

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

2

3 Kathrin Lauss¹, René Wardenaar², Marieke H.A. van Hulten³, Victor Guryev⁴, Joost J.B.

4 Keurentjes³, Maike Stam^{1§}, Frank Johannes^{2,5,6§}

5

6 **1** University of Amsterdam, Swammerdam Institute for Life Sciences, Science Park 904

7 1098XH Amsterdam, The Netherlands.

8 **2** University of Groningen, Groningen Bioinformatics Centre, Faculty of Mathematics and

9 Natural Sciences, Nijenborgh 7, 9747 AG Groningen, The Netherlands.

10 **3** University of Wageningen, Laboratory of Genetics, Droevendaalsesteeg 1, 6708PB

11 Wageningen, The Netherlands.

12 **4** Genome structure aging, European Research Institute for the Biology of Ageing, University

13 Medical Centre Groningen and University of Groningen, Antonius Deusinglaan 1, Building

14 3226, 9713 AV Groningen, The Netherlands

15 **5** *Current address*: Population epigenetics and epigenomics, Department of Plant Sciences,

16 Technical University Munich, Liesel-Beckmann-Str. 2, 85354 Freising, Germany

17 **6** *Current address*: Institute for Advanced Study, Technical University Munich,

18 Lichtenbergstr. 2a, 85748 Garching, Germany

19

20

21 **§** Corresponding co-last author: m.e.stam@uva.nl (MS), frank@johanneslab.org (FJ)

22

23

24 **Abstract**

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

39

40 **Author Summary**

41 Crossing two genetically distinct parents generates hybrid offspring. Sometimes hybrids are
42 performing better than their parents in particular traits and this is referred to as heterosis.
43 Hybridization and heterosis are naturally occurring processes and crop breeders
44 intentionally cross genetically different parental lines in order to generate hybrids with
45 maximized traits such as yield or stress tolerance. So far, the mechanisms behind heterosis
46 are not well understood. In this study we focused on the effect of epigenetic variation onto
47 heterosis in hybrids, and for this purpose we created epigenetic hybrids (epiHybrids) by
48 crossing wildtype plants with a selection of genetically very similar but epigenetically

49 divergent lines. An extensive phenotypic analysis of the epiHybrids and their parental lines
50 showed that epigenetic divergence between parental genomes can be a major determinant
51 of heterosis. Importantly, multiple heterotic phenotypes could be associated with
52 meiotically heritable differentially methylated regions (DMRs) in the parental genomes,
53 allowing us to map epigenetic quantitative trait loci (QTLs) for heterosis. Our results indicate
54 that epigenetic variation can contribute to heterosis and suggests that heritable epigenetic
55 variation could be exploited for the improvement of crop traits.

56

57 **Introduction**

58 Heterosis describes an F1 hybrid phenotype that is superior compared to the phenotype of
59 its parent varieties. The phenomenon has been exploited extensively in agricultural breeding
60 for decades and has improved crop performance tremendously [1,2]. Despite its commercial
61 impact, knowledge of the molecular basis underlying heterosis remains incomplete. Most
62 studies mainly focused on finding genetic explanations, resulting in the classical dominance
63 [1,3,4] and overdominance [4,5] models describing heterosis. In line with genetic
64 explanations it has been observed that interspecies hybrids often show a higher degree of
65 heterosis than intraspecies hybrids, indicating that genetic distance correlates with the
66 extent of heterosis [2,6]. However, genetic explanations do often not sufficiently explain nor
67 predict heterosis. There is growing evidence that also epigenetic divergence plays a role in
68 heterosis [7–9]. It has, for example, been shown that altered epigenetic profiles at genes
69 regulating circadian rhythm play an important role in heterotic *Arabidopsis* hybrids [10].
70 Moreover, heterotic hybrids of *Arabidopsis*, maize and tomato are shown to differ in levels
71 of small regulatory RNAs and/or DNA methylation (5mC) relative to their parental lines [11–
72 14]. Processes such as the transfer of 5mC between alleles (trans chromosomal methylation,
73 TCM), or a loss of 5mC at one of the alleles (trans chromosomal demethylation, TCdM) have
74 been indicated to contribute to the observed remodeling of the epigenome [7,13,15].

