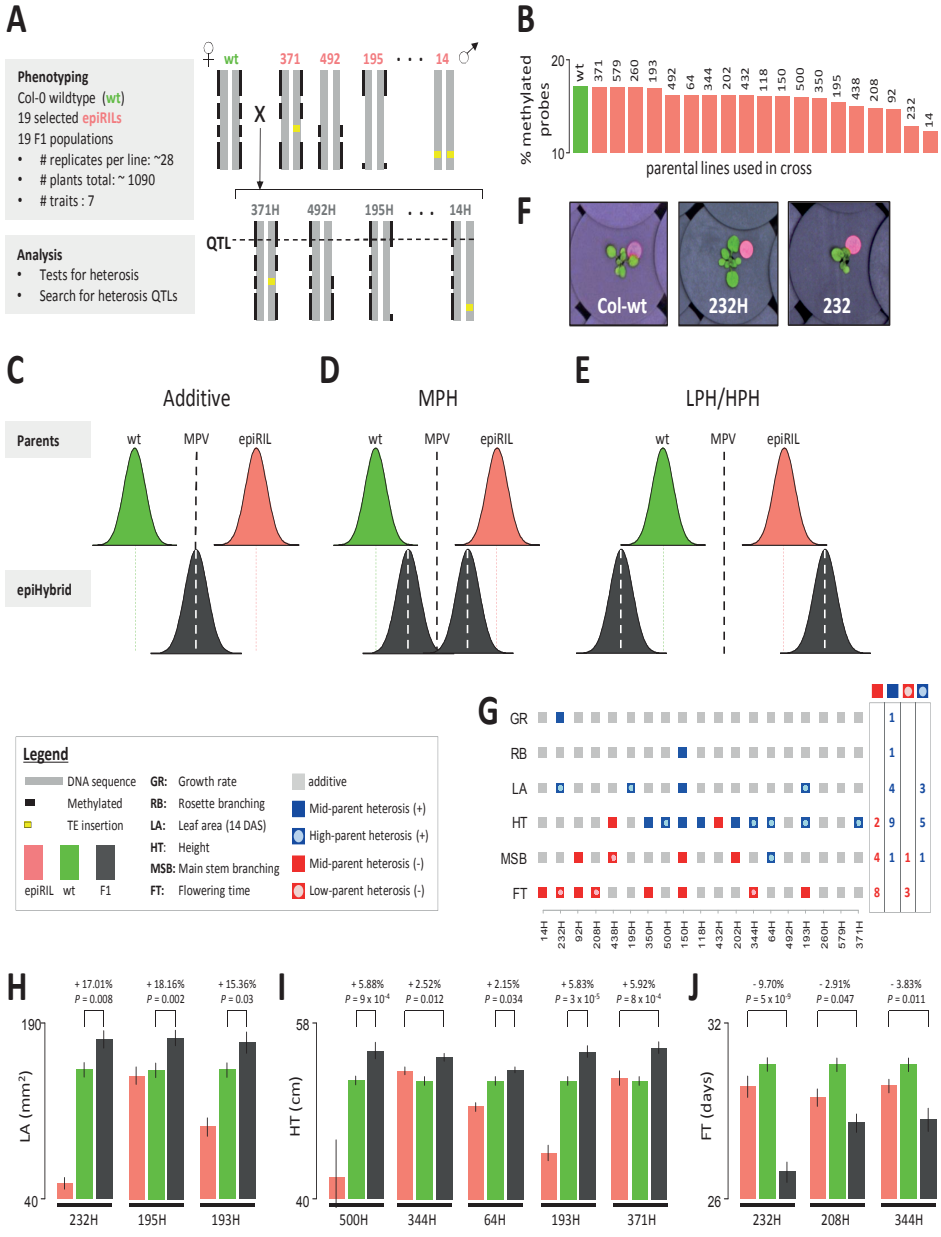

A
Phenotyping
Col-0 wildtype (wt)
19 selected epiRILs
19 F1 populations
• # replicates per line: ~28
• # plants total: ~ 1090
• # traits : 7
Analysis
• Tests for heterosis
• Search for heterosis QTLs
wt 371 492 195 14
371H 492H 195H 14H
QTL
B
% methylated probes
parental lines used in cross
wt 371 579 260 193 492 64 344 202 432 118 150 500 350 195 438 208 92 232 14
F
Col-wt
232H
232
C
Additive
D
MPH
E
LPH/HPH
Parents
epiHybrid
wt
MPV
epiRIL
Legend
DNA sequence
Methylated
TE insertion
epiRIL wt F1
GR: Growth rate
RB: Rosette branching
LA: Leaf area (14 DAS)
HT: Height
MSB: Main stem branching
FT: Flowering time
additive
Mid-parent heterosis (+)
High-parent heterosis (+)
Mid-parent heterosis (-)
Low-parent heterosis (-)
G
GR
RB
LA
HT
MSB
FT
14H 232H 92H 208H 438H 195H 350H 500H 150H 118H 432H 202H 344H 64H 492H 193H 260H 579H 371H
H
LA (mm²)
+ 17.01% P = 0.008
+ 18.16% P = 0.002
+ 15.36% P = 0.03
232H 195H 193H
I
HT (cm)
+ 5.88% P = 9 x 10⁻⁴
+ 2.52% P = 0.012
+ 2.15% P = 0.034
+ 5.83% P = 3 x 10⁻⁵
+ 5.92% P = 8 x 10⁻⁴
500H 344H 64H 193H 371H
J
FT (days)
- 9.70% P = 5 x 10⁻⁹
- 2.91% P = 0.047
- 3.83% P = 0.011
232H 208H 344H

Figure 1: 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

1g). 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.

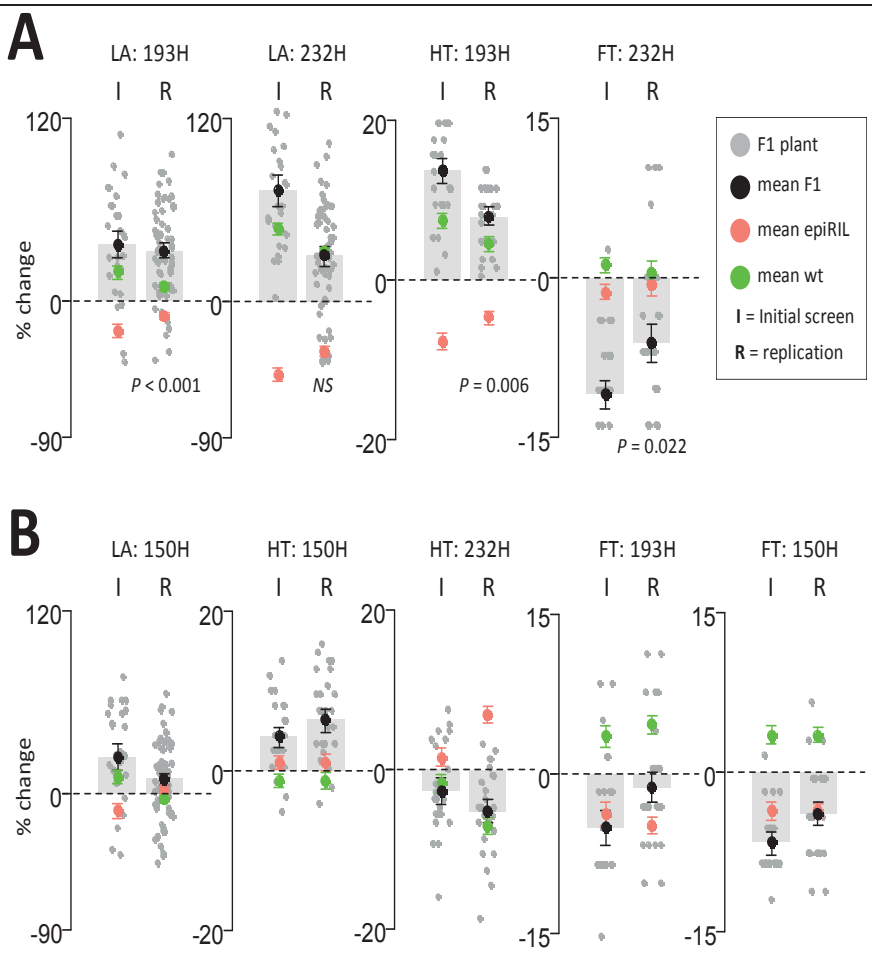
We observed cases of HPH for LA, HT and MSB, and cases of LPH for FT and MSB. HPH for LA occurred in epiHybrids 232H, 195H and 193H (3/19 epiHybrids). Those epiHybrids significantly exceeded their best parent (Col-wt) by 17%, 18% and 15%, respectively (Fig 1h, S19 Table). Interestingly, although growth 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].


A
LA: 193H
LA: 232H
HT: 193H
FT: 232H
I R
% change
120
0
-90
20
-20
15
-15
P < 0.001
NS
P = 0.006
P = 0.022
F1 plant
mean F1
mean epiRIL
mean wt
I = Initial screen
R = replication
B
LA: 150H
HT: 150H
HT: 232H
FT: 193H
FT: 150H

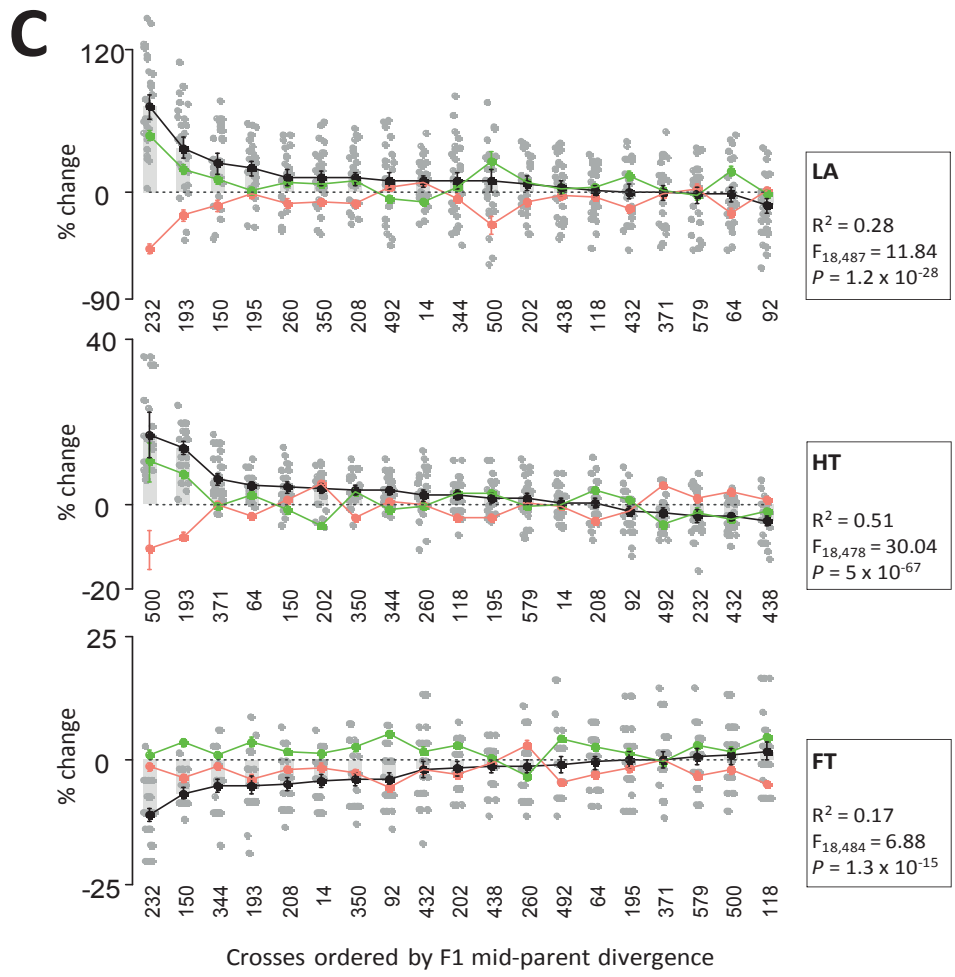


Figure 2: Confirmation of mid-parent (MP) divergence in the initial screen and replicate experiment for epiHybrids 150H, 193H and 232H. A) Results for cases of HPH and LPH for LA, HT and FT in initial experiment. B) Results for traits showing less eminent phenotypic effects for LA, HT and FT. The mid-parent 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

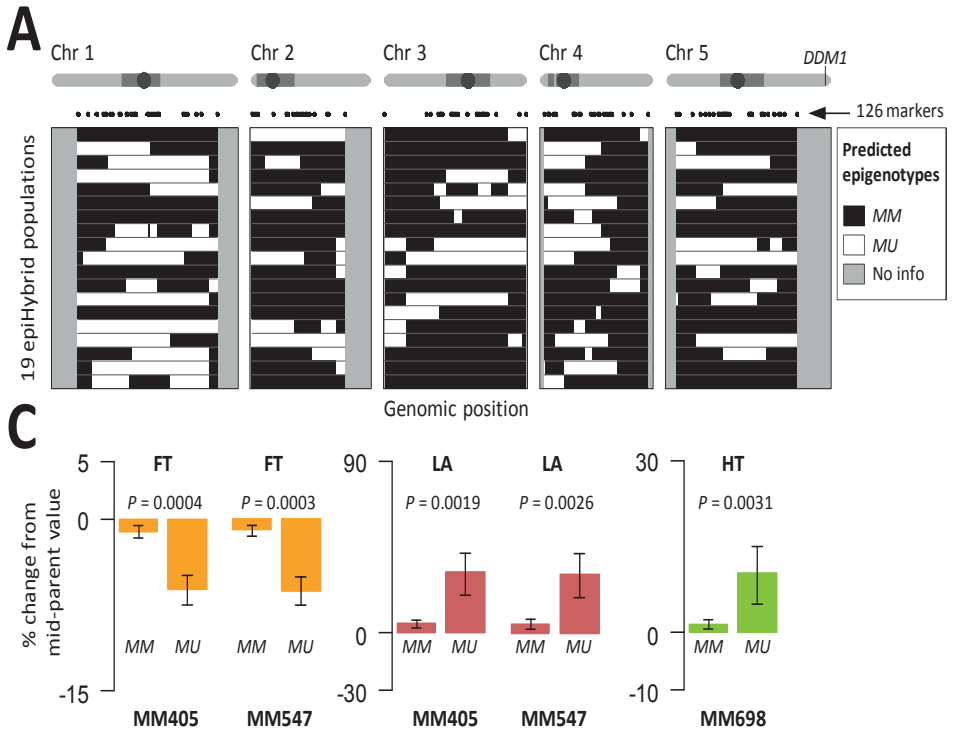
epiHybrids. The MPV is shown as a horizontal dashed line and MP divergence is shown as change from MPV in percent. The epiHybrids are ordered from highest (left) to lowest (right) F1 MP divergence. To illustrate the F1 epiHybrid distribution for each trait, the individual replicate plants are depicted 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^{epi} for root length in the epiRILs[56]. The same study identified $QTLs^{epi}$ 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.