75 Strikingly, some of these changes in 5mC levels have been shown to be stable over multiple
76 generations [15,16].

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

82

83 **Results and Discussion**

84 **Construction of epigenetic Hybrids**

85 Hybrids are usually generated from parental lines that vary at both the genomic and
86 epigenomic level and disentangling those two sources of variation is challenging. To
87 overcome this limitation, we generated epigenetic *A. thaliana* F1 hybrids (epiHybrids) from
88 near-isogenic but epigenetically divergent parental lines by crossing Col-0 wildtype (Col-wt)
89 as maternal parent to 19 near-isogenic *ddm1-2*-derived epigenetic recombinant inbred lines
90 (epiRILs) [17] as the paternal parents (Fig 1a).

91 *DDM1 (DECREASE IN DNA METHYLATION 1)* is a nucleosome remodeler and a *ddm1-*
92 *2* deficiency leads to a severe loss of 5mC [18], primarily in long transposable elements and
93 other repeat sequences [19]. EpiRILs carry chromosomes that are a mosaic of Col-wt and
94 hypomethylated *ddm1-2*-derived genomic regions [17,20,21] (Fig 1a). Nineteen epiRIL
95 parental lines were selected that sample a broad range of 5mC divergence from the Col-wt
96 reference methylome (Fig 1b, S1 Table). Besides, lines were chosen that have a wildtype
97 methylation profile at *FWA* (S1 Fig, S1 Table), as loss of DNA methylation at the *FWA*
98 (*FLOWERING LOCUS WAGENINGEN*) locus is known to affect flowering time [22].
99 Furthermore, we selected for a range of phenotypic variation in two traits that have
100 previously been monitored in the epiRILs, flowering time and root length (S1 Table); outliers

101 were excluded [17]. With our experimental design we could demonstrate, as proof-of-
102 principle, the extent to which divergence in 5mC profiles in parental lines can contribute to
103 heterosis.

104

105 **Heterotic phenotypes occur in the epiHybrids**

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

117 The extent of heterosis was evaluated by comparing the hybrid performance with its
118 parental lines. We distinguished five effects (Fig 1c-e): additivity, positive mid-parent
119 heterosis (positive MPH), negative mid-parent heterosis (negative MPH), high-parent
120 heterosis (HPH) and low-parent heterosis (LPH). An additive effect describes a hybrid
121 performance that is equal or close to the average performance of the two parents (the mid-
122 parent value, MPV). MPH refers to deviations in percent from the MPV in positive or
123 negative direction. Hybrids displaying MPH are further tested for HPH and LPH, which
124 describe hybrid performance exceeding the high parent, or falling below the lowest parent,
125 respectively. In crop breeding, the focus is usually on obtaining HPH and LPH as these
126 present novel phenotypes that are outside the parental range. Depending on the trait

127 monitored and commercial application, either HPH or LPH can be considered superior. For
128 instance, early flowering may be preferable over late flowering; in such cases maximizing
129 LPH may be desirable. For other traits, such as yield or biomass, it is more important to
130 maximize HPH. However, in order to obtain a comprehensive view of hybrid performance it
131 is informative to also track MPH in addition to LPH and HPH, because many mature traits
132 may be affected by other traits that do not display fully penetrant heterotic effects.

133 We observed a remarkably wide range of heterotic phenotypes among the
134 epiHybrids (Fig 1g, S2-19 Tables). The magnitude of these phenotypic effects was substantial
135 (Fig 1h-j, S2 Fig, S8-19 Tables) and similar to that typically seen in hybrids of Arabidopsis
136 natural accessions[23,24]. Many epiHybrids (16/19) exhibited significant MPH in at least one
137 of the six monitored traits (FDR = 0.05, Fig 1g). Across all hybrids and traits, we observed 30
138 cases of positive MPH and negative MPH. Among those, four cases show LPH and nine cases
139 show HPH (Fig 1g). Interestingly, in 11 out of the 17 cases of MPH the phenotypic means of
140 the epiHybrids were in the direction of the phenotypic means of the epiRIL parent rather
141 than in the direction of the Col-wt parent (S2-7 Tables, F1 trend). Also all four LPH and two
142 of the HPH cases were in the direction of the epiRIL parent (Fig 1i-j, S2 Fig). This observation
143 illustrates that *ddm1-2*-derived hypomethylated epialleles are often (partially) dominant
144 over wild-type epialleles, which contrasts the situation seen in EMS screens where novel
145 mutations typically act recessively.