A
Chr 1
Chr 2
Chr 3
Chr 4
Chr 5
DDM1
126 markers
19 epiHybrid populations
Predicted epigenotypes
MM
MU
No info
Genomic position
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
MM
MU
MM405
MM547
MM405
MM547
MM698

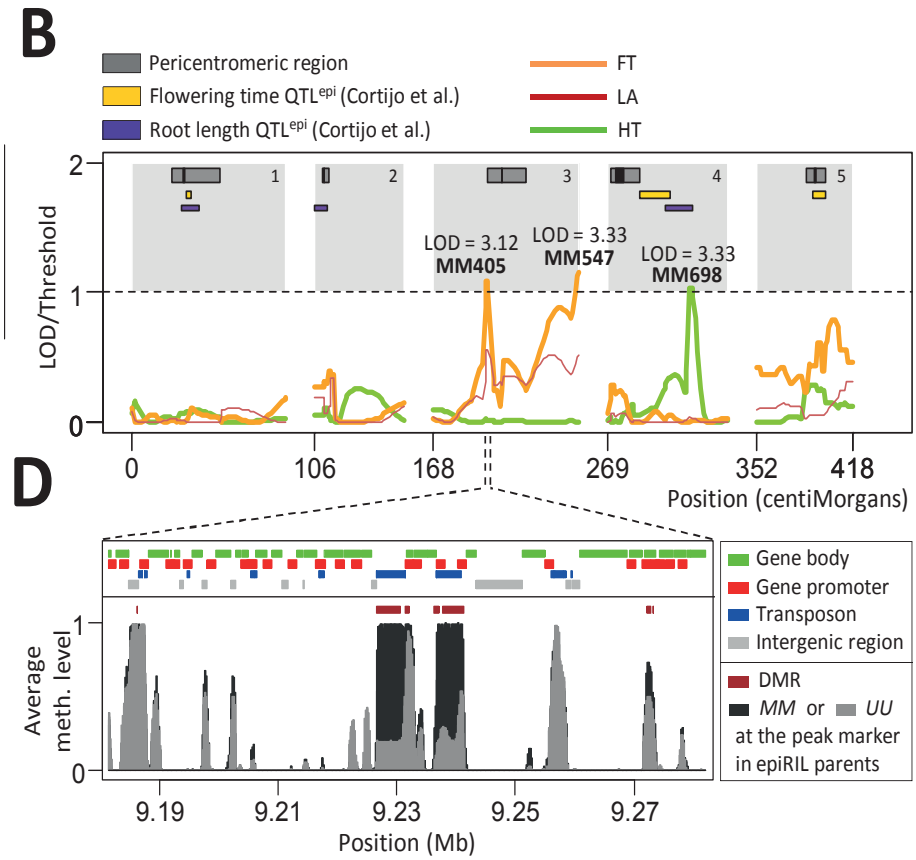


Figure 3: Interval mapping approach detects significant QTLs for mid-parent divergence. A) Genome-wide patterns of Col-wt and ddm1-2 inherited epi-haplotypes in the (epi)genomes of the parental epiRILs used in this study. B) QTL profiles for FT, HT and LA. Published QTLsepi for root length and flowering 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.

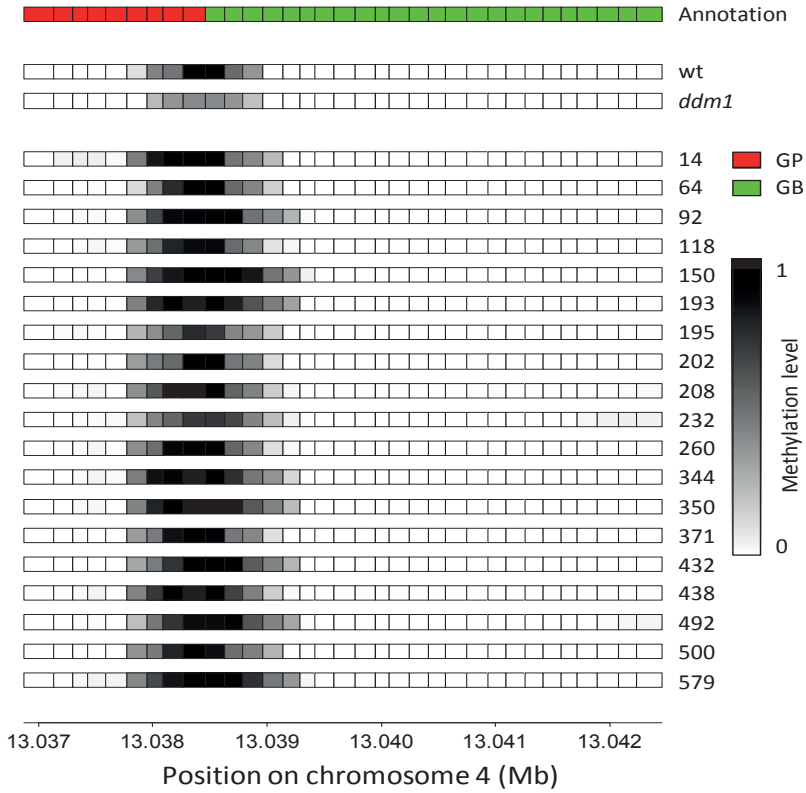
Funding

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).

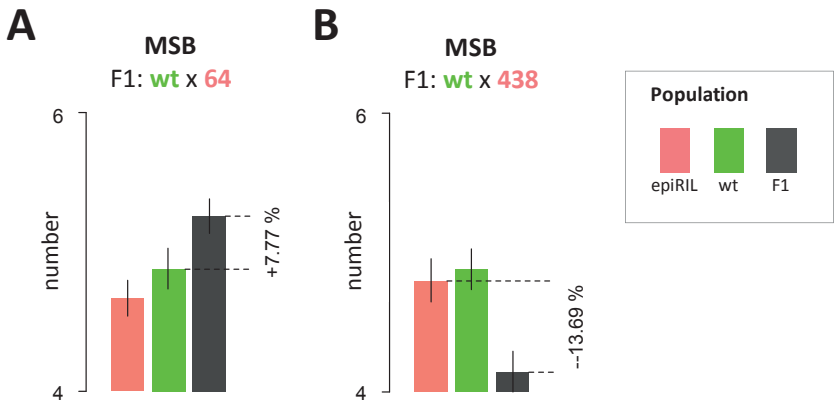
Author Contributions

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.

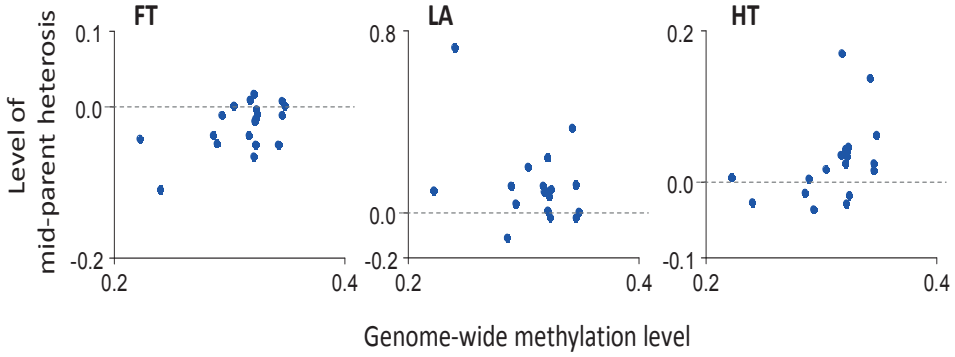
Supporting Information



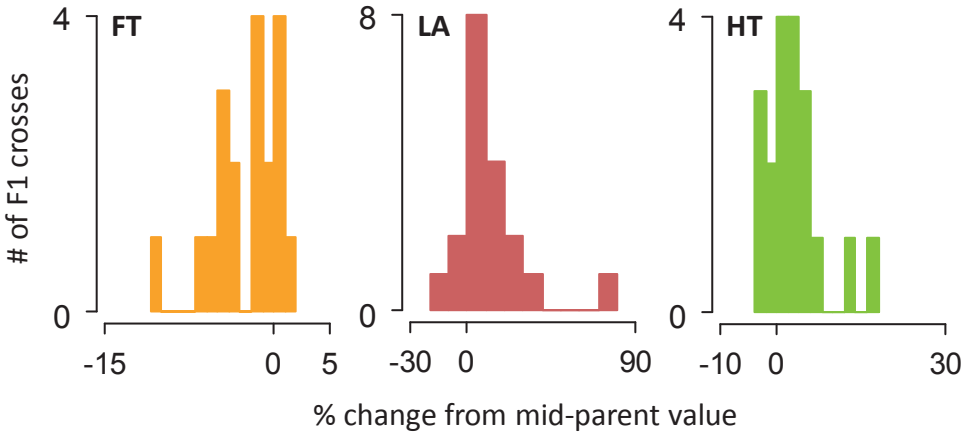
S1 Fig. Methylation profile at the *FWA* locus. Methylation level of tiling array probes located within the gene promoter (GP; red rectangles) and gene body (GB; green rectangles) of the *FWA* gene. The methylation profiles are shown for the wild-type parent, the *ddm1* mutant parent and the 19 selected 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).



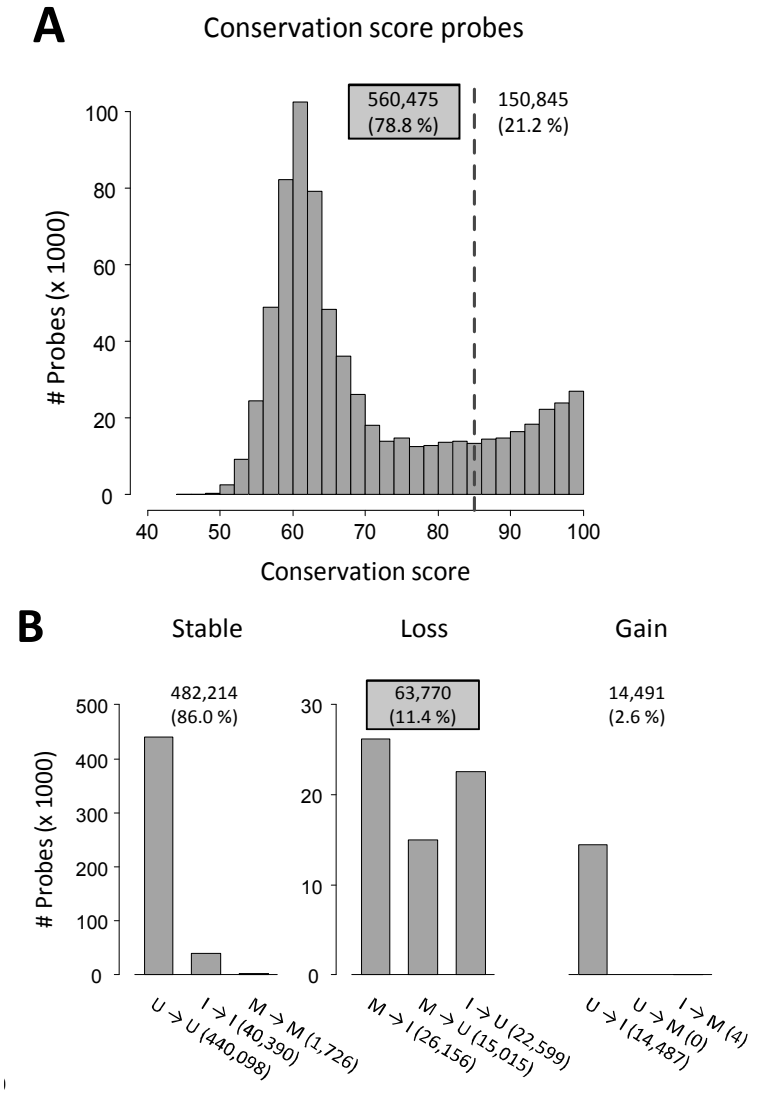
S2 Fig. Detected cases of high-parent heterosis (HPH; A) and low-parent heterosis (LPH; B) for MSB. In case of HPH percent increase is calculated with respect to the parent with higher phenotypic values. In case of LPH percent increase is calculated with respect to the parent with lower phenotypic values. 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).

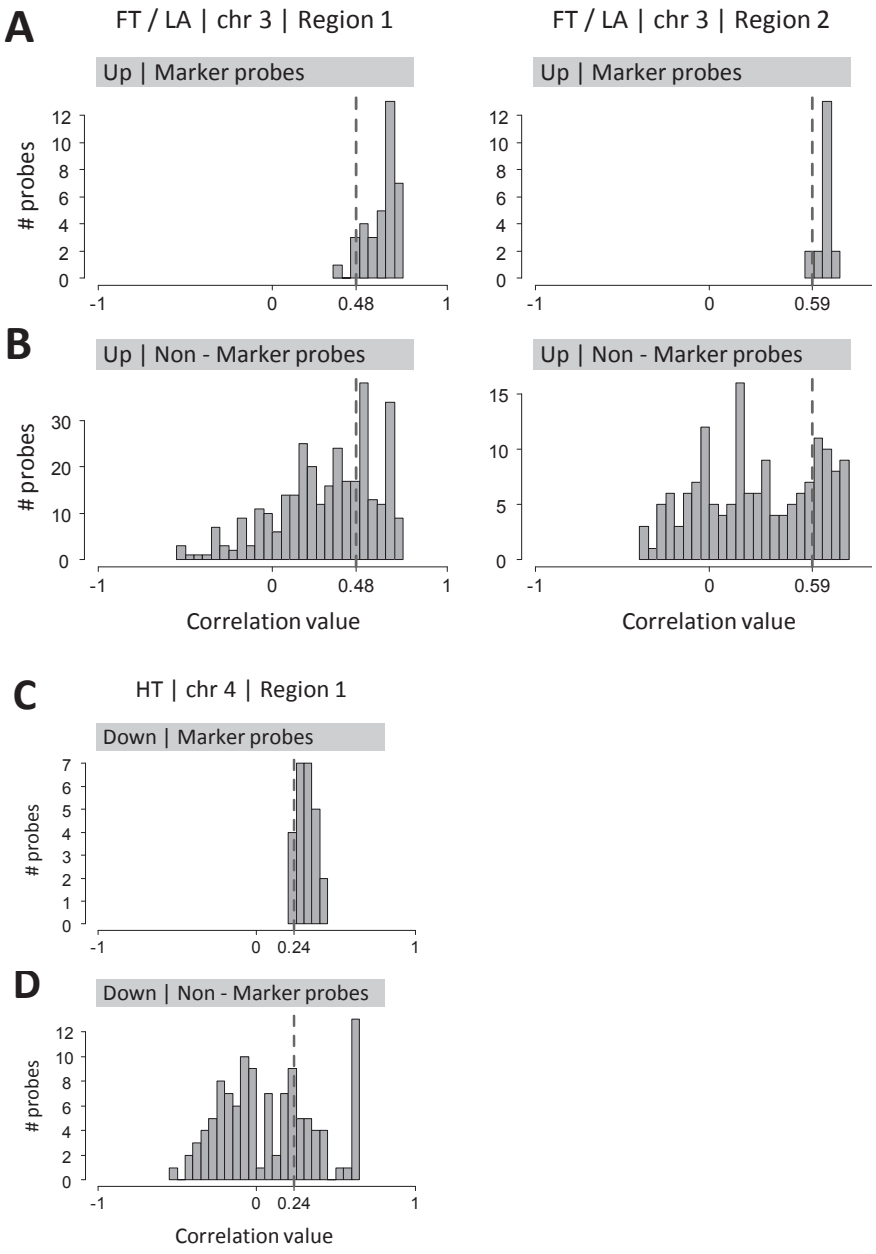


S4 Fig. Shown are frequency histograms of the percent change from mid-parent value for the 19 epiHybrid crosses. The percent change values are quantitatively distributed among the 19 epiHybrid crosses and can be treated as a phenotype for QTL mapping.



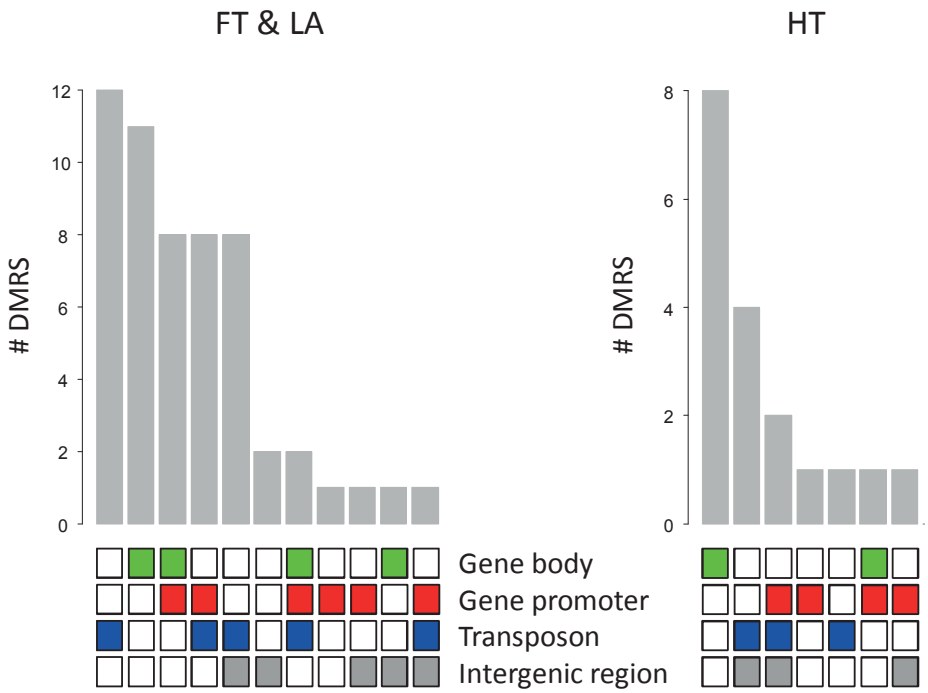
S5 Fig. Conservation score of probes and methylation differences between epiRIL, Col-wt and *ddm1-2* founder lines. The conservation score distribution of all 711,320 probes (A) and methylation differences between the epiRIL Col-wt and *ddm1-2* founder lines (B). Shown is the conservation score cutoff 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.

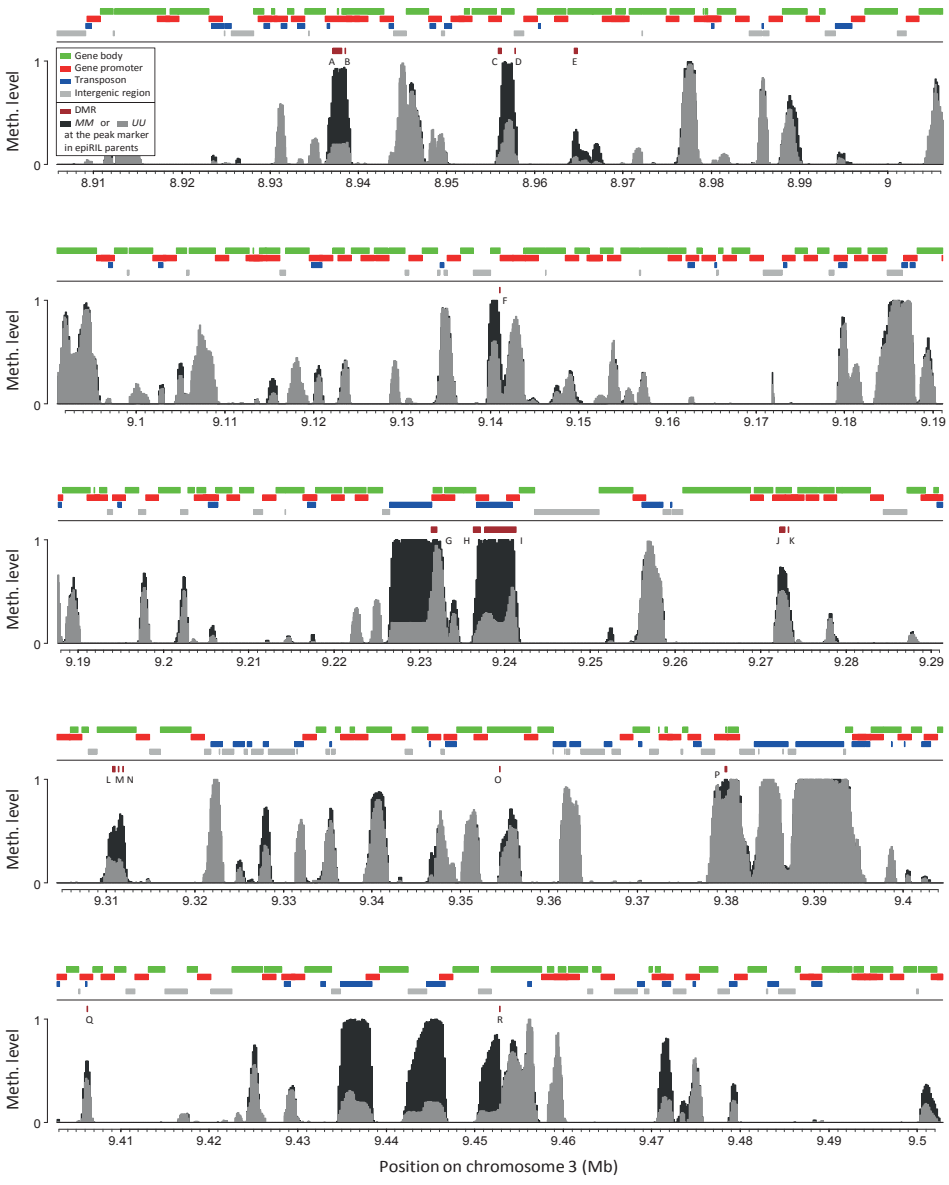


S6 Fig. Selection of candidate probes based on correlation with peak marker. Shown are the distributions of the correlation values of marker (A and C) and

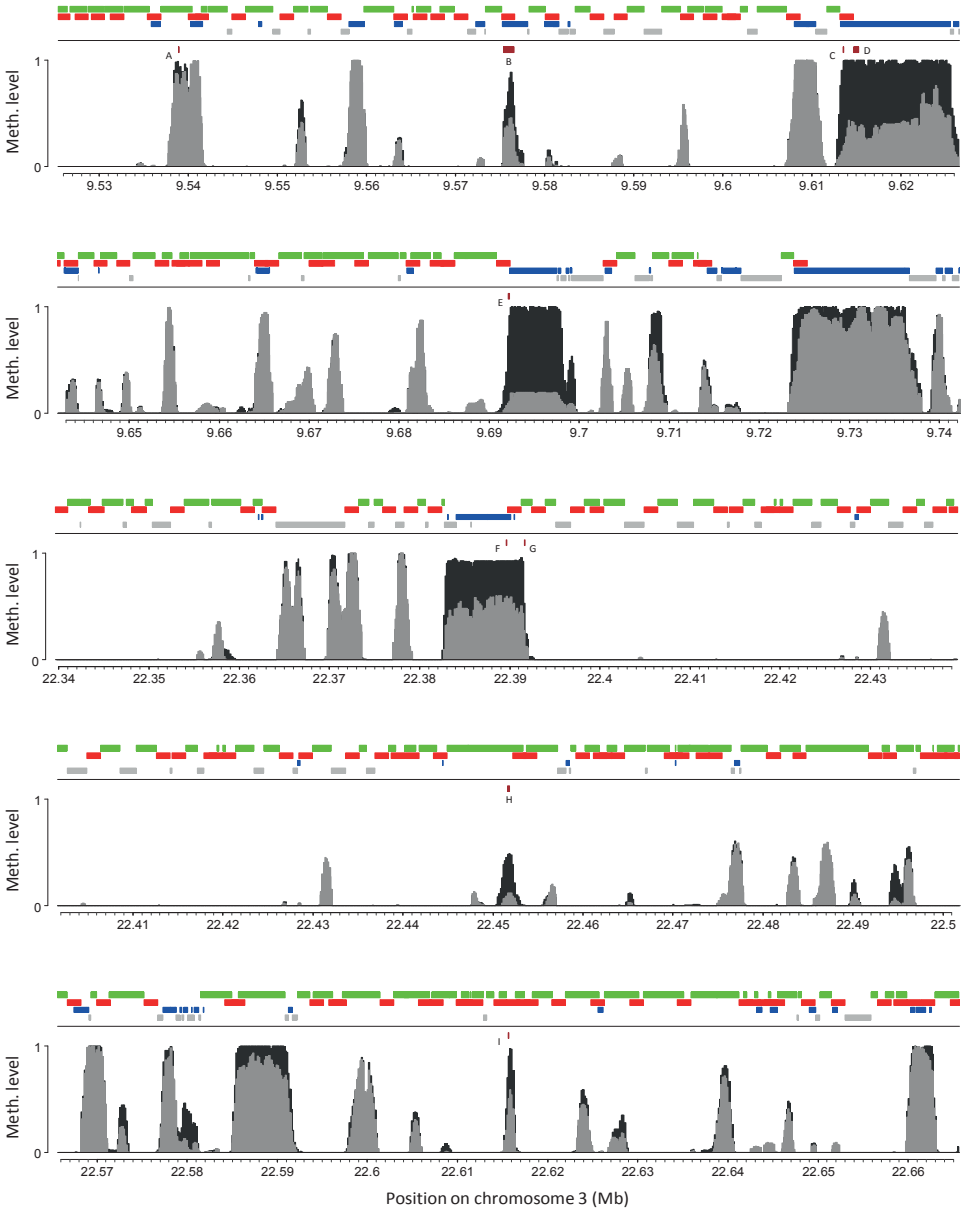
non-marker probes (B and D) upstream or downstream from the peak marker, and the cutoff that was used for the selection of candidate probes for the Flowering Time (FT), Leaf Area (LA) and Height (HT) QTL intervals. All marker probes were selected. The cutoff for non-marker probes was based on the 5th 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 nonmarker probes upstream from the QTL interval. For all QTL intervals the interval started or ended with a peak marker.



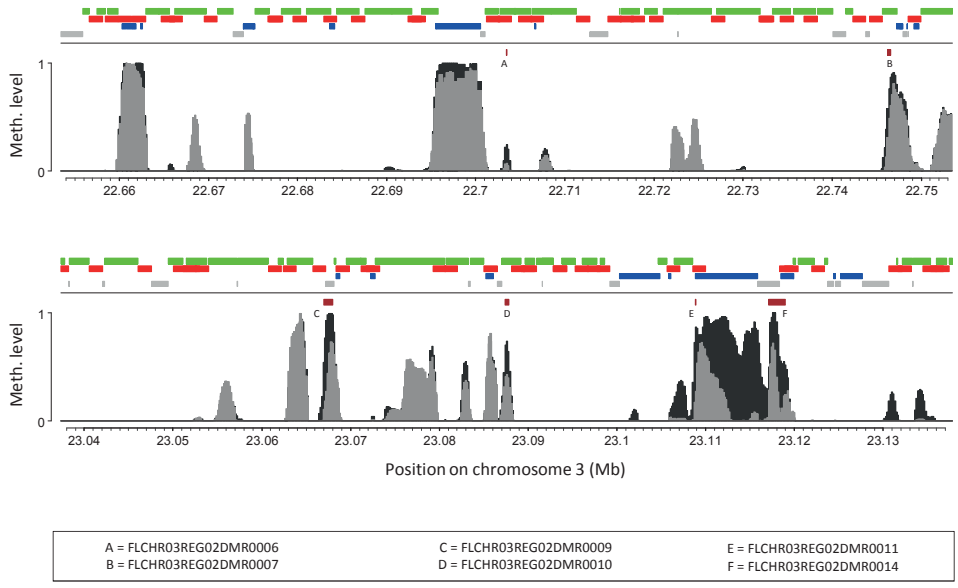
S7 Fig. Annotation categories of the DMRs. Shown are the number of DMRs that have an overlap with the different combinations of annotations indicated by the colored rectangles. The left barplot shows the results for Flowering Time (FT) and Leaf Area (LA). The right barplot shows the results for Height (HT).



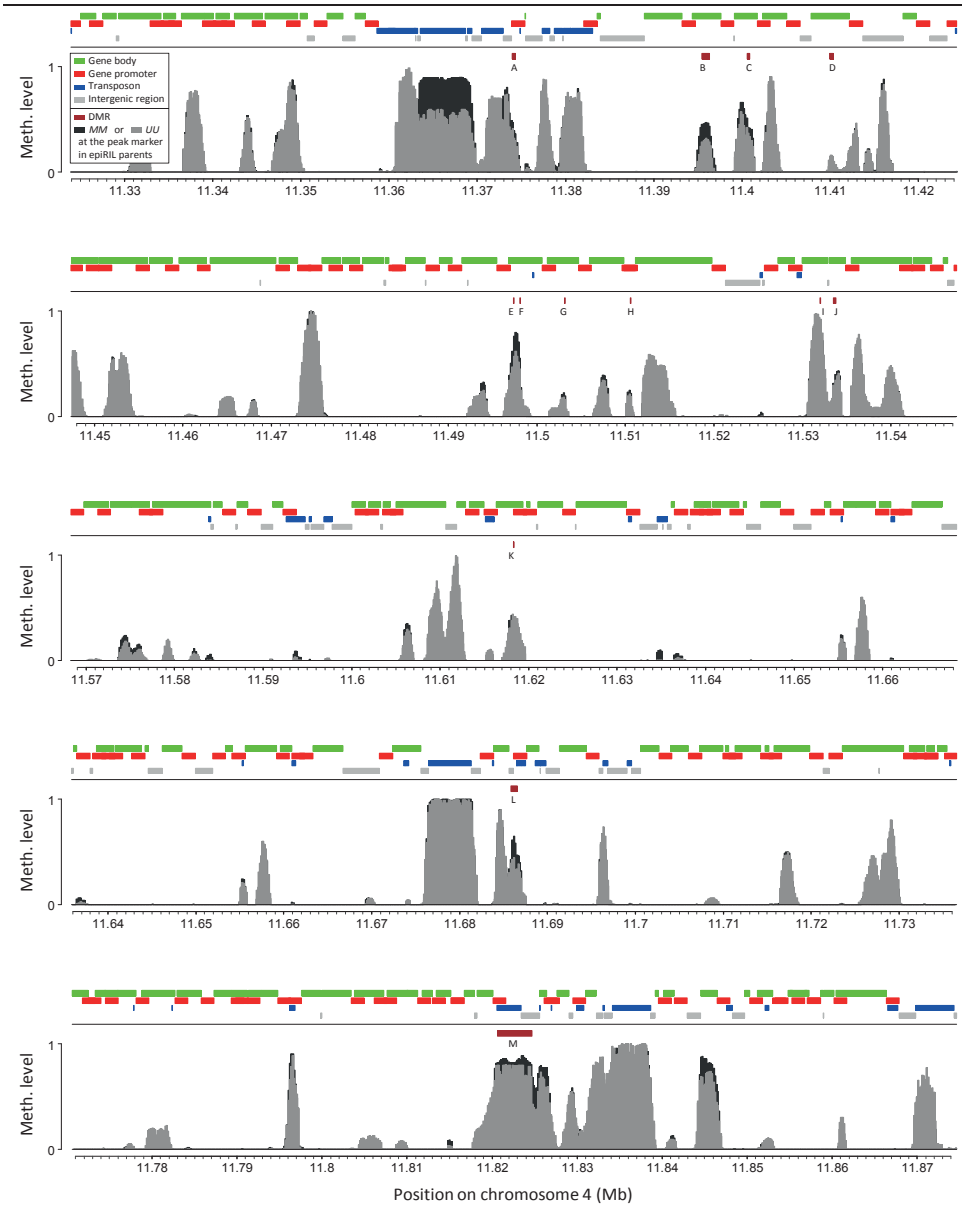
A = FLCHR03REG01DMR0001
B = FLCHR03REG01DMR0002
C = FLCHR03REG01DMR0003
D = FLCHR03REG01DMR0004
E = FLCHR03REG01DMR0005
F = FLCHR03REG01DMR0008
G = FLCHR03REG01DMR0011
H = FLCHR03REG01DMR0012
I = FLCHR03REG01DMR0013
J = FLCHR03REG01DMR0014
K = FLCHR03REG01DMR0015
L = FLCHR03REG01DMR0016
M = FLCHR03REG01DMR0017
N = FLCHR03REG01DMR0018
O = FLCHR03REG01DMR0021
P = FLCHR03REG01DMR0022
Q = FLCHR03REG01DMR0023
R = FLCHR03REG01DMR0026



A = FLCHR03REG01DMR0027	D = FLCHR03REG01DMR0030	G = FLCHR03REG02DMR0003
B = FLCHR03REG01DMR0028	E = FLCHR03REG01DMR0037	H = FLCHR03REG02DMR0004
C = FLCHR03REG01DMR0029	F = FLCHR03REG02DMR0002	I = FLCHR03REG02DMR0005



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.