146 We observed cases of HPH for LA, HT and MSB, and cases of LPH for FT and MSB.
147 HPH for LA occurred in epiHybrids 232H, 195H and 193H (3/19 epiHybrids). Those epiHybrids
148 significantly exceeded their best parent (Col-wt) by 17%, 18% and 15%, respectively (Fig 1h,
149 S19 Table). Interestingly, although growth rate (GR) is developmentally related to LA, hybrid
150 effects in GR were only moderately, albeit positively, correlated with LA ($\rho = 0.57$, $P =$
151 0.02), which implies that LA heterosis is determined by other traits besides GR.

152 For HT we detected five cases of significant HPH with up to 6% increases in HT (Fig
153 1i, S14 Table). One may expect LA HPH to strongly correlate with HT HPH, as the rosette is
154 providing nutrients for the developing shoot[25]. However, HPH for both LA and HT occurred
155 only in one epiHybrid (193H; Fig 1g).

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

157 Besides positive heterosis, our phenotypic screen revealed strong negative heterotic
158 effects for FT (earlier flowering) and MSB (less main stem branching). Significant LPH
159 occurred in the epiHybrids 232H, 208H and 344H (FT) and 438H (MSB) (Fig 1j, S2 Fig, S15 and
160 S17 Tables). In the most prominent case for FT (232H), FT was about 10% earlier than that of
161 the earliest flowering parent. 208H and 244H flowered 3% and 4% earlier than their lowest
162 parent (epiRIL 208 and epiRIL 344), respectively. 438H showed 14% less MSB than the lowest
163 parent (S2 Fig).

164 The reproducibility of our findings was tested by performing replicate experiments,
165 using seeds from newly performed crosses and the same climate controlled growth chamber
166 as before. We focused on epiHybrids that exhibited relatively strong positive or negative
167 heterotic phenotypes in the initial screen (193H, 150H, 232H; Fig 1g), and measured LA, FT
168 and HT. We found that the direction of the heterotic effects in LA, FT and HT was
169 reproducible in all cases tested (Fig 2a and b). Importantly, the LA and HT HPH observed for
170 193H, and the strong FT LPH for 232H were perfectly reproducible, while LA HPH observed
171 for 232H became positive MPH (Fig 2a). Taken together, these results show that the
172 heterotic effects observed in the epiHybrids are relatively stable for LA, HT and FT, even
173 across fresh parental seed batches and independently performed crosses, which is not
174 always the case for Arabidopsis phenotypes [26].

175

176

177

178 **Heterotic phenotypes are associated with QTLs**

179 To understand the sources of the LA, HT and FT heterotic effects observed among the ~530
180 epiHybrid plants, we calculated the phenotypic divergence of each epiHybrid plant from its
181 respective mid-parent value. Using variance component analysis we estimated that 17%,
182 28% and 51% of the total variation in mid-parent divergence for FT, LA and HT, respectively,
183 can be attributed to (epi)genomic differences between the Col-wt and epiRILs used for the
184 crosses (Fig 2c, S20 Table , SI text). Global 5mC divergence between the Col-wt and the
185 epiRILs parental lines could not account for this variation (S3 Fig). We therefore reasoned
186 that heterotic phenotypes are due to (partial) dominance effects caused by specific regions
187 being epi-heterozygous for an epiRIL-inherited hypomethylated epiallele (*U*) and a Col-wt-
188 inherited methylated epiallele (*M*). To test this possibility, we used the methylomes of Col-
189 wt and the epiRIL parents[20] to predict epi-homozygous (*MM*) and epi-heterozygous (*MU*)
190 regions in the genomes of the epiHybrids (Fig 3a, SI text), and assessed whether heritable
191 epigenetic differences at specific loci could explain the variation in MPH among crosses (S4
192 Fig). The analysis revealed two QTLs on chromosome (chr) 3 contributing to the between-
193 cross variation in MPH in FT (QTL 1: LOD=3.12, 37.62 cM; QTL 2: LOD=3.33, 101.44 cM, Fig
194 3b; S21 Table). EpiHybrids epi-heterozygous (*MU*) at these loci showed significant negative
195 MPH compared to their epi-homozygous (*MM*) counterparts (Fig 3c). While not significant at
196 the genome-wide scale (Fig 3b), the same two QTLs had substantial suggestive effects on LA
197 heterosis in the opposite direction than FT (Fig 3b and c), indicating that both QTLs act
198 pleiotropically.