A = HTCHR04REG01DMR0006	E = HTCHR04REG01DMR0010	I = HTCHR04REG01DMR0014	M = HTCHR04REG01DMR0018
B = HTCHR04REG01DMR0007	F = HTCHR04REG01DMR0011	J = HTCHR04REG01DMR0015	
C = HTCHR04REG01DMR0008	G = HTCHR04REG01DMR0012	K = HTCHR04REG01DMR0016	
D = HTCHR04REG01DMR0009	H = HTCHR04REG01DMR0013	L = HTCHR04REG01DMR0017	

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.

Line ID	Number of probes			Percentage of probes			Loss meth.	FT	RL	Outlier
	U	I	M	U	I	M	at *FWA*	(days)	(cm)	
14	501188	123340	86792	70.46	17.34	12.20	NO	43.02	59.51	NO
232	507776	111848	91696	71.39	15.72	12.89	NO	39.47	57.38	NO
92	478672	128169	104479	67.29	18.02	14.69	NO	38.92	54.28	NO
208	511974	93980	105366	71.98	13.21	14.81	NO	41.48	54.05	NO
438	523734	81102	106484	73.63	11.40	14.97	NO	39.47	56.18	NO
195	475069	126735	109516	66.79	17.82	15.40	NO	41.50	65.73	NO
350	513998	84165	113157	72.26	11.83	15.91	NO	40.15	55.62	NO
500	491867	106096	113357	69.15	14.92	15.94	NO	39.88	50.58	NO
150	468401	128733	114186	65.85	18.10	16.05	NO	40.01	57.67	NO
118	497059	99933	114328	69.88	14.05	16.07	NO	40.85	60.40	NO
432	479227	117478	114615	67.37	16.52	16.11	NO	40.91	54.92	NO
202	503503	93006	114811	70.78	13.08	16.14	NO	40.18	58.73	NO
344	504283	92213	114824	70.89	12.96	16.14	NO	39.96	65.21	NO
64	496686	99596	115038	69.83	14.00	16.17	NO	39.96	59.48	NO
492	493583	102625	115112	69.39	14.43	16.18	NO	41.06	59.86	NO
193	493558	97478	120284	69.39	13.70	16.91	NO	39.00	NA	NO
260	499492	90694	121134	70.22	12.75	17.03	NO	43.38	57.36	NO
579	497286	92790	121244	69.91	13.04	17.04	NO	41.02	65.69	NO
371	483236	106279	121805	67.94	14.94	17.12	NO	39.84	56.31	NO
wt	499673	89513	122134	70.25	12.58	17.17	NO			

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.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	μ_{Pl}	μ_{Ph}	μ_{F1}	σ^2_{Pl}	σ^2_{Ph}	σ^2_{F1}	Ph	F1 trend
HT	14H	25	27	27	51.87	52.08	52.31	7.20	4.97	3.73	wt	wt
HT	232H	25	24	26	52.08	53.81	51.58	4.97	30.54	8.45	epi	wt
HT	92H	25	27	27	50.56	52.08	50.67	3.12	4.97	6.85	wt	epi
HT	208H	25	24	26	48.29	52.08	50.48	7.65	4.97	5.93	wt	wt
HT	438H	25	25	27	52.08	53.68	51.00	4.97	4.85	6.46	epi	wt
HT	195H	25	23	26	49.17	52.08	51.52	11.79	4.97	4.07	wt	wt
HT	350H	25	25	24	48.90	52.08	52.38	6.44	4.97	3.72	wt	wt
HT	500H	25	19	28	42.16	52.08	55.14	282.03	4.97	18.40	wt	wt
HT	150H	25	25	27	52.08	53.30	55.00	4.97	13.88	5.46	epi	epi
HT	118H	25	26	26	49.15	52.08	51.88	5.12	4.97	4.27	wt	wt
HT	432H	25	26	27	52.08	55.71	52.44	4.97	6.80	3.79	epi	wt
HT	202H	25	25	24	52.08	57.64	57.10	4.97	5.99	5.96	epi	epi
HT	344H	25	27	25	52.08	53.17	54.50	4.97	4.02	3.38	epi	epi
HT	64H	25	26	25	49.46	52.08	53.20	5.38	4.97	2.02	wt	wt
HT	492H	25	26	27	52.08	57.23	53.74	4.97	8.82	4.81	epi	wt
HT	193H	25	24	26	44.79	52.08	55.12	14.45	4.97	8.43	wt	wt
HT	260H	25	26	27	52.08	52.31	53.52	4.97	6.76	8.89	epi	epi
HT	579H	25	23	28	52.08	52.48	53.11	4.97	5.35	7.30	epi	epi
HT	371H	25	26	25	52.08	52.40	55.50	4.97	13.22	8.54	epi	epi

S2 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 $N_{Col\text{-}wt}$, N_{epiRIL} and N_{F1}, respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with μ_{Pl} and $\sigma 2_{Pl}$, μ_{Ph} and $\sigma 2_{Ph}$, and μ_{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 > ± 2 SD from the mean were removed.

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

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 $N_{Col\text{-}wt}$, N_{epiRIL} and N_{F1}, respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with μ_{Pl} and $\sigma 2_{Pl}$, μ_{Ph} and $\sigma 2_{Ph}$, and μ_{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 > ± 2 SD from the mean were removed.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	μ_{Pl}	μ_{Ph}	μ_{F1}	σ^2_{Pl}	σ^2_{Ph}	σ^2_{F1}	Ph	F1 trend
RB	14H	23	28	27	3.13	4.07	3.89	0.75	2.66	1.33	epi	epi
RB	232H	23	22	27	3.13	4.68	4.30	0.75	1.66	1.29	epi	epi
RB	92H	23	27	26	3.13	4.44	3.73	0.75	1.26	1.08	epi	wt
RB	208H	23	26	27	3.13	3.42	3.56	0.75	2.09	1.79	epi	epi
RB	438H	23	25	27	3.13	4.24	3.63	0.75	1.86	1.24	epi	wt
RB	195H	23	26	28	3.13	3.54	2.86	0.75	2.18	1.39	epi	wt
RB	350H	23	26	25	3.13	4.04	3.68	0.75	2.44	1.81	epi	epi
RB	500H	23	19	28	3.13	3.89	3.50	0.75	2.54	1.74	epi	wt
RB	150H	23	25	26	3.13	4.12	4.35	0.75	0.86	1.12	epi	epi
RB	118H	23	26	28	3.13	3.31	3.50	0.75	1.34	1.59	epi	epi
RB	432H	23	27	28	3.13	3.93	3.86	0.75	2.23	2.35	epi	epi
RB	202H	23	25	25	3.13	4.20	3.56	0.75	1.17	0.92	epi	wt
RB	344H	23	27	28	3.13	3.70	3.64	0.75	0.91	1.94	epi	epi
RB	64H	23	27	28	3.13	3.85	3.18	0.75	1.13	2.00	epi	wt
RB	492H	23	26	28	3.13	4.23	3.93	0.75	0.90	1.33	epi	epi
RB	193H	23	24	27	3.13	5.88	4.30	0.75	1.77	1.52	epi	wt
RB	260H	23	26	28	3.13	3.65	3.11	0.75	1.92	1.06	epi	wt
RB	579H	23	23	27	3.13	4.17	3.44	0.75	1.24	2.41	epi	wt
RB	371H	23	26	25	3.13	4.04	3.76	0.75	1.64	2.61	epi	epi

S4 Table. Phenotype summary for Rosette Branching (RB). 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 $N_{Col\text{-}wt}$, N_{epiRIL} and N_{F1}, respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with μ_{Pl} and $\sigma 2_{Pl}$, μ_{Ph} and $\sigma 2_{Ph}$, and μ_{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 > ± 2 SD from the mean were removed.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	μ_{Pl}	μ_{Ph}	μ_{F1}	σ^2_{Pl}	σ^2_{Ph}	σ^2_{F1}	Ph	F1 trend
FT	14H	24	26	26	29.77	30.58	28.92	5.70	1.30	2.31	wt	epi
FT	232H	24	22	27	29.82	30.58	26.93	2.73	1.30	3.30	wt	epi
FT	92H	24	27	27	27.44	30.58	27.96	0.72	1.30	2.58	wt	epi
FT	208H	24	24	25	29.46	30.58	28.60	2.09	1.30	2.50	wt	epi
FT	438H	24	25	26	30.32	30.58	30.12	2.56	1.30	2.19	wt	epi
FT	195H	24	25	28	29.68	30.58	30.18	4.48	1.30	4.67	wt	wt
FT	350H	24	24	26	29.04	30.58	28.69	2.30	1.30	2.78	wt	epi
FT	500H	24	18	27	29.50	30.58	30.33	5.32	1.30	3.62	wt	wt
FT	150H	24	27	25	28.48	30.58	27.60	4.64	1.30	0.92	wt	epi
FT	118H	24	26	28	27.85	30.58	29.75	3.10	1.30	4.94	wt	wt
FT	432H	24	26	28	29.46	30.58	29.50	3.22	1.30	5.59	wt	epi
FT	202H	24	25	24	28.84	30.58	29.25	1.56	1.30	2.63	wt	epi
FT	344H	24	26	27	29.85	30.58	28.70	1.50	1.30	3.91	wt	epi
FT	64H	24	27	26	28.96	30.58	29.69	2.96	1.30	2.78	wt	epi
FT	492H	24	25	27	28.00	30.58	29.04	2.08	1.30	4.73	wt	epi
FT	193H	24	24	27	28.42	30.58	28.04	6.60	1.30	4.19	wt	epi
FT	260H	24	28	28	30.58	32.61	31.25	1.30	4.03	3.60	epi	wt
FT	579H	24	21	27	28.81	30.58	29.96	3.26	1.30	3.81	wt	wt
FT	371H	24	28	24	30.58	30.61	30.67	1.30	4.25	3.88	epi	epi

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 $N_{Col\text{-}wt}$, N_{epiRIL} and N_{F1},, respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with μ_{Pl} and $\sigma 2_{Pl}$, μ_{Ph} and $\sigma 2_{Ph}$, and μ_{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 > ± 2 SD from the mean were removed.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	μ_{Pl}	μ_{Ph}	μ_{F1}	σ^2_{Pl}	σ^2_{Ph}	σ^2_{F1}	Ph	F1 trend
GR	14H	25	27	26	0.4015	0.4162	0.4120	0.000804	0.002092	0.001200	epi	epi
GR	232H	25	23	26	0.3417	0.4015	0.4010	0.003646	0.000804	0.002400	wt	wt
GR	92H	25	26	26	0.4015	0.4036	0.4020	0.000804	0.000968	0.001400	epi	wt
GR	208H	25	24	26	0.3908	0.4015	0.3990	0.001424	0.000804	0.001100	wt	wt
GR	438H	25	25	27	0.3903	0.4015	0.3930	0.001582	0.000804	0.002600	wt	epi
GR	195H	25	24	26	0.3981	0.4015	0.4100	0.000548	0.000804	0.000700	wt	wt
GR	350H	25	26	26	0.3995	0.4015	0.4050	0.002061	0.000804	0.000800	wt	wt
GR	500H	25	17	25	0.3902	0.4015	0.4060	0.002656	0.000804	0.002000	wt	wt
GR	150H	25	24	25	0.4015	0.4096	0.3900	0.000804	0.002040	0.001300	epi	wt
GR	118H	25	26	26	0.3964	0.4015	0.4000	0.001640	0.000804	0.000800	wt	wt
GR	432H	25	25	26	0.4015	0.4029	0.4090	0.000804	0.001296	0.001500	epi	epi
GR	202H	25	24	22	0.4015	0.4078	0.4070	0.000804	0.002066	0.001000	epi	epi
GR	344H	25	27	27	0.4015	0.4099	0.4100	0.000804	0.001282	0.000900	epi	epi
GR	64H	25	25	26	0.3969	0.4015	0.3950	0.002778	0.000804	0.001400	wt	epi
GR	492H	25	25	26	0.4015	0.4211	0.4180	0.000804	0.002629	0.001500	epi	epi
GR	193H	25	24	27	0.3885	0.4015	0.4000	0.002730	0.000804	0.001100	wt	wt
GR	260H	25	27	27	0.4015	0.4147	0.3960	0.000804	0.001070	0.001800	epi	wt
GR	579H	25	22	26	0.4015	0.4185	0.4080	0.000804	0.001496	0.001800	epi	wt
GR	371H	25	25	24	0.3845	0.4015	0.4020	0.000976	0.000804	0.001300	wt	wt

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 $N_{Col\text{-}wt}$, N_{epiRIL} and N_{F1}, respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with μ_{Pl} and $\sigma 2_{Pl}$, μ_{Ph} and $\sigma 2_{Ph}$, and μ_{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 > ± 2 SD from the mean were removed.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	μ_{Pl}	μ_{Ph}	μ_{F1}	σ^2_{Pl}	σ^2_{Ph}	σ^2_{F1}	Ph	F1 trend
LA	14H	25	26	27	150.09	176.23	179.36	927.61	1505.19	971.14	epi	epi
LA	232H	25	23	27	53.07	150.09	175.61	552.72	927.61	1509.10	wt	wt
LA	92H	25	28	27	150.09	154.11	135.48	927.61	706.68	1800.42	epi	wt
LA	208H	25	24	26	123.62	150.09	153.21	382.60	927.61	1089.11	wt	wt
LA	438H	25	25	28	142.03	150.09	152.19	2081.07	927.61	1244.92	wt	wt
LA	195H	25	27	27	144.97	150.09	177.34	1449.86	927.61	1171.99	wt	wt
LA	350H	25	27	27	129.15	150.09	156.55	932.72	927.61	1190.17	wt	wt
LA	500H	25	17	26	87.13	150.09	129.95	3232.36	927.61	1513.60	wt	wt
LA	150H	25	26	27	119.98	150.09	168.37	2463.93	927.61	1849.20	wt	wt
LA	118H	25	27	27	139.70	150.09	146.63	1004.92	927.61	887.87	wt	wt
LA	432H	25	26	27	113.77	150.09	132.64	1073.30	927.61	954.32	wt	wt
LA	202H	25	25	24	126.99	150.09	148.70	837.43	927.61	1146.08	wt	wt
LA	344H	25	26	27	136.45	150.09	157.36	648.07	927.61	2165.23	wt	wt
LA	64H	25	28	27	104.96	150.09	125.15	1455.38	927.61	1146.89	wt	epi
LA	492H	25	26	27	150.09	165.14	173.98	927.61	1144.50	2107.41	epi	epi
LA	193H	25	25	27	101.53	150.09	173.14	1512.20	927.61	2062.76	wt	wt
LA	260H	25	26	26	125.74	150.09	154.98	1686.51	927.61	996.05	wt	wt
LA	579H	25	21	28	150.09	158.62	151.49	927.61	2232.32	1484.02	epi	wt
LA	371H	25	27	24	145.79	150.09	148.37	1217.65	927.61	1615.85	wt	wt

S7 Table. Phenotype summary for Leaf Area (LA). Provided are sample sizes, means and variances for the Colwt parents, epiRIL parents and the epiHybrids. The sample sizes for the Col-wt parents, epiRIL parents and the epiHybrids (F1) are denoted with $N_{Col\text{-}wt}$, N_{epiRIL} and N_{F1}, respectively. The means and variances for the low parents (Pl), high parents (Ph) and epiHybrids (F1) are denoted with μ_{Pl} and $\sigma 2_{Pl}$, μ_{Ph} and $\sigma 2_{Ph}$, and μ_{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 > ± 2 SD from the mean were removed.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_A	$df\ (l_A)$	*P*- value
HT	14H	25	27	27	-175.087	6	-175.313	5	0.500977
HT	232H	25	24	26	-193.748	6	-195.101	5	0.099987
HT	92H	25	27	27	-171.983	6	-172.622	5	0.258605
HT	208H	25	24	26	-172.526	6	-172.648	5	0.621807
HT	438H	25	25	27	-172.738	6	-177.973	5	0.001213
HT	195H	25	23	26	-170.167	6	-171.357	5	0.122811
HT	350H	25	25	24	-162.601	6	-169.200	5	0.000280
HT	500H	25	19	28	-215.088	6	-222.375	5	0.000135
HT	150H	25	25	27	-183.603	6	-190.431	5	0.000220
HT	118H	25	26	26	-167.886	6	-170.941	5	0.013438
HT	432H	25	26	27	-172.157	6	-176.276	5	0.004103
HT	202H	25	25	24	-167.340	6	-174.378	5	0.000176
HT	344H	25	27	25	-161.824	6	-169.780	5	6.64E-05
HT	64H	25	26	25	-157.053	6	-173.241	5	1.27E-08
HT	492H	25	26	27	-178.752	6	-180.089	5	0.102026
HT	193H	25	24	26	-184.729	6	-227.271	5	2.86E-20
HT	260H	25	26	27	-183.575	6	-185.550	5	0.046847
HT	579H	25	23	28	-173.515	6	-174.444	5	0.172689
HT	371H	25	26	25	-186.759	6	-197.020	5	5.89E-06

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 freedom given by $df\ (l_F)$; l_A denotes the log-likelihood of the additive model with degrees of freedom given by $df\ (l_A)$.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_A	$df\ (l_A)$	*P*- value
MSB	14H	25	26	28	-86.633	6	-87.452	5	0.200506
MSB	232H	25	22	27	-84.846	6	-84.949	5	0.649279
MSB	92H	25	25	25	-66.653	6	-77.920	5	2.06E-06
MSB	208H	25	25	26	-94.913	6	-95.558	5	0.256165
MSB	438H	25	25	28	-86.590	6	-94.303	5	8.59E-05
MSB	195H	25	25	27	-76.313	6	-76.371	5	0.733909
MSB	350H	25	25	23	-76.077	6	-77.302	5	0.117456
MSB	500H	25	19	28	-81.993	6	-82.743	5	0.220798
MSB	150H	25	26	27	-78.908	6	-83.193	5	0.003419
MSB	118H	25	28	28	-77.750	6	-79.707	5	0.047878
MSB	432H	25	26	26	-74.565	6	-74.696	5	0.608896
MSB	202H	25	26	24	-71.846	6	-76.181	5	0.003233
MSB	344H	25	27	27	-82.741	6	-83.566	5	0.198965
MSB	64H	25	27	27	-80.756	6	-85.391	5	0.002330
MSB	492H	25	25	27	-82.578	6	-84.399	5	0.056305
MSB	193H	25	25	27	-75.564	6	-77.719	5	0.037905
MSB	260H	25	25	28	-84.092	6	-85.150	5	0.145780
MSB	579H	25	23	27	-80.000	6	-80.261	5	0.469619
MSB	371H	25	25	23	-66.936	6	-66.966	5	0.807561

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 freedom given by $df\ (l_F)$; l_A denotes the log-likelihood of the additive model with degrees of freedom given by $df\ (l_A)$.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_A	$df\ (l_A)$	*P*- value
RB	14H	23	28	27	-123.530	6	-124.040	5	0.312689
RB	232H	23	22	27	-106.450	6	-107.470	5	0.154046
RB	92H	23	27	26	-107.240	6	-107.270	5	0.819323
RB	208H	23	26	27	-120.610	6	-121.020	5	0.365240
RB	438H	23	25	27	-112.350	6	-112.370	5	0.836748
RB	195H	23	26	28	-119.220	6	-120.670	5	0.088788
RB	350H	23	26	25	-119.270	6	-119.320	5	0.767054
RB	500H	23	19	28	-111.220	6	-111.220	5	0.968836
RB	150H	23	25	26	-99.803	6	-104.160	5	0.003173
RB	118H	23	26	28	-114.860	6	-115.370	5	0.314406
RB	432H	23	27	28	-128.700	6	-129.180	5	0.327151
RB	202H	23	25	25	-99.779	6	-99.877	5	0.658879
RB	344H	23	27	28	-113.940	6	-114.240	5	0.441264
RB	64H	23	27	28	-117.340	6	-117.880	5	0.298122
RB	492H	23	26	28	-107.190	6	-107.670	5	0.328254
RB	193H	23	24	27	-112.790	6	-113.040	5	0.473836
RB	260H	23	26	28	-113.820	6	-114.450	5	0.261998
RB	579H	23	23	27	-113.210	6	-113.400	5	0.532890
RB	371H	23	26	25	-118.660	6	-118.780	5	0.623956

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 $df\ (l_A)$.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_A	$df\ (l_A)$	*P*- value
FT	14H	24	26	26	-143.000	6	-147.990	5	0.001582
FT	232H	24	22	27	-132.370	6	-164.510	5	1.07E-15
FT	92H	24	27	27	-120.600	6	-125.380	5	0.001990
FT	208H	24	24	25	-125.480	6	-132.940	5	0.000112
FT	438H	24	25	26	-129.960	6	-130.420	5	0.338015
FT	195H	24	25	28	-151.190	6	-151.200	5	0.921234
FT	350H	24	24	26	-129.930	6	-134.270	5	0.003208
FT	500H	24	18	27	-131.930	6	-132.120	5	0.535302
FT	150H	24	27	25	-129.100	6	-149.140	5	2.44E-10
FT	118H	24	26	28	-149.340	6	-149.990	5	0.253293
FT	432H	24	26	28	-151.590	6	-152.150	5	0.290411
FT	202H	24	25	24	-122.340	6	-123.110	5	0.215080
FT	344H	24	26	27	-134.510	6	-141.130	5	0.000276
FT	64H	24	27	26	-138.830	6	-138.850	5	0.833507
FT	492H	24	25	27	-139.610	6	-139.770	5	0.578003
FT	193H	24	24	27	-150.030	6	-154.540	5	0.002682
FT	260H	24	28	28	-152.570	6	-152.910	5	0.413305
FT	579H	24	21	27	-134.240	6	-134.430	5	0.544378
FT	371H	24	28	24	-145.990	6	-146.000	5	0.877055

S11 Table. Test for mid-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_A denotes the log-likelihood of the additive model with degrees of freedom given by $df\ (l_A)$.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_A	$df\ (l_A)$	*P*- value
GR	14H	25	27	26	150.551	6	150.493	5	0.733068
GR	232H	25	23	26	128.591	6	125.538	5	0.013471
GR	92H	25	26	26	157.357	6	157.349	5	0.898575
GR	208H	25	24	26	151.884	6	151.826	5	0.733505
GR	438H	25	25	27	142.269	6	142.223	5	0.762863
GR	195H	25	24	26	168.894	6	167.574	5	0.104150
GR	350H	25	26	26	154.404	6	154.245	5	0.572709
GR	500H	25	17	25	123.444	6	123.021	5	0.357796
GR	150H	25	24	25	143.266	6	141.757	5	0.082317
GR	118H	25	26	26	157.024	6	157.017	5	0.911717
GR	432H	25	25	26	150.435	6	150.126	5	0.431830
GR	202H	25	24	22	140.537	6	140.505	5	0.800766
GR	344H	25	27	27	163.081	6	162.924	5	0.575091
GR	64H	25	25	26	141.311	6	141.232	5	0.690526
GR	492H	25	25	26	141.750	6	141.525	5	0.501849
GR	193H	25	24	27	145.561	6	145.379	5	0.547169
GR	260H	25	27	27	156.114	6	155.174	5	0.170241
GR	579H	25	22	26	140.752	6	140.719	5	0.796252
GR	371H	25	25	24	152.041	6	151.537	5	0.315467

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 $df\ (l_A)$.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_A	$df\ (l_A)$	*P*- value
LA	14H	25	26	27	-382.560	6	-384.760	5	0.036036
LA	232H	25	23	27	-361.760	6	-400.260	5	1.70E-18
LA	92H	25	28	27	-390.460	6	-392.140	5	0.066879
LA	208H	25	24	26	-352.600	6	-355.030	5	0.027650
LA	438H	25	25	28	-389.870	6	-390.120	5	0.477701
LA	195H	25	27	27	-389.670	6	-396.390	5	0.000247
LA	350H	25	27	27	-383.920	6	-386.240	5	0.031481
LA	500H	25	17	26	-344.270	6	-344.830	5	0.290000
LA	150H	25	26	27	-397.660	6	-403.140	5	0.000935
LA	118H	25	27	27	-380.970	6	-381.000	5	0.808696
LA	432H	25	26	27	-377.930	6	-377.930	5	0.923223
LA	202H	25	25	24	-357.570	6	-358.360	5	0.208724
LA	344H	25	26	27	-382.430	6	-383.470	5	0.149790
LA	64H	25	28	27	-394.490	6	-394.530	5	0.767753
LA	492H	25	26	27	-389.460	6	-390.820	5	0.098866
LA	193H	25	25	27	-387.710	6	-398.830	5	2.42E-06
LA	260H	25	26	26	-379.510	6	-381.790	5	0.032656
LA	579H	25	21	28	-372.110	6	-372.150	5	0.761378
LA	371H	25	27	24	-376.310	6	-376.310	5	0.962719

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 $df\ (l_A)$.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_{FD}	$df\ (l_{FD})$	*P*- value
HT	14H	25	27	27	-175.090	6	-175.170	5	0.685930
HT	232H	25	24	26	-193.750	6	-193.990	5	0.487088
HT	92H	25	27	27	-171.980	6	--	--	--
HT	208H	25	24	26	-172.530	6	--	--	--
HT	438H	25	25	27	-172.740	6	-174.070	5	0.102791
HT	195H	25	23	26	-170.170	6	--	--	--
HT	350H	25	25	24	-162.600	6	-162.720	5	0.620032
HT	500H	25	19	28	-215.090	6	-220.570	5	0.000933
HT	150H	25	25	27	-183.600	6	-185.510	5	0.050756
HT	118H	25	26	26	-167.890	6	--	--	--
HT	432H	25	26	27	-172.160	6	--	--	--
HT	202H	25	25	24	-167.340	6	--	--	--
HT	344H	25	27	25	-161.820	6	-164.970	5	0.012123
HT	64H	25	26	25	-157.050	6	-159.300	5	0.034208
HT	492H	25	26	27	-178.750	6	--	--	--
HT	193H	25	24	26	-184.730	6	-193.540	5	2.70E-05
HT	260H	25	26	27	-183.570	6	-184.820	5	0.114777
HT	579H	25	23	28	-173.510	6	-173.920	5	0.370676
HT	371H	25	26	25	-186.760	6	-192.430	5	0.000761

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.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_{FD}	$df\ (l_{FD})$	*P*- value
MSB	14H	25	26	28	-86.633	6	-86.726	5	0.666353
MSB	232H	25	22	27	-84.846	6	--	--	--
MSB	92H	25	25	25	-66.653	6	-67.158	5	0.314777
MSB	208H	25	25	26	-94.913	6	--	--	--
MSB	438H	25	25	28	-86.590	6	-91.527	5	0.001678
MSB	195H	25	25	27	-76.313	6	--	--	--
MSB	350H	25	25	23	-76.077	6	--	--	--
MSB	500H	25	19	28	-81.993	6	--	--	--
MSB	150H	25	26	27	-78.908	6	--	--	--
MSB	118H	25	28	28	-77.750	6	--	--	--
MSB	432H	25	26	26	-74.565	6	--	--	--
MSB	202H	25	26	24	-71.846	6	-71.944	5	0.657954
MSB	344H	25	27	27	-82.741	6	--	--	--
MSB	64H	25	27	27	-80.756	6	-82.700	5	0.048647
MSB	492H	25	25	27	-82.578	6	--	--	--
MSB	193H	25	25	27	-75.564	6	--	--	--
MSB	260H	25	25	28	-84.092	6	--	--	--
MSB	579H	25	23	27	-80.000	6	--	--	--
MSB	371H	25	25	23	-66.936	6	--	--	--

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.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_{FD}	$df\ (l_{FD})$	*P*- value
RB	14H	23	28	27	-123.530	6	--	--	--
RB	232H	23	22	27	-106.450	6	--	--	--
RB	92H	23	27	26	-107.240	6	--	--	--
RB	208H	23	26	27	-120.610	6	-120.670	5	0.729701
RB	438H	23	25	27	-112.350	6	--	--	--
RB	195H	23	26	28	-119.220	6	-119.670	5	0.340869
RB	350H	23	26	25	-119.270	6	--	--	--
RB	500H	23	19	28	-111.220	6	--	--	--
RB	150H	23	25	26	-99.803	6	-100.130	5	0.415976
RB	118H	23	26	28	-114.860	6	-115.030	5	0.559295
RB	432H	23	27	28	-128.700	6	--	--	--
RB	202H	23	25	25	-99.779	6	--	--	--
RB	344H	23	27	28	-113.940	6	--	--	--
RB	64H	23	27	28	-117.340	6	--	--	--
RB	492H	23	26	28	-107.190	6	--	--	--
RB	193H	23	24	27	-112.790	6	--	--	--
RB	260H	23	26	28	-113.820	6	-113.820	5	0.930222
RB	579H	23	23	27	-113.210	6	--	--	--
RB	371H	23	26	25	-118.660	6	--	--	--

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.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_{FD}	$df\ (l_{FD})$	*P*- value
FT	14H	24	26	26	-143.000	6	-144.160	5	0.127591
FT	232H	24	22	27	-132.370	6	-149.350	5	5.60E-09
FT	92H	24	27	27	-120.600	6	--	--	--
FT	208H	24	24	25	-125.480	6	-127.450	5	0.047087
FT	438H	24	25	26	-129.960	6	-130.070	5	0.635624
FT	195H	24	25	28	-151.190	6	--	--	--
FT	350H	24	24	26	-129.930	6	-130.230	5	0.438016
FT	500H	24	18	27	-131.930	6	--	--	--
FT	150H	24	27	25	-129.100	6	-130.970	5	0.053643
FT	118H	24	26	28	-149.340	6	--	--	--
FT	432H	24	26	28	-151.590	6	--	--	--
FT	202H	24	25	24	-122.340	6	--	--	--
FT	344H	24	26	27	-134.510	6	-137.740	5	0.011081
FT	64H	24	27	26	-138.830	6	--	--	--
FT	492H	24	25	27	-139.610	6	--	--	--
FT	193H	24	24	27	-150.030	6	-150.200	5	0.562762
FT	260H	24	28	28	-152.570	6	--	--	--
FT	579H	24	21	27	-134.240	6	--	--	--
FT	371H	24	28	24	-145.990	6	-146.000	5	0.915332

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.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_{FD}	$df\ (l_{FD})$	*P*- value
GR	14H	25	27	26	150.551	6	--	--	--
GR	232H	25	23	26	128.591	6	--	--	--
GR	92H	25	26	26	157.357	6	157.357	5	0.995782
GR	208H	25	24	26	151.884	6	--	--	--
GR	438H	25	25	27	142.269	6	--	--	--
GR	195H	25	24	26	168.894	6	168.268	5	0.263104
GR	350H	25	26	26	154.404	6	154.317	5	0.677049
GR	500H	25	17	25	123.444	6	123.345	5	0.656684
GR	150H	25	24	25	143.266	6	142.465	5	0.205608
GR	118H	25	26	26	157.024	6	--	--	--
GR	432H	25	25	26	150.435	6	150.255	5	0.547971
GR	202H	25	24	22	140.537	6	--	--	--
GR	344H	25	27	27	163.081	6	--	--	--
GR	64H	25	25	26	141.311	6	141.304	5	0.90742
GR	492H	25	25	26	141.750	6	--	--	--
GR	193H	25	24	27	145.561	6	--	--	--
GR	260H	25	27	27	156.114	6	155.930	5	0.543461
GR	579H	25	22	26	140.752	6	--	--	--
GR	371H	25	25	24	152.041	6	--	--	--

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; 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.