199 We also detected a single QTL locus on chr 4 (LOD=3.33, 56.00 cM) that contributes
200 to the between-cross variation in MPH for HT (Fig 3b, S21 Table). In this case, *MU* epiHybrids
201 showed significant positive MPH compared to *MM* epiHybrids (Fig 3c). Interestingly, the HT
202 QTL overlaps with a previously identified QTL^{epi} for root length in the epiRILs[21]. The same

203 study identified QTLs^{epi} associated with FT [21] that we did not detect here (Fig 3b), implying
204 that different regions may play a role in FT trait variation than in FT heterosis.

205 **Heterotic phenotypes are associated with DMRs in the parental genomes**

206 The detection of heterosis QTLs for FT, LA and HT provided a rationale to search for causal
207 variants in the QTL confidence intervals. TE-associated structural variants (TEASVs) are
208 known to occur at low frequency in a *ddm1-2*-derived DNA hypomethylated background
209 [17,21,27,28], hence we re-analyzed whole-genome sequencing data from the epiRIL
210 parents [21] for TEASVs but did not detect any that could account for the QTL effects,
211 suggesting that the QTLs most likely have an epigenetic basis (SI text). Indeed, a thorough
212 analysis of the methylomes of the parental epiRILs, using the available MeDIP tiling array
213 data [20], identified 55 and 18 potentially causal differentially methylated regions (DMRs) in
214 the FT, LA and HT QTL regions, mapping to 35 and 14 unique genes, respectively (Fig 3d, S5–
215 S9 Figs, S22-S26 Tables, SI text). Potentially interesting genes in the candidate regions of the
216 FT/LA QTLs (S25 Table) include for example RPL5A, which was shown to affect development
217 through regulating auxin and influencing leaf shape and patterning [29,30], and AT3G26480,
218 a protein that shows partial homology to GTS1, which has been implemented in biomass
219 accumulation [31]. Another potentially interesting candidate is Chup1, which is crucial for
220 chloroplast movement in leaves in response to light [32]. These candidate genes provide
221 excellent targets for follow-up studies.

222

223 **Conclusions**

224 In a recently published study, heterosis for rosette area was reported in an epigenetic F1
225 hybrid generated by crossing a *met1*-derived epiRIL with Col-wt [33]. *DNA-*
226 *METHYLTRANSFERASE1* (*MET1*) is involved in maintenance of DNA methylation at cytosines
227 in CG sequence context and a mutation in this gene causes a severe loss of DNA methylation
228 in the CG and CHH context [34]. Heterosis was observed in a parent-of origin manner; the

229 reciprocal cross did not result in heterosis [33]. This suggests that the heterosis detected
230 may be due to an effect of the maternal cytoplasm rather than differences in epigenetic
231 marks in the parental genomes. Here, we used Col-wt as maternal parent in all crosses to
232 specifically monitor phenotypic effects associated with the epiRIL methylomes. We observed
233 a wide range of heterotic effects, and our proof-of-principle QTL mapping approach
234 indicated that these phenotypic effects are very likely attributable to methylation
235 differences between Col-wt and the epiRILs. Moreover, our results, together with those of
236 Dapp et al. [9], indicate that heterosis in F1 hybrids generated from epigenetically divergent
237 lines may be a more general phenomenon. A more recent study described widespread DNA
238 methylation changes in an epiHybrid derived from Col-wt and a *met1*-mutant [34].
239 Remarkably, the formation of spontaneous non-parental epialleles was observed in the
240 epiHybrid, mostly at pericentromeric transposon sequences, but also at genic loci [34]. This
241 demonstrates that novel epigenetic variation, which is not readily predictable from the
242 parental methylomes, can be created during hybridization. Future research needs to address
243 if and how these methylome changes relate to phenotypic variation. This study also stresses
244 that for a refined understanding of the effect of epigenetic QTLs as described in this study,
245 methylation changes should be thoroughly analyzed.