Phenotype	epiHybrid ID	$N_{Col\text{-}wt}$	N_{epiRIL}	N_{F1}	l_F	$df\ (l_F)$	l_{FD}	$df\ (l_{FD})$	*P*- value
LA	14H	25	26	27	-382.560	6	-382.610	5	0.746755
LA	232H	25	23	27	-361.760	6	-365.260	5	0.008124
LA	92H	25	28	27	-390.460	6	-391.490	5	0.151534
LA	208H	25	24	26	-352.600	6	-352.660	5	0.725336
LA	438H	25	25	28	-389.870	6	-389.900	5	0.815725
LA	195H	25	27	27	-389.670	6	-394.280	5	0.002390
LA	350H	25	27	27	-383.920	6	-384.180	5	0.472832
LA	500H	25	17	26	-344.270	6	--	--	--
LA	150H	25	26	27	-397.660	6	-399.250	5	0.075198
LA	118H	25	27	27	-380.970	6	--	--	--
LA	432H	25	26	27	-377.930	6	--	--	--
LA	202H	25	25	24	-357.570	6	--	--	--
LA	344H	25	26	27	-382.430	6	-382.660	5	0.502031
LA	64H	25	28	27	-394.490	6	--	--	--
LA	492H	25	26	27	-389.460	6	-389.780	5	0.423605
LA	193H	25	25	27	-387.710	6	-390.050	5	0.030435
LA	260H	25	26	26	-379.510	6	-379.670	5	0.573184
LA	579H	25	21	28	-372.110	6	--	--	--
LA	371H	25	27	24	-376.310	6	--	--	--

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.

Phenotype	R^2.adj	*F*- value	df.numerator	df.denominator	*P*- value
GR	0.026272	1.732991	18	471	3.11E-02
RB	0.017461	1.505481	18	494	8.27E-02
LA	0.278716	11.84114	18	487	1.15E-28
HT	0.512668	30.04658	18	479	5.01E-67
MSB	0.163212	6.439624	18	484	2.07E-14
FT	0.174109	6.879348	18	484	1.31E-15

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.

Phenotype	DMR	DMR*	Type	LOD	LOD/Thr	Chr	Region	Position (cM)	DMR* start (bps)	DMR* stop (bps)
Flowering	MM399	MM399	lower			3	1	34.80	8937118	8938596
Time (FT)	**MM405**	**MM405**	**peak**	**3.12**	**1.08**	**3**	**1**	**37.62**	**9692187**	**9698029**
	c3.loc40	MM405	upper			3	1	40.00	9692187	9698029
	MM546	MM546	lower			3	2	95.64	22232528	22235548
	MM547	**MM547**	**peak**	**3.33**	**1.16**	**3**	**2**	**101.44**	**23204533**	**23207428**
	MM547	MM547	upper			3	2	101.44	23204533	23207428
Leaf	**MM405**	**MM405**	**peak**	**2.40**	**0.55**					
Area (LA)	**MM547**	**MM547**	**peak**	**2.26**	**0.52**					
Height (HT)	MM698	MM698	lower			4	1	54.68	11363438	11369209
	c4.loc56	**MM698**	**peak**	**3.33**	**1.03**	**4**	**1**	**56.00**	**11363438**	**11369209**
	c4.loc62	MM699	upper			4	1	62.00	11820298	11824662

S21 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.

	Flowering Time & Leaf Area			Height	
	Chr 3		Total	Chr 4	Total
	Region 1	Region 2		Region 1	
# Probes in QTL interval	4608	5883	**10491**	2794	**2794**
# Probes pass 1 - quality	3719	4891	**8610**	2400	**2400**
# Probes pass 2 - loss meth	387	190	**577**	143	**143**
# Probes pass 3 - correlation	171	70	**241**	62	**62**
# DMRs	39	16	**55**	18	**18**
# Unique genes	24	11	**35**	14	**14**
# Gene promoters (GP)	16	7	**23**	6	**6**
# Gene bodies (GB)	14	5	**19**	8	**8**
# Transposable elements (TE)	18	9	**27**	13	**13**
# Intergenic regions (IGR)	9	5	**14**	6	**6**

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.

DMR	Unit Class			
ID	GP	GB	TE	IGR
FLCHR03REG01DMR0001	X	X		
FLCHR03REG01DMR0002	X	X		
FLCHR03REG01DMR0003	X	X		
FLCHR03REG01DMR0004		X		X
FLCHR03REG01DMR0005		X		
FLCHR03REG01DMR0006			X	
FLCHR03REG01DMR0007			X	X
FLCHR03REG01DMR0008	X	X		
FLCHR03REG01DMR0009				X
FLCHR03REG01DMR0010			X	X
FLCHR03REG01DMR0011	X	X	X	
FLCHR03REG01DMR0012	X	X	X	
FLCHR03REG01DMR0013	X		X	
FLCHR03REG01DMR0014	X	X		
FLCHR03REG01DMR0015	X	X		
FLCHR03REG01DMR0016		X		
FLCHR03REG01DMR0017		X		
FLCHR03REG01DMR0018		X		
FLCHR03REG01DMR0019			X	X
FLCHR03REG01DMR0020			X	X

DMR	Unit Class			
ID	GP	GB	TE	IGR
FLCHR03REG01DMR0021		X		
FLCHR03REG01DMR0022	X	X		
FLCHR03REG01DMR0023	X		X	
FLCHR03REG01DMR0024			X	
FLCHR03REG01DMR0025			X	X
FLCHR03REG01DMR0026		X		
FLCHR03REG01DMR0027		X		
FLCHR03REG01DMR0028	X		X	
FLCHR03REG01DMR0029	X		X	
FLCHR03REG01DMR0030	X		X	
FLCHR03REG01DMR0031			X	
FLCHR03REG01DMR0032			X	
FLCHR03REG01DMR0033			X	
FLCHR03REG01DMR0034			X	
FLCHR03REG01DMR0035			X	
FLCHR03REG01DMR0036			X	
FLCHR03REG01DMR0037	X		X	
FLCHR03REG01DMR0038			X	
FLCHR03REG01DMR0039			X	X

DMR	Unit Class			
ID	GP	GB	TE	IGR
FLCHR03REG02DMR0001			X	
FLCHR03REG02DMR0002	X		X	
FLCHR03REG02DMR0003		X		
FLCHR03REG02DMR0004		X		
FLCHR03REG02DMR0005	X			
FLCHR03REG02DMR0006	X	X		
FLCHR03REG02DMR0007		X		
FLCHR03REG02DMR0008			X	X
FLCHR03REG02DMR0009	X			X
FLCHR03REG02DMR0010		X		
FLCHR03REG02DMR0011	X		X	
FLCHR03REG02DMR0012			X	
FLCHR03REG02DMR0013			X	X
FLCHR03REG02DMR0014	X		X	X
FLCHR03REG02DMR0015				X
FLCHR03REG02DMR0016			X	

Annotation Combination	# DMRs
Gene Promoter *only*	1
Gene Body *only*	11
Transposon *only*	12
Intergenic *only*	2
Gene Promoter & Gene Body	8
Gene Promoter & Transposon	8
Gene Promoter & Intergenic	1
Gene Body & Transposon	0
Gene Body & Intergenic	1
Transposon & intergenic	8
Gene Promoter & Gene Body & Transposon	2
Gene Promoter & Gene Body & Intergenic	0
Gene Promoter & Transposon & Intergenic	1
Gene Body & Transposon & Intergenic	0
Gene Promoter & Gene Body & Transposon & Intergenic	0

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" (**F**lowering Time and ***L***eaf 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.

DMR ID	Unit Class GP	GB	TE	IGR	DMR ID	Unit Class GP	GB	TE	IGR	DMR ID	Unit Class GP	GB	TE	IGR
HTCHR04REG01DMR0001		X		X	HTCHR04REG01DMR0007	X				HTCHR04REG01DMR0013	X			
HTCHR04REG01DMR0002		X		X	HTCHR04REG01DMR0008	X				HTCHR04REG01DMR0014		X		
HTCHR04REG01DMR0003		X		X	HTCHR04REG01DMR0009	X				HTCHR04REG01DMR0015		X		
HTCHR04REG01DMR0004		X			HTCHR04REG01DMR0010	X				HTCHR04REG01DMR0016	X	X		
HTCHR04REG01DMR0005		X		X	HTCHR04REG01DMR0011	X				HTCHR04REG01DMR0017	X		X	X
HTCHR04REG01DMR0006	X			X	HTCHR04REG01DMR0012	X				HTCHR04REG01DMR0018	X		X	X

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.

DMR				Unit						Gene			Transposon	
ID	Chr	Start	Stop	Class	ID-1	ID-2	Start	Stop	Strand	Type	Name	Description	Family	Super family
FLCHR03REG01DMR0001	3	8936998	8938240	GP	AT3G24515		8936280	8937779	-	protein_coding_gene	UBC37	Ubiquitin-conjugating enzyme 37		
FLCHR03REG01DMR0001	3	8936998	8938240	GB	AT3G24516		8937588	8937776	-	protein_coding_gene		unknown protein		
FLCHR03REG01DMR0001	3	8936998	8938240	GP	AT3G24516		8937777	8939276	-	protein_coding_gene		unknown protein		
FLCHR03REG01DMR0001	3	8936998	8938240	GB	AT3G24517		8938106	8938641	-	protein_coding_gene		unknown protein		
FLCHR03REG01DMR0002	3	8938419	8938718	GP	AT3G24516		8937777	8939276	-	protein_coding_gene				
FLCHR03REG01DMR0002	3	8938419	8938718	GB	AT3G24517		8938106	8938641	-	protein_coding_gene				
FLCHR03REG01DMR0002	3	8938419	8938718	GB	AT3G24518		8938260	8942488	-	other_RNA		RNA gene linked to cell surface glycosylphosphatidylinositol (GPI)-anchored proteins (GAPs)		
FLCHR03REG01DMR0002	3	8938419	8938718	GP	AT3G24517		8938642	8940141	-	protein_coding_gene				
FLCHR03REG01DMR0003	3	8955763	8956377	GP	AT3G24542		8954757	8956256	+	protein_coding_gene		Beta-galactosidase related protein		
FLCHR03REG01DMR0003	3	8955763	8956377	GB	AT3G24542		8956257	8957667	+	protein_coding_gene		Beta-galactosidase related protein		
FLCHR03REG01DMR0004	3	8957618	8957917	GB	AT3G24542		8956257	8957667	+	protein_coding_gene				
FLCHR03REG01DMR0004	3	8957618	8957917	IGR			8957668	8958757						
FLCHR03REG01DMR0005	3	8964380	8964951	GB	AT3G24560		8963770	8966616	+	protein_coding_gene	RSY3	Raspberry 3		
FLCHR03REG01DMR0006	3	9011517	9015459	TE	AT3TE37690	CT00000552	9011213	9017190	-				TA11	LINE/L1
FLCHR03REG01DMR0007	3	9015686	9017273	TE	AT3TE37690	CT00000552	9011213	9017190	-				TA11	LINE/L1
FLCHR03REG01DMR0007	3	9015686	9017273	IGR			9017191	9018803						
FLCHR03REG01DMR0008	3	9140932	9141231	GB	AT3G25090		9140039	9141094	-	protein_coding_gene		F-box associated ubiquitination effector family protein		
FLCHR03REG01DMR0008	3	9140932	9141231	GP	AT3G25090		9141095	9142594	-	protein_coding_gene		F-box associated ubiquitination effector family protein		
FLCHR03REG01DMR0009	3	9186026	9186325	IGR			9184793	9186526						
FLCHR03REG01DMR0010	3	9226433	9230751	IGR			9225711	9226540						
FLCHR03REG01DMR0010	3	9226433	9230751	TE	AT3TE38560	CT00027716	9226541	9226632	+				ATCOPIA65A	LTR/Copia
FLCHR03REG01DMR0010	3	9226433	9230751	TE	AT3TE38565	CT00000926	9226633	9231377	+				ATCOPIA65	LTR/Copia
FLCHR03REG01DMR0011	3	9231282	9232185	TE	AT3TE38565	CT00000926	9226633	9231377	+				ATCOPIA65	LTR/Copia
FLCHR03REG01DMR0011	3	9231282	9232185	TE	AT3TE38570	CT00030923	9231378	9231407	+				ATCOPIA65	LTR/Copia
FLCHR03REG01DMR0011	3	9231282	9232185	TE	AT3TE38575	CT00030607	9231408	9231453	+				ATCOPIA45	LTR/Copia
FLCHR03REG01DMR0011	3	9231282	9232185	GP	AT3G25470		9231419	9232918	+	protein_coding_gene		Bacterial hemolysin-related		
FLCHR03REG01DMR0011	3	9231282	9232185	GB	AT3G25460		9231581	9232666	-	protein_coding_gene		F-box and associated interaction domains-containing protein		
FLCHR03REG01DMR0012	3	9236235	9237302	GB	AT3G25480		9235196	9236607	-	protein_coding_gene		Rhodanese/Cell cycle control phosphatase superfamily protein		
FLCHR03REG01DMR0012	3	9236235	9237302	GP	AT3G25480		9236608	9238107	-	protein_coding_gene		Rhodanese/Cell cycle control phosphatase superfamily protein		
FLCHR03REG01DMR0012	3	9236235	9237302	TE	AT3TE38590	CT00028056	9236680	9236764	-				ATLINE1_3A	LINE/L1
FLCHR03REG01DMR0012	3	9236235	9237302	TE	AT3TE38595	CT00001168	9236767	9240922	-				ATLINEIII	LINE/L1
FLCHR03REG01DMR0013	3	9237526	9241479	GP	AT3G25480		9236608	9238107	-	protein_coding_gene		Rhodanese/Cell cycle control phosphatase superfamily protein		
FLCHR03REG01DMR0013	3	9237526	9241479	TE	AT3TE38595	CT00001168	9236767	9240922	-				ATLINEIII	LINE/L1
FLCHR03REG01DMR0013	3	9237526	9241479	GP	AT3G25490		9240225	9241724	+	protein_coding_gene		Protein kinase family protein		
FLCHR03REG01DMR0014	3	9272151	9272965	GP	AT3G25520		9271401	9272900	-	protein_coding_gene	ATL5	ribosomal protein L5		
FLCHR03REG01DMR0014	3	9272151	9272965	GB	AT3G25530		9271800	9273615	-	protein_coding_gene	GLYR1	glyoxylate reductase 1		
FLCHR03REG01DMR0014	3	9272151	9272965	GP	AT3G25540		9272835	9274334	+	protein_coding_gene	LAG1	Longevity assurance gene 1		
FLCHR03REG01DMR0015	3	9273133	9273432	GB	AT3G25530		9271800	9273615	-	protein_coding_gene	GLYR1	glyoxylate reductase 1		

FLCHR03REG01DMR0015	3	9273133	9273432	GP	AT3G25540		9272835	9274334	+	protein_coding_gene	LAG1	Longevity assurance gene 1			
FLCHR03REG01DMR0016	3	9310602	9311075	GB	AT3G25610		9308942	9313353	-	protein_coding_gene		ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein			
FLCHR03REG01DMR0017	3	9311267	9311566	GB	AT3G25610		9308942	9313353	-	protein_coding_gene		ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein			
FLCHR03REG01DMR0018	3	9311756	9312055	GB	AT3G25610		9308942	9313353	-	protein_coding_gene		ATPase E1-E2 type family protein / haloacid dehalogenase-like hydrolase family protein			
FLCHR03REG01DMR0019	3	9327277	9327749	IGR			9326438	9327693							
FLCHR03REG01DMR0019	3	9327277	9327749	TE	AT3TE38960	CT00012214	9327694	9328269	+					ATCOPIA42	LTR/Copia
FLCHR03REG01DMR0020	3	9327942	9328547	TE	AT3TE38960	CT00012214	9327694	9328269	+					ATCOPIA42	LTR/Copia
FLCHR03REG01DMR0020	3	9327942	9328547	IGR			9328270	9331275							
FLCHR03REG01DMR0021	3	9354321	9354620	GB	AT3G25690		9353010	9357953	+	protein_coding_gene	CHUP1	Chloroplast unusual positioning 1			
FLCHR03REG01DMR0022	3	9379770	9380200	GP	AT3G25720		9378734	9380233	+	protein_coding_gene		RNA-directed DNA polymerase (reverse transcriptase)-related family protein			
FLCHR03REG01DMR0022	3	9379770	9380200	GB	AT3G25719		9379916	9380068	-	protein_coding_gene		unknown protein			
FLCHR03REG01DMR0022	3	9379770	9380200	GP	AT3G25719		9380069	9381568	-	protein_coding_gene		unknown protein			
FLCHR03REG01DMR0023	3	9406125	9406424	GP	AT3G25770		9405373	9406872	+	protein_coding_gene	AOC2	Allene oxide cyclase 2			
FLCHR03REG01DMR0023	3	9406125	9406424	TE	AT3TE39280	CT00023996	9406026	9406190	-					RathE3_cons	RathE3_cons
FLCHR03REG01DMR0024	3	9435048	9436453	TE	AT3TE39395	CT00001462	9434862	9438342	-					ATLINE2	LINE/L1
FLCHR03REG01DMR0025	3	9444444	9444743	IGR			9442508	9444590							
FLCHR03REG01DMR0025	3	9444444	9444743	TE	AT3TE39430	CT00002685	9444591	9446673	-					ATLINEIII	LINE/L1
FLCHR03REG01DMR0026	3	9452707	9453006	GB	AT3G25840		9451861	9457554	-	protein_coding_gene		Protein kinase superfamily protein			
FLCHR03REG01DMR0027	3	9538793	9539092	GB	AT3G26100		9536923	9540453	+	protein_coding_gene		Regulator of chromosome condensation (RCC1) family protein			
FLCHR03REG01DMR0028	3	9575250	9576723	GP	AT3G26170		9575149	9576648	-	protein_coding_gene	CYP71B19	cytochrome P450, family 71, subfamily B, polypeptide 19			
FLCHR03REG01DMR0028	3	9575250	9576723	TE	AT3TE39965	CT00012132	9575237	9575819	+					ATREP19	DNA
FLCHR03REG01DMR0028	3	9575250	9576723	TE	AT3TE39970	CT00029107	9575820	9575888	+					ATREP3	RC/Helitron
FLCHR03REG01DMR0028	3	9575250	9576723	TE	AT3TE39975	CT00019302	9575900	9576191	+					ATLINE1_1	LINE/L1
FLCHR03REG01DMR0028	3	9575250	9576723	TE	AT3TE39980	CT00004243	9576220	9577646	+					ATREP11	RC/Helitron
FLCHR03REG01DMR0029	3	9613421	9613720	GP	AT3G26250		9613169	9614668	-	protein_coding_gene		Cysteine/Histidine-rich C1 domain family protein			
FLCHR03REG01DMR0029	3	9613421	9613720	TE	AT3TE40145	CT00000079	9613221	9625546	-					VANDAL13	DNA/MuDR
FLCHR03REG01DMR0030	3	9614571	9615361	GP	AT3G26250		9613169	9614668	-	protein_coding_gene		Cysteine/Histidine-rich C1 domain family protein			
FLCHR03REG01DMR0030	3	9614571	9615361	TE	AT3TE40145	CT00000079	9613221	9625546	-					VANDAL13	DNA/MuDR
FLCHR03REG01DMR0031	3	9616005	9616304	TE	AT3TE40145	CT00000079	9613221	9625546	-					VANDAL13	DNA/MuDR
FLCHR03REG01DMR0032	3	9617036	9617335	TE	AT3TE40145	CT00000079	9613221	9625546	-					VANDAL13	DNA/MuDR
FLCHR03REG01DMR0033	3	9617711	9618010	TE	AT3TE40145	CT00000079	9613221	9625546	-					VANDAL13	DNA/MuDR
FLCHR03REG01DMR0034	3	9618204	9618503	TE	AT3TE40145	CT00000079	9613221	9625546	-					VANDAL13	DNA/MuDR
FLCHR03REG01DMR0035	3	9618823	9619122	TE	AT3TE40145	CT00000079	9613221	9625546	-					VANDAL13	DNA/MuDR
FLCHR03REG01DMR0036	3	9621997	9622296	TE	AT3TE40145	CT00000079	9613221	9625546	-					VANDAL13	DNA/MuDR
FLCHR03REG01DMR0037	3	9691951	9692367	GP	AT3G26480		9690828	9692327	-	protein_coding_gene		Transducin family protein / WD-40 repeat family protein			
FLCHR03REG01DMR0037	3	9691951	9692367	TE	AT3TE40420	CT00000677	9692214	9697529	-					ATCOPIA82	LTR/Copia
FLCHR03REG01DMR0038	3	9692617	9697164	TE	AT3TE40420	CT00000677	9692214	9697529	-					ATCOPIA82	LTR/Copia
FLCHR03REG01DMR0039	3	9697364	9698151	TE	AT3TE40420	CT00000677	9692214	9697529	-					ATCOPIA82	LTR/Copia
FLCHR03REG01DMR0039	3	9697364	9698151	IGR			9697530	9697667							
FLCHR03REG01DMR0039	3	9697364	9698151	TE	AT3TE40425	CT00018512	9697668	9697985	+					ATLINE1_3A	LINE/L1

FLCHR03REG01DMR0039	3	9697364	9698151	IGR			9697986	9698527														
FLCHR03REG02DMR0001	3	22232404	22236691	TE	AT3TE90530	CT00001057	22232308	22236743	+											ATCOPIA23	LTR/Copia	
FLCHR03REG02DMR0002	3	22389505	22389804	TE	AT3TE91200	CT00001135	22385751	22390053	-											ATCOPIA51	LTR/Copia	
FLCHR03REG02DMR0002	3	22389505	22389804	GP	AT3G60570		22389747	22391246	+	protein_coding_gene	EXPB5	Expansin B5										
FLCHR03REG02DMR0003	3	22391514	22391813	GB	AT3G60570		22391247	22392463	+	protein_coding_gene	EXPB5	Expansin B5										
FLCHR03REG02DMR0004	3	22451538	22451974	GB	AT3G60740		22447245	22453364	-	protein_coding_gene	TTN1	Titan 1										
FLCHR03REG02DMR0005	3	22615528	22615827	GP	AT3G61113		22615416	22616915	-	protein_coding_gene	URM12	Ubiquitin-related modifier 12										
FLCHR03REG02DMR0006	3	22703336	22703635	GP	AT3G61340		22702482	22703981	-	protein_coding_gene		F-box and associated interaction domains-containing protein;										
FLCHR03REG02DMR0006	3	22703336	22703635	GB	AT3G61350		22702684	22704481	+	protein_coding_gene	SKIP4	SKP1 interacting partner 4										
FLCHR03REG02DMR0007	3	22746086	22746660	GB	AT3G61470		22745653	22747282	+	protein_coding_gene	LHCA2	Photosystem I light harvesting complex gene 2										
FLCHR03REG02DMR0008	3	22767859	22768158	IGR			22767798	22767980														
FLCHR03REG02DMR0008	3	22767859	22768158	TE	AT3TE92810	CT00030896	22767981	22768018	-											VANDAL6	DNA/MuDR	
FLCHR03REG02DMR0008	3	22767859	22768158	IGR			22768019	22768916														
FLCHR03REG02DMR0009	3	23066834	23068104	GP	AT3G62330		23065722	23067221	-	protein_coding_gene		Zinc knuckle (CCHC-type) family protein										
FLCHR03REG02DMR0009	3	23066834	23068104	IGR			23067222	23068159														
FLCHR03REG02DMR0010	3	23087276	23087932	GB	AT3G62390		23087017	23089206	-	protein_coding_gene	TBL6	Trichome Birefringence-like 6										
FLCHR03REG02DMR0011	3	23108707	23109006	GP	AT3G62470		23108512	23110011	-	protein_coding_gene		Pentatricopeptide repeat (PPR) superfamily protein										
FLCHR03REG02DMR0011	3	23108707	23109006	TE	AT3TE94210	CT00020361	23108857	23109114	-											ATLANTYS3	LTR/Gypsy	
FLCHR03REG02DMR0012	3	23110389	23110688	TE	AT3TE94220	CT00000498	23109446	23115857	+											ATLANTYS3	LTR/Gypsy	
FLCHR03REG02DMR0013	3	23115853	23116152	TE	AT3TE94220	CT00000498	23109446	23115857	+											ATLANTYS3	LTR/Gypsy	
FLCHR03REG02DMR0013	3	23115853	23116152	IGR			23115858	23118343														
FLCHR03REG02DMR0014	3	23116989	23119117	IGR			23115858	23118343														
FLCHR03REG02DMR0014	3	23116989	23119117	GP	AT3G62499		23118344	23119843	+	protein_coding_gene		YTH family protein										
FLCHR03REG02DMR0014	3	23116989	23119117	TE	AT3TE94230	CT00013434	23118448	23118546	-											HELITRONY1B	RC/Helitron	
FLCHR03REG02DMR0014	3	23116989	23119117	TE	AT3TE94235	CT00004882	23118547	23119829	+											HELITRONY1E	RC/Helitron	
FLCHR03REG02DMR0014	3	23116989	23119117	GP	AT3G62500		23118923	23120422	+	protein_coding_gene		unknown protein										
FLCHR03REG02DMR0015	3	23203261	23203560	IGR			23202914	23203576														
FLCHR03REG02DMR0016	3	23205264	23207553	TE	AT3TE94580	CT00001270	23203577	23207484	-											ATLINEIII	LINE/L1	
FLCHR03REG02DMR0016	3	23205264	23207553	TE	AT3TE94585	CT00026247	23207485	23207582	-											ATLINEIII	LINE/L1	

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.

DMR				Unit						Gene			Transposon	
ID	Chr	Start	Stop	Class	ID-1	ID-2	Start	Stop	Strand	Type	Name	Description	Family	Super family
HTCHR04REG01DMR0001	4	11363314	11363722	IGR			11363211	11363559						
HTCHR04REG01DMR0001	4	11363314	11363722	TE	AT4TE52315	CT00000765	11363560	11368654	-				ATCOPIA10	LTR/Copia
HTCHR04REG01DMR0002	4	11368555	11369002	TE	AT4TE52315	CT00000765	11363560	11368654	-				ATCOPIA10	LTR/Copia
HTCHR04REG01DMR0002	4	11368555	11369002	IGR			11368655	11368905						
HTCHR04REG01DMR0002	4	11368555	11369002	TE	AT4TE52320	CT00013475	11368906	11369421	+				ATLINE2	LINE/L1
HTCHR04REG01DMR0003	4	11369210	11369673	TE	AT4TE52320	CT00013475	11368906	11369421	+				ATLINE2	LINE/L1
HTCHR04REG01DMR0003	4	11369210	11369673	IGR			11369422	11370469						
HTCHR04REG01DMR0004	4	11370887	11371186	TE	AT4TE52330	CT00023036	11370801	11370918	+				ARNOLDY2	DNA/MuDR
HTCHR04REG01DMR0004	4	11370887	11371186	TE	AT4TE52335	CT00018296	11370919	11370988	+				ATHPOGON1	DNA/Pogo
HTCHR04REG01DMR0004	4	11370887	11371186	TE	AT4TE52340	CT00003442	11370989	11372669	+				ARNOLDY1	DNA/MuDR
HTCHR04REG01DMR0005	4	11372657	11373127	TE	AT4TE52340	CT00003442	11370989	11372669	+				ARNOLDY1	DNA/MuDR
HTCHR04REG01DMR0005	4	11372657	11373127	TE	AT4TE52345	CT00003963	11372670	11372756	+				ARNOLDY1	DNA/MuDR
HTCHR04REG01DMR0005	4	11372657	11373127	TE	AT4TE52350	CT00023994	11372757	11372836	+				ARNOLDY1	DNA/MuDR
HTCHR04REG01DMR0005	4	11372657	11373127	TE	AT4TE52355	CT00027192	11372837	11372917	+				ARNOLDY2	DNA/MuDR
HTCHR04REG01DMR0005	4	11372657	11373127	IGR			11372918	11373863						
HTCHR04REG01DMR0006	4	11373846	11374462	IGR			11372918	11373863						
HTCHR04REG01DMR0006	4	11373846	11374462	GP	AT4G21362		11373864	11375363	+	miRNA	MIR867A	microRNA867A		
HTCHR04REG01DMR0007	4	11395329	11396453	GB	AT4G21390		11394355	11397583	-	protein_coding_gene	B120	B 120		
HTCHR04REG01DMR0008	4	11400435	11401010	GB	AT4G21400		11399131	11401709	-	protein_coding_gene	CRK28	Cysteine-rich RLK (RECEPTOR-like protein kinase) 28		
HTCHR04REG01DMR0009	4	11409790	11410443	GB	AT4G21430		11407804	11412159	-	protein_coding_gene	B160	B160		
HTCHR04REG01DMR0010	4	11497250	11497549	GB	AT4G21640		11496834	11500618	-	protein_coding_gene		Subtilase family protein		
HTCHR04REG01DMR0011	4	11497935	11498234	GB	AT4G21640		11496834	11500618	-	protein_coding_gene		Subtilase family protein		
HTCHR04REG01DMR0012	4	11503018	11503317	GB	AT4G21650		11501198	11504678	-	protein_coding_gene		Subtilase family protein		
HTCHR04REG01DMR0013	4	11510467	11510766	GP	AT4G21670		11509615	11511114	+	protein_coding_gene	CPL1	C-terminal domain phosphatase-like 1		
HTCHR04REG01DMR0013	4	11510467	11510766	GP	AT4G21660		11509843	11511342	-	protein_coding_gene		Proline-rich spliceosome-associated (PSP) family protein		
HTCHR04REG01DMR0014	4	11531886	11532185	GB	AT4G21700		11529956	11532844	+	protein_coding_gene		unknown protein		
HTCHR04REG01DMR0015	4	11533374	11533865	GB	AT4G21705		11532971	11534847	-	protein_coding_gene		unknown protein		
HTCHR04REG01DMR0016	4	11618186	11618485	GB	AT4G21900		11616346	11619350	-	protein_coding_gene	PRORP3	Proteinaceous RNase P 3		
HTCHR04REG01DMR0016	4	11618186	11618485	GP	AT4G21902		11618286	11619785	+	protein_coding_gene		unknown protein		
HTCHR04REG01DMR0017	4	11685712	11686680	IGR			11685645	11686119						
HTCHR04REG01DMR0017	4	11685712	11686680	GP	AT4G22060		11686120	11687619	+	protein_coding_gene		unknown protein		
HTCHR04REG01DMR0017	4	11685712	11686680	TE	AT4TE53985	CT00031613	11686456	11686481	+				ATREP10D	RC/Helitron
HTCHR04REG01DMR0017	4	11685712	11686680	TE	AT4TE53990	CT00028593	11686482	11686531	-				ATREP9	RC/Helitron
HTCHR04REG01DMR0017	4	11685712	11686680	TE	AT4TE53995	CT00012414	11686532	11687098	-				ATTIR16T3A	DNA
HTCHR04REG01DMR0018	4	11820473	11824785	GP	AT4G22410		11820136	11821635	-	protein_coding_gene		Ubiquitin C-terminal hydrolases superfamily protein		
HTCHR04REG01DMR0018	4	11820473	11824785	TE	AT4TE54700	CT00001830	11820547	11823366	-				ATGP3	LTR/Gypsy
HTCHR04REG01DMR0018	4	11820473	11824785	TE	AT4TE54705	CT00002032	11823367	11823389	-				ATGP3	LTR/Gypsy
HTCHR04REG01DMR0018	4	11820473	11824785	IGR			11823390	11825590						