246

247 **Material and Methods**

248 **Plant Material**

249 The epigenetic recombinant inbred lines (epiRILs) in our study were generated by Johannes
250 et al [17]. The epiRILs were constructed as follows: An *Arabidopsis thaliana* Col-0 line
251 deficient for *ddm1-2* (*DECREASE IN DNA METHYLATION 1*) was crossed to an isogenic Col-0
252 wildtype line (Col-wt) and the resulting F1 was backcrossed as female parent to Col-wt.
253 Subsequently about 500 progeny plants with a wildtype *DDM1* allele were selected and
254 propagated through six more rounds of selfing, generating a population of 500 different

255 epiRILs. We selected 19 different epiRILs as paternal plants for generating epiHybrids (Line
256 IDs: 14, 232, 92, 208, 438, 195, 350, 500, 150, 118, 432, 202, 344, 64, 193, 508, 260, 579,
257 371). Our selection criteria were as follows: 1) Wide range of DNA methylation divergence
258 from Col-wt and among the selected lines; 2) Wildtype DNA methylation state at the FWA
259 locus in order to avoid that differences in DNA methylation at this locus give rise to
260 differences in flowering time [22] in the hybrids; 3) Wide range of phenotypic variation in
261 flowering time and root length among the selected lines. The epiRIL lines were purchased
262 from the Arabidopsis Stock center of INRA Versailles (<http://publiclines.versailles.inra.fr/>).

263

264 **Crosses**

265 To generate F1 hybrids from the selected epiRIL lines and Col-wt, all parental plants were
266 grown in parallel in soil (Jongkind 7 from Jongkind BV, <http://www.jongkind.com/>) in pots
267 (Danish size 40 cell, Desch Plantpak, <http://www.desch-plantpak.com/en/Home.aspx>). The
268 plants were grown at 20°C, 60% humidity, in long day conditions (16h light, 8h dark), and
269 were watered 3 times per week. All crosses were performed in parallel in a time frame of
270 two weeks to avoid phenotypic effects in the F1 progeny due to differences in growing
271 conditions. To exclude that differences in maternal cytoplasm affect the phenotypes of the
272 F1 plants, Col-wt plants were used as a maternal parent and the epiRILs as paternal parents.
273 In parallel, all parental lines, Col-wt and epiRILs, were propagated by manual selfing. This to
274 1) ensure that parental and F1 hybrid seeds were generated under the same growing
275 conditions and 2) exclude potential phenotypic effects derived from hand pollination[35].

276

277 **Phenotypic Screen**

278 The seeds were stratified at 4°C for 3 days on petri-dishes containing filter paper and water
279 before transferring them onto Rockwool/Grodan blocks (soaked in Hyponex NPK: 6.5 – 6.19
280 medium) in a climate controlled chamber (20°C, 70% humidity, long day conditions (16h

281 light, 8h dark)). The transfer of the seeds onto the Rockwool blocks is defined as time point 0
282 days after sowing (DAS). Seeds from each parental and hybrid line were sown in 28
283 replicates and their positions were randomized throughout the growth chamber to level out
284 phenotypic effects caused by plant position. The plants were watered two or three times per
285 week depending on their size. After the plants started flowering, they were transferred to
286 the greenhouse (20°C, 60% humidity, long day conditions (16h light, 8h dark)). In the
287 greenhouse, the plants were watered 3 times per week and stabilized by binding them to
288 wooden sticks at later developmental stages. The plants were harvested once the siliques of
289 the main inflorescence and its side branches were ripe.

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

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

298 Height (HT): HT was scored manually in cm on dried plants. The measurement was
299 taken at the main inflorescence, from the rosette to the highest flowerhead.

300 Branching: Branching was scored on the dried plants by counting the branches
301 emerging from the rosette (RB) and from the main stem (MSB).

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

305

306 **Data analysis**

307 For the data analysis see the Supplementary Information.

308

309 **Replication experiment with selected hybrids**

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

316

317 **Figure Legends**

318 **Fig 1. Heterosis occurs in epiHybrids.** (a) Experimental setup. Lines are depicted
319 schematically as one chromosome with the numbers indicating the epiRIL ID (e.g. 371 & 492)
320 and the respective epiHybrid (e.g. 371H & 492H). (b) Genome-wide 5mC levels (y-axis) of the
321 Col-wt line in green and the epiRIL parental lines in salmon. Numbers indicate the epiRIL IDs.
322 The 5mC levels were calculated as the proportion of methylated MeDIP probes with respect
323 to the total amount of probes. (c-e) Three classes of phenotypic effects monitored in the
324 epiHybrids. The black dashed line indicates the mid-parent value. The green and salmon
325 dashed lines indicate the mean performance of the parental lines. The white dashed lines
326 indicate the mean performance of the epiHybrids. (f) Col-wt, epiHybrid 232H and epiRIL 232
327 at 13 days after sowing as an example for high-parent heterosis. (g) Phenotypic effects in six
328 traits monitored across the 19 epiHybrids. The right panel summarizes positive and negative
329 heterotic effects per trait. (h-j) Examples of epiHybrids exhibiting high-parent heterosis in
330 leaf area and height (LA and HT; h and i), and low-parent heterosis in flowering time (FT; j)
331 Error bars, ± 1 SEM. Deviation from high parent or low parent is shown in percent.