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 Lauss[1], René Wardenaar[2], Marieke H.A. van Hulten[3], Victor Guryev[4], Joost J.B. Keurentjes[3], Maike Stam[1§], Frank Johannes[2,5,6§]

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, 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, frank@johanneslab.org

Contents

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 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 [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*s 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-

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

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

where o_t are the observed leaf area measurements in 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 anlayses 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 ith individual ($i = 1, 2, 3, \ldots, N$). Individual i can belong to either one of the two parental populations ($P1, P2$) or the hybrid offspring population $F1$. We arbitrarily assign $P1 = 1, P2 = 2$ and $F1 = 3$. To keep track of population membership, let $\vec{Z}_i$ be a 3-dimensional component-label vector, where the j element is defined to be one or zero, according to whether the component of orgin of Y_i is equal to j or not ($j = 1, 2, 3$). The $\vec{Z}_j$ is distributed as a multinomial distribution consisting of one draw of 3 categories with probabilities $\lambda_1, \lambda_2, \lambda_3$. Hence,

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

In our case

$$\lambda_1 = \frac{N_{P1}}{N}, \lambda_2 = \frac{N_{P2}}{N}, \lambda_3 = \frac{N_{F1}}{N},$$

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

$$\log L(\vec{\Psi}|\vec{y}) = \sum_{i=1}^{N} \sum_{j=1}^{3} z_{ij} (\log \lambda_j + \log f_j(y_i; \vec{\Omega}_j)).$$

This log likelihood function can be partioned more intuitively according to the contributions of each of the three populations ($P1, P2, F1$):

$$\begin{aligned}
\log L(\vec{\Psi}|\vec{y}) &= N_{P1} \log\left(\frac{N_{P1}}{N}\right) + \sum_{i=1}^{N_{P1}} \log f_{P1}(y_i; \vec{\Omega}_{P1}) \\
&+ N_{P2} \log\left(\frac{N_{P2}}{N}\right) + \sum_{i=1}^{N_{P2}} \log f_{P2}(y_i; \vec{\Omega}_{P2}) \\
&+ N_{F1} \log\left(\frac{N_{F1}}{N}\right) + \sum_{i=1}^{N_{F1}} \log f_{F1}(y_i; \vec{\Omega}_{F1}) \\
&\propto \sum_{i=1}^{N_{P1}} \log f_{P1}(y_i; \vec{\Omega}_{P1}) + \sum_{i=1}^{N_{P2}} \log f_{P2}(y_i; \vec{\Omega}_{P2}) + \sum_{i=1}^{N_{F1}} \log f_{F1}(y_i; \vec{\Omega}_{F1}).
\end{aligned}$$

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.

Model	log lik	df	$\vec{\Omega}_{P1k}$	$\vec{\Omega}_{P2k}$	$\vec{\Omega}_{F1k}$
Full	l_F	6	μ_{P1k}, σ^2_{P1k}	μ_{P2k}, σ^2_{P2k}	μ_{F1k}, σ^2_{F1k}
Additive	l_A	5	μ_{P1k}, σ^2_{P1k}	μ_{P2k}, σ^2_{P2k}	$\mu_{F1k} = \frac{\mu_{P1k}+\mu_{P2k}}{2}$, σ^2_{F1k}

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 χ^2 random variable with degress of freedom equal to the differences in the number of parameters of the full compared to the additive model. In total we performed $19 \cdot 7 = 133$ 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

Model	log lik	df	$\vec{\Omega}_{P1k}$	$\vec{\Omega}_{P2k}$	$\vec{\Omega}_{F1k}$
Full	l_F	6	μ_{P1k}, σ^2_{P1k}	μ_{P2k}, σ^2_{P2k}	μ_{F1k}, σ^2_{F1k}
FPD	l_{FPD}	5	μ_{P1k}, σ^2_{P1k}	μ_{P2k}, σ^2_{P2k}	$\mu_{F1k} = \mu_{Phk}$, σ^2_{F1k}

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

Model	log lik	df	$\vec{\Omega}_{P1k}$	$\vec{\Omega}_{P2k}$	$\vec{\Omega}_{F1k}$
Full	l_F	6	μ_{P1k}, σ^2_{P1k}	μ_{P2k}, σ^2_{P2k}	μ_{F1k}, σ^2_{F1k}
FND	l_{FND}	5	μ_{P1k}, σ^2_{P1k}	μ_{P2k}, σ^2_{P2k}	$\mu_{F1k} = \mu_{Plk}$, σ^2_{F1k}

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 kth 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 teated the z_{ik}s ($k = 1, 2, 3, \cdots, 19; i = 1, 2, \cdots, 30$) as a quantitative phenotype. Assume the value for the ith plant is given by

$$z_i = \mu_0 + \beta_1 C_{i1} + \beta_2 C_{i2} + \cdots + \beta_p C_{ip} + \epsilon_i, \tag{1}$$

where μ_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 partioned as

$$\sigma^2(z) = \sigma^2(C) + \sigma^2(\epsilon), \tag{2}$$

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:

$$R^2 = 1 - \frac{\sigma^2(\epsilon)}{\sigma^2(y)}. \tag{3}$$

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

$$R^2 = H^2 = \frac{\sigma^2(C)}{\sigma^2(y)}, \tag{4}$$

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:

$$\begin{aligned} R^2_{adj} = &= 1 - \frac{\frac{1}{n-(p+1)} s^2(\epsilon)}{\frac{1}{n-1} s^2(y)} \\ &= 1 - \frac{n-1}{n-(p+1)} \frac{\sum_i^n (y_i - \hat{y}_i)^2}{\sum_i^n (y_i - \overline{y})^2} \\ &= 1 - \frac{n-1}{n-(p+1)} \frac{\sum_i^n (y_i - [\hat{\mu}_0 + \sum_{k=1}^p \hat{\beta}_j C_{ik}])^2}{\sum_i^n (y_i - \overline{y})^2} \end{aligned} \quad . \tag{5}$$

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

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

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 line k (z_k) by

$$z_k = \overline{y_k} - mid_k,$$

where $\overline{y}_k$ is the phenotypic mean of the kth epiHybrid population. As show in **Fig. 2A**, the z_ks ($k = 1, 2, 3, ..., 19$) 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 epiHyrids. 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 infomration, the different epiHyrids 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:

$$z_k = \beta_0 + \beta_1 g_{k1} + \beta_2 g_{k2} + \cdots + \beta_k g_{kq} + \epsilon_k, \tag{6}$$

where $\beta_j (j = 1, \ldots, q)$ are the QTL effects, β_0 is the intercept, $g_{kj} (j = 1, \ldots, q; k = 1, \ldots, h)$ are the q epigenotypes measured for epiHybrid line k, and ϵ_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 partioned as:

$$\begin{aligned} \sigma^2(z) &= \sum_{j}^{q} \beta_j^2 \sigma^2(g_j) + 2 \sum_{m<j}^{q} \beta_j \beta_m \sigma(g_m, g_j) + \sigma^2(\epsilon) \\ \sigma^2(z) &= \sigma^2(G) + \sigma^2(\epsilon), \end{aligned} \tag{7}$$

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:

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

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 phenotpic 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

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

and conditional on locus l being UU it is

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

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

$$\begin{aligned} QTL_{F1,l} &= 2\beta_{l:D}N(\beta_{l:A\times A}(p-1) - 2(p(\beta_{l:A\times D} \\ &+ \beta_{l:D\times A} - 2\beta_{l:D\times D} - \beta_{l:A\times D} + \beta_{l:D\times D}))). \end{aligned}$$

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

$$\begin{aligned} QTL^{\bullet}_{F1,l} &= \int_0^1 QTL_{F1,l})dp \\ &= 2\beta_{l:D} - \frac{1}{2}\beta^{\bullet}_{l:A\times A} + 2\beta^{\bullet}_{l:A\times D} - 2\beta^{\bullet}_{l:D\times A}. \end{aligned}$$

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 $\beta^{\bullet}_{l:A\times A}$, $\beta^{\bullet}_{l:A\times D}$ and $\beta^{\bullet}_{l:D\times A}$). Here $A \times A$, $A \times D$ and $A \times D$ refer to additive × additive, additive × dominance and dominance × 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 moblization 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:

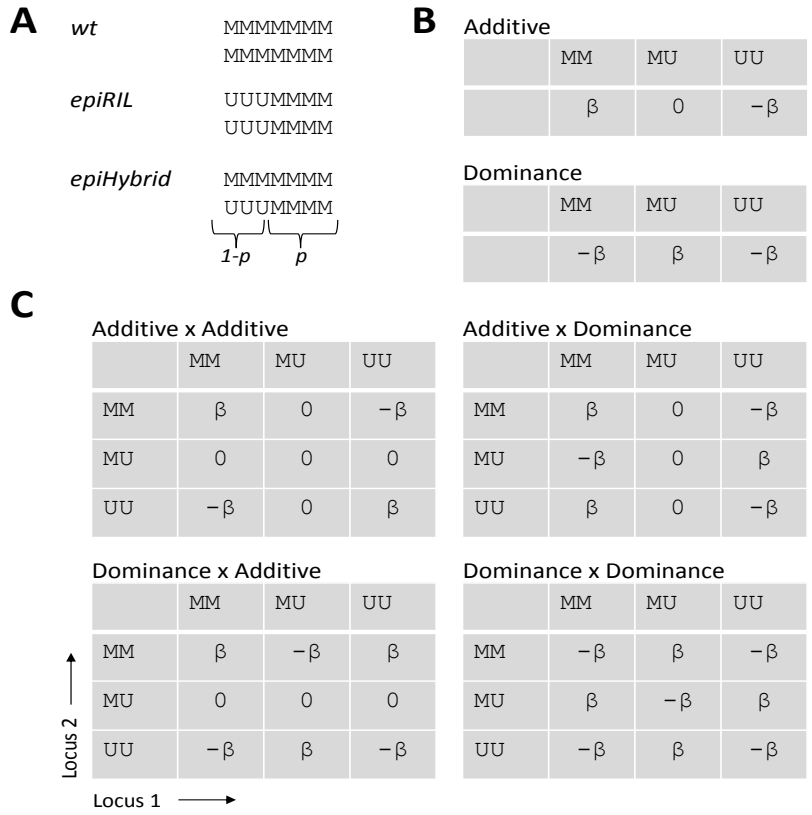


Figure 2.1: (A) We consider N background loci (in this example $N = 7$). In Col-wt all loci are assumed to have epigenotype MM, while in the epiRILs a proportion $1-p$ are assumed to be UU and the remainder (p) are MM. Hence, in the hybrids $1-p$ of the loci are epi-heterozygous MU and the remainder 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 $\rightarrow$ U, M $\rightarrow$ I, I $\rightarrow$ 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 $post(P_i = U)$ and $post(P_i = M)$, respectively. Using this we define the methylation level of probe I as $ML = post(P_i = U) \cdot (-1) + post(P_i = M) \cdot 1$. 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 *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.

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-deafult parameters: –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

[1] Johannes F, Porcher E, Teixeira F, Saliba-Colombani V, Simon M, Agier N, Bulski A, Albuisson J, Heredia F, Bouchez D, Dillmann C, Guerche P, Hospital F, Colot V (2009). Assessing the impact of transgenerational epigenetic variation on complex traits. PLoS Genetics 5: e1000530.

[2] Soppe WJ, et al (2000) Molecular Cell 6, 791–802.

[3] Meyer RC (2004). Plant Physiology 134, 1813–1823.

[4] Colomé-Tatché M, Cortijo S, Wardenaar R, Morgado L, Lahouze B, Sarazin A, Etcheverry M, Martin A, Feng S, Duvernois-Berthet E, Labadie K, Wincker P, Jacobsen SE, Jansen RC, Colot V, Johannes F (2012). Features of the Arabidopsis recombination landscape resulting from the combined loss of sequence variation and DNA methylation. Proc. Natl. Acad. Sci. USA doi:10.1073/pnas.1212955109.

[5] Benjamini Y, and Hochberg Y (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 57, 289–300.

[6] Lander E, Botstein D (1989) Mapping Mendelian Factors Underlying Quantitative Traits Using RFLP Linkage Maps Genetics 121, 185–199.

[7] Broman KW, Wu H, Sen Ś, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. Bioinformatics 19:889-890

[8] Cortijo S, Wardenaar R, M. Colomé-Tatché M, Johannes F, Colot V (2014). Genome-wide analysis of DNA methylation in Arabidopsis using MeDIP-chip. Methods in Molecular Biology 1112:125-49.

[9] Cortijo S, Wardenaar R, Colomé-Tatché M, Gilly A, Etcheverry M, Labadie K, Caillieux E, Hospital F, Aury J-M, Wincker P, Roudier F, Jansen RC, Colot V, Johannes F (2014). Mapping the epigenetic basis of complex traits. Science doi:10.1126/science.1248127.

[10] Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T (2009). Bursts of retrotransposition reproduced in Arabidopsis. Nature 461, 423-426.

[11] Langmead B, Salzberg S (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357–359.

[12] Kloosterman WP, Guryev V, van Roosmalen M, Duran KJ, de Bruijn E, Bakker SCM, Letteboer T, van Nesselrooij B, Hochstenbach R, Poot M, et al. (2011). Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. Hum Mol Genet 20: 1916–24.