332

333 **Fig 2. Confirmation of mid-parent (MP) divergence in the initial screen and replicate**
334 **experiment for epiHybrids 150H, 193H and 232H. (a)** Results for cases of HPH and LPH for
335 LA, HT and FT in initial experiment. **(b)** Results for traits showing less eminent phenotypic
336 effects for LA, HT and FT. The mid-parent value (MPV) is shown as a dashed horizontal line
337 and the MP divergence is shown as change from MPV in percent. To illustrate the F1
338 epiHybrid distribution for each trait, the individual replicate plants are depicted as dots. **(c)**
339 F1 MP divergence for LA, HT and FT for all epiHybrids. The MPV is shown as a horizontal
340 dashed line and MP divergence is shown as change from MPV in percent. The epiHybrids are
341 ordered from highest (left) to lowest (right) F1 MP divergence. To illustrate the F1 epiHybrid
342 distribution for each trait, the individual replicate plants are depicted as dots. Variance
343 component analysis was used to estimate how much of the total variation in MP divergence
344 can be explained by between-cross variation. The F-statistic from this analysis is shown in
345 the boxes.

346

347 **Fig 3. Interval mapping approach detects significant QTLs for mid-parent divergence. (a)**
348 Genome-wide patterns of Col-wt and *ddm1-2* inherited epi-haplotypes in the (epi)genomes
349 of the parental epiRILs used in this study. **(b)** QTL profiles for FT, HT and LA. Published
350 QTLs^{epi} for root length and flowering time are shown. **(c)** Effect direction of the QTLs. Error
351 bars, ± 1 SE of the Estimate (SEE). **(d)** Zoom in of one of the QTL intervals of FT. The top
352 panel shows the annotations along the genome. The bottom panel shows the locations of
353 candidate DMRs and the average methylation level along the genome for epiRIL parents that
354 are either methylated (*MM*) or unmethylated (*UU*) at the peak marker.

355

356 **Acknowledgements**

357 This work was supported by the Centre for Improving Plant Yield (CIPY)(part of the
358 Netherlands Genomics Initiative and the Netherlands Organization for Scientific Research).

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

363 **Author Contributions**

364 K.L., M.S. and F.J. designed the study, interpreted the data and wrote the manuscript with
365 contributions from J.J.B.K. and R.W. K.L. and M.H.A.v.H. planned and performed the
366 phenotypic screen. F.J. and R.W. performed the data analysis. V.G. analyzed sequencing data
367 of the epiRIL parents.

368 **References**

- 369 1. Schnable PS, Springer NM. Progress toward understanding heterosis in crop plants.
370 *Annu Rev Plant Biol.* 2013;64: 71–88. doi:10.1146/annurev-arplant-042110-103827
- 371 2. Chen ZJ. Molecular mechanisms of polyploidy and hybrid vigor. *Trends Plant Sci.*
372 2010;15: 57–71. doi:10.1016/j.tplants.2009.12.003
- 373 3. Jones DF. Dominance of linked Factors as a means of accounting for Heterosis.
374 *Genetics.* 1917;2: 466 – 479.
- 375 4. Crow JF. Anecdotal , Historical and Critical Commentaries on Genetics 90 Years Ago :
376 The Beginning of Hybrid Maize. *Genetics.* 1998;148: 923–928.
- 377 5. Crow JF. Alternative Hypotheses of Hybrid Vigor. *Genetics.* 1948;33: 477–487.
378 Available:
379 [http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1209419&tool=pmcentr](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1209419&tool=pmcentrez&rendertype=abstract)
380 [ez&rendertype=abstract](http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1209419&tool=pmcentrez&rendertype=abstract)
- 381 6. East EM. Heterosis. *Genetics.* 1936;21: 336 – 397.
- 382 7. Groszmann M, Greaves IK, Fujimoto R, James Peacock W, Dennis ES. The role of
383 epigenetics in hybrid vigour. *Trends Genet.* 2013;29: 684–690. Available:
384 <http://linkinghub.elsevier.com/retrieve/pii/S016895251300125X>
- 385 8. Springer NM. Epigenetics and crop improvement. *Trends Genet.* Elsevier Ltd;
386 2013;29: 241–247. doi:10.1016/j.tig.2012.10.009
- 387 9. Dapp M, Reinders J, Bédiée A, Balsera C, Bucher E, Theiler G, et al. Heterosis and
388 inbreeding depression of epigenetic Arabidopsis hybrids. *Nat Plants.* 2015;1: 1–8.
389 doi:10.1038/nplants.2015.92
- 390 10. Ni Z, Kim E-D, Ha M, Lackey E, Liu J, Zhang Y, et al. Altered circadian rhythms regulate
391 growth vigour in hybrids and allopolyploids. *Nature.* 2009;457: 327–333.
392 doi:10.1038/nature07523
- 393 11. Groszmann M, Greaves IK, Albertyn ZI, Scofield GN, Peacock WJ, Dennis ES. Changes
394 in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to

- 395 hybrid vigor. *Proc Natl Acad Sci U S A.* 2011;108: 2617–2622. Available:
396 <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=2126654>
397 [5&retmode=ref&cmd=prlinks](http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=2126654)
- 398 12. Barber WT, Zhang W, Win H, Varala KK, Dorweiler JE, Hudson ME, et al. Repeat
399 associated small RNAs vary among parents and following hybridization in maize. *Proc*
400 *Natl Acad Sci U S A.* 2012;109: 10444–10449. Available:
401 <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=2268999>
402 [0&retmode=ref&cmd=prlinks](http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=2268999)
- 403 13. Shivaprasad P V, Dunn RM, Santos BA, Bassett A, Baulcombe DC. Extraordinary
404 transgressive phenotypes of hybrid tomato are influenced by epigenetics and small
405 silencing RNAs. *EMBO J.* 2012;31: 257–266. doi:10.1038/emboj.2011.458
- 406 14. Shen H, He H, Li J, Chen W, Wang X, Guo L, et al. Genome-Wide Analysis of DNA
407 Methylation and Gene Expression Changes in Two Arabidopsis Ecotypes and Their
408 Reciprocal Hybrids. *Plant Cell.* 2012;24: 875–892. Available:
409 <http://www.plantcell.org/cgi/doi/10.1105/tpc.111.094870>
- 410 15. Greaves IK, Groszmann M, Ying H, Taylor JM, Peacock WJ, Dennis ES. Trans
411 chromosomal methylation in Arabidopsis hybrids. *Proc Natl Acad Sci U S A.* 2012;109:
412 3570–3575. doi:10.1073/pnas.1201043109
- 413 16. Greaves IK, Groszmann M, Wang A, Peacock WJ, Dennis ES. Inheritance of Trans
414 Chromosomal Methylation patterns from Arabidopsis F1 hybrids. *Proc Natl Acad Sci U*
415 *S A.* 2014;111: 2017–2022. doi:10.1073/pnas.1323656111
- 416 17. Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, et al.
417 Assessing the impact of transgenerational epigenetic variation on complex traits.
418 *PLoS Genet.* 2009;5. doi:10.1371/journal.pgen.1000530
- 419 18. Vongs A, Kakutani T, Martienssen RA, Richards EJ. Arabidopsis thaliana DNA
420 methylation mutants. *Science.* 1993;260: 1926–1928. doi:10.1126/science.8316832
- 421 19. Zemach A, Kim MY, Hsieh P-H, Coleman-Derr D, Eshed-Williams L, Thao K, et al. The
422 Arabidopsis Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access
423 H1-Containing Heterochromatin. *Cell.* 2013;153: 193–205. Available:
424 <http://dx.doi.org/10.1016/j.cell.2013.02.033>
- 425 20. Colomé-Tatché M, Cortijo S, Wardenaar R, Morgado L, Lahouze B, Sarazin A, et al.
426 Features of the Arabidopsis recombination landscape resulting from the combined
427 loss of sequence variation and DNA methylation. *Proc Natl Acad Sci. National Acad*
428 *Sciences.* 2012;109: 16240–16245. Available:
429 <http://www.pnas.org/content/109/40/16240.short>
- 430 21. Cortijo S, Wardenaar R, Colomé-Tatché M, Gilly A, Etcheverry M, Labadie K, et al.
431 Mapping the epigenetic basis of complex traits. *Science.* 2014;343: 1145–1148.
432 doi:10.1126/science.1248127
- 433 22. Soppe WJ, Jacobsen SE, Alonso-Blanco C, Jackson JP, Kakutani T, Koornneef M, et al.
434 The late flowering phenotype of *fwa* mutants is caused by gain-of-function epigenetic
435 alleles of a homeodomain gene. *Mol Cell.* 2000;6: 791–802. Available:
436 <http://www.ncbi.nlm.nih.gov/pubmed/11090618>
- 437 23. Groszmann M, Gonzalez-Bayon R, Greaves IK, Wang L, Huen AK, Peacock WJ, et al.
438 Intraspecific Arabidopsis hybrids show different patterns of heterosis despite the
439 close relatedness of the parental genomes. *Plant Physiol.* 2014;166: 265–280.

- 440 doi:10.1104/pp.114.243998
- 441 24. Wang L, Greaves IK, Groszmann M, Wu LM, Dennis ES, Peacock WJ. Hybrid mimics
442 and hybrid vigor in Arabidopsis. *Proc Natl Acad Sci.* 2015;112: E4959–E4967.
443 doi:10.1073/pnas.1514190112
- 444 25. Bennett E, Roberts J a., Wagstaff C. Manipulating resource allocation in plants. *J Exp*
445 *Bot.* 2012;63: 3391–3400. doi:10.1093/jxb/err442
- 446 26. Massonnet C, Vile D, Fabre J, Hannah M a., Caldana C, Lisee J, et al. Probing the
447 Reproducibility of Leaf Growth and Molecular Phenotypes: A Comparison of Three
448 Arabidopsis Accessions Cultivated in Ten Laboratories. *Plant Physiol.* 2010;152: 2142–
449 2157. doi:10.1104/pp.109.148338
- 450 27. Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. Bursts of
451 retrotransposition reproduced in Arabidopsis. *Nature*; 2009;461: 423–427.
452 doi:10.1038/nature08351
- 453 28. Kooke R, Johannes F, Wardenaar R, Becker F, Etcheverry M, Colot V, et al. Epigenetic
454 Basis of Morphological Variation and Phenotypic Plasticity in Arabidopsis thaliana.
455 *Plant Cell.* 2015;27: 337 – 348. doi:10.1105/tpc.114.133025
- 456 29. Rosado A, Li R, van de Ven W, Hsu E, Raikhel N V. Arabidopsis ribosomal proteins
457 control developmental programs through translational regulation of auxin response
458 factors. *Proc Natl Acad Sci U S A.* 2012;109: 19537–19544.
459 doi:10.1073/pnas.1214774109
- 460 30. Pinon V, Etchells JP, Rossignol P, Collier S a, Arroyo JM, Martienssen R a, et al. Three
461 PIGGYBACK genes that specifically influence leaf patterning encode ribosomal
462 proteins. *Development.* 2008;135: 1315–1324. doi:10.1242/dev.016469
- 463 31. Gachomo EW, Jimenez-Lopez JC, Baptiste LJ, Kotchoni SO. GIGANTUS1 (GTS1), a
464 member of Transducin/WD40 protein superfamily, controls seed germination,
465 growth and biomass accumulation through ribosome-biogenesis protein interactions
466 in Arabidopsis thaliana. *BMC Plant Biol.*; 2014;14: 1 – 17. doi:10.1186/1471-2229-14-
467 37
- 468 32. Oikawa K, Kasahara M, Kiyosue T, Kagawa T, Suetsugu N. CHLOROPLAST UNUSUAL
469 POSITIONING1 Is Essential for proper Chloroplast positioning. *The Plant Cell*, 2003;15:
470 2805–2815. doi:10.1105/tpc.016428.tions.
- 471 33. Dapp M, Reinders J, Bédiée A, Balsera C, Bucher E, Theiler G, et al. Heterosis and
472 inbreeding depression of epigenetic Arabidopsis hybrids. *Nat Plants.* 2015;1: 1-8.
473 doi:10.1038/nplants.2015.92
- 474 34. Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, et al. Non-CG methylation patterns
475 shape the epigenetic landscape in Arabidopsis. *Nat Struct Mol Biol.* 2014;21: 64–72.
476 doi:10.1038/nsmb.2735
- 477 35. Meyer RC. Heterosis of Biomass Production in Arabidopsis. Establishment during
478 Early Development. *PLANT Physiol.* 2004;134: 1813–1823. Available:
479 <http://www.plantphysiol.org/cgi/doi/10.1104/pp.103.033001>

480



